APPLICATION OF ARTIFICIAL NEURAL NETWORK MODELING IN THERMAL PROCESS CALCULATIONS OF CANNED FOODS

by

MAHTAB KHODAVERDI AFAGHI

Department of Food Science and Agricultural Chemistry
Macdonald Campus of McGill University
Montreal, Canada

February, 2000

A Thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements of the degree of Masters of Science

Suggested short title:

ANN modeling of thermal process calculation methods

ABSTRACT

The feasibility of using Artificial Neural Network (ANN) models for application in thermal process calculations was studied. As a preliminary study, ANN models were developed based on tabulated data for Ball and Stumbo methods of process calculations. The ANN models for Ball method related *g-value*, a measure of process time, and f_h/U, a measure of process lethality. Optimizing training data set size, number of hidden layers and PEs in each hidden layer as well as learning parameters is important in obtaining an efficient ANN model. Development of ANN model for Stumbo method followed the same procedures as the Ball's method, except that j_{cc} value (cooling lag factor) was included as an additional input. The developed ANN models for Ball and Stumbo methods were validated using a set of processing conditions, resulting in a new set of g and f_h/U values. A range of retort temperatures (RT), initial temperatures (IT), heating rate indexes (f_h), heating lag factors (j_{ch}) and cooling lag factors (j_{cc}) were used to calculate the related process time and process lethality. The developed ANN models were recalled with the new set of parameters. The relative prediction errors of the ANN models were 1% and 3% for Ball and Stumbo method ANN models, respectively. The higher error of ANN models of Stumbo method could be possibly due to the smaller number of training data and wider range of parameters in tables of this method than the tables in Ball method. In general, the ANN models were able to simulate the Ball and Stumbo methods of process calculations reasonably well.

For a better understanding of the effect of process parameters on the evaluation of thermal process, the accuracy of several formula methods (Steele & Board, Ball, Stumbo and Pham) were studied over a wide range of commercial conditions. A computer simulation based on finite difference method of numerical solutions of heat transfer to packaged foods in cylindrical containers was applied to obtain the time-temperature data for designed conditions (retort and initial temperatures, thermal diffusivity, package sizes and processing time). Moreover, the process time and process lethality from this simulation were used as the reference values for the purpose of comparison. The accuracy of methods was evaluated based on the variation of each parameter over the range of conditions employed in the study. Retort temperature had the most significant effect on

calculated process deviations, and Stumbo and Pham methods had the best performance. In a more specific evaluation, the comparison of the methods was carried out based on the can dimensions and *g-value*. Higher *g-values* resulted in higher errors and *H/D* near unity had the highest relative error.

As the final goal of the study, a multi-layer ANN model was developed as an alternative to thermal process calculations. In developing this model, the time-temperature data from the finite difference simulation was used to compute the process lethality, process time as well as heat penetration parameters: f_h, j_{ch}, f_c and j_{cc}, which were needed for training and testing of the models. The ANN predicted process time or process lethality was compared to respective values from finite difference model. The performance of ANN models was also compared to the different formula methods (Ball, Stumbo and Pham). ANN model was able to predict the process times with a mean average error of 2 minutes, which was comparable to Pham method. The mean prediction error of process lethality was 2.74%, which was comparable to Stumbo method.

RÉSUMÉ

L'applicabilité des réseaux d'intelligence artificielle (RIA) aux calculs impliqués dans les traitements thermiques a été étudiée. Les travaux préliminaires consistaient à développer des modèles RIA basés sur les données des méthodes de calcul du traitement de Ball et Stumbo. Le RIA basé sur la méthode de Ball tentait d'établir la relation entre la *valeur-g*, mesurant le temps de traitement, et f_h/U, mesurant le point d'asepsie du traitement. L'optimisation de la taille des données d'entraînement du réseau, le nombre de couches cachées, les noeuds dans chaques ces couches ainsi que les paramètres d'apprentissage était cruciale pour l'obtention d'un modèle efficace. Pour le RIA basé sur la méthode de Stumbo, le cheminement était similaire à celui utilisant la méthode de Ball, mais celle-ci intégrait aussi la facteur de retardement du refroidissement, j_{cc}. Les modèles RIA développés pour les méthodes Ball et Stumbo ont été validés en utilisant une série de valeurs de g et f_h/U obtenue en variant la température du milieu (TM), la température initiale (TI), l'index de vitesse d'échauffement (f_h), et les facteurs de retardement de l'échauffement (j_{ch}) et du refroidissement (j_{cc}). L'erreur relative de prédiction était de 1% pour la méthode basée sur l'approche de Ball et 3% pour celle de Stumbo. La valeur relativement élevée de l'erreur pour la méthode RIA de Stumbo peut être attribuée au nombre restreint de données d'entraînement par rapport au nombre de paramètre. En général, les modèles RIA ont simulé de façon acceptable les méthodes de Ball et Stumbo pour le calcul du traitement.

Pour mieux comprendre l'effet des paramètres de traitement sur l'évaluation du processus thermique, l'exactitude de plusieurs formules (Steele & Board, Ball, Stumbo and Pham) a été étudiée pour une variété de conditions industriellement utilisées. Une simulation basée sur la méthode de différence finie à solution numérique du transfert de chaleur dans un produit emballé dans un contenant cylindrique a été appliquée pour obtenir les données de temps et de température reliées aux conditions d'étude sélectionnées (la température du milieu, la température initiale, diffusion thermique, taille de l'emballage, et longueur du traitement). De plus, la longueur de traitement et le point d'asepsie du traitement ainsi prédits ont fait l'objet de comparaison pour évaluer l'effet de chaque paramètre. La température du milieu affectait de façon significative l'erreur sur la

prédiction. Les méthodes de Stumbo et Pham ont produit les meilleurs résultats. Une étude plus poussée démontre qu'une augmentation de la *valeur-g* entraîne une erreur plus élevée, alors qu'un rapprochement du paramètre *H/D*, relié à la dimension de l'emballage, à l'unité se traduit par une erreur relative supérieure.

Pour but final de cette étude, un modèle RIA multi-couches a été développé. Cette alternative aux calculs du processus thermique, utilise les données de temps et de températures obtenues par la simulation basée sur la méthode de différence finie pour prédire le point d'asepsie du traitement, la longueur de traitement ainsi que les paramètres de pénétrations de chaleurs : f_h, j_{ch}, f_c and j_{cc}. Ces valeurs sont nécessaires pour entrainer et tester les modèles. Les valeurs prédites par les RIA ont été comparées aux valeurs obtenues par les calculs basés sur la différence finie ainsi qu'un nombre de formules acceptées (Ball, Stumbo and Pham). Les modèles RIA développés ont pu prédire le temps de traitement avec une erreur moyenne de 2 mins, une valeur comparable à celle obtenue avec la méthode de Pham. L'erreur moyenne de prédiction pour le point d'asepsie du traitement était de 2.74%, valeur comparable à l'approche Stumbo.

ACKNOWLEDGEMENTS

I would like to extend my sincere appreciation and thanks to my thesis supervisor, Dr. Hosahalli Ramaswamy for his guidance, valuable advice, encouragement, keen interest and support throughout the course of the study.

Also I also thank Dr. I. Alli, Chairman of the Department of Food Science and Agricultural Chemistry and Dr. F.R.van de Voort for their great advice at the start of the study. Many thanks are also extended to Dr. S.O. Prasher, research committee member, for his helpful suggestions on some aspects of the project.

I thank Dr. S. Sreekanth for his very valuable help and advice in starting the project. Dr. Alain LeBail's technical assistance for modification of the finite difference program of Dr. S. Sablani (Ph.D. Thesis, 1995) for adoption to this work is gratefully acknowledged. Many thanks to Dr. Farhad Salehi for his precious insight and guidance throughout the study. Also, I would like to thank Ms. Aline Dimitri for the French translation of the abstract.

My sincere thanks to all my colleagues in the Food Science Department, especially the Food Processing group: Dr. Tatiana Koutchma, Dr. Reza Zareifard, Dr. Michelle Marcotte, Dr. Dina Mussa, Dr. Sasithorn Tajchakavit, Ms. Sarmista Basak, Mr. Pramod Pandey, Ms. Farideh Nourian, Mr. Harish Krishnamurthy, Mr. Curien Chen, Ms. Lawrende Chiu, Mr. Olivier Turpin and Ms. Nada Houjaij.

Above all, I thank my wonderful parents, my dearest brothers, Keyhan and Soheyl for their encouragement, love and support.

NOMENCLATURE

a	Significant dimension, the radius of cylinder (m)
D	Decimal reduction time (min)
f_c	Cooling rate index (min)
f_h	Heating rate index (min)
F_o	Process lethality (min)
g	Temperature difference between product and heating medium (retort) at end of heating (°C)
h	Vertical position in cylinder (m)
h_f	Fluid heat transfer coefficient ($W/m^2.K$)
IT	Initial Temperature (°C)
j_{cc}	Cooling lag factor
j_{ch}	Heating lag factor
J_0	Bessel function of order zero
k_c	Cylinder thermal diffusivity (W/m.K)
L	Lethal rate: Thickness or half thickness of a slab depending on it being heated or cooled from one side or both sides, respectively
n	Number of records
r	Radial position in cylinder (m)
RT	Retort Temperature (°C)
T	Temperature (°C)
t	Time (min)
t_c	Cooling time (min)
T_{cw}	Cooling water temperature (°C)
t_g	Heating time (min)
T_g	Product temperature at the end of heating (°C)
T_o	Reference temperature (°C)
U	Dimensionless temperature ratio = (T-RT)/(IT-TR)
U	The equivalent of all lethal heat received by some designed point in the container during process at the retort temperature (min)
X	Distance from the coldest plane of a slab
Y	ANN predicted value
Y_o	Desired value
z	Temperature sensitivity indicator (C°)
Δ	Step size (used with space or time)

Subscripts

1,2	Refer to two levels with respect to temperature
i	Initial condition, space index
IT	Initial condition of product
j	Space index
ref	Reference
RT	(Retort) heating medium

Greek symbols

α	Thermal diffusivity (m^2/s)
β	Root of the characteristic equation (4.3)
γ	Root of characteristic equation (4.4)

Dimension less numbers

Bi	Biot number
Fo	Fourier number

Abbreviations

ANN	Artificial neural network
CER	Carbon dioxide evolution rate
FTIR	Fourier transform infrared
LDS	Linear discriminate analysis
MAE	Mean absolute error
MIMO	Multiple input and multiple output
MLNN	Multiple layer neural network
MRE	Mean relative error
NIR	Near infra red
NUR	Nitrogen utilization rate
OCR	Oxygen consumption rate
PCA	Principle component analysis
PE	Processing element
RMS	Root mean square

TABLE OF CONTENTS

ABSTRACT i
RÉSUMÉ iii
ACKNOWLEDGMENTS v
NOMENCLATURE vi
LIST OF TABLES x
LIST OF FIGURES xii
CONNECTING STATEMENTS xiv
CONTRIBUTION OF AUTHORS xv

CHAPTER 1 INTRODUCTION 1

CHAPTER 2 LITERATURE REVIEW 5

Principles of thermal processing 5
Kinetics of microbial destruction 6
Thermal process calculations 9
Thermal process calculation methods 10
General method 10
Improved General method 11
Formula methods 12
Ball's method 12
Stumbo's method 17
Hayakawa's method 17
Steele and Board's method 18
Pham's method 18
Evaluation of Formula methods 19
Numerical models of thermal process calculations 21
Artificial neural networks 24
Development of an ANN model 26
Applications of ANN in Food Science and Agriculture 29

CHAPTER 3 ANN SIMULATION OF BALL AND STUMBO FORMULA METHODS OF THERMAL PROCESS CALCULATIONS 35

ABSTRACT 35
INTRODUCTION 36
Artificial neural network models 37
MATERIAL AND METHODS 39
ANN modeling of Ball and Stumbo methods 39
Development of ANN models 39
Ball model 39
Stumbo model 41
Validating the ANN models 42
RESULTS AND DISCUSSION 43
ANN based Ball model 43
Model development and optimization 43
Model verification 49
ANN based Stumbo model 49
CONCLUSIONS 54

CHAPTER 4 COMPARISON OF FORMULA METHODS OF THERMAL PROCESS CALCULATIONS FOR PACKAGED FOODS IN CYLINDRICAL CONTAINERS 55

ABSTRACT 55
INTRODUCTION 56
Brief review of formula methods 57
Ball's method 57
Stumbo's method 58
Steele and Board's method 58
Pham's method 58
MATERIAL AND METHODS 59
Finite difference program 59
The basis of finite difference models 59
Overall accuracy evaluation of the formula methods 64
Evaluation of formula methods under specific conditions 65
RESULTS AND DISCUSSIONS 65
Optimization of finite difference model 65
Overall evaluation of formula methods 67
Specific evaluation of formula methods 73

CONCLUSIONS 79

CHAPTER 5 ARTIFICIAL NEURAL NETWORK MODELS AS ALTERNATIVES TO THERMAL PROCESS CALCULATION METHODS 80

ABSTRACT 80
INTRODUCTION 81
MATERIAL AND METHODS 82
Data generation 82
ANN model development 84
Optimization of ANN models 86
Comparison of ANN models with formula methods 87
RESULTS AND DISCUSSION 87
ANN models optimization 87
Performance of the ANN models 89
Comparison of ANN models 92
CONCLUSIONS 99

CHAPTER 6 GENERAL CONCLUSIONS 100

REFERENCES 102

LIST OF TABLES

Table 3.1	Combinations of training and testing data for network optimization	41
Table 3.2	Range of network architecture and learning parameters used in the study	41
Table 3.3	Range of parameters for validating ANN based Ball and Stumbo models	43
Table 3.4	Effect of PEs and hidden layer(s) on the performance of ANN based Ball models	46
Table 3.5	Effect of epoch size on prediction ability: ANN based Ball models	46
Table 3.6	Effect of momentum on prediction ability of ANN based Ball Models	47
Table 3.7	Optimal levels of architectural parameters for ANN based Ball and Stumbo models	47
Table 3.8	Error parameters for ANN based Ball and Stumbo models	49
Table 4.1	Errors in process time calculation influenced by processing conditions and thermal diffusivity of product	69
Table 4.2	Errors in process lethality calculation influenced by processing conditions and thermal diffusivity of product	69
Table 4.3	Means and Standard Deviations (SD) of relative errors in process time calculations for selected process calculation methods	71
Table 4.4	Means and Standard Deviations (SD) of relative errors in process lethality calculations for selected process calculation methods	72
Table 4.5	Maximum. Minimum. Mean and Standard Deviations of errors for different methods	72
Table 5.1	Range of parameters used in finite difference program	83
Table 5.2	Combinations of training and testing data for network optimization	86
Table 5.3	Range of network architecture and learning parameters used in the study	87
Table 5.4	Effect of momentum and epoch size on the performance of Model A and B	89

Table 5.5	Error parameters for performance and verification of ANN models A and B	92
Table 5.6	Means and Standard Deviations (SD) of relative errors in process time calculations for selected process calculation methods	95
Table 5.7	Means and Standard Deviations (SD) of relative errors in process lethality calculations for selected process calculation methods	96
Table 5.8	Range. mean and standard deviation of errors for all the combination of parameters	96

LIST OF FIGURES

Figure 2.1	Typical semi logarithmic heating and cooling curves	13
Figure 2.2	Grid format for temperature calculation using a numerical method	23
Figure 2.3	Basic structure of a biological neuron	25
Figure 2.4	Typical diagram of ANN model	25
Figure 2.5	Diagram of a processing element and related calculations	27
Figure 3.1	Learning curve for ANN model A (prediction *log g*)	45
Figure 3.2	Learning curve for ANN model B (prediction f_h/U)	45
Figure 3.3	Comparison of ANN predicted values and Ball table parameters respect to input variable	48
Figure 3.4	Validation of ANN models of Ball method	50
Figure 3.5	Comparison of ANN predicted values and Stumbo table parameters respect to input variable	52
Figure 3.6	Validation of ANN models of Stumbo's method	53
Figure 4.1	Optimization of finite difference model for a) number of nodes in each axis and b) time step	66
Figure 4.2	Validation of finite difference model against the analytical solution	68
Figure 4.3	Errors in process time calculations for different g-values (a: g=0.05 C°, b: g=1 C°; c: g=5 C°; d: g=10 C°; e: g=15 C°)	74
Figure 4.4	Errors in process lethality calculations for different g-values (a: g=0.05 C°, b: g=1 C°; c: g=5 C°; d: g=10 C°; e: g=15 C°)	75
Figure 4.5	Errors for process lethality calculations with respect to different g-values (a: Steele & Board's method, b: Ball's method; c: Stumbo's method)	77
Figure 4.6	Errors for process time calculations with respect to different g-values (a: Steele & Board's method, b: Ball's method; c: Stumbo's method; d: Pham's method)	78
Figure 5.1	Arrangement and procedure of developing ANN models	85

Figure 5.2	Learning curve for ANN model A	88
Figure 5.3	Learning curve for ANN model B	88
Figure 5.4	Effect of PEs and hidden layers on the performance of model A and B	90
Figure 5.5	Performance of ANN models A and B with respect to predicted output	91
Figure 5.6	Performance of ANN models A and B in process time and process lethality prediction	93
Figure 5.7	Performance of ANN model in process time calculation in comparison with the other selected Formula methods	97
Figure 5.8	Performance of ANN model in process lethality calculation in comparison with the other selected Formula methods	98

CONNECTING STATEMENTS

This thesis research is divided into three parts. The first part is a feasibility study demonstrating the potential of ANN modeling in simulating two simple (and most popular) process calculation methods (Ball and Stumbo methods). In this part, input for the ANN models were obtained from tables developed by Ball and Stumbo, which largely simplifies the situation for generating input data. The development of such simplified ANN models is described in Chapter 3, and is part of the manuscript number 4 listed on next page.

The accuracy of existing formula methods against the data predicted by a finite difference computer model under a range of commercial operating conditions were evaluated in the next chapter as a prelude to developing a more general ANN model, which is the final objective of the study. This aspect is detailed in Chapter 4 demonstrating the discrepancy of some process calculation methods in predicting accurate values under certain circumstances. This forms the basis for another publication (number 5).

The final Chapter 5 is the principal focus of this study. However, the basis for the developed final ANN models came from Chapter 3 (for optimizing the ANN performance) and Chapter 4 (for generating the input data using the finite difference model). The different methods are ultimately compared in Chapter 5, demonstrating the utility of ANN models. This aspect forms the basis of the final manuscript (number 6). The details of the manuscripts/presentations arising from the thesis are highlighted in the next section with an explanation to the role of co-authors.

CONTRIBUTION OF AUTHORS

The following papers have been prepared for different presentations and/or publications:

1. Afaghi, M. and Ramaswamy, H.S. 1997. A neural network approach for thermal process calculations. Presented at the CIFST Annual Conference. Radisson Hotel. Montreal. QC. September 22-25.

2. Afaghi, M and Ramaswamy, H.S. 1999. Artificial neural network models as alternatives to formula methods of thermal process calculations. Presented at the CIFST Annual Conference. Grand Okanagan Hotel. Kelowna. BC. June 6-9.

3. Afaghi, M. and Ramaswamy, H.S. 1999. Comparison of formula methods of thermal process calculations for foods in cylindrical containers. Presented at the Annual Meeting of the Institute of Food Technologists (IFT). McCormick Place South. Chicago. IL USA. July 24-28.

4. Afaghi, M. and Ramaswamy, H.S. 1999. ANN simulation of Ball and Stumbo formula methods of thermal process calculations. (Submitted. Food Research International).

5. Afaghi, M. and Ramaswamy, H.S. 1999. Comparison of formula methods of thermal process calculations for cylindrical packaged foods. (To be submitted).

6. Afaghi, M. Ramaswamy, H.S. 1999. Artificial neural network models as alternative of thermal process calculation methods. (To be submitted)

In the above papers. all the experimental (modeling) work. analysis of results and preparation of manuscripts were done by the candidate (M. Afaghi). under the supervision of Dr. H.S. Ramaswamy.

CHAPTER 1

INTRODUCTION

Thermal processing is one of the most important methods of food preservation of the twentieth century. Since the innovation of this method by Nicholos Appert in 1810, thermal processing of packaged foods has been improved extensively in all related aspects of the process. Design of an effective process requires a sound knowledge of the destruction kinetics of the concerned microorganism and temperature history of the product. Process calculation methods are commonly designed to compute the required processing time for a target sterilization value or to evaluate the sterilization value for a given process. Accurate process calculation methods are required with respect to both safety and quality consideration of product.

Bigelow *et al.* (1920) established the first graphical procedure of thermal process determination, referred as General method. Ball (1923) introduced a mathematical method, known as the first Formula method. Ball's method has broadly served the food industry despite some of its sweeping assumptions, which results in some inaccuracies. Ball's formula method is based on the equation for the straight-line portion of the semi-logarithmic heating curve at the can center. To calculate the process time for defined process lethality or to calculate the lethality of a given process, Ball developed some tables and graphs. Stumbo's method (1966), developed as the revised version of Ball's method, results in more accurate process calculations (Smith and Tung, 1982). However, the procedures of process calculation in these two methods are quite similar and the accuracy of each method depends on the accuracy of the evaluated parameters and addition of correct cooling lethality to process calculation through the tables and graphs. The application of these tables and graphs can be time consuming and may become a source of error in process calculations. In addition to these two methods, several other methods exist in literature as alternatives to the existing process calculation methods developed through modifying procedures (Hayakawa, 1970; Pham, 1987, 1990; Vinters, 1975; Steele and Board, 1979a,b).

Smith and Tung (1982) evaluated the accuracy of some formula methods as compared to a numerical method for a range of processing conditions and can sizes, and pointed out differences in their performances. Studies have been also carried out to check the accuracy of formula methods in relation to a computer simulation of thermal processing using finite difference numerical solutions for thin profile packaged foods (Ghazala *et al.*, 1990). Stoforos *et al.* (1997) reviewed in detail the basis of different methods of process calculations and the accuracy of these methods as compared to a finite difference simulation of heat transfer for one set of processing conditions and can size.

Advent of computers and ease of programming has provided the potential application of mathematical models in process design, validation, control, and optimization (Hayakawa, 1970, 1978; Manson *et al.*, 1970; Stumbo, 1973; Teixeira *et al.*, 1969a, b; Teixeira, 1978). These models are mainly used in temperature profile prediction of product undergoing thermal processes, as they provide a more versatile alternative to time consuming and expensive heat penetration tests. In addition to time-temperature prediction, with these models the retort temperature need not be held constant and can be varied in any prescribed manner throughout the process. The rapid evaluation of an unscheduled process deviation is another important application of these models (Heldman and Lund, 1992).

The application of artificial neural network (ANN) methods has been growing in the several areas of food technology and agriculture. ANN is a powerful technique for correlating data using a number of processing elements. Using the ANN technique, the computer learns to make intelligent decisions using known input-output data and adjusting some internal parameters of the network through repetitive introduction of known examples. The strength of ANNs is in their ability to handle complex nonlinear relationships with ease and without any prior knowledge of their relationships. The ANN has potential advantages of adaptation and learning ability, fault tolerance of noisy or incomplete data, and high computational speed. Eerikainen *et al.* (1993) introduced the early application of neural networks in food related subjects. ANN has shown a promising application in extrusion process control (Eerikainen *et al.*, 1994). As an alternative to statistical models in data analysis of FTIR, GS-MS and sensory evaluation, ANN models had a higher performance (Bochereau *et al.*, 1992; Tomlins and Gay, 1994; Vallejo-

Cordoba *et al.*, 1995). Sablani *et al.* (1995) investigated the potential of ANN models for prediction of optimal sterilization temperatures.

One typical type of ANN structure in known as back-propagation networks, which has shown promising results in prediction modeling and classification (Sreekanth *et al.*, 1998; Lacroix, *et al.*, 1997; Bochereau *et al.*, 1992; Freeman, 1993). A back-propagation network consists of a sequence of layers with full connection between the layers. Three required layers in these networks are input layer, hidden layer(s) and output layer. Input layer transfers the input information to the network to be processed by hidden layer(s). The processed information is passed to an external source through the output layer. During training, the internal parameters are adjusted to produce the possible closest ANN output to the desired output. The adequacy of a trained network depends on the nature and size of the training data set as well as selecting the optimal internal parameters. In other words, the performance of the ANN model greatly depends on the training data with respect to both quantity and quality (Swingler, 1996). Once trained, the ANN model presents rapid answers to any input variable in the domain of training data set. If the conditions change in such a way that deprives performance of the network, the ANN model can be trained further under the new conditions to correct its performance (Baughman and Liu, 1995). Considering these abilities, ANN models render themselves as a possible alternative to mathematical models and regression techniques (Ni and Gunasekaran, 1998; Tomilins and Gay, 1994).

During the sterilization process, there are several parameters, which affect the accuracy and efficiency of heating process evaluation. The most relevant factors in evaluation of a heat treatment are type and heat resistance of microorganisms, pH of food product, heating conditions, thermo-physical properties of food and package size and type (Valentas *et al.*, 1991). An accurate thermal process calculation model takes into account the significance of producing a high quality product while ensuring the minimum required quality.

The following were the objectives of this research:

1. Development of ANN models based on input data from Ball and Stumbo's tables as a preliminary study to evaluate the feasibility of ANN models in thermal process calculations.
2. Evaluating the accuracy of different formula methods over a wide range of processing conditions and can sizes against a reference computer simulation model based on numerical solution of partial differential equations related to heat conduction equation involving finite cylinders.
3. Development of an ANN model using the data obtained from the finite difference simulation under a wide range of conditions and comparison of the performance of ANN model with traditional formula methods.

Successful thermal processing encompasses the assurance of safety and high quality of the food products. Several methods developed to establish thermal processes have been improved with respect to innovation and application of computers in food technology. In this study, artificial neural network technique is being evaluated to develop a possible more versatile thermal process calculation method. Due to the fact that ANN technique has the ability in modeling non-linear systems, it is hoped that ANN based thermal process calculation model will be capable of pursuing the objectives of an accurate thermal process calculation model.

CHAPTER 2

LITERATURE REVIEW

Principles of thermal processing

Thermal processing of packaged foods basically involves heating of food products for a selected time at a selected temperature to destroy pathogenic microorganisms endangering public health as well as those microorganisms and enzymes that deteriorate the food during storage. For the first time, Nicolas Appert, in 1810, introduced the concept of in-container thermal processes. Extensive emphasis has been given to improvements in thermal processing, as this is one of the most important methods of food preservation. Although the heat labile nature of microorganisms is the basis of thermally preservation of food products, yet the same but undesirable effect can destroy part of nutrients and quality factors. Therefore, an efficient thermal processing requires to be accurately designed to ensure both the safety and quality of food products.

An effective thermal processing is defined based on the definition of commercial sterilization. Commercial sterilization of food product inhibits the growth of both microorganisms and their spores under normal storage conditions in the container. A commercially sterile food product may contain viable spores, such as thermophilic spores, which will not develop under normal storage conditions. The US Food and Drug Administration in 1977 defined the concept of "minimal thermal process" as "the application of heat to food, either before or after sealing in a hermetically sealed container, for a period of time and at a temperature, scientifically determined to be adequate to ensure the destruction of microorganisms of public health concern."

Several important factors determine the extent of thermal processing, such as:

1) Type and heat resistance of the target microorganism, spore or enzyme

2) pH of food product

3) Heating conditions

4) Thermophysical properties of food product and container shape and size

5) Storage conditions following the process

The primary step in thermal process establishment, which is defining and selecting the target microorganism or enzyme, is directly related to food product conditions. Temperature and oxygen are important factors in optimum growth of microorganisms. Based on appropriate temperature for growth, microorganisms are classified into psychrophiles with rapid growth between 0-5°C, mesophiles with optimum growth between 5-40°C and thermophiles with optimum growth at temperatures higher than 40°C (Rose, 1965). With respect to oxygen requirement for growth, microorganisms are classified as obligate aerobes, facultative anaerobes and obligate anaerobes. Packaged foods under vacuum in sealed containers provide low levels of oxygen, therefore, these conditions do not support the growth of obligate aerobes, and further the spores of obligate aerobes are less heat resistant to heat as compared to the spores of facultative and obligate anaerobes. The growth and activity of anaerobic microorganisms are highly pH dependent. From a thermal processing standpoint, foods are divided into three groups based on pH:

1) High acid foods (pH<3.7)

2) Acid or medium acid foods (3.7<pH<4.5)

3) Low acid foods (pH>4.5)

In thermal processing a special attention is devoted to *Clostridium botulinum* which is a highly heat-resistant, spore-forming, anaerobic pathogen that produces botulism toxin. *Clostridium botulinum* does not generally grow and produce toxin at pH below 4.6. Therefore, in thermal processing, a pH of 4.5 is considered as dividing line between the acid and low acid food products. Molds, yeasts and bacteria, which tolerate high acidic conditions are targeted in thermal processing of high-acid food products. *Bacillus coagulase* and *Saccharomyces cerevisiae* are important in high-acid foods. Microorganisms such as *Bacillus Stearothermphilus*, *Bacillus thermoacidurans* and *Clostridium thermodaccolyticum* are more heat resistant than *C. botulinum*, but they are mostly thermophilic in nature and in case of storage of cans at temperatures below 30°C, they are not of much concern.

Kinetics of microbial destruction

Evaluating the thermal resistance of target microorganisms is required in thermal processing design. Thermal destruction of microorganisms generally follows a first-order

reaction indicating a logarithmic order of death. Therefore, if the logarithm of number of microorganisms surviving a given heat treatment at a particular temperature is plotted against heating time, it will result in a straight line, called the survivor curve. The microbial destruction rate is generally defined in terms of a decimal reduction time (*D-value*), which is heating time that results in 90% destruction of the existing microbial population. This is concept is represented mathematically as:

$$D = \frac{t_2 - t_1}{[\log(a) - \log(b)]} \tag{2.1}$$

where:

a: number of survivors at time t_1

b: number of survivors at time t_2

t_2-t_1: heating time

Defined as such, the *D-value* represents the negative reciprocal slope of the survivor curve. Thermal death time (*TDT*) is another approach reflecting the relative resistance of bacteria to different temperatures. These data are obtained by subjecting a microbial population to a series of heat treatment at a given temperature and testing for survivors. *TDT* is the measurement with respect to an initial microbial load and it simply represents a certain multiple of *D-values*. The temperature sensitivity of *D-value* is defined by the term *z-value*, which is temperature range resulting in ten-fold change in *D-value*:

$$z = \frac{(T_2 - T_1)}{[\log(D_1) - \log(D_2)]} \tag{2.2}$$

where:

D_1: D-value at temperature T_1

D_2: D-value at temperature T_2

Also the z-value can be obtained from *TDT* curve using TDT_1 and TDT_2 instead of D_1 and D_2-*values*. Thus, z represents the negative reciprocal slope of *D* value or *TDT* curve.

In order to compare the relative sterilization capacities of thermal processes, the term lethality (*F-value*) is introduced. The *F-value* is defined as the number of minutes required at a specific temperature to destroy a specified number of microorganisms with a specific *z-value*. For convenience a unit of lethality is defined as equivalent heating of one minute at a reference temperature of 121°C (250°F) for the sterilization process. Thus, the *F-value* represents a certain multiple or fraction of *D-value* depending on the type of microorganism. Mathematical equivalent of this definition is:

$$F = nD = F_{T_0} \cdot 10^{\frac{(T_0 - T)}{z}} \tag{2.3}$$

where:

T_0: Reference temperature

F_0: Lethality at T_0

Lethal rate is defined for comparing different processes in terms of achieved lethality and it is considered to be the heating time at the reference temperature relative to an equivalent heating of one minute at the given temperature:

$$L = 10^{\frac{T - T_{ref}}{z}} \tag{2.4}$$

In a real process in which the food product undergoes a time-temperature profile, the lethal rate is integrated over the processing time to result in an overall process lethality (also defined by F_{T_0}):

$$F_{T_0} = \int_0^t L.dt \tag{2.5}$$

In terms of food product safety, assurance of a minimum lethality at the thermal center of food product is required. However it is desirable to minimize the overall destruction of quality factors.

Thermal process calculations

The purpose of thermal process calculations is to determine an appropriate process time under specified heating conditions required to achieve a given process lethality or estimating the lethality for a given process. Thermal process determination through a physical-mathematical approach requires the basic information of thermal destruction kinetics of microorganisms and quality factors conjoined with time-temperature data of product to integrate the lethal effects of thermal processing.

Therefore, an efficient process design requires sound information on the heat penetration data and their characteristics. Heat penetration data are function of several factors and different combinations of factors can result in same process lethality. These factors can be summarized as follows:

1) Method of heating process (eg. still process vs. agitated process)
2) Type of heating medium (steam, water)
3) Heating conditions (retort temperature, initial temperature of product)
4) Product type (solid, liquid, particulate)
5) Container type, shape and size

Obtaining accurate data regarding the heating and/or cooling of a food product in container is important for accurate determination of time and temperature with respect to sterilization of a given product. However, it is impractical to obtain heat penetration profiles for the whole range of conditions. Accordingly, thermal process calculation procedures are developed with ability of time-temperature prediction with respect to some experimentally determined parameters. Obviously, the applicability and restrictions of each method are defined by the assumptions taken into account to obtain temperature prediction model.

Following each heating phase of thermal processing operation is a cooling phase to control and terminate the lethal effect of thermal processing. In order to account the cooling effect, the lethality equation (2.5) can be rearranged as follow:

$$F_{o} = \int_{0}^{t_{h}} 10^{(T - T_{ref})/z} dt + \int_{0}^{t_{c}} 10^{(T - T_{ref})/z} dt \qquad (2.6)$$

where:

t_g: total heating time

dt: small time interval

t_c: total cooling time

The duration of cooling process longer than bringing the product temperature to a level enough to stop the lethality is not important. However, the achieved lethality during the initial cooling phase is accounted in process calculation methods. For a known time-temperature profile the solution of equation (2.6) will result in a relationship between F_n (process lethality) and t_c (process time).

Thermal process calculation methods

The methods of process calculations are divided into two broad groups: General methods and Formula methods. General methods apply the real time-temperature data from test containers to integrate graphically or numerically their lethal effects over the process. Therefore, this group of methods is the most accurate method for given experimented conditions. Conversely, Formula methods apply the time-temperature data in the form of parameters, with use of mathematical procedures, to integrate the lethal effects. Process calculation methods have been extensively reviewed and evaluated (Hayakawa, 1977, 1978; Merson *et al.*, 1978; Stumbo and Longley, 1966; Smith and Tung, 1982; Ghazala *et al.*, 1990; Larkin and Berry, 1991; Stoforos, 1997).

General Method

For the first time, Bigelow *et al.* (1920) introduced the Original General method, which is the fundamental of all other process calculation methods. General method is the most accurate method of process evaluation as it relies on discrete, experimental or numerical time-temperature data to determine the sterilization value of the process. Once the time-temperature is known, the sterilization value is determined by graphical or numerical integrating of equation (2.5). Overall, General method has the advantage of versatility in applicability for any kind of heat transfer. However, the application is restricted to the condition under which the time-temperature data is obtained, and any

change in either processing conditions (retort temperature, initial temperature), product or package size requires a related temperature profile.

The original General method is based on lethal rate as the reciprocal of *TDT* value. The area under the lethal rate curve vs. heating time will result in a sterilization value of the thermal process. In order to determine the process sterilization value ($F_{,,}$) with General method, different works has been carried out, both in graphical and numerical integration to make this method less laborious. Schultz and Olson (1940) developed lethal rate papers and dimensionless temperature differences in forms of T_{RT}-T/T_{RT}-T_{IT} to account for variation in retort temperature and initial temperature of product independent of experimentally obtained data. Extended works have been carried out for different z-values and processing temperatures (Cass, 1947; Hayakawa, 1973; Leonhardt, 1978). Simpson's rule, trapezoidal rule and Gaussian integration are improved techniques for numerical integration. Patashink (1953) used a trapezoidal rule by considering some assumptions in this technique. Hayakawa (1968) developed a method using Gaussian integration formula, involving a template, which could be used with the graphical method.

The limitation of General method is in determining the process time for target process lethality. For this purpose a trial and error method is proposed. Also it is tedious to consider a lethal rate curve during cooling which depends on the heating time. Often a single shape of cooling curve is imposed irrespective of the heating curve. Obviously, this cannot be right for all situations as the lethal rate is related to temperature difference of the product and heating medium at the end of heating (*g-value*) and this assumption applies to only to cases with g-value close to zero.

Improved General Method

Ball (1923) introduced improved General method with a graphical approach for sterilization value determination. A hypothetical thermal destruction curve was constructed parallel to thermal death time curve, resulting in *F-value* equal to 1 min at 121.1°C (250°F). This modification permitted the comparison of different processes in terms of target lethality. According to this method the lethal rate could be determined as follow:

$$L = 10^{(T-121.1)/z} \tag{2.7}$$

The area under the curve, resulting from plotting L vs. heating time, represents the equivalent minutes at 121.1°C. The area under the curve can be determined by counting the number of squares, using a planimeter or by approximating to standard shapes. Also for this method, process time determination for target process lethality requires a trial and error technique.

Formula Methods

The procedure that Ball (1923) applied in thermal process calculation was the basis for a new set of methods termed "Formula methods". Process determination using Formula methods is considerably faster due to applying heat penetration data in form of heating and cooling rate indexes (f_h and f_c) and lag factors (j_{ch} and j_{cc}). Formula method determines the process time for a pre-selected process lethality or alternatively, process lethality of a given process. Hence using the parametric form of heat penetration data with appropriate mathematical procedures, the effect of processing conditions (retort temperature and initial temperature of the product) as well as the effect of different package sizes. Following is a brief review of some of the most applied Formula methods.

Ball's method: As mentioned previously, Ball's method (1923) is considered to be the first developed Formula method and it has a great contribution to the value of mathematics in food processing (Merson *et al.*, 1978). Ball developed equations for predicting the temperature at the critical point inside a can. This equation was applied in general equation of (2.5) to determine the process lethality. The temperature prediction equations were based on the experimental observations that a semilogarithmic plot of the difference between medium and product temperatures vs. time, after some initial time lag, was usually a straight line. The same approach is applied for cooling portion. Such typical heating and cooling curves are shown in Figure 2.1. They basically comprise of a hyperbolic heating lag portion, a logarithmic straight-line heating, a hyperbolic cooling lag and a logarithmic straight line cooling. Ball ignored the effect heating lag portion, as in usual conditions the product temperature in below the lethal temperature and the accumulated achieved lethality is negligible. However, the effect of heating lag can be included for processes with high IT, large *z-value*, or processes with low heating required

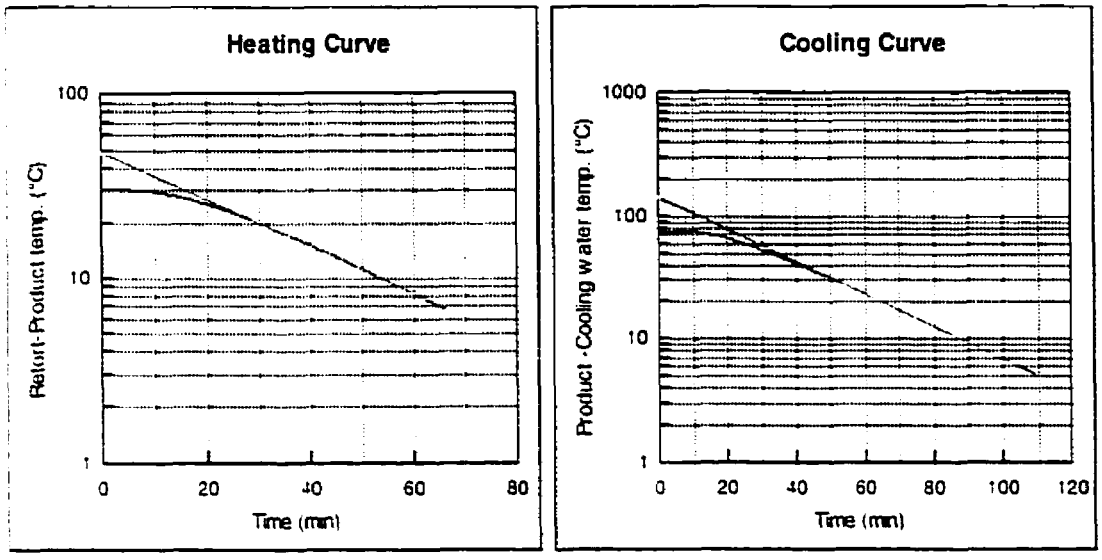


Figure 2.1 Typical semi-logarithmic heating and cooling curves

lethality, F_{ch}. Therefore, only the equation for the straight-line portion was considered for the heating portion. This equation with respect of heating rate index (f_h) and heating lag factor (j_h) was described as follows:

$$T = T_{RT} - j_{ch}(T_{RT} - T_{IT})10^{-t_h / f_h} \tag{2.8}$$

where:

T_{RT}: retort temperature (°C)

T_{IT}: initial temperature of the product (°C)

t_h: heating time (min)

f_h: heating rate index (min)

j_{ch}: heating lag factor

For the cooling curve, the effect of cooling lag was considered in calculation of sterilization value. After start of cooling, the critical center of the package is still at lethal temperatures and the effect of this lag should be considered in addition to the logarithmic straight portion of the cooling. In order to account for cooling, Ball used two separate equations to predict the temperature during cooling.

For the first portion or when $0 \le t_c \le f_c \log(j_c / 0.657)$:

The temperature in the critical point of the package was predicted with Equation (2.9):

$$T = T_g - 0.3(T_g - T_w) \times \left[1 - \sqrt{1 + \left(\frac{1}{0.5275 \log(j_c / 0.657)}\right)^2 \left(\frac{t_c}{f_c}\right)^2}\right] \tag{2.9}$$

where:

t_c: cooling time

j_c: cooling lag factor

f_c: cooling rate index

T_g: product temperature at the end of heating

T_{cw}: cooling medium temperature

For the straight-line portion of the cooling curve when the equation was same as heating portion with related cooling $f_c \log(\ j_c / 0.657\ ) \le t_c$:

$$T = T_{CW} + j_c(T_c - T_{CW})10^{-t_c / f_c} \qquad (2.10)$$

After defining the temperature prediction equations, Ball substituted them in the equation (2.5). The resulting equation could not be integrated analytically and concluded in a direct relation between F_o and t_g (process time). Hence Ball solved the equation graphically by evaluating the resulting integrals. The results of integrals were presented in table and graph formats. These tables and graphs related *g-value*, as a measure of process time, to f_h/U ratio, as a measure of sterilization value. Ball applied parameter U as the F value of the process at the retort temperature. These tables and graphs could be applied for any *z-value*.

During development of the tables and graphs, Ball assumed a constant cooling lag factor (j_c), equal to 1.41. Also the heating rate (f_h) was applied for the cooling portion, or in other words f_c was equal to f_h. Further, in the development of the method (Ball and Olson, 1957), with respect to two dimensionless parameters of P_h and P_c for determining the sterilization value, they accounted for variation of f_h and f_c. Obviously the conditions of all processes do not comply with these assumptions and result in errors in process calculations. Merson *et al.* (1978) and Hayakawa (1978) evaluated this method and reported the inaccuracies of the method. Ball method overestimates the process time, which provides a safety factor (Merson *et al.*, 1978) for process calculations.

The time duration until retort reaches the processing temperature is termed come-up time (CUT) and a portion of this time can hold lethality effect. Ball assumed 42% of this time should be added to process time, which holds the product at retort temperature until the steam off. He implemented this time duration in his method by shifting the zero of the heating time axis by 0.58 of CUT. It should be mentioned that this percentage is not always same and several researchers have taken into account to carefully evaluate the effect of CUT (Hayakawa and Ball, 1971; Ramaswamy, 1993).

Further there are some food products that do not follow a constant mode of heat transfer, and due to modification of the nature of the food product during the heating, the mode of heat transfer will change. This condition of heating is known as broken heating and is more common of those kinds of food products initially heat by convection: then, due to the activity of some thickening agents as starch gelatinization, the mode of heat transfer changes to conduction. Ball and Olson (1957) accounted for this effect by considering two straight-line segment of the heating curve with different slopes instead of one straight line (broken heating) in semi-logarithmic heating graph.

Hicks (1958) found several mathematical errors related to the lethality of heating phases, F_{sh}, in Ball and Olson's parametric values. He prepared a numerical table of recalculated parametric values. Gillespy (1951) defined a method for estimating the *F-value* for the whole can by using Ball's asymptotic approximation for heating and developed an approximate method for cooling. Herndon *et al.* (1968) prepared computer-derived tables based on Ball's method. Griffin *et al.* (1969; 1971) improved these tables for broken-heating curves and for cooling curves. Pflug (1968) compiled abridged tables from Ball and Olson's tables and also from Hick's table in parametric values and considerably simplified the calculation required for process evaluations.

Spinak and Wiley (1982) reviewed Ball's method and considered the accuracy of method in comparison with a General method. In this work they compared the process times for processing selected food products in retort pouch with respect to Ball's method assumptions and real values from the process. The result showed the process time determined by Ball's formula method requires the effect of retort come-up time lethality and the actual cooling lag factor (f_{cc}). However, actual f_c value showed no significant effect in accuracy of Ball's formula method.

In a further review of Ball's method (Steele and Board, 1979a), the inaccuracy of Ball's method was considered to be in a safe region for underestimating the process sterilization ratio. Therefore, it was concluded in this work that related inaccuracy introduces a safety factor to some calculated processes and of course the production of safe products is one of the primary criteria of thermal processing.

Vinters *et al.* (1975) parameterized the data from the tables in Ball's formula method, replaced tables and graphs in this method by algebraic, regression equations. Therefore, this method could be used in a programmable calculator.

Stumbo's method: While attempting to revise the tables developed by Ball, Stumbo and Longley (1966) developed a new set tables to account for the variability of j_{cc} values. The values of these tables were obtained by graphical measurement of hand drawn temperature profiles plotted on the lethal rate papers and subsequent interpolation of these tables. As the authors implied, these tables were applicable for cases in which the difference between f_h and f_c were less than 20% of the respective f_h values.

Later, the revised form of these tables were published (Stumbo, 1973). The revised tables were based on computer integration of temperature profiles generated from heat transfer equations, using finite difference simulations. However, it should be noted that these tables are applicable only when $f_h \neq f_c$ (Hayakawa, 1978). A correction for $f_h \neq f_c$ can be made as long as the value for the heating lethality can be obtained through a different method (Stoforos, 1997). The rest of procedures for process calculation were same as Ball's method. Computer implementation of Stumbo's method was introduced first by Manson and Zahradnik (1967), using Stumbo and Longley's tables and after by Tung and Garland (1978), using Stumbo's revised table values.

Hayakawa's method: For the first time Hayakawa (1970) applied a set of empirical formulas, each of them concerned to a specific range of j values. The curvilinear parts of the heating and cooling curve were represented by exponential cotangent and cosine functions. The minimum and maximum j values that are related to these tables are 0.045 and 3.0, respectively. He also prepared a table of new parametric values by using his empirical formulas, and developed a procedure for the evaluation of heat processes. These tables were applicable to almost any *z-value*, resulted in eliminating the required interpolation for the *z-value*. Therefore, the process calculation was significantly simplified.

Steele and Board's method: Steele and Board (1979b) and Steele *et al.* (1979) reviewed Ball's method and improved this method by introducing sterilization ratios to be used instead of temperature differences. This ratio was defined as the temperature difference between the heating (or cooling) medium and product by the slope of the thermal death time curve for the concerned microorganism. Using these dimensionless ratios was advantageous in four main aspects. Firstly, the method could be applied for any temperature scale whilst the same scale was used in determining the sterilization ratios. Secondly, the number of associated tables were less compare to Ball's method as z-value was incorporated in the sterilization ratio. Thirdly, the tables could be approximated more readily for a programmable calculator, as one variable was eliminated and finally, the limits of integration were selected in such a way that errors in tabulated values were negligible.

Pham's method: In an attempt to revise and improve the Stumbo's method, Pham (1987) introduced two serious of algebraic equations for thermal process calculation. The method relied on the conduction heat transfer equation of a finite cylinder. Heat transfer coefficient at the container walls was infinite and the initial temperature distribution was uniform. Two ranges of sterilization values were considered: high sterilization value or in cases when the product temperature is very close to heating medium temperature at the end of heating ($U/f>1$) and low sterilization value ($U/f<1$). For high sterilization value range an analytical solution was applied for the temperature prediction equation, the resulting solution of equations where simple algebraic equations, relating *g-value* to U directly within 3% error (Pham, 1987). For the low sterilization value range a numerical solution was applied and regression equations were generated. However, the result of these solutions was introduced in tabulated format, in form of dimensionless parameters. Pham considered the variability of T_{IT} and T_{CW} in these tables and equations. It should be noted that none of the previous mentioned methods took into account the variability of these parameters. Also, Pham (1990) incorporated the variability of f_h and f_c in the development studies of a new method. As the author implies, the applied method for considering this variability can be used in many other methods. In addition, Pham (1990) compared his method (with and without equality of f_h and f_c) with other methods using Smith and Tung

(1982) methodology. In case of $f_h=f_c$ the method was as accurate as the Stumbo's method, and in cases of $f_h \neq f_c$, 1% error was reported for a 20% difference between these parameters. This method has the advantage of being applied in on-line computer control and optimization of the thermal process, for being introduced in form of algebraic equations relating U and g directly together.

Evaluation of Formula methods

The accuracy evaluation of formula methods is essential for an efficient and successful thermal processing determination. Smith and Tung (1982) necessitated such a study for the importance of an accurate thermal process resulting in the minimum nutritional loss and improving quality of the product. Five major formula methods: Ball's table method, Ball's equation method, Stumbo's method, Steele & Board's method and Hayakaw's method were examined in this study. The reference model was based on a numerical general method to calculate the accumulated achieved lethality of the process for conduction heating food products in cylindrical packaging. A set of processing condition, product thermal diffusivity and can dimensions were used as the initial inputs of the model to obtain achieved process lethality. In an initial study, unachieved temperature difference at the end of heating (*g-value*) and can dimensions (*H/D*) were assigned as the most significant factors in deviations in calculation. A general increase in deviation for *H/D* near to unity was observed and the magnitude of *g-value* had a direct effect on deviations. Stumbo method had the highest accuracy in process lethality calculations. All the methods underestimated the process lethality, which represented an extra safety limit.

Pham (1990) used the same methodology as Smith and Tung (1982) after developing a formula method accommodating the variability of f_h and f_c. Pham's formula method (1987) had about the same accuracy as the method of Stumbo and had the better accuracy than Ball, Steele & Board and Hayakawa's methods. The same accuracy of Pham and Stumbo's method is for similarity in their derivation, however the major difference is the algebraic expressions used in Pham method.

Ghazala *et al.* (1990) examined the accuracy of five formula methods in comparison with a finite difference model based on conduction heat transfer of thin profile packaged food products. The comparison was performed based on a set of processing

conditions, food properties and a range of packaging size. A fixed value of process lethality was assumed to arrive at different process times with respect to selected conditions. Although Pham and Stumbo had the smaller errors compare to Steele & Board methods (equation and table) and Ball method, within the range of experimental conditions the overall difference between the methods were small. It was believed that finite difference models based on a thin profile packaging result in a better estimation of process parameters (especially j_{cc}), which formula methods are based on.

Larkin and Berry (1991) had a more specific estimation of different formula methods based on cooling lethality. A range of cooling rates (resulted from the variation in thermal diffusivity of product or can size), cooling lag factors (by changing the relative position along the radius within the can), can dimensions and temperature differences between final heating and cooling temperatures were considered for this study. All the selected formula methods rather than Pham's and Ball's method with a constant j_{cc} of 1.41, underestimated the cooling lethality of the process for a complete range of selected can dimensions. However with an increase in j_{cc} value, Pham's method began to underestimate the cooling lethality also. The temperature difference at the beginning of cooling and cooling medium temperature affected the cooling lethality but the differences between the lethalities predicted by different formula methods did not change. Retort temperature had only effect on Pham's method in cooling lethality calculation and the other formula methods were independent of this variable. Also this study showed that lower jcc values than 1.41 results in more overestimation of cooling lethality. Nevertheless, smaller j_{cc} values showed lower contributions of cooling lethality to total process lethality.

Stoforos *et al.* (1997) carried out a broad and comprehensive study on thermal process calculation methods. The performance of some of the methods was compared against temperature prediction using a finite difference model for conduction heat transfer equations. Ball method has inability of temperature prediction at the beginning of the heating, showed as a comparison with Hayakawa's method and finite difference model. However, this initial lag in temperature prediction is negligible as the product temperature is below the lethal temperature and the achieved lethality in the initial portion of heating is insignificant. Besides, the accuracy of each model at the beginning of the heating determines which particular model can be used in handling time-varying medium

temperatures. Numerical solutions of heat conduction equations, which are capable of accommodating the medium temperature variability, are among the most preferred methods for handling time-varying medium temperatures.

Numerical models of thermal process calculations

Numerical methods are based on simulating the conduction heat transfer in packaged foods. Numerical solutions to Fourier's partial differential equation of conduction heat transfer results in a temperature profile of product during the thermal processes. For a regular shaped container such as a cylinder or a rectangular slab, which are more common in food industry, a finite difference solution based on a regular gridwork of the nodes is applied. However, for irregular shapes, a more complex technique of finite element method is required.

The equation, which expresses transient temperature at the center of a finite cylinder, is as follows in cylindrical coordinates:

$$\frac{\partial T}{\partial t} = \alpha \left(\frac{\partial^2 T}{\partial r^2} + \frac{1}{r}\frac{\partial T}{\partial r} + \frac{\partial^2 T}{\partial h^2} \right) \tag{2.11}$$

where:

T: temperature

t: time

α: thermal diffusivity

r: radial position in cylinder

h: vertical position in cylinder

This equation is the partial differential equation for two-dimensional unsteady state conduction heat transfer in a finite cylinder. This equation can be written in finite differences to be solved with a numerical solution:

$$T_{(i,j)}^{(t+\Delta t)} = T_{(i,j)}^{(t)} + \frac{\alpha \Delta t}{\Delta r^2}\left[T_{(i-1,j)} - 2T_{(i,j)} + T_{(i+1,j)}\right]^{(t)} + \frac{\alpha \Delta t}{2r\Delta r}\left[T_{(i-1,j)} - T_{(i+1,j)}\right]^{(t)} + \frac{\alpha \Delta t}{\Delta h^2}\left[T_{(i,j-1)} - 2T_{(i,j)} + T_{(i,j+1)}\right]^{(t)} \tag{2.12}$$

Finite differences are discrete increments of time and space defined as small fractions of process time and container space (Δt, Δh and Δr, respectively). For convenience in calculations and based on symmetry, usually half or a quarter of the container is considered in calculations. The temperature nodes are assigned for this selected volume as shown in Figure 2.2. Appropriate boundary and initial conditions are required to calculate the new temperature at each node, after a small time interval. At the beginning of the process, the interior nodes are set equal to initial temperature of the product and nodes in surface are set equal to retort temperature (when the associated surface heat transfer coefficient is large). After each time interval the new temperature at each node is calculated, which replaces the previous temperature. This procedure is continued until the end of heating when the boundary conditions change from heating to cooling and computation continues usually with a finite heat transfer coefficient associated at the container surface. In this way, the temperature at the center of can is calculated after each time interval, which results in a temperature profile from which the process sterilization value can be computed.

Numerical models can be used instead of heat penetration test to obtain the temperature profile of the product, when the thermal properties of the product are known. Also the retort temperature need not be held constant and can vary during the process. Therefore these models can be used in continuous processes, when the cans pass from one chamber to another, and the heating medium temperature changes during heating process. This advantage has provided the potential of these models in on-line control of thermal processing. Teixeira *et al.* (1969) applied such numerical models in optimizing the nutrient retention in conduction-heated foods. The main importance of these models was providing the temperature at any position inside the can. The application of these models in on-line control has been extensively studied (Datta *et al.*, 1986; Teixeira and Manson, 1982; Tucker, 1991; Teixeira and Tucker, 1997).

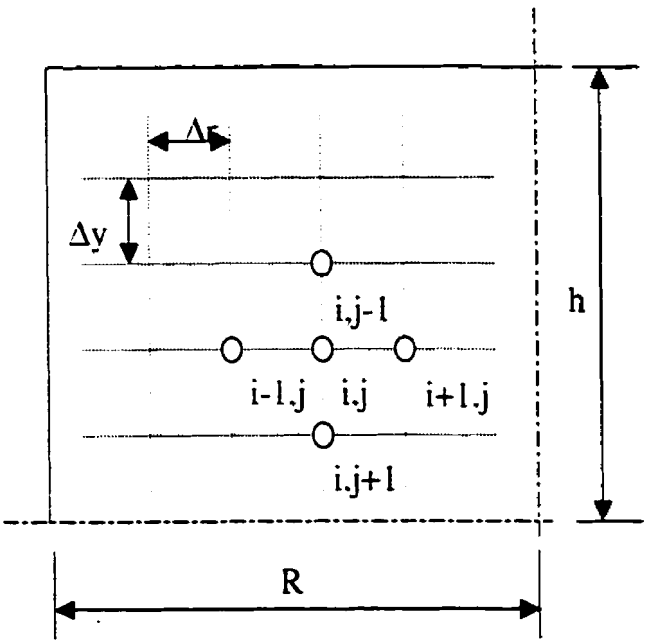


Figure 2.2 Grid format for temperature calculation using a numerical method

Artificial Neural Networks

Artificial neural networks (ANNs) are computing systems built up of interconnected processing elements, which are able to map information between a set of input variables to related output variables. A very fundamental component of brain or nerve cells, neurons, is also the basis of the ANN computing.

Regard to this similarity, there are three main sections of biological neurons, which are important in understanding the structure of ANNs. A biological neuron consists of three main components (Figure 2.3). Dendrites (input paths) receive the information as signals from other neurons. Signals are chemicals in nature, but they have electrical side effects, which can be measured. The soma or cell body sums the incoming signals and after receiving sufficient inputs, it will fire signal through its axons (output paths). The axon of a neuron splits up and connects to dendrites of other neurons through a junction referred to as a synapse. The strength or synaptic efficiency depends on released chemicals from axon and the amount that is received by dendrites. The synaptic efficiency is what is modified when the brain learns.

In simulating the brain neural network, a processing element (PE) with its connections is the ANN equivalent of a neuron. The PE acts as the cell body. There are input connections to the PE, like dendrites, which transfers the signals to the PE. Output connections from PE, like the axon, transfer the signals to the other PEs. The interconnections possess parameters known as weights, which will be modified through the learning procedure (as the change in synaptic strength). The information transformation is performed through the interconnections between all the PEs. In a typical ANN structure, processing elements are arranged through layers (Figure 2.4).

Usually an ANN consists of one input layer, at least one hidden layer and one output layer. Input layer receives the information from an external source and sends them to the network. Hidden layers process the information from input layer and transfer them to the output layer. Output layer receives the processed information and sends them to an external source. When the PEs in input layer send the information in the form of signals to the network, the strengths of interconnections are altered based on the amplitude of the signals. These changes are distributed all over the network through connections and, at the

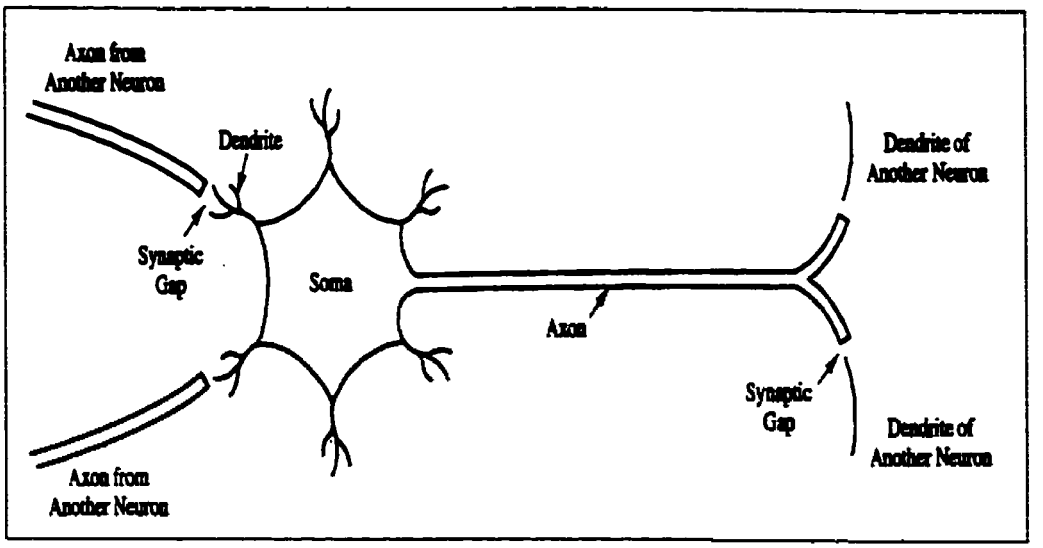


Figure 2.3 Basic structure of a biological neuron

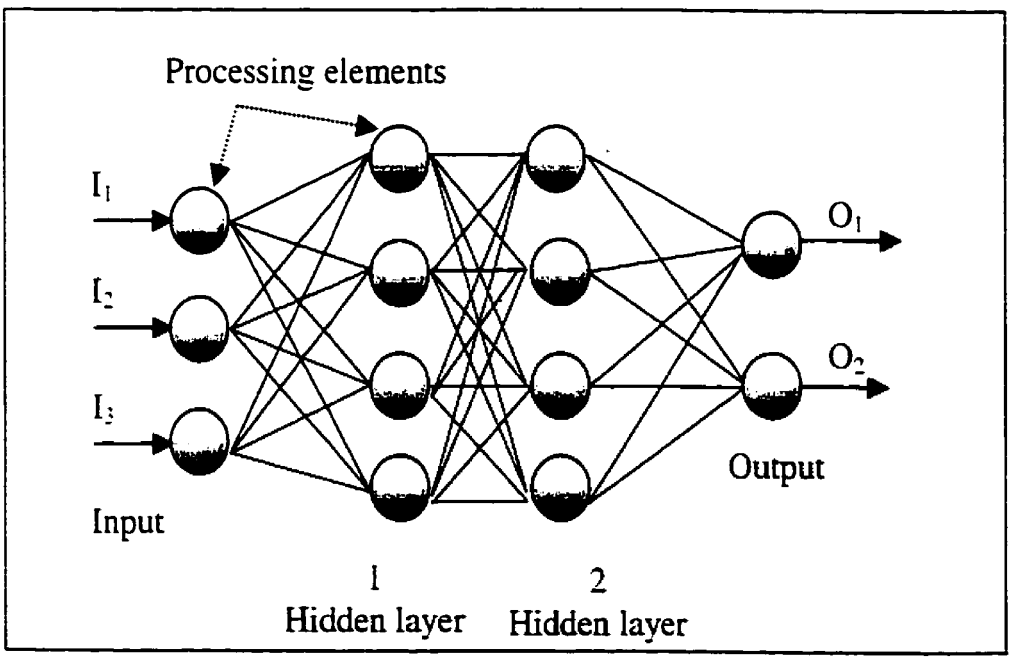


Figure 2.4 Typical diagram of ANN model

end. manifest themselves in the form of outputs. One important characteristic of ANN is that it processes information numerically rather than symbolically. The network retains its information through the magnitude of the signals passing through the network and interconnections between processing elements.

Figure 2.5 shows the basic structure of a processing element and its connections. The inputs into j^{th} layer are basically the outputs of the previous layer processing elements as an input vector. a. with components a_i (i=1 to n). The output of processing element (b_j) is the result of a transfer function (f) over a summation of inputs multiplied by weight parameters and addition of a bias value (T_j). Transfer function can be varied depending on the nature of the data. The most common transfer functions in solving non-linear problems are sigmoid function and hyperbolic tangent. A sigmoid (S-shaped) function is depicted mathematically as follows:

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2.13}$$

This function varies between 0 (at x_i=-∞) and 1 (at x_i=+∞). Sigmoid functions. due to their limiting values. are known as threshold functions. At very low input values. the threshold-function output is zero. At very high input values. the output value is one.

Hyperbolic tangent functions also typically produce well-behaved networks. This function also has two limiting boundaries of +1 and −1:

$$f(x) = \tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{2.14}$$

As the response of this function includes both positive and negative region. their application are recommended for data set with negative and positive output range (Swingler. 1996).

Development of an artificial neural network model

An important factor. which distinguishes different neural networks. is the method of setting the values of the weights or training the network. In training, ANNs generally can

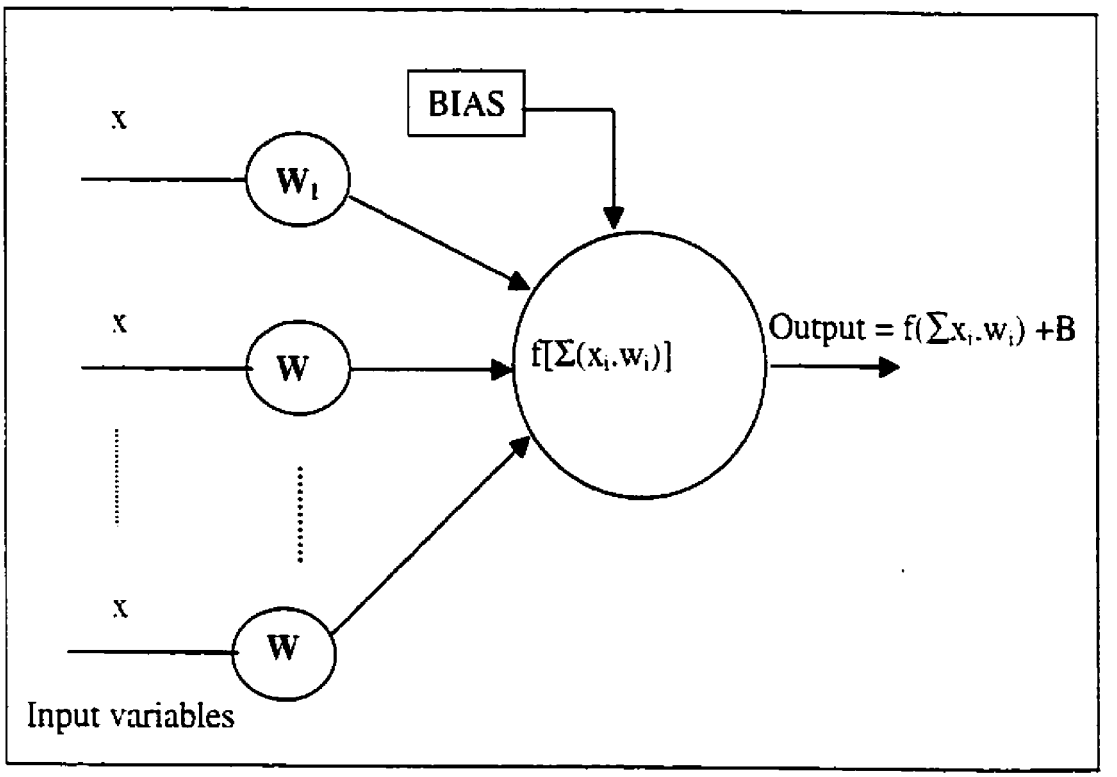


Figure 2.5 Diagram of a processing element and related calculations

either be supervised or unsupervised. In supervised training, there is an associated output along with any input vector in training data set. Therefore, the weights are adjusted according to the target output. The most common application of this method is in classification, prediction, and pattern association problems. On the other hand, in an unsupervised learning, a sequence of input vector is provided but no target outputs are specified. The network will act in a self-organizing manner to modify the weights so that most similar input vectors are assigned to the same output (or cluster) unit. As the definition reveals, unsupervised learning is mostly applicable in data clustering problems.

Besides the mentioned classification, in each group, there are several algorithms of training. Backpropagation is one of the most applicable algorithms in problems involving mapping of a given set of inputs to a specified set of target outputs (Fausett, 1994). Backpropagation algorithm is simply a gradient descent method to minimize the total squared error between the output computed by the network and target output. There are three main stages in backpropagation algorithm: a feedforward of the input training pattern, calculation and backpropagation of associated error and adjustment of weights with respect to calculated error. Although the training of ANN model is time consuming, a trained network will only apply a feedforward phase to compute the associated output and renders this phase very rapidly.

The major task in training a network is to bring the network to a balance between memorization and generalization. Memorization is the ability of the network in correctly responding to input vectors used in training. Generalization is the ability of network to respond to input vectors in the domain of training vectors but not identical to input vectors. A proposed procedure for achieving this goal applies two disjoint sets of data known as training and testing data sets. Training data set is used to updating of weights; however, at intervals during training, the network is tested with testing data. As long as the network error for testing data is decreasing, the training continues. When this error begins to increase the network starts to memorize the features in input pattern too specifically and loses its generalization ability therefore at this point the training is terminated.

To develop an optimized network, the number of hidden layers and processing elements in each layer should be selected carefully. In addition to structure of network, related parameters to improve the learning task should be optimized. Two of these

important factors are momentum and learning rate, which control how effectively backpropagation trains the network (Baughman and Liu, 1995). The learning rate is a positive constant and controls the rate at which the new weight factors are adjusted based on the calculated gradient descent correction term. The momentum coefficient is an extra weight added onto the weight factors that accelerates the rate at which the weight factors are adjusted, helping the network to avoid the local minimas. A well-optimized network has the ability of extracting the relationship between the training data set and adapting its internal structure based on related information.

Applications of ANN in Food Science and Agriculture

Many problems encountered in food, agricultural and biological industries lend themselves to a neural network solution. The application of ANN in different problems is introduced.

One of the first applications of neuro-computing in food technology was introduced by Thai and Shewfelt (1991). The combination of neuro-computing and statistical techniques was used to determine the most relevant factors in relating sensory judgment by human and physical measurements of external color of tomato and peach. In this study, the application of neural networks was faster compared to statistical techniques for having fewer steps in analysis step; however, statistical techniques were slightly more advantageous for higher accuracy. Consequently, neuro-computing technique were applied to get the basic features of the relation; and statistical techniques were used to improve the numerical accuracy of resulting mathematical functions. With respect to this study and above mentioned procedure, the relationship between sensory evaluations and physical measurements of tomato and peach were identified as particularly linear.

Bochereau *et al.* (1992) applied neural network technique as a part of procedure for prediction of apple quality using Near Infra-Red (NIR) spectra data. In order to reduce the number of input parameters and excluding uncorrelated inputs to the neural network model, principal component analysis (PCA) was applied. A multiple regression analysis was then carried out to derive the best linear estimator and finally the neural network model was developed to extract the non-linear function input and output components. The final result

was a model for predicting the apple quality from NIR spectra. ANN modeling resulted in 5 % increase in accuracy of the linear model with R^2 equal to 0.82

Linko and Zhu (1992) studied the potential of ANN techniques to construct an appropriate real-time state estimation and prediction model applicable to process control in glucoamylase fermentation. In comparison with conventional modeling, ANN had the ability of handling uncertainties, complexity, noise and unavailability of data, which are the common cases in the biochemical reaction analyses. The most related on-line data as oxygen consumption rate (OCR), carbon dioxide evolution rate (CER) and the nitrogen utilization rate (NUR) along with other inputs were applied to estimate the enzyme activity and biomass in on-line control. The ANN models had the advantages of ease of use without a prior knowledge on the complex mechanism of microbial physiology or on the interrelationships of the variables. Once trained on a typical example data, it ANN model could suppress the noisy data. The developed ANN model performed very satisfactory and compared well with real values from off-line analyses.

ANN models were applied as a quality prediction model for black tea and resin samples based on data from chromatography and sensory information (Tomlins and Gay, 1994). Although multiple regression techniques were more accurate compared to ANN modeling, the results of ANN model were improved by reducing the training variables assigned by stepwise multiple regression. In addition, ANNs for cases in which the accuracy is not important (for example in quality control) provides more advantageous application in predicting several parameters simultaneously. The combination of statistical procedures and ANN were suggested as useful tools in pattern recognition and regression of chromatographic (GC and HPLC) and sensory data.

Park *et al.* (1994) applied an ANN model for predicting and classifying beef characteristics using ultrasonic spectral feautures as the input data for training the models. ANN models provided better results compared to statistical regression models for predicting the sensory attributes of beef as juiciness, muscle fiber tenderness, connective tissue, overall tenderness and flavor intensity. The relationships between physical attributes and sensory properties were mostly linear as an ANN model developed without hidden layers performed best. The accuracy of ANN model for classification were a function of learning schedules, number of processing elements in hidden layer and number

of input variables. Increasing the ultrasonic spectral features as input variables increased the accuracy of classification.

Sayeed *et al.* (1995) investigated the possibility of an ANN model development for sank quality evaluation. For this purpose a machine vision technique was applied to quantify the quality features of typical snack products in the form of texture (reflecting the internal structure) along with size and shape features. The relationship between the mentioned variables and sensory attributes were not adequately linear as examined by linear regression. therefore the input variables were applied to train a backpropagation ANN model to learn the non-linearity between input and output variables. The developed ANN model was validated with a set of untrained data and the prediction results were compared to those from a taste panel. The developed methodology had a potential application in snack production industry as a non-destructive method of quality evaluation.

The potential ANN models for milk shelf-life prediction was evaluated by Vallejo-Cordoba *et al.* (1995). Dynamic headspace gas chromatographic data collected during the storage of pasteurized milk were applied as input to be related to flavor-based shelf-life of milk in days as dependent variable. ANN and Principle Component Regression (PCR) techniques were used and compared for their predictability. ANN. with 2-day standard error in shelf-life prediction. performed better that the 4-day error PCR technique. and indicated a higher predictability for this concept.

Sablani *et al.* (1995) investigated the potential of ANN models for prediction of optimal sterilization temperatures. A finite difference model generated the data for a set of assigned conditions (can size. thermal diffusivity and kinetic parameters of quality factor) to obtain the associated sterilization temperatures. Afterwards the optimum conditions and related quality factors were applied as inputs to train the ANN model. The trained model was able to predict the optimal sterilization temperatures with less than 5% error. In attempt to predict the overall heat transfer coefficient and fluid to particle heat transfer coefficient using ANN modeling technique, Sablani *et al.* (1997) obtained significantly superior results compared to dimensionless correlation techniques. Beside the higher accuracy of prediction. ANN models were more versatile than dimensionless number models.

Rauan *et al.* (1997) compared the predictability of ANN models for rheoligical properties of a cookie dough. The raw input data were obtained from mixing power consumption curves. The investigated rheological properties were farinograph peak, extensibility and maximum resistance. Two spectrum analysis techniques, Fast Fourier Transform (FFT) and power spectral density (PSD) were used for data preprocessing as a part of ANN model development. Using these techniques the raw data (time domain data) were converted into frequency domain data. The above-mentioned techniques improved the training ability of ANN models for reducing the noise and size of the initial raw data and were recognized as successful techniques of data preprocessing in this study.

Ghazanfari *et al.* (1996) studied the suitability of multi-structure neural networks (MSNN) over a single multi-layer neural network (MLNN) for classification of pistachio nuts. The MSNN classifier was constructed from four single ANN models, which were trained separately with one output variable from each network representing one variety (class) of pistachio nuts. The input data for these models were physical attributes of nuts derived from their images. The classification accuracy of developed MSNN was compared with MLNN model. The MSNN technique significantly increased the classification performance. In addition, MSNN had the advantage of smaller network architecture, which is more preferable for hardware and software implementation.

Angerosa *et al.* (1996) developed an ANN model as alternative to sensory evaluation performed by a taste panel for virgin olive oil by panel test. A wide range of samples (covering different variety, quality, ripeness, sanitary and geographical origin) was subjected to a sensory evaluation by a taste panel. The results of quantification of volatile fractions using headspace gas chromatographic technique were used as input data to predict the panel test scores (output variables). The ANN model was able to generalize the relationship between these two sets of data successfully with a high degree of accuracy, suggesting the substitution of developed ANN model with panel test.

Briandet *et al.* (1996) applied ANN technique as one of alternative statistical approaches for correlating data from Fourier transform infrared (FTIR) for detection of adulterated freeze dried instant coffee. This FTIR technique was applied as a rapid alternative to wet chemistry methods; however, data analysis of this technique were more complicated. ANN models were compared to principle component regression and partial

least square regression, and the superior performance and generalization potential of ANN model was demonstrated with an independent set of data with 100% correct classifications ability.

Pamer *et al.* (1997) developed an ANN model to evaluate the factors involved in the contamination process and also the significant factors in the magnitude of aflatoxin contamination of in preharvest peanuts. Alternative to ANN model, a stepwise traditional regression model was developed for this purpose. The input variables (soil temperature, drought duration, crop age and accumulated heat units) were related to aflatoxin levels. This study showed the ability of ANN models for this purpose in addition to better performance than linear regression technique.

Kim and Cho (1997) applied three developed ANN models as a part of a fuzzy controller in bread baking process. For this purpose three main parameter in bread baking process namely volume, browning and temperature were measured in 3 seconds and 2 minutes intervals, resulting in training data for 3 ANN model development. The neural networks showed a good performance for predicting temperature, volume and browning. The knowledge from developed models and an experienced operator were used to define 11 rules for fuzzy controller. Using a fuzzy controller instead of a human operator can reduce the cost of heating of the oven, without the loss of bread quality.

ANN models as an alternative in data-processing technique to partial least squares (PLS) and principle component analysis (PCA) showed better prediction ability (Horimoto *et al.*, 1997). In order to classify the flavor quality of milk as a part of quality control, dynamic headspace gas chromatographic was used to quantify the off-flavor components. UHT milk was inoculated with various bacterial species and storage times. The content of training and testing data as well as optimized parameters, which were related to ANN model development (learning rate, momentum factor and hidden PEs), influenced the predictive ability model.

During isobaric-isothermal inactivation of α-amylase enzyme, pressure and temperature had a complex effect (Geeraerd *et al.*, 1998). A model based on Arrhenius law presented a non-linear description of the system. Due to lack of data, an ANN model was applied with respect to its ability in prediction of complex systems. The ANN model was able to predict the complex effect of pressure and temperature on the inactivation rate

constant using a limited set of experimental data. The effect of transfer function of each neuron was believed to be important in a low complex ANN model development.

Gebri *et al.* (1998) applied ANN modeling in predicting the origin of white vinegar. Among different chemico-physical and sensory analysis, selected parameters such as polyalcohols, pH, tartaric acid and proline had the most reliable rule in discrimination of the product. A broad data set covering products from a variety of raw materials and from various countries subjected to multivariate statistical analysis were used to develop the ANN model. The ANN model was able to predict the botanical origin of the product and also was able to re-classify products with unknown origin.

The above review indicates the status of available methods of process calculations (with sufficient scope for developing new method, which can reduce process calculation errors), as well as the potential of ANN models for use in process calculations. The present thesis aims at developing an ANN based process calculation model, which would be more versatile and accurate than the existing methods.

CHAPTER 3

ANN SIMULATION OF BALL AND STUMBO FORMULA METHODS OF THERMAL PROCESS CALCULATIONS

ABSTRACT

Application of ANN models to simulate Ball and Stumbo methods of thermal process calculations is presented in this study. ANN models were developed based on relating Ball and Stumbo table parameters to facilitate process calculations. The ANN models related g value (as a measure of process time) to f_h/U (as a measure of process lethality). Tables developed by Stumbo accommodate the j_{cc} (cooling lag factor) while relating g and f_h/U, therefore, for developing ANN models of Stumbo method, j_{cc} was also considered as an additional input variable. The developed ANN models for Ball and Stumbo methods were validated using a new set of processing conditions, involving a range of retort temperatures, initial temperatures, heating rates and heating lag factors [and additionally, for Stumbo method, cooling lag factors (j_{cc})]. The prediction efficiency of ANN models were a function of the size of training data set, number of hidden layers and PEs in each hidden layer as well as other learning parameters. ANN based Ball models had an average error of 1% for the validation data set while the ANN based Stumbo models had a slightly higher 3% average error for process time and process lethality calculations. The smaller size of data set and a wider range of parameters associated with Stumbo tables were considered to be the reason for the associated higher errors with the ANN based Stumbo models. In general, the ANN models were considered to simulate Ball and Stumbo methods of process calculations well providing a basis for further exploration of the concept.

INTRODUCTION

Thermal processing of packaged foods is one of the most widely used methods of preservation in twentieth century (Teixeira and Tucker, 1997). Nicholos Appert introduced this method for the first time in 1810. The concept of thermal processing is based on heating of packaged foods for a certain length of time to obtain a safe product complying with public health standards. Thermal processing is based on established time-temperature profiles. Associated with thermal processing is always some degradation of heat-sensitive quality factors that is undesirable. Since much demand is on safe and shelf-stable food products along with a high quality attributes, processing schedules are designed to keep the process time to the required minimum.

The main objective of thermal process calculations is to determine the process time for achieving a pre-selected process lethality or evaluating the lethality of a given process. For the first time, Bigelow *et al.* (1920) introduced a graphical procedure of evaluating the efficiency of heat treatment process for packaged foods. Ball (1923) refined the concept and developed a mathematical model (referred to as Ball method) for calculating the required process time of canned foods. Besides some limitations in this method, the Ball method is considered a milestone in canning industry and still is a widely used method in industry. This method is positioned as a classical tribute to the value of mathematics in food processing.

Ball method is based on the observation that a semi-logarithmic graph of temperature difference between product and medium vs. time is a straight line after an initial curved portion (lag time). Ball (1923) calculated the process time based on the equation of this straight line and integrating the effect of process time over the kinetic data of microbial destruction. Usually the effect of curved portion of heating is not considered, as the lethal effect of this portion is negligible compared to the total achieved lethality. However, under conditions of high initial temperature, large z-values or processing for low required lethality, the efficiency of this portion of heating should be taken into consideration.

The estimation of proper heat processes is essentially based on the calculation of process lethality, F_o, computed as follows:

$$F_o = \int_0^t 10^{\frac{T_{ref}-T}{z}} dt \qquad (1)$$

where:

T_{ref}: reference temperature

T: center point temperature

z: temperature sensitivity indicator of F value.

For low acid food products, 250°F (121.1°C) is widely used as a reference temperature. Ball developed tables and graphs to relate process time and process lethality. These tables and graphs were based on relating two parameters: g (temperature difference between product and heating medium at the end of heating) as a measure of process time to f_h/U (ratio of heating rate to sterilization value) as the measure of process lethality. Some restrictive assumptions were made during the development of these tables and graphs. One of the most important parameters, which can sometimes cause errors in calculations, was a constant cooling lag factor, j_{cc} value, equal to 1.41. Therefore, Ball method overestimates process lethality when $j_{cc}<1.41$ and underestimates when $j_{cc}>1.41$ (Ball and Olson, 1957; Hayakawa, 1978). Also, Ball (1923) assumed an equal heating and cooling rate ($f_h=f_c$). In practice, the values of these two parameters differ from each other. It has been found that for steam heating and water cooling $f_c \sim 1.3 f_h$ (Pham, 1990) and obviously disregarding this inequity interferes with the accuracy of process calculations. Ball method has been widely reviewed and evaluated by several researchers (Merson *et al.*, 1978; Flambert *et al.*, 1977; Steele and Board, 1979; Spinak and Wiley, 1982; Stoforos, 1991).

Stumbo and Longley (1966) reviewed and revised tables developed by Ball (1923). The variability of j_{cc} value was considered in these tables for more reliable and accurate thermal processes calculations (Stumbo, 1973). These tables are obtained from hand drawn heat-penetration curves and relate f_h/U and g values with respect to variability of j_{cc}. Also the variability of f_h and f_c was considered in developing these tables, but it applies only to cases where the difference of f_h and f_c is not more than 20% of respective f_h value (Hayakawa, 1978, Stoforos *et al.*, 1997).

Artificial Neural Network Models

Artificial Neural Network (ANN) models are computational models made up of several processing elements (PEs), which are connected through weighted links. Development of model

is basically based on adjusting the weights through presenting adequate number of examples consisting of input and output variable pairs.

Generally ANN models are constructed from an input layer, an output layer and one or more hidden layer(s). Input layer transfers the information to the network for processing. Output layer receives the processed information and sends it to external receptor. Hidden layers process the received information from input layer, extracts the existing features in input variables and predict an output. Among different networks, back-propagation network has the most applicability in prediction and classification problems. Backpropagation network is a feed-forward network, which propagates back a portion of error between desired and ANN predicted output to correct the weight parameters.

Recently, ANNs have shown a potential for applications in food science and technology areas. Eerikainen *et al.* (1993) presented a series of examples of potential application of ANN models in food-related applications. Parmer *et al.* (1997) applied neural network modeling for estimation of aflatoxin contamination in peanuts which performed better than traditional linear regression techniques. Ni and Gunasekaran (1998) simulated the complex task of food quality prediction using ANN models. Boucherau *et al.* (1992) developed a neural-network-based method for prediction of apple juice quality using near infrared spectra. Neural networks were used to determine the significant variables in harvesting and processing effects on Surimi quality by Peters *et al.* (1996). Kim and Cho (1997) developed an ANN model for the bread baking process, using three quality factors of volume, browning and temperature. In thermal processing, Sablani *et al.* (1995) applied neural network models to predict optimum thermal processing conditions. As an application in drying, Sreekanth *et al.* (1998) developed an ANN model for prediction of psychometric parameters.

The global objective of this study was to evaluate the potential of ANN models as an alternative to conventional process calculation methods. As a first step in this direction, data from two classical methods: Ball and Stumbo methods, were used to develop the ANN models and assess their performance. Hence, the specific objective of this paper is ANN simulation of Ball and Stumbo formula methods of process calculations.

MATERIAL AND METHODS

ANN modeling of Ball and Stumbo methods

Ball and Stumbo methods relate process time and process lethality by employing two parameters, g and f_h/U, in the form of tables or graphs. In tables developed by Stumbo, the effect of j_{cc} on this relation is also considered. For a known process time along with the other processing conditions, g-value will be known; alternatively for a given process lethality, f_h/U will be known. For developing ANN based process calculation models, Ball and Stumbo methods were simulated based on relating the parameters of these tables. Data (f_h/U vs. g or g vs. f_h/U) for both training and verification were obtained from tables published in Lopez (1987) for Ball method and Stumbo (1973) for Stumbo method. Detailed tables were available in Lopez (1987) and hence 1050 data pairs were used in Ball model for training. For Stumbo method, only 44 data sets relating f_h/U vs. g (and vice versa), were available at 9 different values of j_{cc}. A z-value of 10C° was used for both models. For each method, two models were developed one for predicting f_h/U based on g, and vice versa.

Development of ANN models

Ball model

NeuralWorks Professional II/Plus, version 5.23 (NeuralWare Inc., Pittsburgh, PA) was employed for ANN modeling. A standard back-propagation algorithm with tangent hyperbolic transfer function and normalized cumulative delta learning rule was applied as the basic architecture of network. As the input and output variables had a broad range, for a better presentation of input and outputs to the network and a more uniform data set, a mathematical transform function was applied. Therefore, *log g* was predicted from *log*(f_h/U), and f_h/U in the form of *arctangent*(*log* f_h/U) was predicted from *log g*. For a more convenient recalling of models, from now prediction of *log g* will be denoted, as Model A and prediction of f_h/U will be denoted as Model B.

It has been recognized that the network predictability and performance should be optimized with respect to the size of the training data set, structure of the network, number of hidden layers, number of processing elements in the hidden layer, type of learning rule and transfer function. Several other variables also affect the ANN model development and

performance. Among them momentum and epoch size have the most significant effect on the network performance. The momentum represents a fractional value from the previous weight change added to the present weight correction. Its value can change between 0 and 1, however, higher momentum values speed up the learning process and prevent the model from being trapped in a local minimum. At the same time, large values of momentum will cause large error oscillations (NerualWare, 1996; Baughman and Liu, 1995). Epoch is the number of training presentations between weight updates. This method of updating weights which increases the convergence speed is known as cumulative delta rule learning. However, using cumulative delta-rule learning for updating weights will require more calculation and, if the epoch is too large, the advantage of using overall error function can be lost. Therefore, depending on the problem, the epoch size should be optimized.

The performance of an ANN model is influenced greatly by the quantity and quality of data used as the training set (Linko and Zhu, 1992). The training data should have sufficient information to describe relation between input and output variables. A correct format of input and output variables could increase the learning ability of ANNs by detecting and extracting the relation between input and output variables (NeuralWare, 1996). This stage of ANN models development is thus regarded as data preprocessing, which could be performed in different ways to improve the model efficiency (Lacroix *et al.*, 1997).

Keeping the above constraints in mind, to select the required number of examples in training set, a learning curve was developed. The original data (1050 data) were sorted from the lowest to highest value. To have six equal groups of data, every seventh data row was selected and removed from the original set. These formed the first (Group A) data set of 175 data points. The same procedure was repeated with the remaining data until six homogenous groups (A to F) of data each with 175 data points were obtained from the original data. Each of the group thus covered the entire range of data in original set. Different combinations of these groups were used to achieve different sizes of training and testing data sets. These arrangements are shown in Table 3.1. The data sets that were not used for training were used for testing during development of model. The developed models were tested also with all the available data from the original data set to evaluate the consistency of performance. This procedure was followed for both models (A and B).

Various network architectures and network learning parameters (momentum and epoch size) were examined, as summarized in Table 3.2. To avoid over-training of models, the original data were separated into two groups as training and testing data sets. Testing data were applied to cross-validate the model during the training stage. For this purpose, the NeuralWare software provides an option called "save best". Function of this option is to stop the training stage after a certain number of cycles, testing with testing data and saving the results of the network in a file. Each time, the result of most recent and better network replaces the last saved network. This procedure is continued until a pre-selected error of network is reached or after number of iterations without any improvement in learning of network. Finally, the "save best" network is retained, and the model is ready for verification.

Table 3.1 Combinations of training and testing data for network optimization

Group(s) in Training set	Group(s) in Testing set	Number of data pairs in Training set	Number of data pairs in Testing set
A	B, C, D, E, F	175	875
A,F	B, C, D, E	350	700
A, C, F	B, D, E	525	525
A, B, C, F	D, E	700	300
A, B, C, D, F	E	875	175

Table 3.2 Range of network architecture and learning parameters used in the study

Number of Pes		Momentum	Epoch
1 hidden layer	2 hidden layers		
2	2 & 2	0.2	4
5	2 & 5	**0.3**	8
10	5 & 5	0.4	**16**
15	10 & 5	0.6	20
-	-	0.8	32

* *Bold numbers with underscore are default values provided in software.*

Stumbo model

The number of data pairs available for this model was 396 (44 f_h/U vs. g combinations for 9 j_{cr} values) and covered a larger range of f_h/U and g as compared to Ball table. Input variables

in developing ANN model for Stumbo method were f_h/U and j_{cc} for predicting g value or alternatively g and j_{cc} to predict f_h/U. The j_{cc} variability was considered as the second input. As in Ball models, notations of Model A and Model B were used for prediction of *log* g and f_h/U values, respectively. A part of original data was used for testing. The selection of the ANN structure and related learning parameters were followed as defined for the Ball model.

Validating the ANN models

After developing ANN models for Ball and Stumbo methods, a new set of data covering a specified range of processing conditions, as shown in Table 3.3, was applied. A Lotus® based spreadsheet program for process calculations by formula methods, developed by Ramaswamy (1991), was applied to calculate the corresponding process time or process lethality. The calculated g (or f_h/U) from the processing conditions and process time/process lethality were used to recall the ANN models and obtain the corresponding ANN value. After, these recalled ANN values were applied to calculate the ANN based process time or process lethality, the accuracy of ANN predicted values were compared with the respective values from each method. The following criteria were used for evaluating ANN models performance:

Root Mean Square (RMS) of error= $\sqrt{\sum \frac{(Y_0 - Y)^2}{n}}$

Mean Absolute Error (MAE) = Mean of $|Y_0 - Y|$

Mean Relative Error (MRE, %) = Mean of $\frac{|Y_0 - Y|}{Y_0} * 100$

where:

Y_0: Desired value

Y: ANN predicted value

n: number of records

For validating ANN models of Stumbo method, the variability of j_{cc} was also considered between the processing conditions. Table 3.3 shows the range of conditions used to calculate Stumbo process time and process lethality.

Table 3.3 Range of parameters for validating ANN based Ball and Stumbo models

Parameters	Ball	Stumbo
RT (°C)	115-125	110-130
IT (°C)	65	70-90
f_h (min)	10-20	23-185
j_{ch}	1.2-2.0	1.80-2.04
j_{cc}	1.41	1.70-2.00
Process lethality (min)	5-64	5-15
Process time (min)	13.7-147	25-400

RESULTS AND DISCUSSION

ANN based Ball Model

Model development and optimization

It has been recognized that the network predictability and performance is dependent on the nature and size of training data set, structure of the network and some relevant parameters during training. Training data set should ideally cover the desired range of problem and consist of enough records to let the network learn the features. ANN models learn trend from the data in training data set. This capacity is based on the number of processing elements (PE) or more specifically number of connections. If the number of PEs are more than what is needed with respect to training of the model, the network may lose the generalization ability and begin to memorize the data which is mostly undesirable (NerualWare, 1996).

A well-trained network is optimized between generalization ability and learning the trends in training data set. To prevent the memorizing problem, a separate data set from training data should normally be selected randomly and this data set will be applied during training stage. After certain number of cycles, the training is stopped and the testing data set is used to calculate the error. After the network has converged to or reached a minimum global error, the training can be stopped.

For developing the ANN-based Ball model, the first objective was to choose the correct size of the training set, which ensured the learning and generalization ability of network

simultaneously. Figures 3.1 and 3.2 illustrate the effect of number of examples in training data set based on recalling training data set (as a measure of network ability to learn the data), as well as testing the network with testing data which were not used in training data set (as a measure of generalization ability of network). Also, each network model was tested with the entire original data set and the combined results are shown in these graphs. For both models, the errors associated with training and testing data sets were very close for each training set size, demonstrating the potential ability the of network for generalization. The results also confirm that the generalization ability of network was not affected by training data set. However, the size of the training set had a significant impact on the prediction ability of network. At the lowest level tried, the errors were quite high. For the ANN based Ball-method, a training data set of size containing as many as 175 data points could be accommodated because of the availability over 1000 data points in the original set. And even at this size, the associated errors were high and converged to lower values only after the size was increased. The network had minimum errors with a size of 525 data points which was selected for both models (A and B), the remaining being used exclusively for testing.

The error parameters for selecting the required number of hidden layers and PEs in each hidden layer are shown in Table 3.4. Changing the number of PEs in networks, with one hidden layer, generally resulted only in a small improvement in performance. However, the addition of a second layer, markedly improved the prediction ability, which was more significant for Model B than for Model A. Increasing the number of hidden layers to 3 did not result in any advantage over the model with 2 hidden layers, but increased the computation time. Hence, Model A was selected with 2 hidden layers and 2 PEs in each layer, and a 2-hidden layer network with 5 processing elements in each layer was assigned for Model B. It has been generally recognized that assigning the appropriate number of PEs and hidden layer(s) is mostly a trial and error method, however, two hidden layers for most of problems is sufficient (Neuralware, 1996).

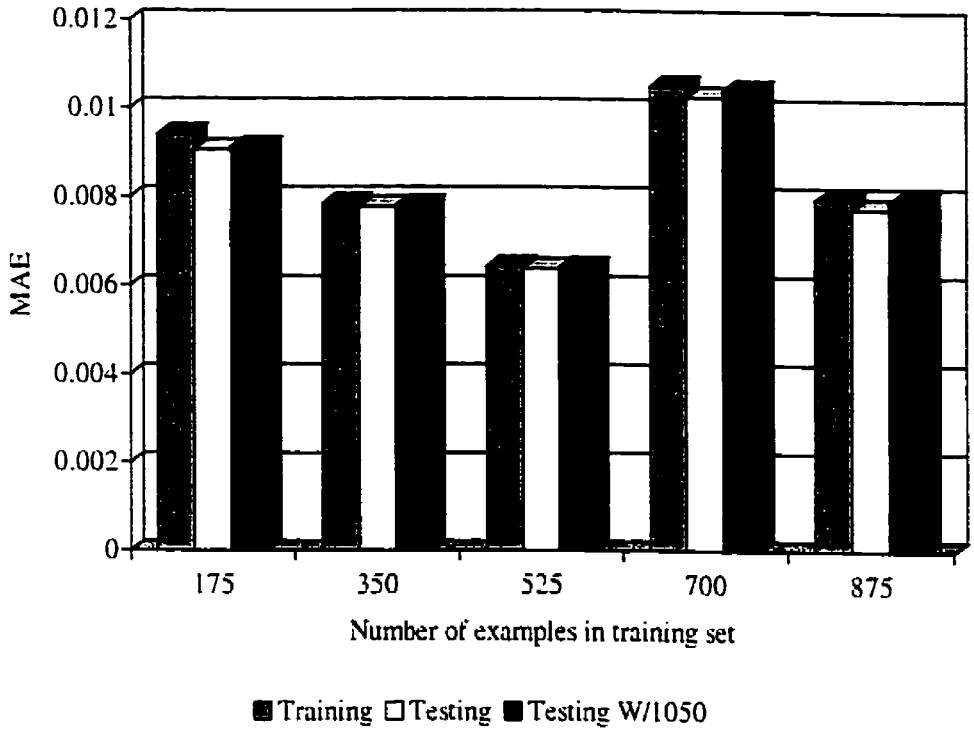


Figure 3.1 Learning curve for ANN model A (prediction *log g*)

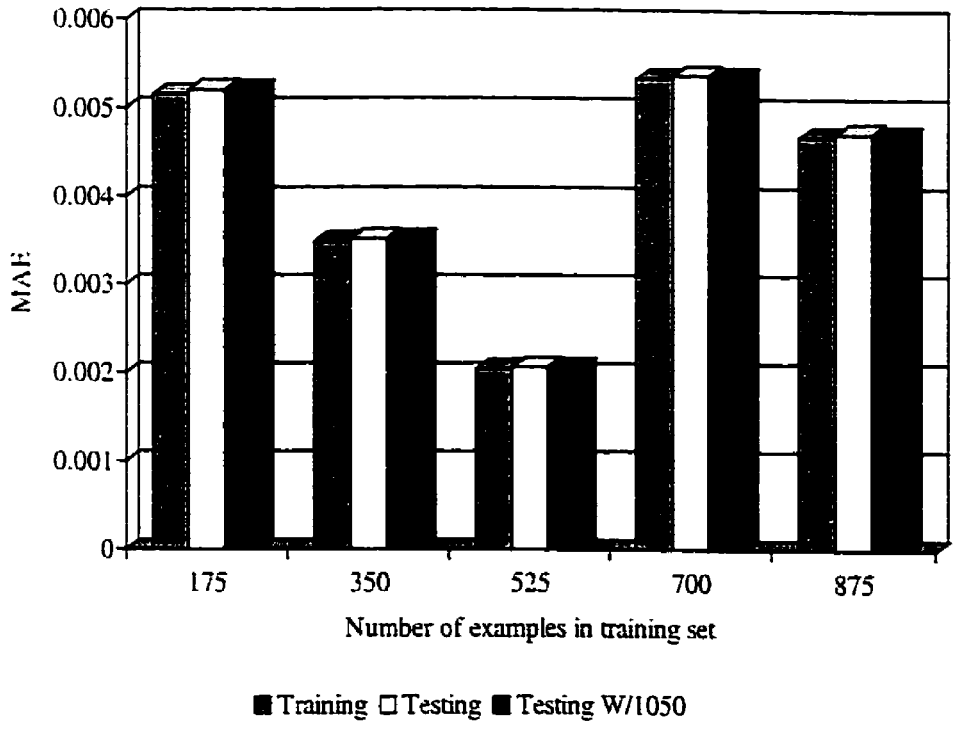


Figure 3.2 Learning curve for ANN model B (prediction f_h/U)

Table 3.4 Effect of PEs and hidden layer(s) on the performance of ANN based Ball models

Number of hidden layers	Number of PEs in layers	Model A (predicting *log g*)		Model B (predicting f_W/U)	
		RMS error	MAE	RMS error	MAE
One hidden layer	2	0.068	0.037	3.584	1.041
	5	0.067	0.038	5.792	1.709
	10	0.067	0.039	2.418	0.715
	15	0.076	0.051	2.722	0.798
2 hidden layers	2 & 2	0.043	0.024	6.332	1.881
	2 & 5	0.042	0.023	1.903	0.589
	5 & 5	0.057	0.032	1.093	0.361
	10 & 5	0.073	0.051	3.106	0.903

Several other parameters have a role in a successful training and increasing the performance of the network: momentum and epoch size, and learning rule. The epoch size was systematically varied between the limits and its effect on the RMS was evaluated (Table 3.5). Epoch size did not affect prediction ability of Model A; however, it affected the prediction ability of ANN model B. After selecting the correct epoch size, this parameter was fixed and the variability of momentum was considered. Table 3.6 shows the impact of momentum on learning of network. Increasing this parameter decreased the error parameters, but increasing it beyond some values increased the error. The increase of momentum value can speed up the learning rate and prevent the network from trapping in to local minima; however this increase can increase the error oscillation and consequently affecting the performance of model.

Table 3.5 Effect of epoch size on prediction ability: ANN based Ball models

Epoch size	Model A (predicting *log g*)		Model B (predicting f_W/U)	
	RMS	MAE	RMS	MAE
4	0.020	0.015	0.712	0.811
8	0.021	0.016	0.542	0.836
16	0.021	0.016	1.093	0.936
20	0.021	0.016	1.481	1.031
32	0.022	0.017	2.505	1.448

Table 3.6 Effect of momentum on prediction ability of ANN based Ball Models

Momentum Value	Model A (predicting *log g*)		Model B (predicting f_h/U)	
	RMS	MAE	RMS	MAE
0.2	0.047	0.024	1.756	1.112
0.3	0.039	0.018	1.426	1.010
0.4	0.043	0.024	1.093	0.936
0.6	0.045	0.028	1.610	1.057
0.8	0.080	0.050	1.721	1.071

Based on this initial exercise, the optimal configuration of the network for the two ANN based Ball models were determined and are shown in Table 3.7. The prediction performance of the ANN models with all learning parameters selected as detailed in Table 3.7 is illustrated in Figure 3.3 as plots of ANN predicted values vs. Ball table values. The predicted values were very close to the Ball table values and were evenly distributed throughout the entire range. Calculated error parameters and R^2 are given in Table 3.8 which show that ANN models were able to accurately predict Ball table parameter.

Table 3.7 Optimal levels of architectural parameters for ANN based Ball and Stumbo models

Parameter	ANN Model A *Ball method*	ANN Model B *Ball method*	ANN Model A *Stumbo method*	ANN Model B *Stumbo method*
Number of hidden layers	2	2	2	2
Number of PEs in hidden layers	2 & 2	5 & 5	2 & 2	2 & 2
Epoch size	4	8	4	8
Momentum	0.3	0.4	0.4	0.2

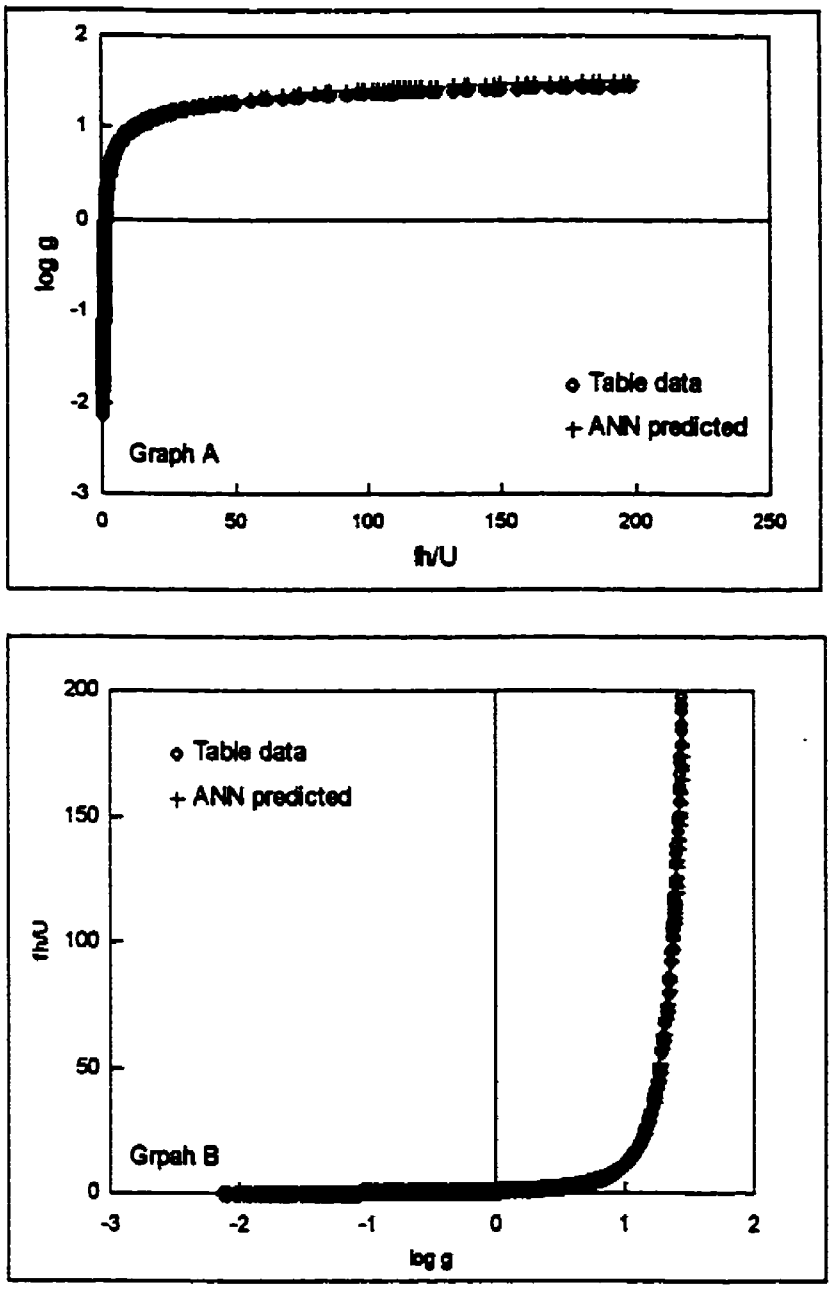


Figure 3.3 Comparison of ANN predicted values and Ball table parameters respect to input variable

Table 3.8 Error parameters for ANN based Ball and Stumbo models

Error parameters	ANN Ball Model A		ANN Ball Model B	
	Log g	Process time	f_h/U	Process lethality
RMS	0.04	0.607	2.48	0.165
MAE	0.02	0.458	0.66	0.084
MRE	4.98	1.025	1.14	1.321
R^2	0.99	0.99	0.99	0.99
Error parameters	**ANN Stumbo Model A**		**ANN Stumbo Model B**	
	Log g	Process time	f_h/U	Process lethality
RMS	0.08	6.43	3.25	0.33
MAE	0.06	5.06	1.35	0.24
MRE	6.13	3.04	2.78	3.03
R^2	0.99	0.99	0.99	0.99

Model verification

Figure 3.4 illustrates the results of validation of the developed models with a new set of data which involved a wide range of operating parameters normally encountered in commercial thermal processing operations (Table 3.3). The prediction error parameters are also summarized in Table 3.8. For both models (process time prediction. Model A and process lethality prediction. Model B) the R^2 was close to unity. which showed excellent correlation between the ANN model and Ball method. The associated mean relative errors for both process time and process lethality predictions were 1.0 and 1.3%. respectively. The mean relative error is more meaningful with respect to process lethality. while for process time. the mean absolute error (0.5 min) which indicates the error in user units (min) is more easily reconcilable. Over the range of process time encountered (up to 150 min). this represents a mean range error of about 0.3% which was considered small.

ANN based Stumbo Model

The various pre-processing constraints were also evaluated with respect to the Stumbo model although some could not be accommodated because of the smaller size of the available data. For example. since only 44 data pairs were available at each j_{cc} value. the size optimization

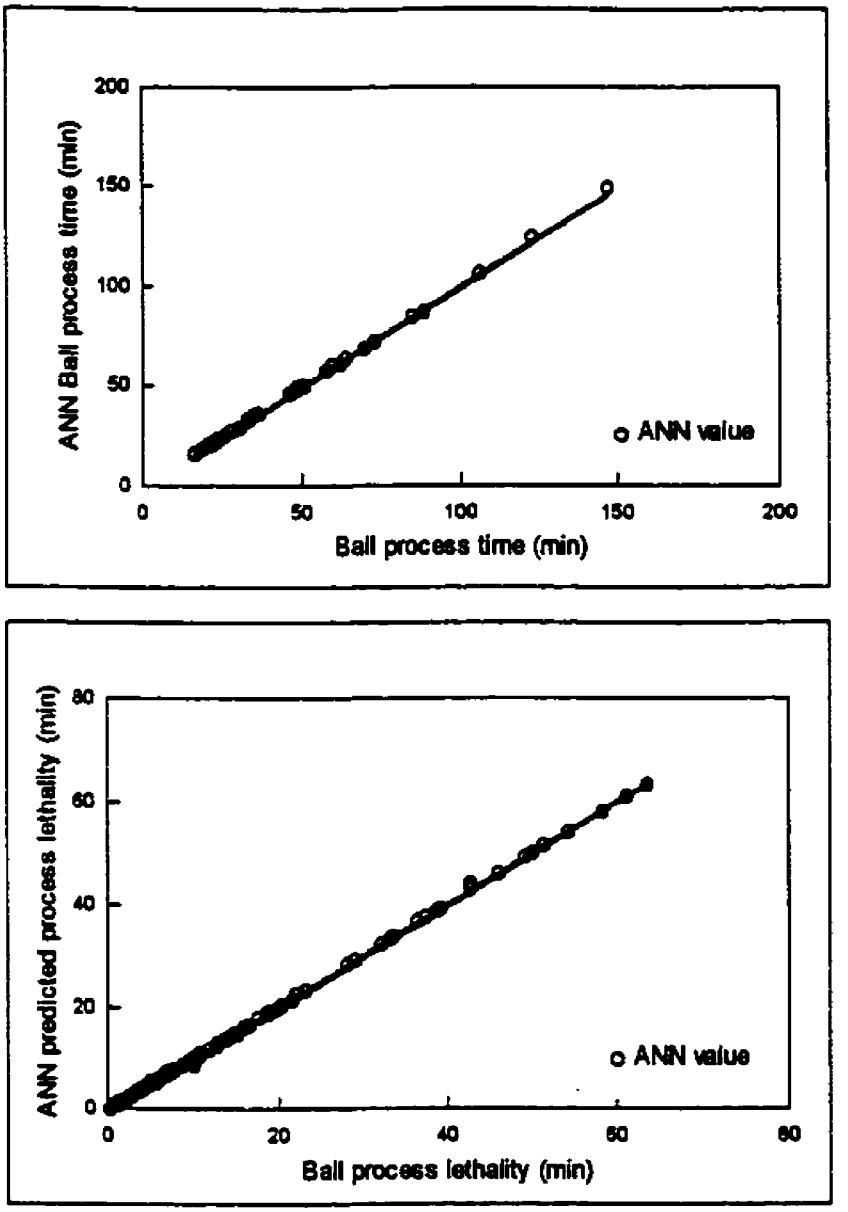


Figure 3.4 Validation of ANN models of Ball method

could not be done. The required number of processing elements and the number hidden layerswere optimized based on the RMS of errors and MAE. Since the quality of data between these two tables (Ball and Stumbo) were similar, optimal conditions with respect to momentum, transfer function, and learning rule were expected to be somewhat similar to those achieved in the ANN based Ball model. Again, 2-hidden-layer network performed considerably better than 1-hidden-layer networks. The optimal conditions for number of PEs were 2 PEs in each layer, while for one of the Ball models, the best model had 5 PEs in each hidden layer. The appropriate architecture of networks was also used to select the proper values for epoch and momentum, and these were also slightly different. The differences in the architectural configurations were considered to be the result of smaller size of data set available for Stumbo model. The optimal conditions are summarized in Table 3.7.

The performance of the ANN based Stumbo models as a plot of ANN predicted value vs. desired output value is shown in Figure 3.5. The computed errors and the associated R^2 values are included in Table 3.8. The figure demonstrates an excellent agreement between predicted and expected values with a high R^2 value. The mean errors were, however, somewhat higher than those observed with ANN based Ball models. For Model A, for example, the MRE was 6.1% for the Stumbo model as compared to 5% for Ball model for log g prediction and 2.8% and 1.1%, respectively for f_h/U predictions. The differences were considered to be primarily the result of differences in the size of available data over a wider range of f_h/U and the addition of j_{cc} as a second input. Accommodating these changes probably requires a larger training set, but the number of data in Stumbo tables were far less than that in Ball table. It was shown earlier with ANN based Ball model that the performance errors were quite high even with the availability of 175 data points. It may thus be logical to assume a higher performance error with Stumbo models.

The results for validating ANN models in the form of predicted vs. expected process time and process lethality is shown in Figure 3.6. With respect to process lethality and process lethality predictions, the associated errors (MRE) for ANN based Stumbo models, were again 3% while they were about 1% with ANN based Ball models. The associated R^2 were again very high and the quality of predictions was excellent. The slightly higher errors with predictions by the ANN based Stumbo models were again ascribed to non-availability of adequate number of data sets to properly train and develop the ANN models. The associated errors were, however,

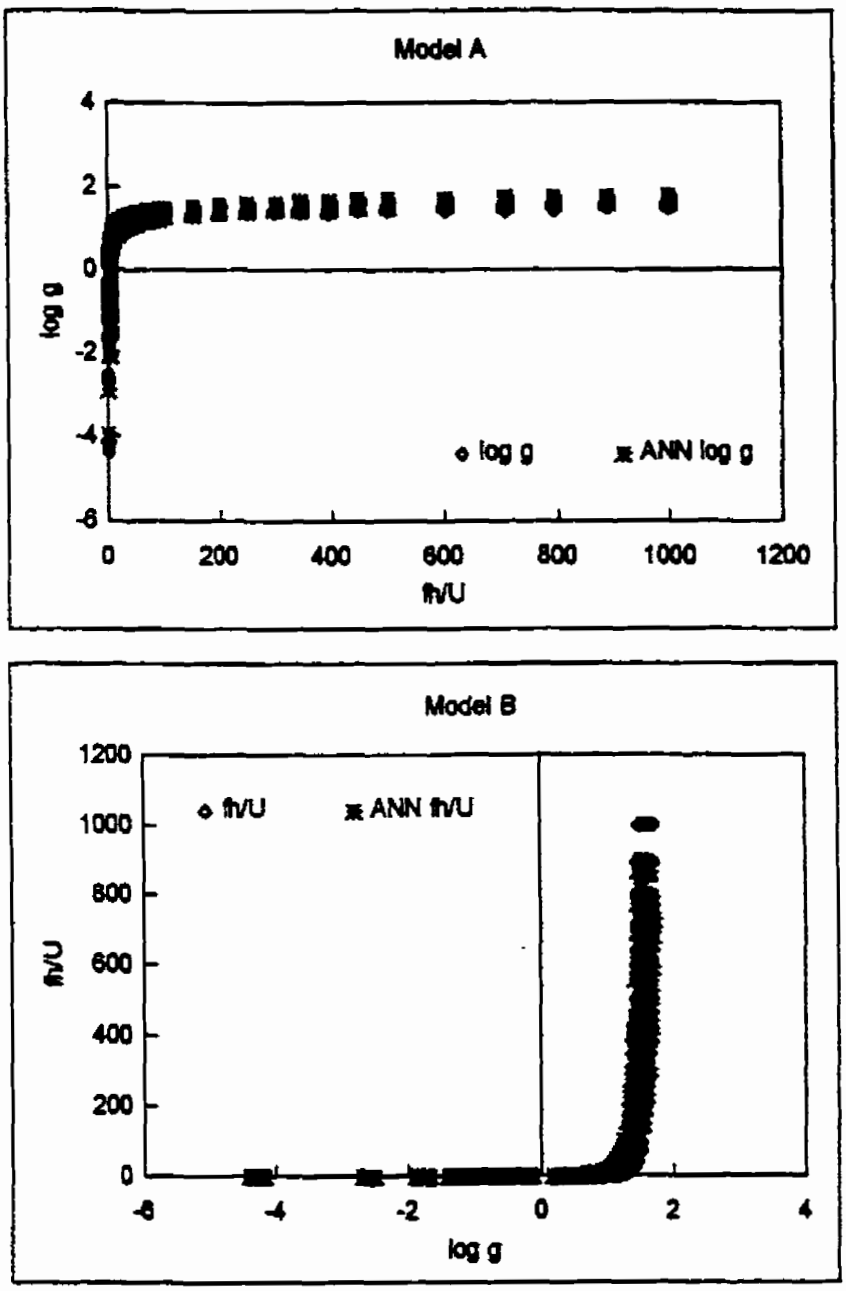


Figure 3.5 Comparison of ANN predicted values and Stumbo table parameters respect to input variable

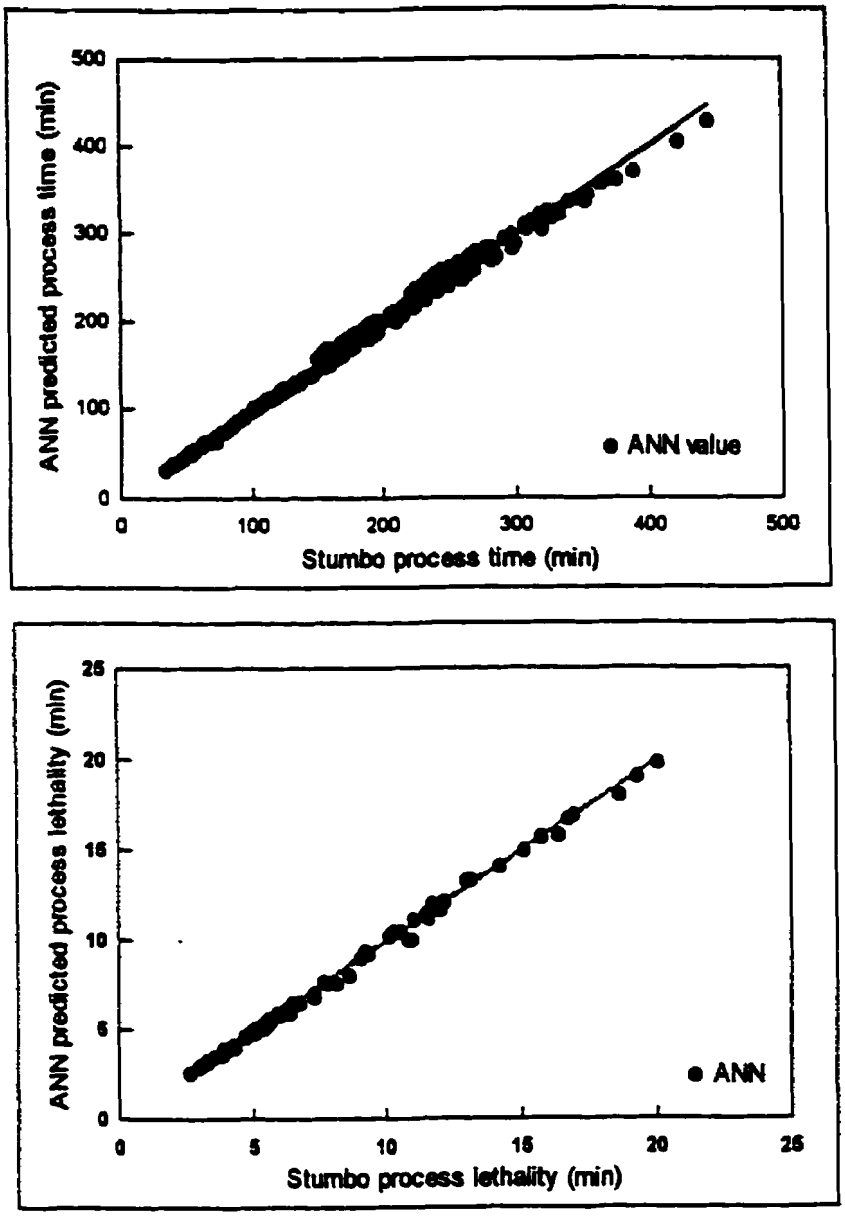


Figure 3.6 Validation of ANN models of Stumbo's method

considered low enough and demonstrate the utility of the ANN concept for simulating Ball and Stumbo process calculation techniques.

CONCLUSIONS

The feasibility of ANN modelling for process calculations was evaluated in this study based on their ability to simulate Ball and Stumbo methods. The ANN models were able to correlate the parameters of Ball and Stumbo tables with good accuracy (mean relative errors of 1-3%). It was found essential to optimize the network configuration with an appropriate size of training data set, and several learning parameters. The developed ANN models were validated with a new set of data for process time and process lethality calculations covering a wide range of commercial processing conditions. ANN models predicted the process time and process lethality with mean relative error of 1% for Ball models and 3% for Stumbo models. The slightly higher errors associated with ANN based Stumbo models were considered to be due to a wider range of parameters and smaller size of available training data set as compared to Ball model. However, the associated low errors and high R^2 value with the ANN based models in simulating Ball and Stumbo data demonstrate a good potential for the development of ANN models for process calculations.

CHAPTER 4

COMPARISON OF FORMULA METHODS OF THERMAL PROCESS CALCULATIONS FOR PACKAGED FOODS IN CYLINDRICAL CONTAINERS

ABSTRACT

Evaluating the accuracy of formula methods provides a basis for their use and a better understanding of the effect of process parameters on the required minimum process and hence on nutritional and sensory qualities of processed food. Selected formula methods were tested for accuracy against a finite difference model accommodating different processing conditions (retort and initial temperatures, thermal diffusivity, process time) and package sizes within the range used in commercial applications. For selected range of process conditions and container sizes, temperature profiles were obtained using an optimized finite difference model for conduction heat transfer involving cylindrical shapes. Time-temperature data gathered were used for computing heating and cooling parameters (f and j). Using these parameters and appropriate process conditions, process lethality/process time was computed using selected formula methods (Ball, Steele & Board, Stumbo and Pham). Prediction errors were computed based on deviations in process time/lethality from the finite difference model. In addition to comparing them over the broad range of testing conditions, the sensitivity of the methods to selected process parameters were evaluated. The parameters included retort temperature (110-130°C), product initial temperature (70-90°C), thermal diffusivity of food product ($1.0*10^{-7}$-$2*10^{-7}$ m^2/s), cans dimension (*H/D*) and the unaccomplished temperature difference (*g-value*). Retort temperature was the most significant parameter affecting the associated errors for the selected methods. Steele & Board's method had the highest error of process time calculation with a maximum 34 min over-processing for a 1-hr process. Stumbo's method had the highest accuracy, however in some cases, under-estimation of process time up to 8 min for a 1-hr process was observed. Higher *g-values* increased the calculation errors and can dimensions with *H/D* near to unity had higher errors.

INTRODUCTION

Methods of process calculations are grouped into two main categories: General and Formula methods. Using graphical or numerical procedures, the General methods integrate the experimentally gathered time-temperature data for lethal effects and therefore, this group of methods is considered the most accurate one for test situations. However, for any change in processing conditions, product formulation or container size, a new temperature profile is required. For this reason their application has been restricted to product, package and process specific conditions, under which test data were gathered. Conversely, Formula methods are developed with less restrictive assumptions on food product, package size and type and heat processes, and within a certain range, they have the flexibility for adaptation to variation in test conditions. Therefore, these methods have become more popular in food industry. The approach that Ball (1923) followed in establishing a method of process calculation is known as the first Formula method. After this work, extensive analysis has been accomplished in enhancement of this type of methods. Formula methods have been reviewed by several researchers (Hayakawa, 1977, 1978; Merson *et al.*, 1978; Stumbo and Longley, 1966; Smith and Tung, 1982; Stoforos, 1997).

The accuracy of these methods has been the core of attention of many studies to assure the safety and economic aspects of the thermal processing. While the destructive effect of thermal processes on microbial population is desirable, their deleterious effect on the nutrients and quality attributes of the food product is undesirable. In addition avoiding over-processing is a significant step in energy conservation and economic aspects of thermal processes. With respect to growing application of computers in retort control, an extensive knowledge in application of different methods of calculation becomes necessary. This concept necessitates evaluation of the accuracy of these methods with different process parameters, packaging size and type and product characteristics. Accuracy of these methods should ideally be tested against data from real time-temperature profiles. However, since such experiments are often expensive and mostly impractical over the broad range of conditions encountered in food industry, alternate methods of data gathering has been sought.

In this regard, computer generation of data based on numerical solution of finite difference models has largely replaced the need to perform extensive heat penetration tests. The study by Texieria *et al.* (1969) perhaps, provided the first successful application of computer assisted finite difference technique in thermal processing research. Since then numerous studies have made use of these technique (Teixeira and Manson, 1982; Datta *et al.*, 1986; Tucker, 1991; Teixeira and Tucker, 1997). Today it is taken for granted that these techniques are as accurate as experimental tests when the appropriate processing conditions, thermo-physical properties and boundary condition are incorporated into the model. Since the performance of these models depends on the accuracy of input data for the above variables, their accuracy should nevertheless be checked to make sure they are compatible with the test situation.

The objective of these study was to evaluate the accuracy of some selected Formula methods against a finite difference model based on conduction heat transfer in canned food products, under a wide range of processing conditions as employed in commercial applications.

Brief review of Formula methods

A few of the existing methods selected for evaluation is briefly discussed in this part. They have been detailed elsewhere (Chapter 2).

Ball's method

Introduced in 1923, Ball's formula method is the most widely used method by the food industry. The temperature prediction equations are based on the observation that the semi-logarithmic plot of temperature difference between product and heating medium temperature is a straight line after an initial lag time. Achieved process lethality and process time are computed from each other, using processing conditions (retort temperature and initial temperature) and table or graphs of related parameters (f_h/U and g). Development of these tables and graphs was carried out with respect to some limiting assumptions, which resulted in some inaccuracies in the method. The most significant assumptions were a constant cooling lag factor (j_{cc}) equal to 1.41, and equal heating and

cooling rate indexes (f_h=f_c). Obviously, for processes deviating from these assumptions, Ball's method will inaccurately compute the process time or process lethality.

Stumbo's method

While revising Ball's method to increase its accuracy, Stumbo and Longley (1966) published a new set of tables with respect to the variability of j_{cc} in real processes. The procedure for process time and process lethality calculation in this method is quite similar to Ball's method. These tables were originally based on data from hand drown heat penetration curves. On subsequent works, the parameters of table were recalculated (Stumbo, 1973) using a finite difference solution to predict the time-temperature data of product. The equation applied in this model was similar to what was used by Gillespy (1953) and identical assumptions were used for solving them.

Steele and Board's method

This method was developed following recognition of inaccuracies in Ball's formula method for thermal processes (Steele and Board, 1979). Based on the method developed by Steele and Board, the tables to be applied in thermal process calculations were based on sterilization ratios rather than temperature differences. These tables were obtained using Simpson's rule to solve the lethality equation. This sterilization value was defined as the temperature difference between the cold point of the can and the heating or cooling medium over the slope of the thermal death time curve for the microorganism of concern. Vinters *et al.* (1975) and Steele *et al.* (1979) introduced some polynomials to be used instead of developed tables in programmable calculations and on-line computers.

Pham's method

This method is a modified version of Stumbo's method (Pham, 1987 and 1990), and is based on two ranges of sterilization values. For high sterilization values (in cases that product temperature was close to retort temperature) an analytical solution was applied for conduction heat transfer in a finite cylinder with infinite heat transfer coefficients at the container walls and uniform initial temperature. For low sterilization values a numerical solution was used to solve these equations. Variability of j_{cc} value also was considered in

this method by calculating temperatures at different positions in the can. The variability of f_h and f_c was also considered in this method. This method was tested against a finite difference model simulating the center temperature in can (Pham, 1990) and the method was reported to be at least as accurate as Stumbo's method. The one table required for the use of this method substituted the 57 tables in Stumbo's method using dimensionless parameters.

MATERIAL AND METHODS

Finite difference program

A finite difference program was written in FORTRAN language and developed primarily by Sablani (1995). This program was modified for conduction heat transfer equation for cylindrical can. The measured process time and process lethality were the reference values against which the selected Formula methods were compared. The finite difference model was applied to generate temperature profiles based on achieving a selected heating lethality from which the actual process time (during which medium temperature is set to heating temperature) and delivered process lethality (total lethality calculated at the end of cooling) were computed.

The basis of finite difference models

The finite difference models were based on numerical solution of unsteady state heat conduction for an object of known geometric shape, providing transient temperature distribution throughout the container. At the beginning of the process time, all the interior points of the cylinder were set to the initial temperature of the product, while the temperature at the surface was set at the retort temperature. With a known set of initial conditions, these equations were solved at each time interval. The new temperature distribution at the end of each time interval was used to set the initial conditions for the following time interval. This procedure was continued for a predetermined process time, during which the temperature profile of product was computed. The same procedure was applied for cooling of the product by changing the ambient temperature to cooling water temperature and continuing the calculation process.

The governing partial differential equation for transient heat transfer into a finite cylinder is as follows:

$$\frac{\partial^2 T}{\partial r^2} + \frac{1}{r}\frac{\partial T}{\partial r} + \frac{\partial^2 T}{\partial h^2} = \frac{1}{\alpha}\frac{\partial T}{\partial t} \qquad (4.1)$$

The initial and boundary conditions are:

$T(r,h,0)=T_i$ at t=0

$\frac{\partial T(0,h,t)}{\partial r} = 0$ at t>0 and 0<h<L

$\frac{\partial T(r,0,t)}{\partial r} = 0$ at t>0 and 0<r<a

$k_c \cdot \frac{\partial T}{\partial r} = h_f \cdot (T_f - T_s)$ at r=a ; 0<h<L and t>0

$k_c \cdot \frac{\partial T}{\partial h} = h_{fb} \cdot (T_f - T_s)$ at h=L ; 0<r<a and t>0

where:

T: center temperature (°C)

t: time (s)

r: radial distance (m)

h: vertical distance (m)

a: radius of cylinder (m)

L: half length of cylinder (m)

h_f: fluid heat transfer coefficient ($W/m^2.K$)

k_c: cylinder thermal diffusivity (W/m.K)

α: thermal diffusivity (m^2/s)

Finite difference method is one of the techniques used to solve these types of equations in which the partial derivatives are transferred to discrete differences under appropriate boundary and initial conditions. Carslaw and Jaeger (1986) indicated in detail the analytical solution of this equation. Two major criteria in solving these equations using

numerical methods are stability and convergence of the solutions. Convergence issue is the convergence of the solutions of the difference equation to the solutions of the partial differential equations as the time and space steps are made smaller. However, stability is the question of whether numerical errors and round-off errors decrease or increase as the computation time increases. Although decreasing time step increases the convergence of the solutions, it can affect the stability of the method. Since in numerical differences are very small, and the subtracted value is divided by another very small value, the round-off error becomes predominant when the step size is reduced beyond an optimum limit. Both mentioned criteria depend on the form of initial and boundary conditions. The Crank-Nicholson scheme was used for first and second order spatial derivatives appeared in the heat flow equations and backward difference scheme for the first order derivative in boundary equations. An implicit method was used for time derivatives. Since transformed equations contain temperature values of next time steps, it was necessary to employ iterative technique in the solution procedure.

To evaluate the accuracy of finite difference solution of heat conduction and in order to optimize the required time step and space steps sizes, the time-temperatures predicted from this model were verified against the data obtained from the analytical solution of heat conduction in a finite cylinder. To evaluate the unsteady temperature in a finite cylinder, the solution of temperature ratios (U) under transient temperature for infinite slab and infinite cylinder is applied as a finite cylinder is considered as intersection of an infinite slab with an infinite cylinder (Ramaswamy *et al.*, 1982).

The equations for infinite slab and infinite cylinder obtained by Carslaw and Jaeger (1959). For an infinite slab the equation is as follows:

$$U = \sum_{n=1}^{\infty} \frac{2\sin\beta_n}{\beta_n + \sin\beta_n \cos\beta_n} .\cos(\beta_n x/L).\exp(-\beta_n^2 F_o) \qquad (4.3)$$

And for an infinite cylinder:

$$U = 2Bi\sum_{n=1}^{\infty} \frac{J_o(\gamma_n r/a)}{(Bi^2 + \gamma_n^2)J_o(\gamma_n)} .\exp(-\gamma_n^2 .F_o) \qquad (4.4)$$

where:

U: dimensionless temperature ratio = (T-RT)/(IT-TR)

RT: retort temperature (°C)

IT: initial temperature (°C)

T: temperature (°C)

Bi: Biot number

Fo: Fourier number

β: root of the characteristic equation (4.3)

γ: root of characteristic equation (4.4)

x: distance from the coldest plane of a slab

L: thickness or half thickness of a slab depending on it being heated or cooled from one side or both sides, respectively (m)

r: radius of a cylinder (m)

a: significant dimension, the radius of cylinder

J_0: Bessel function of order zero

When the heating times are sufficiently long (Fo>0.2), the series expansion of equations 4.3 and 4.4 converges rapidly, and U can be estimated accurately by the first term of the series. Therefore, the above mentioned equation of 4.3 and 4.4 for the center temperature in center of an infinite slab will be simplified as follows:

$$U_{sp} = R_p \exp(-S_p F_o) \tag{4.5}$$

And of infinite cylinder:

$$U_{sc} = R_c \exp(-S_c F_o) \tag{4.6}$$

R_p, R_c, S_p and S_c are the characteristic functions of Biot number and their value is constant for heating condition with infinite heat transfer coefficient (R_p=1.273, S_p=2.467, R_c=1.602 and S_c=5.783). Therefore, the unsteady temperature at the center of a finite cylinder (U_{sfc}) can be obtained from solution for an infinite slab and an infinite cylinder:

$$U_{sfc} = U_{sp} U_{sc} \tag{4.7}$$

The general equation to describe the temperature (T) as a function of time (t) during heating or cooling periods with known initial temperature (IT) and ambient temperature (RT) is as follows known as the Ball equation:

$$\log(|T - RT|) = -t / f + \log(j(|IT - RT|) \tag{4.8}$$

Parameters of f and j are known as the heating or cooling rate index and lag factor, respectively. These parameters can be obtained from following equations:

$$f = \frac{2.303}{(S_p / L^2 + S_c / a^2)\alpha} \tag{4.9}$$

$$j = R_p . R_c \tag{4.10}$$

The center temperature in a finite cylinder was calculated through equation 4.8 to 4.10 (analytical solutions) and the temperature profiles from the finite difference models were compared against the analytical solution of Ball equation. Along with the computation time of the model, the RMS of the error between the finite difference model and analytical solution, calculated as below, were used as two criteria of model optimization:

$$RMS\ of\ errors = \sqrt{\frac{\sum_{i=1}^{n}(T_{model_i} - T_{analytical_i})^2}{n}} \tag{4.11}$$

where:

$T_{analytical}$: temeprature predicted by analytical solution (°C)

T_{model}: temperature predicted by finite difference model (°C)

i: each time step (Sec)

n: number of points

Five different time steps of 1, 5, 10 and 20 seconds and four grid sets in both of horizontal and vertical axis (5&5, 10&10, 15&15 and 20&20) were examined.

Overall accuracy evaluation of the formula method

A wide range of processing conditions and commercial can sizes (15 can sizes) were employed as testing conditions for the finite difference program to obtain the required time-temperature data. The finite difference model was programmed to change the medium temperature from heating to cooling based on a selected heating period lethality (F_o) in the range of 5-15 minutes. Each run of the program would provide the process time, process lethality and the temperature profile of the product at the cold point of the can.

From each temperature profile, heating and cooling curves were separated in the form of logarithmic temperature difference of the product and medium vs. processing time. Determination of j and f values for heating and cooling were based on repetitive regression to locate the straight portion of heating or cooling. These values along with process time or process lethality obtained from finite difference program were compiled in a Lotus based program. Process time and process lethality for selected Formula methods were calculated using a computer program developed for this purpose (Ramaswamy, 1991). Deviations from finite difference model values (reference values) were determined as follows:

$$Error = \frac{F_{(ref)} - F_{(methods)}}{F_{(ref)}} x100 \qquad (4.5)$$

The same method was applied to evaluate the deviations in process time calculations. A negative percentage deviation in process lethality demonstrates an underestimation of the lethality with the respective Formula method (and hence undesirable), where a positive error shows inaccuracy on the conservative side. For process time computations, on the other hand, positive errors show underestimation of process time (and hence undesirable), while negative errors show overestimation of methods, but on the safe side.

The accuracy evaluation of the methods was performed from two perspectives. First, the accuracy of methods was compared as a function of one parameter, while the rest parameters had a constant value. In the second approach, the accuracy of each method was evaluated for each variable parameter over the entire set of data.

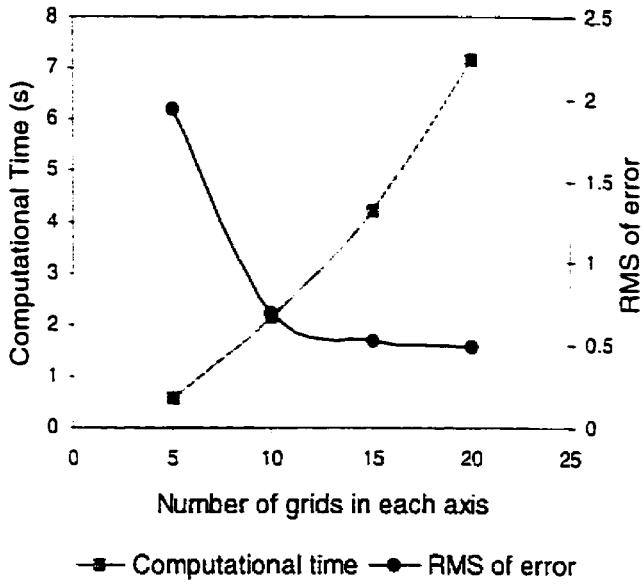
Evaluation of formula methods under specific conditions

Accuracy evaluation of the methods was based on the *g-value* and can sizes in form of height over diameter (*H/D*) as these were reported earlier as conditions under which the methods deviate from reality (Smith and Tung, 1982). A retort temperature of 120°C and initial temperature of 80°C were considered. As the tables in Ball and Stumbo methods were based on the assumption that cooling water was 100°C below retort temperature, the cooling water temperature was fixed to 20°C. The instantaneous come-up time to retort temperature and instantaneous drop in medium temperature to the cooling water temperature was assumed. Heat transfer coefficient of 5000 W/m^2C was assumed for steam as the heating medium and 500 W/m^2C for water as the cooling medium. Thermal conductivity of the product was constant (0.6 W/mC) within the simulation. In order to evaluate the effect of *g-value*, the finite difference model was programmed to change the medium temperature from heating to cooling based on reaching different *g-values* in the range of 0.05-15 C°.

RESULTS AND DISCUSSION

Optimization of finite difference model

The optimization results of the model are shown in Figure 4.1 based on the number of nodes, both in vertical and horizontal axes, as well as time step size. Increasing the number of nodes in each axis increased the computational time due to additional calculations to be performed for center temperature prediction. Smaller increments in the axes increase the accuracy of model in temperature prediction, although as shown in the Figure 4.1, after an initial sharp drop from 5 nodes to 10 nodes, increasing the number of nodes further had less effect. With this respect, a 10 by 10 grid size was selected as the optimum number of nodes in both axes. Same trend was applied to select optimum number of time steps used in computation. As computation time was increasing intensely for time step sizes less than 5 and RMS of error was not changing significantly for variation of time steps below 5 seconds, a 5 second time step size was assigned as the optimum value for this parameter in model.



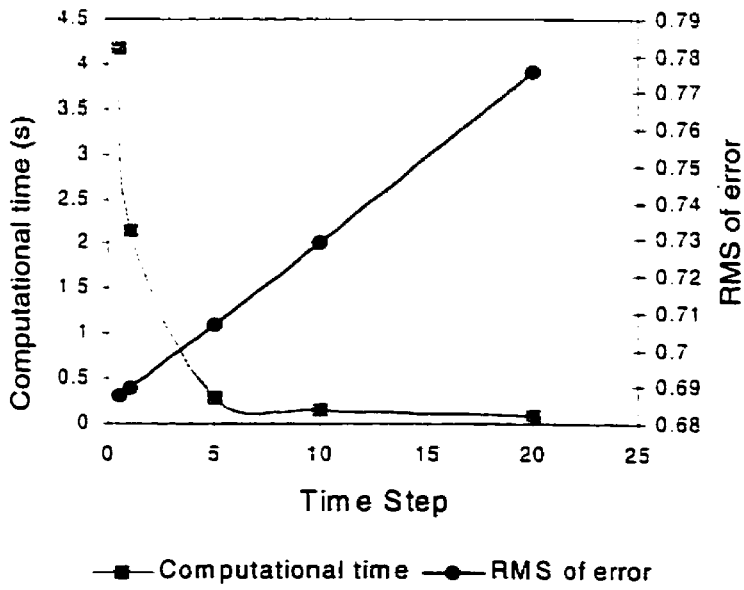


Figure 4.1 Optimization of finite difference model for a) number of nodes in each axis and b) time step

After optimization of the model, the predicted temperature profiles based on finite difference model and analytical equation were compared with each other to validate the accuracy of the finite difference model. The result of validation is depicted in Figure 4.2 in the form of center temperature profiles. The discrepancy at the beginning of the heating is for the situation when Fo<0.2 under which the two procedures were not expected to match. However, the predicted temperatures after this initial deviation coincide with each other. This should be expected as the main basis of heating equations used by each method is in agreement with the experimental data. A perfect overlapping of these two methods in this section of time-temperature curve demonstrate an excellent accuracy of the finite difference model.

Overall evaluation of formula methods

Overall evaluation of selected formula methods was carried out in two directions. First all the considered parameters were maintained at a base condition (RT=120°C, IT=80°C, α=1.5*10^{-7}m^2/s and can size=303 x 509) and one parameter at a time was varied over a selected range. Errors in process time and process lethality calculation were calculated as relative errors (percentage deviation from reference values). Subsequently, each parameter was varied over the entire range with respect to all other conditions and the associated errors were calculated. Tables 4.1 and 4.2 summarize the results of errors associated with different models for each parameter at the base condition as well as the overall. In process time calculations (Table 4.1), retort temperature was the most influencing factor compared to product initial temperature and thermal diffusivity for a can size 303 x 509. Associated errors varied from a high –7 to –21% for Steele & Board method to –0.7 to –1.5% for Stumbo's method. Ball's method was close to Steele & Board modification, while the method of Pham had errors varying from –1.5 to –8.3%. The higher error in process time using the Pham method was the result of using processing conditions and can size which resulted in high sterilization values (U/f_h<0.04), below the range of studied by Pham. Since the developed process calculation equations (Pham, 1987) were for a defined range of sterilization ratios (U/f_h: 0.04-1.5), processing conditions, which do not satisfy this range, can result in high errors in process calculations. With reference to initial temperature and methods of Ball and Steele & Board

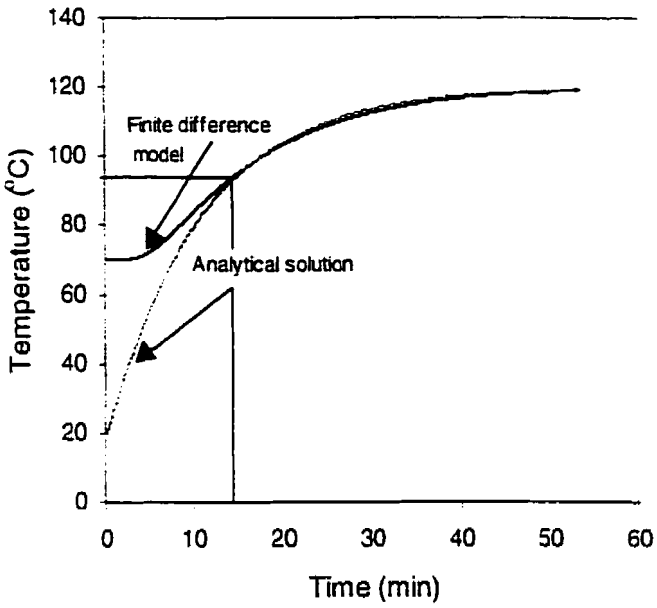


Figure 4.2 Validation of finite difference model against the analytical solution

had errors between –5 and –10%. With process lethality calculations (Table 4.2), retort temperature was the major factor; however, thermal diffusivity and initial temperature were also factors influencing the accuracy of methods. An increase in thermal diffusivity improved the accuracy of method, while an increase in retort temperature increased the associated errors.

Table 4.1 Errors in process time calculation influenced by processing conditions and thermal diffusivity of product

F_o (min)	$\alpha*10^7$ (m^2/s)	IT (°C)	RT (°C)	Can Size	Process time Mean Relative Error (%)			
					Steele	Ball	Stumbo	Pham
5	1	70	**110**	303 x 509	-7	-6.8	-1.5	-1.5
5	1	70	**120**	303 x 509	-15.5	-11.5	-0.8	-2.2
5	1	70	**130**	303 x 509	-21	-17.5	-0.7	-8.3
5	1	**70**	110	303 x 509	-8	-6.8	-1.5	-1.6
5	1	**80**	110	303 x 509	-9	-7.5	-0.9	-0.8
5	1	**90**	110	303 x 509	-10	-8.7	-1.5	-0.4
5	**1**	70	110	303 x 509	-8	-6.8	-1.5	-1.6
5	**1.5**	70	110	303 x 509	-6.7	-5.5	-0.3	-0.7
5	**2.0**	70	110	303 x 509	-6	-5.5	-0.5	-0.5

Table 4.2 Errors in process lethality calculation influenced by processing conditions and thermal diffusivity of product

F_o (min)	$\alpha*10^7$ (m^2/s)	IT (°C)	RT (°C)	Can Size	F_o Relative error (%)		
					Steele	Ball	Stumbo
5	1	70	**110**	303 x 509	20.3	17.0	3.8
5	1	70	**120**	303 x 509	56.6	47.6	4.3
5	1	70	**130**	303 x 509	80.0	73.0	4.2
5	1	**70**	110	303 x 509	20.3	16.0	3.8
5	1	**80**	110	303 x 509	20.7	17.0	2.8
5	1	**90**	110	303 x 509	21.5	18.0	3.3
5	**1**	70	110	303 x 509	20.3	17.0	3.8
5	**1.5**	70	110	303 x 509	13.2	11.5	0.5
5	**2.0**	70	110	303 x 509	10.1	10.0	1.1

Compared to the process time, the mean relative errors associated with process lethalities were much higher because of prediction of fixed low process lethality of 5.0 minutes (while the process time varied from 25 to 400 minutes). Ball and Steele & Board methods over estimated the process lethality by 17 to 80% with respect to the retort temperature, while the error range with respect to initial temperature and thermal diffusivity remained at 10-20%. Stumbo's method was fairly accurate with error ranges up to 4%. Initial temperature and thermal diffusivity have been reported to only have minimum effect on calculation accuracy (Ghazala *et al.*, 1990; Smith and Tung, 1982). In general the relative order with decreasing accuracy was Stumbo, Pham, Ball and Steele & Board. It is interesting to note that all mean relative errors were negative for process time and positive for process lethality, indicating errors generally on the conservative side for all the methods.

In order to study the overall effect of each parameter on process calculation over a broader range of processing conditions, mean and standard deviation of errors for each specific parameter was calculated and summarized in Tables 4.3 and 4.4 for process time and process lethality calculations, respectively.

In process time calculation, retort temperature and length of processing time were more obvious parameters affecting the accuracy of methods. An increase in retort temperature, increased the error terms for all the methods, except for Stumbo's method, which had the lowest error without any specific trend in mean values. All methods overestimated the process time, as shown by the negative sign for mean errors except for Stumbo's method, which underestimated process time under some conditions. Some earlier researchers have observed similar results. However, it should also be noted that Stumbo's method was in fact most accurate amongst the different methods. The mean errors were less than 1% for all situations. On an average, this accounts for a 2 minutes error over a 200 minutes process time, obviously small. In comparison, Ball and Steele & Board methods had up to 15 and 20% over estimation, respectively. Pham's method was close to Stumbo's method, with mean errors mostly in the conservative side. However, as indicated by the accompanying standard deviations of the mean errors (up to 2%), it is logical to assume situations did exist for over estimation of process time. As discussed

earlier. larger errors of high retort temperatures using Pham method is due to resulted sterilization values. which were not defined by Pham (Pham, 1987. 1990)

Table 4.3 Means and Standard Deviations (SD) of relative errors in process time calculation for selected process calculation methods

Parameter	Range	Steele		Ball		Stumbo		Pham	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Retort temperature	110 (°C)	-6.03	2.44	-6.20	3.06	0.71	2.43	-0.25	0.51
	120 (°C)	-12.93	3.66	-9.91	2.72	-0.55	0.64	-0.35	0.90
	130 (°C)	-19.40	4.43	-15.64	3.46	0.48	1.19	-3.36	2.62
Initial temperature	70 (°C)	-11.3	5.25	-9.55	3.92	0.48	1.56	-1.76	2.20
	80 (°C)	-12.45	6.00	-10.57	4.48	0.32	1.58	-1.51	2.18
	90 (°C)	-14.63	7.62	-12.48	5.66	-0.16	1.85	-0.69	2.01
Thermal diffusivity	$1*10^{-7}$ (m^2/s)	-14.45	6.90	-11.51	5.20	-0.21	1.34	-1.83	2.75
	$1.5*10^{-7}$ (m^2/s)	-12.58	6.34	-10.93	4.97	0.24	1.65	-1.25	1.95
	$2*10^{-7}$ (m^2/s)	-11.34	5.97	-10.08	4.31	0.61	1.93	-0.87	1.56
Can Size	Small	-11.22	5.85	-1.011	4.39	0.43	2.08	-1.05	1.58
	Medium	-10.43	6.60	-9.65	5.16	0.82	2.12	-0.87	1.76
	Large	-13	6.51	-10.38	4.74	0.21	1.19	-1.12	2.41
Processing time	Low	-15.78	5.05	-12.71	4.28	-0.03	1.14	-1.80	2.10
	Medium	-11.8	5.57	-9.34	3.67	0.51	1.31	-0.79	2.10
	Long	-8.83	5.06	-7.87	3.76	0.58	2.06	-0.41	1.17

Based on fill weight: small: (<0.5 kg). Medium: (0.5-1 kg). Large: (>1 kg): Processing time: Low: (<100 min). Average: (100-200 min). Long: (>200 min)

Trends observed with process lethality were similar (Table 4.4). Conditions which. under estimated process time. obviously over estimated the process lethality and vice versa. With process lethality. Ball and Steele & Board methods had up to 70% over estimation. Methods of Stumbo and Pham performed better with mean errors less than 7%.

Table 4.5 summarizes the parametric range. mean and standard deviation of errors for combination of parameters. The maximum error in process time calculation belonged to the method of Steele & Board with about 37% mean relative error. while method of Stumbo had only 5% maximum mean relative error. However, as it is been reported

(Smith and Tung, 1982; Ghazala *et al.*, 1990) there were some cases that Stumbo's method under estimates the process time as shown in Table 4.3 with a positive sign.

Table 4.4 Means and Standard Deviations (SD) of relative errors in process lethality calculations for selected process calculation methods

Parameter	Range	Steele		Ball		Stumbo		Pham	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Retort temperature	110 (°C)	11.36	7.20	11.12	5.70	-0.81	3.95	-0.23	4.40
	120 (°C)	40.77	11.66	33.03	10.70	1.88	2.33	7.07	4.18
	130 (°C)	70.9	8.15	57	6.66	-3.2	6.86	29.85	10.12
Initial temperature	70 (°C)	40.37	25.66	31.22	19.06	-2.06	5.32	-0.25	0.35
	80 (°C)	40.85	25.85	31.62	19.19	-1.09	4.81	-0.22	0.30
	90 (°C)	41.43	26.34	31.73	19.45	0.95	5.06	-0.15	0.25
Thermal diffusivity	$1*10^{-7}$ (m^2/s)	45.68	26.13	31.85	19.36	0.36	5.20	-0.23	0.29
	$1.5*10^{-7}$ (m^2/s)	40.49	25.87	31.35	19.16	-0.75	4.91	-0.25	0.38
	$2*10^{-7}$ (m^2/s)	36.87	25.21	31.51	19.16	-1.85	5.32	-0.19	0.24
Can Size	Small	36.46	25.21	31.85	19.61	-0.69	5.44	-0.29	0.39
	Medium	41.54	25.82	30.58	19.43	-0.66	4.81	-0.16	0.23
	Large	50.76	25.31	32.71	17.65	-0.94	5.37	-0.17	0.22
Processing time	Low	55.33	18.86	42.26	16.30	-0.63	5.82	18.46	13.77
	Medium	36.83	25.78	25.62	16	-0.62	4.83	-0.33	0.42
	Long	16.63	17.58	15.62	13.69	-1.09	4.22	-0.19	0.24

Based on fill weight: small: (<0.5 kg), Medium: (1-0.5 kg), Large: (>1 kg); Processing time: Low: (<100 min), Average: (100-200 min), Long: (>200 min)

Table 4.5 Maximum, minimum, mean and standard deviation of errors for different methods

Methods	%Error in process time calculations				%Error in process lethality calculations			
	Min	Max	Mean	SD	Min	Max	Mean	SD
Steele	1.75	37.60	18.44	4.21	2.02	89.09	41.01	25.99
Ball	1.63	26.31	15.61	3.29	0.75	68.02	31.57	19.22
Stumbo	0.01	5.18	1.03	0.69	0.06	19.05	2.60	3.50
Pham	0.00	8.64	2.71	2.18	0.04	6.05	0.60	0.60

Specific evaluation of formula methods

Errors based on the can shape or *H/D* ratio had a consistent pattern for different Formula methods. The results are presented in two different formats for comparative purposes. First the different methods are compared at a given *g-value*. Then each method is compared at different *g-values*. In both situations *H/D* ratio was used as the basis.

The comparison of different Formula methods at a given *g-value* is shown in Figures 4.3 and 4.4 for the process time and process lethality calculations, respectively. The maximum error for almost all the methods was observed for *H/D* near to unity. Smith and Tung (1989) and Pham (1990) observed the same trend in error patterns. Steele & Board's method had the largest errors (-2 to -20%) among evaluated methods and Stumbo's and Pham's methods were the most accurate methods. Stumbo's method predicted process time with accuracy of 1% for *g-value* of 0.05°C and 9% for *g-value* of 15°C and Pham's method had similar accuracy in process time calculation. They appeared to deviate from each other more at g=15°C.

Trends were similar with respect to prediction of process lethality except that the errors were on the opposite side. Steele & Board method had the highest positive error (20-70%) while Stumbo's method had the least (<20%). Pham's method as suggested is only for estimation of process time, and for process lethality, the method had to be solved using a trail and error approach. While this approach was employed for some test situations, it was not possible to extend this approach for all the test situations because of the time consuming iteration process. Since methods of Pham and Stumbo were reported be similar in most cases, similar trends could be probably assumed for predicting the process lethality.

With respect to sterilization values, Lenz and Lund (1977) also considered processes with large *g-values* uncertain. The larger errors for processes at large *g-values* are due to accomplishment of a large part of sterilization during cooling section. In these processing conditions, cooling time and also the product properties during cooling can be the reason for the high errors in total sterilization value of the process. However, for processes with smaller *g-values*, which is more common, the error values are small ranging from 10% for Steele and Board's method to 1% for Stumbo and Pham methods.

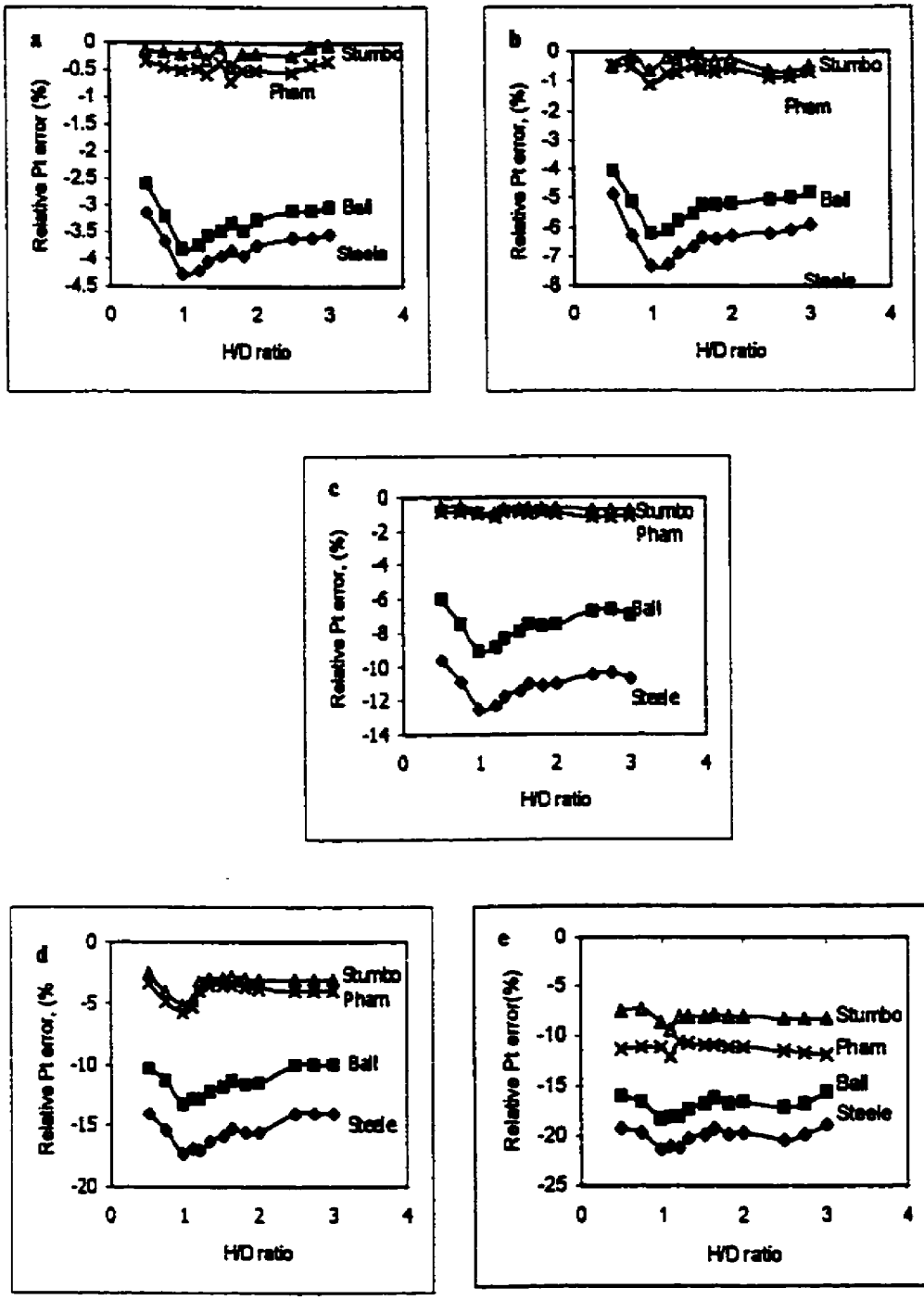


Figure 4.3 Errors in process time calculations for different g-values (a: g=0.05 C°, b: g=1 C°: c: g=5 C°; d: g=10 C°; e: g=15 C°)

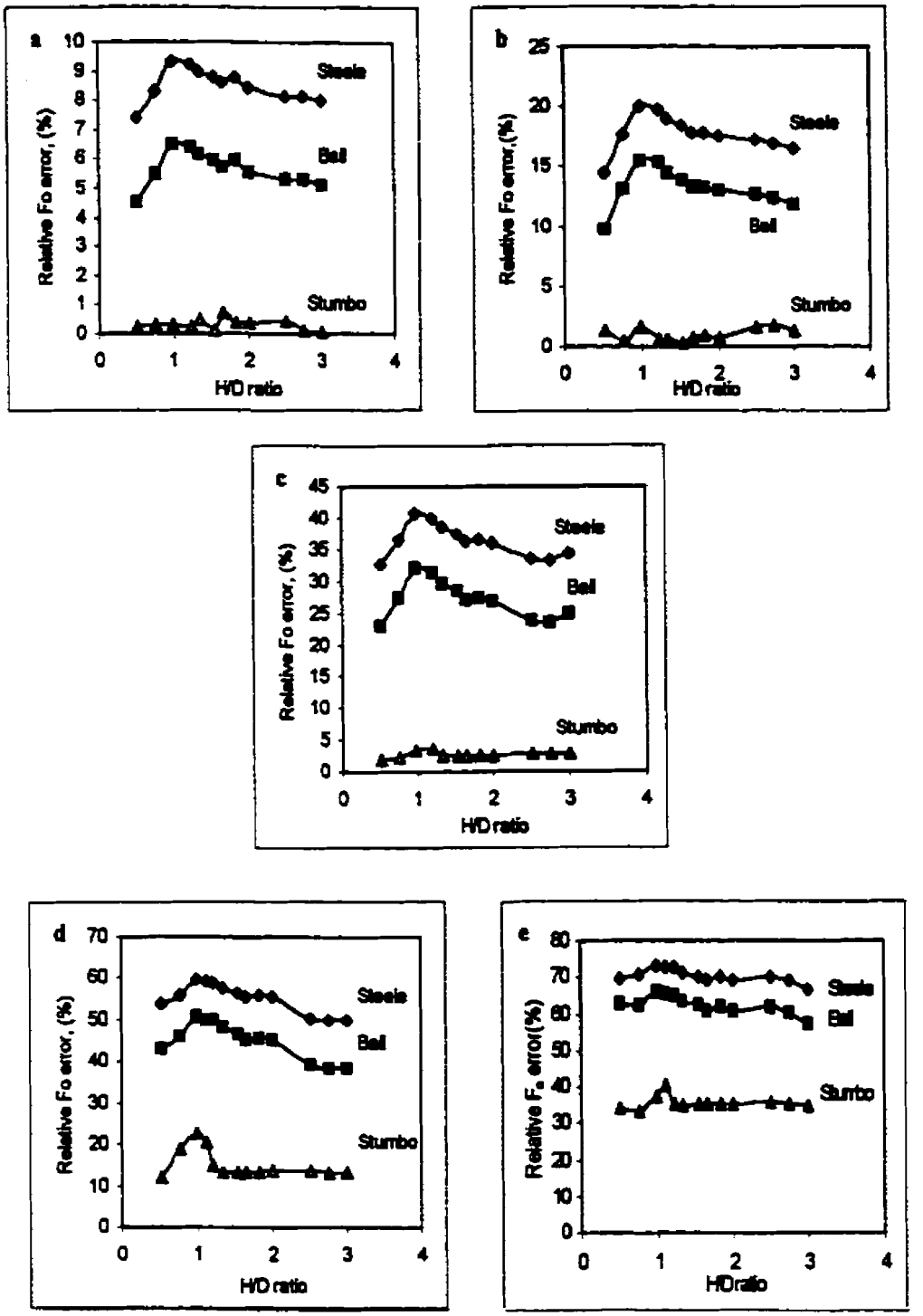


Figure 4.4 Errors in process lethality calculations for different g-values (a: g=0.05 C°, b: g=1 C°; c: g=5 C°; d: g=10 C°; e: g=15 C°)

The effect of *g-value* on deviations in calculation is also due to temperature difference at the end of heating between heating medium and the product. For smaller *g-values*, when the product temperature and retort temperature at the end of heating are close to each other, the effect of temperature gradient is insignificant. However, for larger *g-values*, when this difference is larger, this effect can cause errors in calculation. In this case, while the retort temperature is changed from heating to cooling water, there would be a longer lag for center of the can to receive the effect of cooling water temperature, and even after cooling was began, the center temperature will continue increase for a short while. None of the Formula methods correctly account for this effect and deviate from reality at larger *g-values*.

Figures 4.5 and 4.6 show the performance of each formula method for at different *g-values*. The effect of *g-value* on each method can be seen more clearly in this presentation. With an increase in *g-value*, the associated errors for each method increased, showing a systematic increase in errors for methods of Ball and Steele & Board. With respect to Stumbo and Pham methods, the errors were small up to a *g-value* of 5°C, but then the inaccuracy begins to increase markedly. Again Stumbo and Pham had the highest accuracy. The deviations in process lethality and process time prediction increased for *H/D* close to unity and for larger *H/D* values the errors decreased.

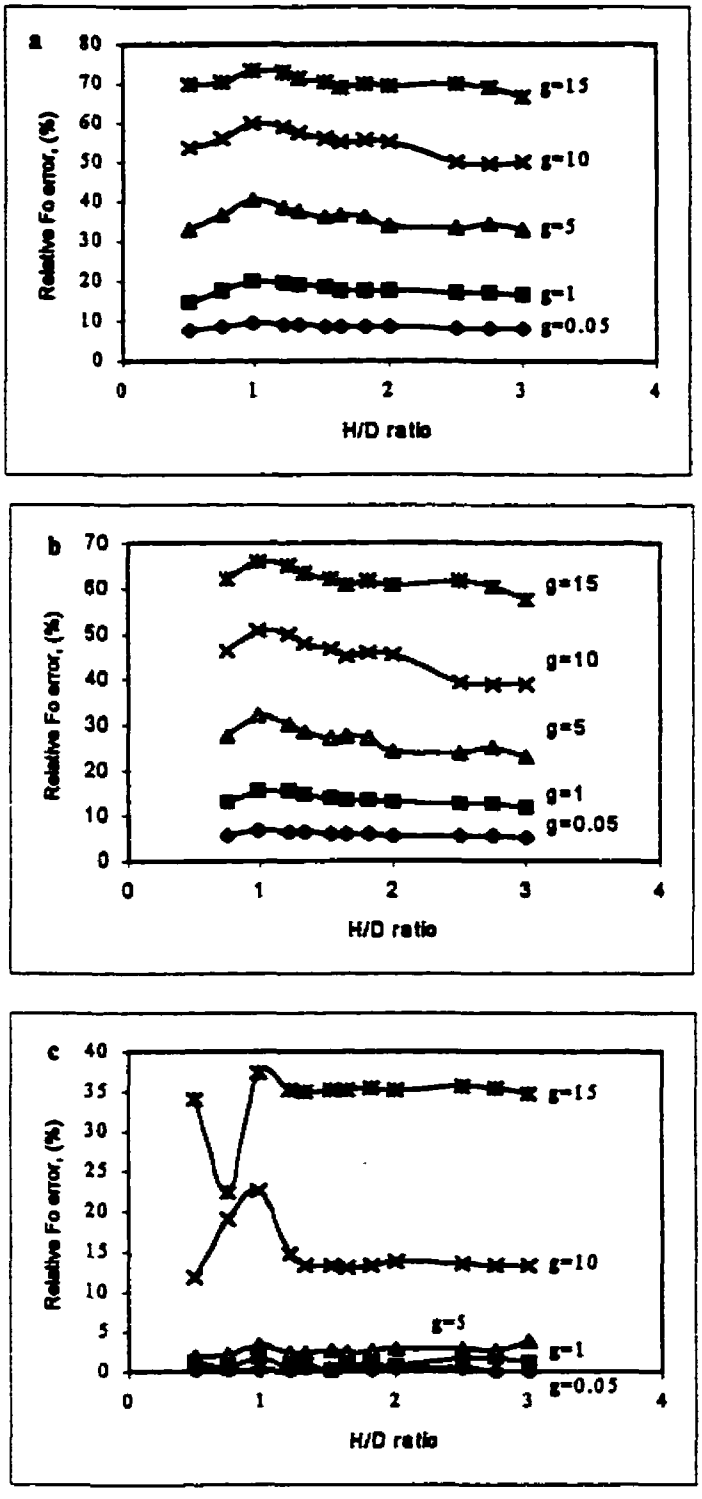


Figure 4.5 Errors for process lethality calculations with respect to different g-values (a: Steele & Board's method, b: Ball's method; c: Stumbo's method)

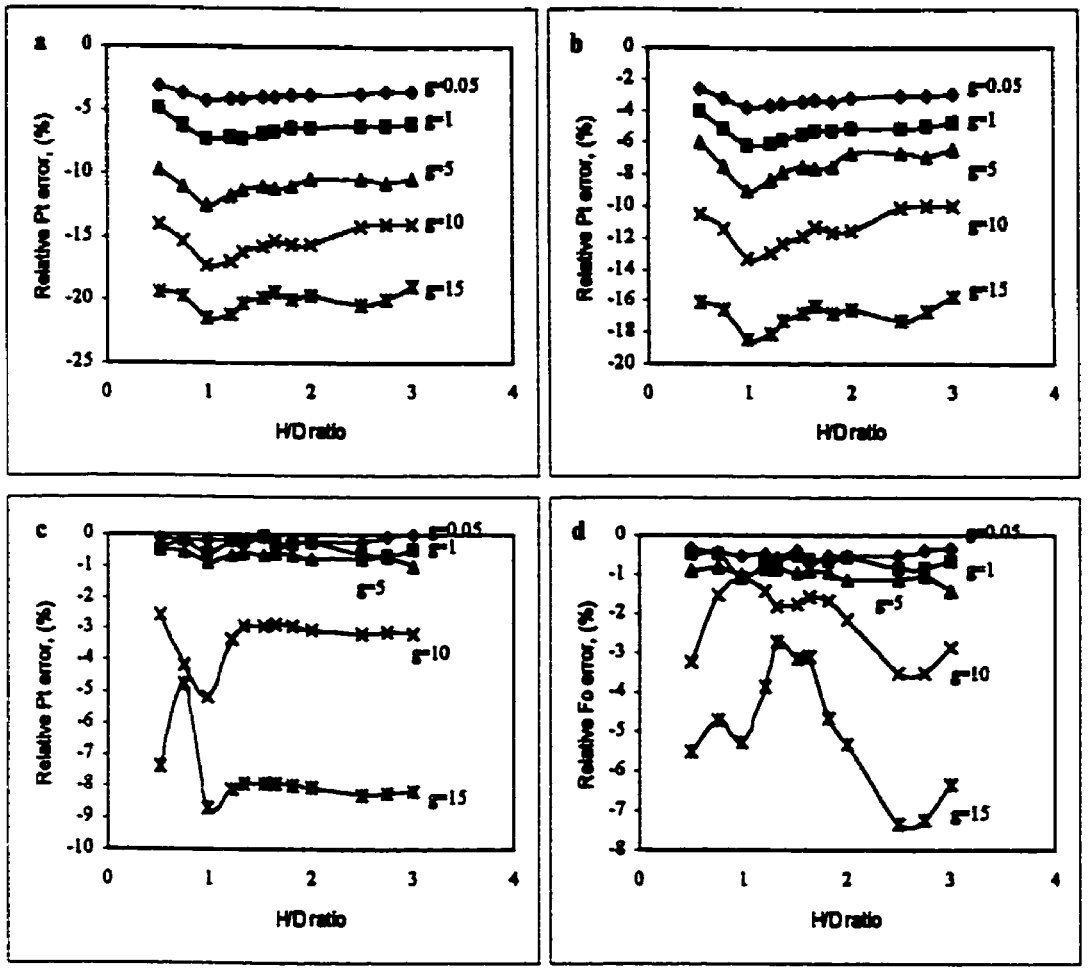


Figure 4.6 Errors for process time calculations with respect to different g-values (a: Steele & Board's method, b: Ball's method; c: Stumbo's method; d: Pham's method)

CONCLUSIONS

The accuracy of selected formula methods was determined with respect to data generated from a finite difference model. Finite difference models have the potential for accommodating variations in processing conditions and can size, and provide data for a wide range of processing conditions. The optimization of the finite difference model against an analytical solution of heat transfer equation with 10x10 grid size predicted accurate temperature profiles. Process calculation errors were evaluated based on the variation of each parameter. Retort temperature was one of the most major parameters influencing the accuracy of the selected process calculation methods. Most methods overestimated the process time with the exception of Stumbo's method for some cases of underestimation. Based on the can shapes, errors had an increase for *H/D* values near to unity. Errors increased at larger *g-values*, where a substantial amount of process sterilization is accomplished during the early cooling stage. This trend in errors was observed for all the methods. Stumbo's method was considered the most accurate method, following it was Pham's method.

CHAPTER 5

ARTIFICIAL NEURAL NETWORK MODELS AS ALTERNATIVES TO THERMAL PROCESS CALCULATION METHODS

ABSTRACT

Artificial Neural Network (ANN) is a computing system capable of processing information by its dynamic state response to external inputs. ANNs learn from examples through iteration by adjusting the internal structure to match the pattern between input and output. In this study, finite difference simulations, which are widely recognized as practical alternatives to experimental methods, were used to generate temperature profiles for a wide range of can sizes under various processing conditions. The time-temperature data were used to compute the process lethality, process time as well as heat penetration parameters: f_h, j_{ch}, f_c and j_{cc}. These data were used for developing the ANN models (both training and testing). The accuracy and ability of ANN models were compared with selected Formula methods, both with respect to process time and process lethality computations. Process calculation results from ANN model were comparable to Pham's and Stumbo's methods, which were previously evaluated as the most accurate amongst the process calculation methods.

INTRODUCTION

Artificial Neural Networks (ANNs) have been the focus of interest in many diverse fields of science and technology. Neural networks are basically computer models, which simulate very simple abilities of our brain. Rather than being programmed for use in a particular application, neural network models generate their own rules by learning from provided examples. ANN has the ability of approximating arbitrary continuous functions based on a set of given observations. As they obtain this ability through the stage of learning, they are known as truly adaptive systems, which do not need any prior knowledge of the nature of relationships between the set of parameters. ANN models are thought to be robust to noise and inconsistencies in data, rendering them more advantageous compared to empirical models. Also, ANN models can be multiple-input and multiple-out (MIMO) systems (Baughman and Liu, 1995). Therefore, variability of multiple parameters in the development of an ANN model is possible. NeuralWare (1995) provides a wide overview of potential applications of neural network as classification, prediction, data association and optimization. Most food related work involves estimation, prediction and control (Eerikainen *et al.*, 1993).

With respect to ANN characteristics and their abilities in modeling, their application has been considered in the development of a general model in thermal process calculations. Accurate determination of thermal processes is required with respect to assurance of safety, as well as retention of quality attributes and nutritional values of food products. For this purpose, different methods have been developed for facilitating thermal process calculations. All these different methods are aimed at determining the lethality (F_o) for a given process or calculating the process time for achieving a given process lethality. In order to establish thermal processes, time-temperature data of the product undergoing the thermal processes is required. More versatile, less expensive and less time consuming alternative of experimental procedures are mathematical models.

Finite difference models of heat transfer into packaged food have been successfully applied in optimization and control (Teixeira *et al. a, b*, 1969; Datta *et al.*, 1986; Teixeira and Manson, 1982; Tucker, 1991; Teixeira and Tucker, 1997). The main feature of these models is the prediction of temperature profile based on the governing heat transfer

equations of packaged food products. Finite difference models require some specification of the food product and system such as thermal diffusivity of the food product, heat transfer coefficient of the heating or cooling medium and thermal conductivity of package. When these conditions are known, the time-temperature data can be obtained at any specific point of the packaged food product for any processing condition and package sizes. The ease of use of finite difference model can be considered in providing an expanded data set, including a wide range of parameters.

In Chapter 3, the potential of ANN to simulate two of the most widely used process calculation methods was demonstrated. In Chapter 4, the prediction errors associated with the available process calculation methods were evaluated using those computed from a finite difference numerical model as reference. It was shown that the prediction errors vary depending on the processing conditions, and by far Stumbo and Pham's method are the most accurate. In Chapter 4, a basis was also established for generating parametric values such as f_h, f_c, j_{cc} and j_{ch} for various operating conditions involving different retort and initial product temperature, thermal diffusivity of the product and can size as well as the associated process time and lethality. This provides a wide range of input-output parameters for establishing and/or verifying useful process calculation models such as they are aimed in the present research-an ANN based process calculation model.

Hence, the specific objective of this chapter was to develop an ANN model as an alternative to existing thermal process calculation methods using the wide range of data set obtained from the appropriate finite difference models and initial conditions.

MATERIAL AND METHODS

Data generation

In order to obtain the training data set for ANN models development, a wide range of processing conditions, product characteristics and packaging sizes were considered as detailed in Chapter 4 and summarized in Table 5.1. The defined range of each parameter covered the common range of processing conditions, and different can sizes were selected from those used in industry. A finite difference program written in Fortran language

(Sablani. 1995) was applied to obtain the related time-temperature profile to each set to conditions. This model was based on the conduction heat transfer for cylindrical packaged food products. The heat transfer coefficient of heating medium (steam) and cooling medium (water) were assumed to be 5000 (W/m^2C) and 500 (W/m^2C), respectively. Each combination of defined values was applied as the initial inputs of finite difference model to obtain the respective temperature profiles.

Each time-temperature profile was divided into two semi-logarithmic curves of heating and cooling. In order to locate the beginning of straight line in each curve an iterative regression technique was applied and f_h, f_c, j_{cc} and j_{ch} were obtained for each set of condition. Hence a data set was obtained which, consisted of processing condition (initial temperature, retort temperature), can size (height and diameter), product characteristic (thermal diffusivity), process time, process lethality and respective heating and cooling parameters. This data set with 1215 data records was considered as the basis for ANN models training.

Table 5.1 Range of parameters used in finite difference program

Retort Temp. (°C)	**Initial Temp. (°C)**	**Thermal diffusivity $*10^7(m^2/S)$**	**Heating lethality (min)**
110	70	1.0	5.0
120	80	1.5	10.0
130	90	2.0	15.0
Can size no.	**Can size**	**Approximate fill wt. (kg)**	**Radius*Hieght/2(cm)**
1	6 z	0.170	2.70 x 4.50
2	8 z tall	0.240	3.41 x 4.13
3	No.1 picnic	0.300	3.41 x 5.08
4	No.211 cylinder	0.370	3.41 x 6.19
5	No. 300	0.410	3.81 x 5.04
6	No.1 tall	0.454	3.89 x 5.95
7	No.303	0.454	4.05 x 5.56
8	No.303 cylinder	0.595	4.05 x 7.07
9	No.2 vacuum	0.340	4.36 x 4.29
10	No.2	0.568	4.36 x 5.80
11	No.2 cylinder	0.738	4.36 x 7.30
12	No.2 ½	0.822	5.16 x 5.95
13	No.3 cylinder	1.420	5.40 x 8.89
14	No.5	1.645	6.51 x 7.15
15	No.10	3.000	7.86 x 8.89

ANN model development

A precise ANN model training requires selection of independent input variables (Baughman and Liu, 1995). If a prior relation between input variables in not available, one of the proposed methods to select the independent variables between input variables, is to discard one variable and train the ANN model based on the remained set of data. Obviously, discarding variables, which are function of the other, will not affect the training. However, when a science of relation between variables is provided, this information should be applied to reduce the number of input variables and increase the accuracy of ANN model. Based on the information obtained, while developing ANN model for methods of Ball and Stumbo, the key variables for relating process time (*g-value*) and process lethality (f_h/U) were considered as the major inputs. As the variation of j_{cc} values is not integrated, therefore j_{cc} was considered as another independent input for each model.

Input variables (*g-value* or f_h/U) were calculated based on the procedure applied in the Ball method (Figure 5.1). Accordingly, the combination of these variables is used in form of *log g* or f_h/U, along with variability of j_{cc}, related the process time to process lethality or vise versa. The procedures of ANN model development were followed much closely to the ANN model development of Stumbo's method. The data set was divided to training subset and testing subset to be used during the training procedure.

NeuralWorks Professional II/Plus, version 5.23 (NeuralWare Inc., Pittsburgh, PA) was employed for ANN modeling. A standard backpropagation algorithm with Tangent hyperbolic transfer function and normalized cumulative delta learning rule were selected for ANN models training. Same formats of input and output variables as detailed in the Stumbo ANN model were applied. For predicting the process time, f_h/U were used as well as j_{cc} to predict *g-value* and subsequently, for predicting the process lethality, *g-value* and j_{cc} were used to predict f_h/U. The predicted parameters from ANN models were applied to calculated the respective process time and process lethality. For convenience Model A and Model B were used to name each model for process time and process lethality prediction, respectively.

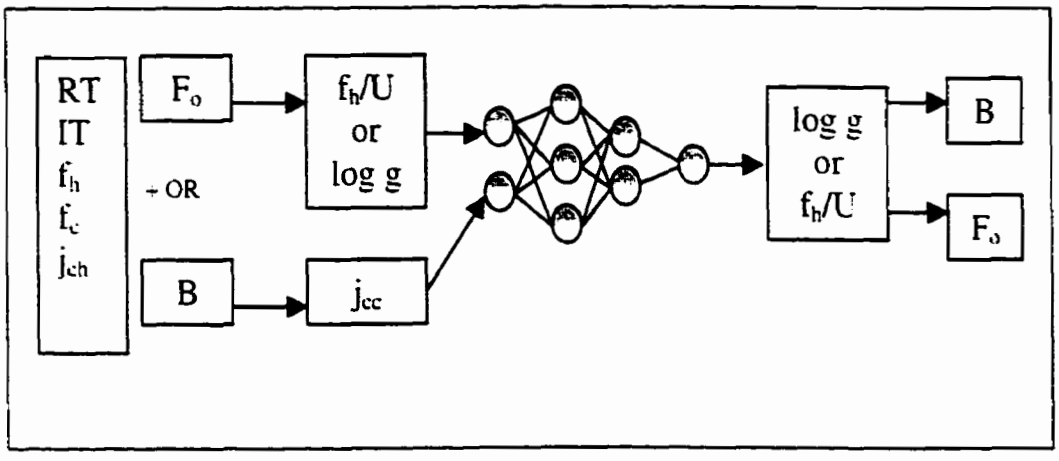


Figure 5.1 Arrangement and procedure of developing ANN models

Optimization of ANN models

The ANN models were optimized for the number of data records in training set, number of processing elements, momentum value and epoch size. The assigned values for each of these parameters are summarized in Tables 5.2 and 5.3. The first step was to select the optimum number of data in training set with respect to default values of network architecture. The selected training set sizes were applied to choose the optimum topology of the networks. Afterwards, two learning parameters, momentum and epoch, were tested.

The following criteria were used for evaluating ANN models performance:

Root Mean Square (RMS) of error= $\sqrt{\sum \frac{(Y_0 - Y)^2}{n}}$

Mean Absolute Error (MAE) = Mean of $|Y_0 - Y|$

Mean Relative Error (MRE, %) = Mean of $\frac{|Y_0 - Y|}{Y_0} * 100$

where:

Y_0: Desired value

Y: ANN predicted value

n: number of records

Table 5.2 Combinations of training and testing data for network optimization

Group(s) in Training set	Group(s) in Testing set	Number of data pairs in Training set	Number of data pairs in Testing set
A	B, C, D, E, F	202	1010
A,F	B, C, D, E	404	404
A, C, F	B, D, E	606	606
A, B, C, F	D, E	808	808
A, B, C, D, F	E	1010	202

Table 5.3 Range of network architecture and learning parameters used in the study

Number of Pes		Momentum	Epoch
1 hidden layer	2 hidden layers		
2	2 & 2	0.2	4
5	2 & 5	**0.3**	8
10	5 & 5	0.4	**16**
15	10 & 5	0.6	20
-	-	0.8	32

** Bold numbers with underscore are default values provided in software.*

Comparison of ANN models

The predicted ANN values of process time and process lethality were compared with respective values of the selected formula methods. The same methodology as the one used in comparison of formula methods was applied to compare the performance of the developed ANN models. All the predicted values from either ANN model or formula methods were compared against the reference model, which was the data from finite difference model. Relative error computed as follows, was applied as the error parameter:

$$Error = \frac{F_{(ref)} - F_{(methods)}}{F_{(ref)}} x100 \qquad (5.1)$$

This error parameter was calculated based on lethality and therefore a positive error of lethality calculation demonstrates an underestimation, while a negative sign shows an overestimation of process time. The error signs for under and overestimation are reversed, while predicting the process time.

RESULTS AND DISCUSSIONS

ANN models optimization

Figures 5.2 and 5.3 shows the required number of data for Model A and Model B, respectively. In Model A, the performance of the network was not a function of training set size, and a training set with 808 data (50%) was selected. However, Model B was more

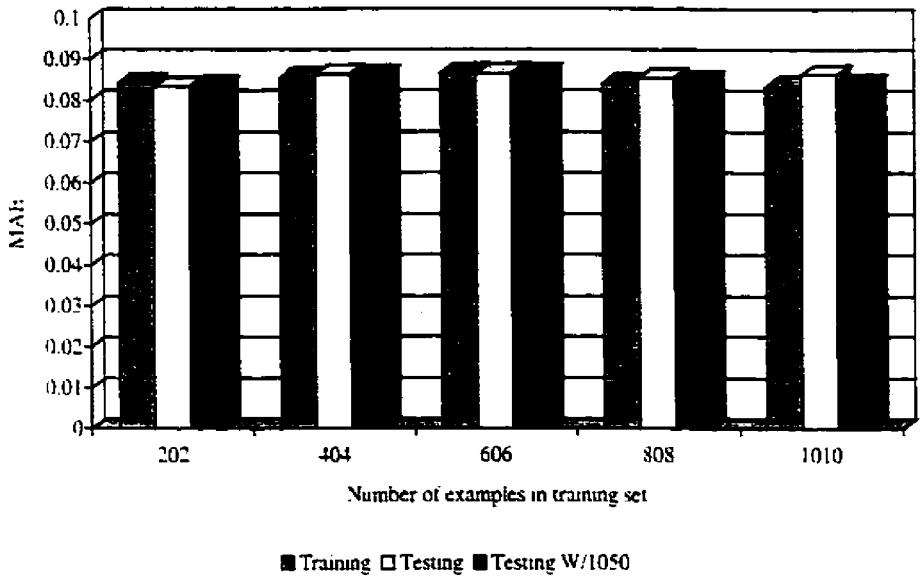


Figure 5.2 Learning curve for ANN model A

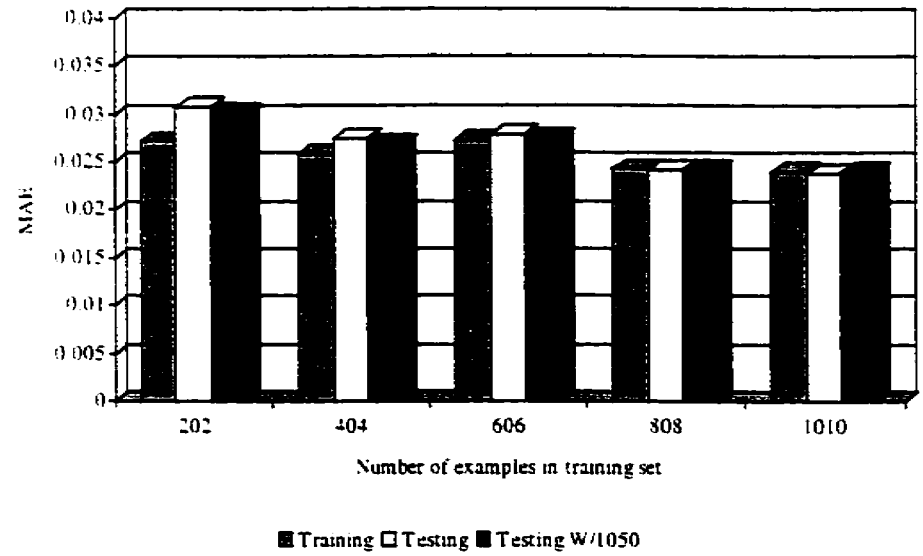


Figure 5.3 Learning curve for ANN model B

dependent on the training set size, and a data set with 1010 data, which had the lowest error of prediction was used. Although the source of data for both models was same, the nature of variables affected the performance of the network and the question how well network can learn depended on the input variable.

Figure 5.4 depicts the effect of PEs and number of hidden layers for Model A and B. In both models, 2 hidden layer networks had a significantly higher performance (Baughman and Liu, 1995; Swingler, 1996). A network with 2 and 5 PEs in each hidden layer had the highest performance for Model A and 2 PEs in each hidden layer for Model B, which were selected as the optimum topology. Table 5.4 shows the result of learning parameters variations for both models. The effects of momentum and epoch were not significant for Model A and increasing momentum increased the error. A momentum value of 0.4 and epoch size of 4 was selected for Model A. The effect of these parameters was even less significant for Model B. Therefore, the default value of momentum and epoch of 4 were selected.

Table 5.4 Effect of momentum and epoch size on the performance of Model A and B

Momentum	**MRE**		**Epoch**	**MRE**	
	Model A	*Model B*		*Model A*	*Model B*
0.2	0.039	0.029	4	0.035	0.028
0.3	0.039	0.029	8	0.037	0.028
0.4	0.041	0.029	16	0.041	0.029
0.6	0.049	0.029	20	0.043	0.029
0.8	0.116	0.030	32	0.053	0.030

Performance of the ANN models

The performance of ANN models A and B are shown in Figure 5.5 as plots of ANN predicted values vs. reference values from the finite difference simulation. The predicted values for both models were very close to desired values and were evenly distributed throughout the entire range. The associated error parameters are summarized in Table 5.5. High R^2 values shows the excellent performance of both ANN models in predicting the output variables.

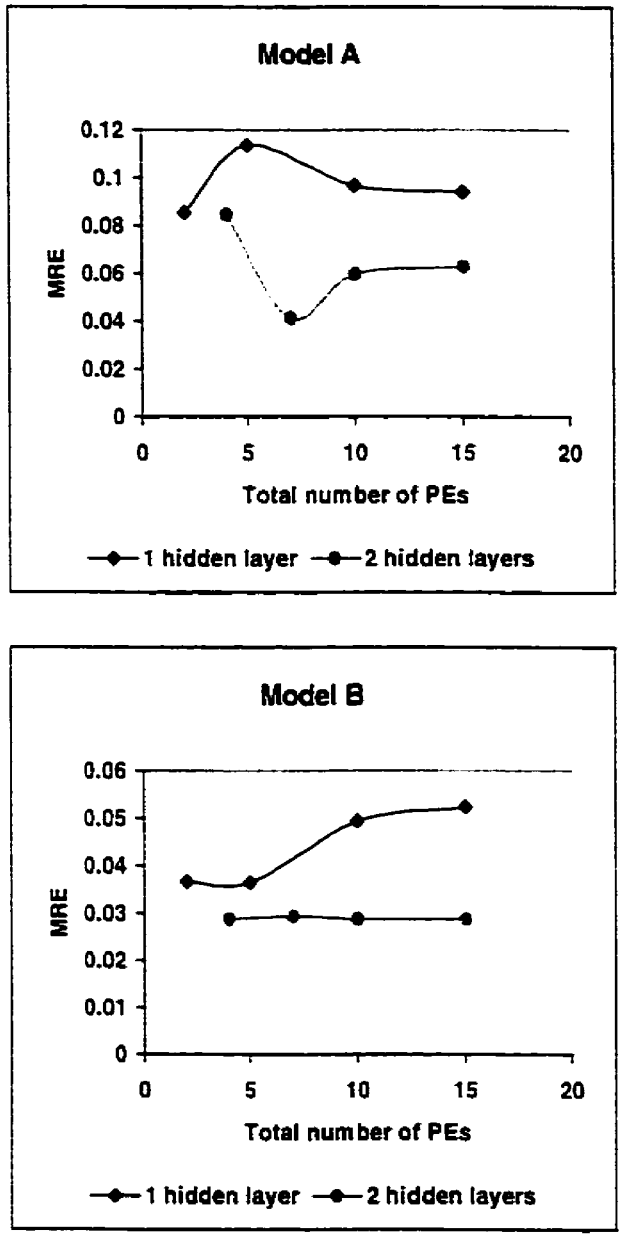


Figure 5.4 Effect of PEs and hidden layers on the performance of model A and B

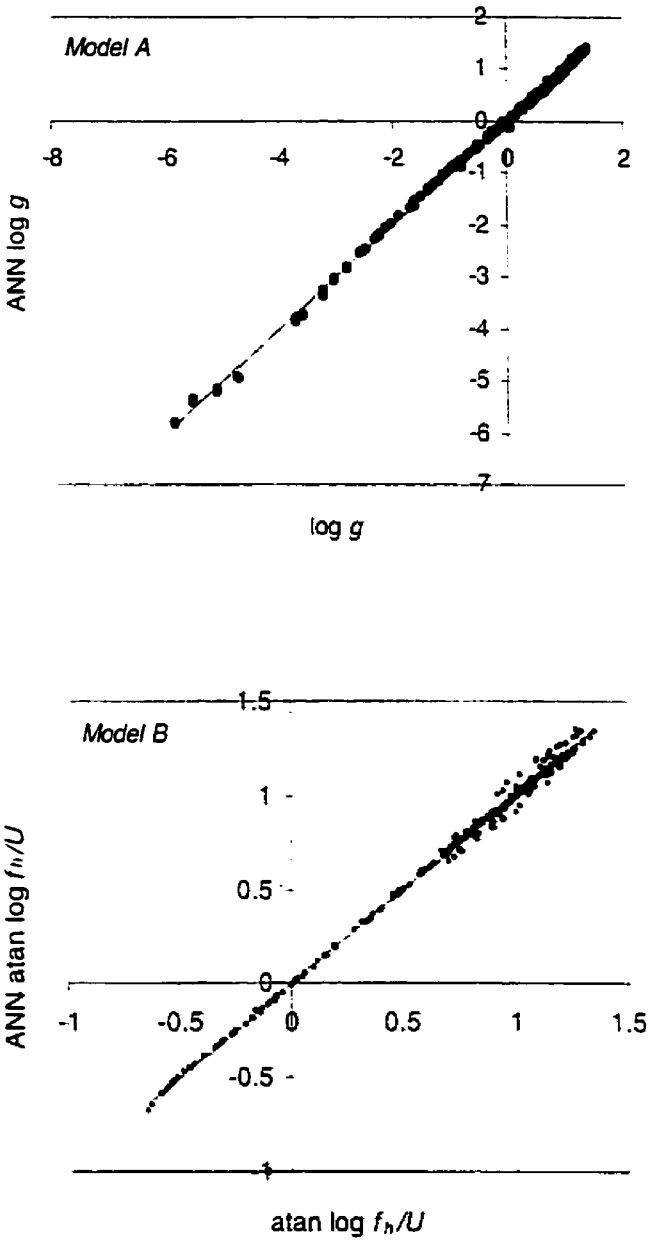


Figure 5.5 Performance of ANN models A and B with respect to predicted output

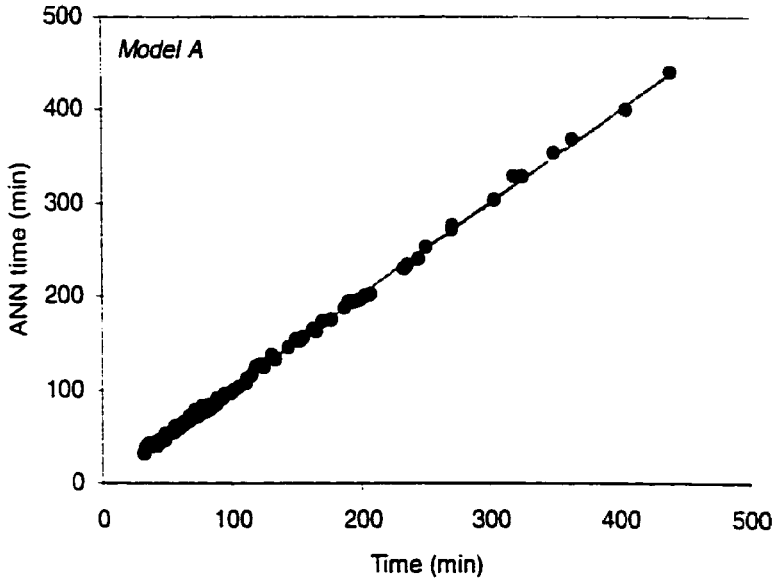
The performance of ANN models is shown in Figure 5.6 in form of process time and process lethality calculated from ANN predicted *g-value* or f_h/U vs. the process time and process lethality obtained from the finite difference model. Both ANN models showed a good performance for predicting process time and process lethality. The error parameters of ANN models prediction are summarized in Table 5.5. Both models had a high R^2 close to unity showing excellent correlation between ANN predicted values and reference values. RMS of errors for process time prediction was higher than process lethality since process time had higher values compare to process lethality, however close mean relative errors of both models shows close performance of models. However, with respect to process lethality mean relative error is more important error parameter and provides better evaluation. In process time calculation, mean average error of 2.7 minutes indicated high accuracy of ANN model in process time prediction applied range of this parameter (between 40–450 minutes)

Table 5.5 Error parameters for performance and verification of ANN models A and B

Error parameters	ANN Model A		ANN Model B	
	log *g*	Process time	Atan log f_h/U	Process lethality
RMS	0.05	4.17	0.02	0.41
MAE	0.03	2.70	0.01	0.27
MRE	2.93	2.14	4.44	2.74
R^2	0.998	0.997	0.997	0.999

Comparison of ANN models with existing formula methods

The performance of ANN models were compared for range of processing conditions, thermal diffusivity of product, can size and processing time with the performance of three selected Formula methods of Ball, Stumbo and Pham. The performance of each method is showed based on the mean of relative errors and standard deviation of error for both process time and process lethality (Tables 5.6 and 5.7). Although no particular trend was observed in mean and SD of errors based on ANN models, the performance of ANN models was fairly in range of methods of Stumbo and Pham, if not better. It should be mentioned these two methods have the highest accuracy of



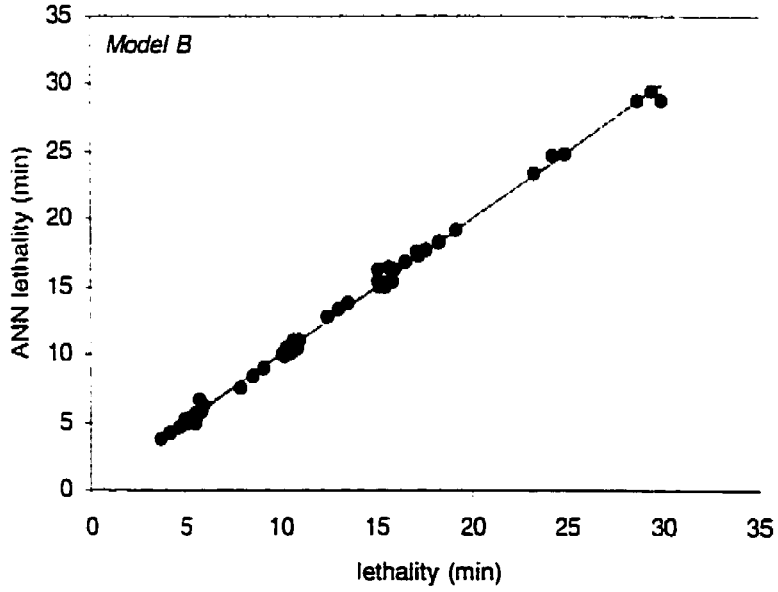


Figure 5.6 Performance of ANN models A and B in process time and process lethality prediction

calculations compared to other formula methods (Smith and Tung, 1982: Pham. 1990; Stoforos. *et al.*, 1997).

The general performance of the selected formula methods and the ANN models are compared in Figures 5.7 and 5.8. Ball method, for considering the restrictive assumptions, had the most deviation while other methods have a close performance (R^2>0.99). The same trend was applied while comparing the performance of other methods and ANN models with respect to process lethality calculations. As the finite difference program was run based on a pre-assigned heating lethality, the process lethality is clustered around three fixed lethality values (one of the reasons for higher errors).

Table 5.8 shows the range of errors for process time and process lethality for formula methods and ANN models. The range of errors as well as the mean and SD of ANN model for process time prediction was in range of Pham's method. Also, in process lethality prediction the range of errors varied between Stumbo's method and Pham's method. The range of errors in process lethality calculations was very close to Pham's method and the same trend in range of errors applied for process time calculations.

Table 5.6 Means and Standard Deviations (SD) of relative errors in process time calculations for selected process calculation methods

Parameter	Range	Ball		Stumbo		Pham		ANN	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Retort temperature	110 (°C)	-6.2	3.06	0.71	2.43	-0.25	0.51	-0.05	2.18
	120 (°C)	-9.91	2.72	-0.55	0.64	-0.35	0.9	-2.47	1.88
	130 (°C)	-15.64	3.46	0.48	1.19	-3.36	2.62	2.48	3.36
Initial temperature	70 (°C)	-9.55	3.92	0.48	1.56	-1.76	2.2	0.18	3.00
	80 (°C)	-10.57	4.48	0.32	1.58	-1.51	2.18	0.07	3.21
	90 (°C)	-12.48	5.66	-0.16	1.85	-0.69	2.01	-0.31	3.53
Thermal diffusivity	$1*10^{-7}$ (m^2/s)	-11.51	5.2	-0.21	1.34	-1.83	2.75	0.32	3.87
	$1.5*10^{-7}$ (m^2/s)	-10.93	4.97	0.24	1.65	-1.25	1.95	-0.15	3.02
	$2*10^{-7}$ (m^2/s)	-10.08	4.31	0.61	1.93	-0.87	1.56	-0.22	2.76
Can Size	Small	-1.011	4.39	0.43	2.08	-1.05	1.58	-0.29	2.52
	Medium	-9.65	5.16	0.82	2.12	-0.87	1.76	0.65	2.92
	Large	-10.38	4.74	0.21	1.19	-1.12	2.41	0.84	3.64
Processing time	Low	-12.71	4.28	-0.03	1.14	-1.80	2.10	0.20	3.03
	Medium	-9.34	3.67	0.51	1.31	-0.79	2.10	1.20	3.66
	Long	-7.87	3.76	0.58	2.06	-0.41	1.17	0.11	3.32

Based on fill weight: small: (<0.5 kg). Medium: (0.5-1 kg). Large: (>1 kg); Processing time: Low: (<100 min). Average: (100-200 min). Long: (>200 min)

Table 5.7 Means and Standard Deviations (SD) of relative errors in process lethality calculations for selected process calculation methods

Parameter	Range	Ball		Stumbo		Pham		ANN	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Retort temperature	110 (°C)	11.12	5.70	-0.81	3.95	-0.23	0.32	-0.23	4.40
	120 (°C)	33.03	10.70	1.88	2.33	-	-	1.74	3.14
	130 (°C)	57	6.66	-3.2	6.86	-	-	-0.23	4.40
Initial temperature	70 (°C)	31.22	19.06	-2.06	5.32	-0.25	0.35	-1.91	5.20
	80 (°C)	31.62	19.19	-1.09	4.81	-0.22	0.30	-1.25	4.89
	90 (°C)	31.73	19.45	0.95	5.06	-0.15	0.25	0.36	4.71
Thermal diffusivity	$1*10^{-7}$ (m^2/s)	31.85	19.36	0.36	5.20	-0.23	0.29	-0.43	5.33
	$1.5*10^{-7}$ (m^2/s)	31.35	19.16	-0.75	4.91	-0.25	0.38	-0.72	4.37
	$2*10^{-7}$ (m^2/s)	31.51	19.16	-1.85	5.32	-0.19	0.24	-1.59	5.36
Can Size	Small	31.85	19.61	-0.69	5.44	-0.29	0.39	-1.25	5.12
	Medium	30.58	19.43	-0.66	4.81	-0.16	0.23	-0.31	3.79
	Large	32.71	17.65	-0.94	5.37	-0.17	0.22	-1.14	6.50
Processing time	Low	42.26	16.30	-0.63	5.82	-	-	-1.78	4.78
	Medium	25.62	16	-0.62	4.83	-0.33	0.42	-0.07	4.93
	Long	15.62	13.69	-1.09	4.22	-0.19	0.24	-0.14	5.49

Based on fill weight: small: (<0.5 kg), Medium: (0.5-1 kg), Large: (>1 kg); Processing time: Low: (<100 min), Average: (100-200 min), Long: (>200 min)

Table 5.8 Range, mean and standard deviation of errors for all the combination of parameters

Methods	Error in process time calculations				Error in process lethality calculations			
	Min	Max	Mean	SD	Min	Max	Mean	SD
Ball	1.63	26.31	15.61	3.29	0.75	68.02	31.57	19.22
Stumbo	0.01	5.18	1.03	0.69	0.06	19.05	2.60	3.50
Pham	0.00	8.64	3.25	1.96	0.00	6.05	0.60	0.60
ANN	0.01	8.92	2.71	2.18	0.06	6.63	1.60	3.10

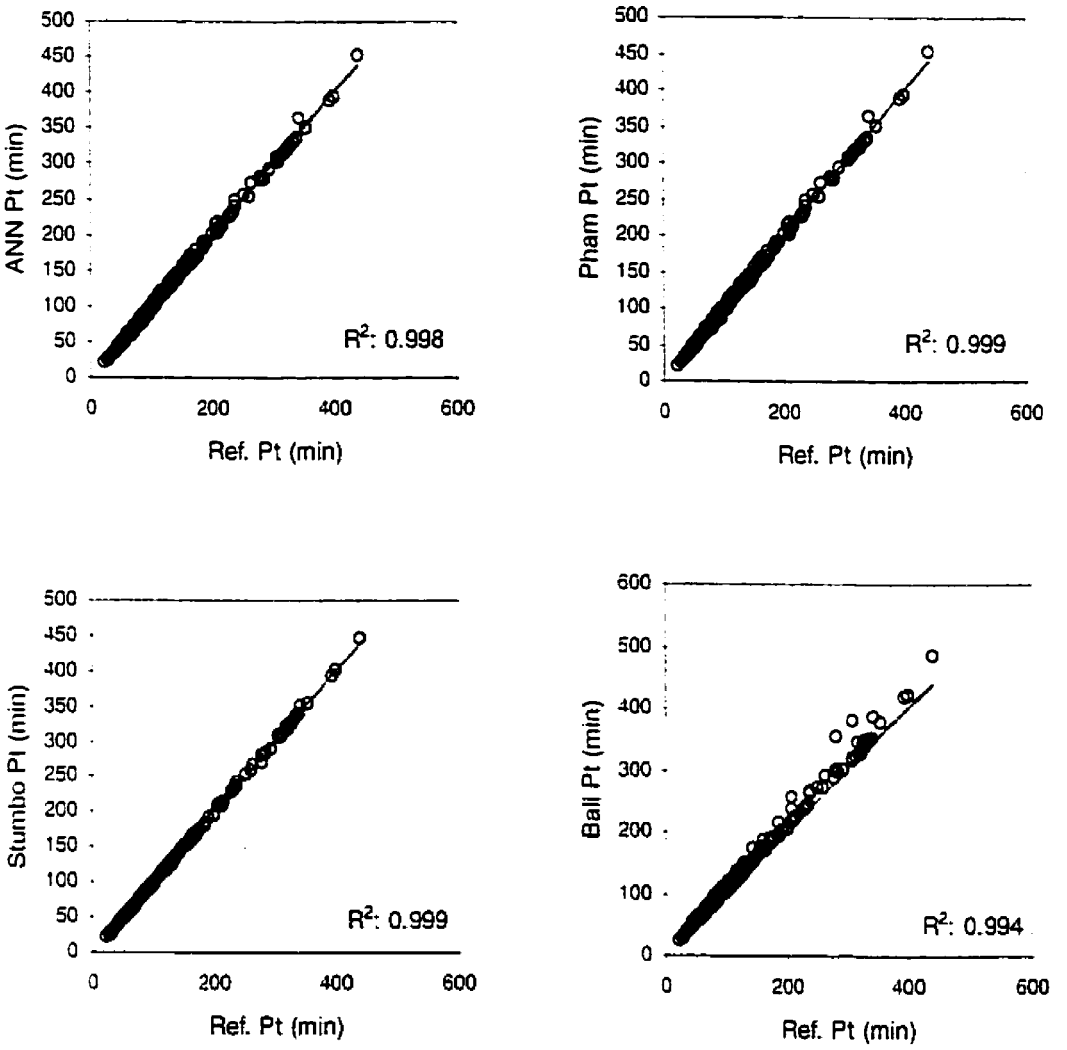


Figure 5.7 Performance of ANN model in process time calculation in comparison with the other selected Formula methods

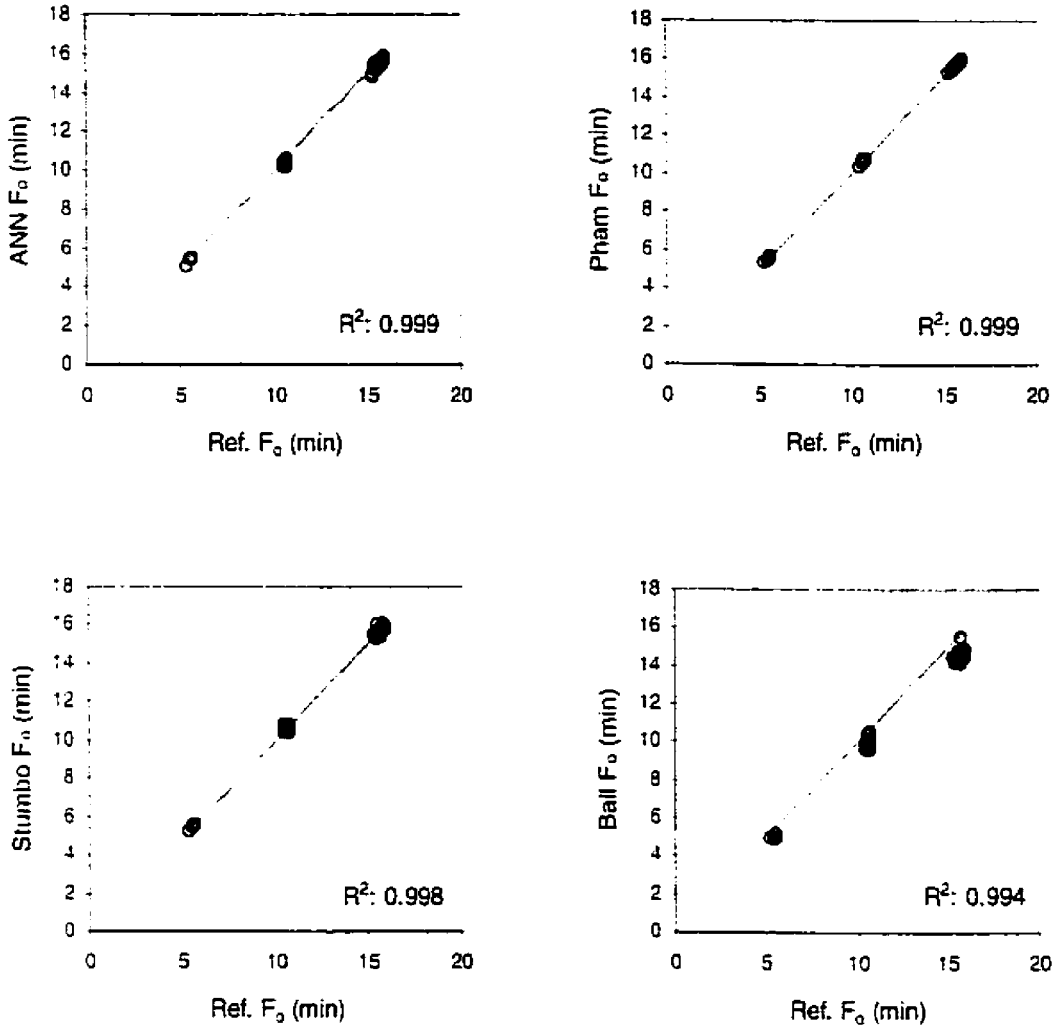


Figure 5.8 Performance of ANN model in process lethality calculation in comparison with the other selected Formula methods

CONCLUSIONS

With respect to preliminary results from ANN modeling of Ball and Stumbo's method, the possibility of ANN model development as an alternative to existent methods of thermal process calculations was studied. A finite difference model was applied to obtain a wide range of data covering the applicable range of processing conditions and can sizes. Considering the accuracy of ANN based Stumbo's model, the same procedure of ANN model development was applied. ANN Model A with 50% data from the total data set, had the best training error, while the ANN Model B required higher number of examples in training (1010 data). Both models had 2 hidden layers, Model A with 2 and 5 PEs in each layer and Model B with 2 PEs in each hidden layer. The initial data was transformed to comply with Stumbo's table and relating of process time and process lethality was performed through g (or f_h/U) and j_{cc} to f_h/U (or g). Using ANN predicted values, process time or process lethality was calculated. ANN Model A with 2.70 minutes MAE was comparable with Pham method in process time prediction. Also, the ANN Model B predicted the process lethality with 2.74% of relative error, which was very close to Stumbo's method. As the methods of Pham and Stumbo were assigned as the most accurate methods of process calculations, the ANN models had the potential of application in process calculation methods.

CHAPTER 6

GENERAL CONCLUSIONS

1. The first phase of the research was development of ANN based on Ball and Stumbo methods, which are the two major and most widely used methods of thermal process calculations. This work was carried out to evaluate the potential for the application of ANN technique for process calculations. A well-trained network required appropriate number of training data set, optimized number of processing elements and hidden layers and internal parameters of the network. For validating these ANN models for each method, a new set of processing conditions were used, and the ANN predicted values were compared to correspondent values from method. ANN models had 1% error in both process time and process lethality prediction based on Ball table data and 3% error for Stumbo table data. The higher error of Stumbo method could be possibly because of the limited number of data and extended range of variables in ANN model development. However, a close to unity value for R^2 showed good potential of ANN technique in thermal process calculations.

2. The accuracy of selected formula methods (Steele and Board, Ball, Stumbo and Pham) were evaluated for a wide range of processing conditions, product thermal diffusivity and can sizes. A finite difference simulation was used to provide the reference values of process time and process calculations as well as the heat penetration parameters (f_h, j_{ch}, f_c and j_{cc}). The performance of the finite difference model was optimized and verified against the analytical solution of conduction heat transfer equation. Retort temperature was the most significant parameter on the performance of methods. Methods of Stumbo and Pham had the higher accuracy as compare to the two other methods. Also in a more specific evaluation, the methods were compared based on the can shapes and *g-values*. For dimension of *H/D* close to unity the errors were highest and they farther increased with *g-values*.

3. As the final goal of this study, ANN models were developed as the alternatives to present process calculation methods. The data from the finite difference simulation was applied to develop ANN models. Training set size was selected based on a developed a learning curve. Using two hidden layers for networks significantly increased the performance of the ANN models; however, learning parameters were less meaningful during development of models. ANN models had excellent ability in process time and process lethality prediction with 2.7 minutes of mean relative errors for process time prediction and 0.27 minutes of mean average error for process lethality prediction. The accuracy of ANN model was compared to performance of existing formula methods and they closely performed to methods of Stumbo and Pham.

REFERENCES

Angerosa. F.. Giacinto. L.D., Vito. R., and Cumitini, S. 1996. Sensory evaluation of virgin olive oil by artificial neural network processing of dynamic head-space gas chromatographic data. Journal of the Science of Food and Agriculture 72:323-328.

Ball. C.O. 1923. Thermal Process Time for Canned Food. Bull. 7-1(37) Natl. Res. Concil. Washington. D.C.

Ball. C.O.. and Olson. F.C.W. 1957. Sterilization in Food Technology-Theory. Practice and Calculation. McGrow Hill. New York.

Baughman. D. R.. and Liu Y. A. 1995. Neural networks in bioprocessing and chemical engineering. Academic press. Inc. San Diego. California.

Bigelow. W.C.. Bohart. G.S.. Richardson. A.C.. and Ball. C.O. 1920. Heat penetration in processing canned foods. National Canners Association. Bull 16L.

Bochereau. L.. Bourgine. P.. and Palagos. B. 1992. A method for prediction by combining data analysis and neural networks: application to prediction of apple quality using near infrared spectra. Journal of Agricultural Engineering and Research 51:207-216.

Briandet. R.. Kemsley. E.K.. and Wilson. R.H. 1996. Approached to adulteration detection in instant coffees using infrared spectroscopy and chemometrics. Journal of the Science of Food and Agriculture 17:359-366.

Carslaw. H. S.. and Jaeger. J. C. 1986. Conduction of heat in solids. 2nd Edition. Oxford University Press. U.K.

Cass. G.V. 1947. A note on the use of Schultz and Olson lethal rate paper for calculation of thermal processes for canned food products in tin containers. Food Research 12:24-26.

Datta. A.K.. Teixeira. A.A.. and Manson. J.E. 1986. Computer-based retort control logic for on-line correction of process deviations. Journal of Food Science 51(2):480-484.

Datta. A.K.. Teixeira. A.A.. and Manson. J.E. 1986. Computer-based retort control logic for on-line correction of process deviations. Journal of Food Science 51(2):480-484.

Eerikainen. T.. Linko. P.. Linko. S.. Siimes. T.. and Zhu. Y-H. 1993. Fuzzy logic and neural network applications in food science and technology. Trends in Food Science and Technology. 4:237-242.

Eerikainen. T.. Zhu. Y-H.. and Linko. P. 1994. Neural networks in extrusion process identification and control. Food Control 5(2):111-119.

Fausett. L. 1994. Fundamentals of neural networks: architectures. algorithms. and applications. Prentice-Hall. Enlewood Cliffs. New Jersey.

Flambert. C.M.F., Deltour, J.. Dickerson. R.W.. and Hayakawa. K. 1977. Lethal effect of food temperature on linear portion of a heating or cooling curve. Journal of Food Science 42(2):545-546.

Freeman. J.A. 1993. Backpropagation in neural network. AI Expert-Neural network special report. January:55-63.

Gebri. V.. Zeppa. G.. Beltramo. R.. Carnacini. A.. and Antonelli. A. 1998. Characterisation of white vinegars of different sources with artificial neural networks. Journal of the Science of Food and Agriculture 78:417-422.

Geeraerd. A.H.. Herremans. C. H.. Ludikhuyze. L.R.. Hendrickx. M.E.. and Van Impe. J.F. 1998. Modeling the kinetics of isobaric-isothermal inactivation of Bacillus subtilis a-amylase with artificial neural networks. Journal of Food Engineering 36:263-279.

Ghazala. S.. Ramaswamy. H.S.. Smith. J.. and Simpson. B. 1990. Thermal process time calculations for thin packages: comparison of formula methods. Journal Food Engineering 13:269-282.

Ghazanfari. A.. Iraudayaraj. J.. and Kusalik. A. 1996. Grading pistachio nuts using a neural network approach. Transactions of the ASAE 39(6):2319-2324.

Gillepsy. T.G. 1953. Estimation of sterilizing values of processes as applies to canned foods. II-Packs heating by conduction: complex processing conditions and value of coming-up time of retort. Journal of the Science of Food and Agriculture 4:553-565.

Gillespy. T.G. 1951. Estimation of sterilization values of process as applied to canned foods. I.-Packs Heating by Conduction. Journal of the Science of Food and Agriculture 2:107-125.

Griffin. R.C. Jr. Henderon. D.H.. and Ball. C.O. 1969. Use of computer-derived tables to calculate sterilizing processed for packaged foods. 2. Application to broken-line heating curves. Food Technology 23(4):519-524.

Griffin. R.C. Jr. Henderon. D.H.. and Ball. C.O. 1971. Use of computer-derived tables to calculate sterilizing processed for packaged foods. 3. Application to cooling curves. Food Technology 25(2):134-143.

Hayakawa. K.-I. 1968. A procedure for calculating sterilizing value of a thermal process. Food Technology 22:905-907.

Hayakawa. K.-I. 1970. Experimental formulas for accurate estimate of food temperature and their application to thermal process evaluation. Food Technology 24:1407-1418

Hayakawa. K.-I. 1973. Modified lethal rate paper technique for the thermal process evaluation. Journal of Canadian Institute of Food Technologists 6(4):295-297.

Hayakawa. K.-I. 1977. Mathematical methods for estimating proper thermal processes and their computer implementation. Advances in Food Research 23:75-141.

Hayakawa. K.-I. 1978. A critical review of mathematical procedures for determining proper heat sterilization process. Food Technology 32(3):59-65.

Hayakawa. K.-I., and Ball. C.O. 1971. Theoretical formulas for temperatures in cans of solid food and for evaluating various heat processes. Journal of Food Science 36:306-310.

Heldman. D.R.. and Lund. D.B. 1992. Handbook of food processing. Marcel Dekker Inc.. New York.

Henderon. D.H.. Griffin. R.C. Jr and Ball. C.O. 1968. Use of computer-derived tables to calculate sterilizing processed for packaged foods. Food Technology 22(4):473-484.

Hicks. E.W. 1958. A revised table of P_h function of Ball and Olson. Food Research 23:396-400.

Horimoto. Y. Lee. K.. and Nakai. S. 1997. Classification of microbial defects in milk using a dynamic headspace gas chromatograph and computer-aided data processing. 2. Artificial neural networks. partial least-squares regression analysis. and principle component regression analysis. Journal of Agriculture and Food Chemistry 45:743-747.

Kim. S.. and Cho. S.I. 1997. Neural network modeling and fuzzy control simulation for bread-baking process. Transactions of the ASAE 40(3):671-676.

Lacroix. R.. Salehi. F.. Yang. X.Z.. and Wade. K.M. 1997. Effects of data preprocessing on the performance of artificial neural networks for dairy yield prediction and cow culling classification. Transactions of the ASAE 40(3):839-846.

Larkin. J.W.. and Berry. M.R. 1991. Estimating cooling process lethality for different cooling j values. Journal of Food Science 56(4):1063-1067.

Lenz . M. K.. and Lund. D. B. 1977. The lethality-Fourier number method: confidence intervals for calculated lethality and mass average retention of conduction-heating canned foods. Journal of Food Science 42:1002-1007.

Leonhardt. G.F. 1978. A general lethal-rate paper for the graphical calculation of processing times. Journal of Food Science 43:660-665.

Linko. P. and Zhu. Y.H. 1992. Neural network modeling for real-time variable estimation and prediction in the control of glucoamylase fermentation. Process Biochemistry 27:275-283.

Lopez. A. 1987. A complete course in canning and related processes. (12th edition). Baltimore. MD: The canning Trade.

Manson. J.E.. and Zahradnik. J.W. 1967. Computer thermal process determination for conduction-heated foods. Food Technology 21(9):1206-1210.

Merson. R. L.. Singh. R. P.. and Carroad. P. A. 1978. An evaluation of Ball's formula method of thermal process calculations. Food Technology 32(30):66-72.

NeuralWare. 1996. Neural Computing: A Technology Handbook for Professional II/Plus and NerualWorks Explorer. Pittsburgh. Pa.: NeuralWare.

Ni. H.. and Gunasekaran. S. 1998. Food quality prediction with neural networks. Food Technology 52(10):60-65.

Park. B.. Chen. Y.R.. Whittaker. A.D.. Miller. R.K.. and Hale. D.S. 1994. Neural network modeling for beef sensory evaluation. Transactions of the ASAE 37(5):1547-1553.

Parmer. R.S.. McClendon. R.W.. Hoogenboom. G.. Blanlenship. P.D.. Cole. R.J.. and Doner. J.W. 1977. Estimation of aflatoxin contamination in preharvest peanuts using neural networks. Transactions of the ASAE 40(3):809-813.

Patashnik. M. 1953. A simplified procedure for thermal process evaluation. Food Technology 7:1-6.

Peters. G.. Morrissey. M.T.. Sylvia. G.. and Bolte. J. 1996. Linear regression. neural network and induction analysis to determine harvesting and processing effects on Surimi quality. Journal of Food Science 61(5):876-880.

Pflung. I.J. 1968. Evaluation the lethality of heat processes using a method employing Hick's table. Food Technology 22:1153-1156.

Pham. Q.T. 1987. Calculation of thermal process lethality for conduction-heated canned foods. Journal of Food Science 52(4):967-974.

Pham. Q.T. 1990. Lethality calculation for thermal processes with different heating and cooling rates. International Journal of Food Science and Technology 25:148-156.

Ramaswamy. H.S. 1993. Come-up time effectiveness for process calculations involving thin-profile packages. Journal of Food Engineering 19:109-117.

Rose. A.H. 1965. Chemical microbiology. Butter Worths & Co. London.

Ruan. R.. Almaer. S. and Zhang. J. 1995. Prediction of dough rheological properties using neural networks. Cereal Chemistry 72(3):308-311.

Ruan. R.. Almaer, S.. Zou. C. and Chen. P.L. 1997. Spectrum analysis of mixing power curves for neural network prediction of dough rheological properties. Transaction of the ASAE 40(3):677-681.

Sablani. S.S.. Ramaswamy. H.S.. and Prasher. S.O. 1995. A neural network approach for thermal processing applications. Journal of Food Processing and Preservation 19:283-301.

Sablani. S.S.. Ramaswamy. H.S.. Sreekanth. S.. and Prasher. S.O. 1997. Neural network modeling of heat transfer to liquid particle mixtures in cans subjected to end-over-end processing. Food Research International 30(2):105-116.

Sayeed. M.S.. Whittaker. A.D.. and Kehtarnavaz. N.D. 1995. Snack quality evaluation method based on image features and neural network prediction. Transactions of the ASAE 38(4):1239-1245.

Smith. T.. and Tung. M. 1982. Comparison of formula methods for calculating thermal process lethality. Journal of Food Science 47:626-630.

Spinak. S.H.. and Wiley. R. C. 1982. Comparison of the General and Ball formula methods for retort pouch process calculations. Journal of Food Science 47:880-888.

Sreekanth. S.. Ramaswamy. H.S.. and Sablani. S. 1998. Prediction of psychometric parameters using neural networks. Drying Technology. 16(3-5):825-837.

Steele. R. J. and Board. P. W. 1979. Thermal process calculations using sterilization values. Food Technology 14:227-235.

Steele. R. J. and Board. P. W. 1979a. Amendments to Ball's formula method for calculating the lethal value of thermal processes. Journal of Food Science 44(1):292-293.

Steele. R. J. and Board. P. W. 1979b. Thermal process calculations using sterilization values. Journal of Food Technology 14:227-235.

Steele. R.J. Board. P.W. Best. B.J. and Willox. M.E. 1979. Revision of the formula method tables for thermal process evaluation. Food Technology 44:954-958.

Stoforos. N. G. 1991. On Ball's formula method for thermal process calculations. Journal of Food Processing and Engineering 13(4):255-269.

Stoforos. N.G.. Noronha. J.. Hendrickx. M.. and Tobback. P. 1997. A critical analysis of methimatical procedures for the evaluatoin and design of in-container thermal processes for foods. Critical Reviews Food Science and Nutrition 37(5):411-441.

Stumbo. C. R.. and Longley. R. E. 1966. New parameters for process calculation. Food Technology 20(3):341-345.

Stumbo. C.R. 1973. Refinement and Extension of $f_h/U:g$ Parameters for Process Calculation. Journal of Food Science 38:726-728.

Swingler. K. 1996. Applying neural networks a practical guide. Academic Press. San Diego. CA.

Teixeira. A. A. Dixon. J. R. Zahradnik. J. W. and Zinsmeister. G. E. 1969a. Computer determination of spore survival distribution in thermally processed conduction heated foods. Food Technology 23:352-354.

Teixeira. A. A. Dixon. J. R. Zahradnik. J. W., and Zinsmeister. G. E. 1969b. Computer optimization of nutrient retention in the thermal processing of conduction-heated foods. Food Technology. 23:845-850.

Teixeira. A.A. 1978. Conduction heating consideration in thermal processing of canned foods. ASME San Francisco Paper # 78-WA/HT-55.

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

Teixeira. A.A.. and Tucker. G.S. 1997. On-line retort control in thermal sterilization of canned foods. Food Control 8(1):13-20.

Thai. C.N.. and Shewfelt. R.L. 1991. Modeling sensory color quality of tomato and peach: neural network and statistical regression. Transaction of the ASAE 34(3):950-955.

Tomlins. K.I. and Gay. C. 1994. Prediction of quality and origin of black tea and pine resin samples from chromatographic and sensory information: evaluation of neural networks. Food Chemistry 30:157-165.

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

Tung. M.A.. and Garland. T.D. 1978. Computer Calculation of Thermal Processes. Journal Food Science 43:365-396.

Valentas. K.J.. Levine. L. and Clark. J.P. 1991. Food processing operations and scale-up. Marcel Dekker Inc.. New York.

Vallejo-Cordoba. B. Arteaga. G.E. and Nakai. S. 1995. Predicting milk shelf-life based on artificial neural networks and headspace gas chromatographic data. Journal of Food Science 60(5):885-888.

Vinters. J.E. Patel. R.H. and Halaby. G.A. 1975. Thermal Process Evaluation by Programmable Calculator. Food Technology 29(3):42-44.