A Neural Model of Time to Toxin Production by Non-Proteolytic *Clostridium botulinum*

Gavin C. Cawley
School of Information Systems
University of East Anglia
Norwich
United Kingdom
gcc@sys.uea.ac.uk

Michael W. Peck
Institute of Food Research
Norwich Research Park
Norwich
United Kingdom
mike.peck@bbsrc.ac.uk

Pablo S. Fernández
Centro de Edafología y Biología
Aplicada del Segura
Murcia
Spain
pfernand@natura.cebas.csic.es

Abstract

Clostridium botulinum *is a bacterium present in the raw ingredients of many foods. It produces a powerful neurotoxin as part of its growth process, that can prove fatal when doses as small as a few micrograms are consumed. It is therefore vital to be able to accurately determine the food processing and storage conditions where toxin production is prevented. This paper describes a new model of the effect of heat treatment and subsequent incubation temperature on time to toxin production by non-proteolytic* C. botulinum. *A multi-layer perceptron network is trained, using a cross-entropy error metric, to estimate the* a posteriori *probability of toxin production given the heat treatment applied and incubation time and temperature. The results obtained compare favourably with those obtained using a conventional model based on a quadratic response surface.*

1. Introduction

Clostridium botulinum is an anaerobic bacterium that produces one of the most powerful toxins known to science as a by-product of its growth processes. Ingestion of only a few micrograms of the toxin can result in severe illness and even death. It is therefore vital that steps should be taken to ensure that the toxin is not present in food. As *C. botulinum* spores are ubiquitous in raw ingredients, food must be processed to ensure that all of these spores are destroyed, or so that the spores are prevented from germinating, leading to cell division and subsequent toxin production. Growth of *C. botulinum* is, in most cases, principally dependent on environmental factors such as temperature, pH, NaCl concentration and gas atmosphere. It is important then to be able to define the conditions under which the spores are prevented from germinating, and giving rise to toxin production. This is especially true in the case of minimally processed chilled foods, as non-proteolytic strains of *C. botulinum* are capable of growth and toxin production at chill temperatures. The safety of these foods with respect to non-proteolytic *C. botulinum* is likely to rely on a combination of heat treatment and subsequent incubation at low temperatures (Lund and Notermans [8], Peck [10].

This paper describes a neural model of toxin production by *C. botulinum* in a meat-based medium containing lysozyme, following a range of heat treatments. The network is trained, using the backpropagation algorithm with a cross-entropy error metric, to estimate the *a posteriori* probability of toxin production, given inputs describing the heat treatment applied and incubation conditions. The cumulative density function encoded within the neural network provides an improvement over a more conventional model based on a quadratic response surface.

The remainder of the paper is structured as follows: Sections 2 and 3 describe the data set, collected by the Institute of Food Research, and a conventional growth model based on quadratic response surface. The proposed neural model is described in section 4 and the results of initial experiments are presented in section 5. Section 6 discusses the conclusions drawn from the results of our initial research and suggests avenues for further work.

2. The Dataset

The neural model of toxin production in *C. botulinum* described here is based on the dataset described in Fernández and Peck [5]. Tubes containing a sterile meat-based medium containing lysozyme, an enzyme found to increase the measured heat resistance of spores of non-proteolytic *C. botulinum* (Lund and Peck [9], Peck [10]), were inoculated with a suspension of the spores of eight strains of non-proteolytic *C. botulinum*, at a final concentra-

Table 1. Heat treatments applied to a meat-based medium containing spores of *C. botulinum*

Temperature (°C)	Duration (min)		
70°C	104.9	529.1	998.9
	1596.3	2065.9	2544.5
75°C	284.6	463.1	734.2
	1071.5	1376.5	1793.0
80°C	11.4	69.7	98.0
	127.9	183.8	229.6
	294.9	362.7	
85°C	23.3	35.7	52.0
	57.8	83.8	
90°C	10.3	10.9	15.3
	23.5	33.5	

tion of 10^6 spores per tube, and subjected to a range of heat treatments, shown in table 1. The tubes were then cooled and incubated at temperatures of 5, 8, 12, 16 and 25°C for 90 days. Five replicates were performed at each incubation temperature, for each heat treatment regime. The tubes were inspected every 2–3 days for signs of growth, indicated by obvious formation of gas. At the end of the experiment, samples from each heat treatment regime, showing growth at the lowest incubation temperature and for the highest incubation temperature that did not show growth, were tested for toxin (Peck *et al.* [11], Potter *et al.* [12], Carlin and Peck [4]). This type of dataset is known as *time to growth* data, as the results are presented in terms of a table showing the number of days after which each tube showed signs of growth. Full details of the experimental method are recorded in Fernández and Peck [5].

3. Modelling of Time to Growth Data Using a Response Surface

The effects of environmental conditions on the growth of microbial populations are frequently modelled using a quadratic response surface. For a system with two independent variables x_1 and x_2, a quadratic response surface is represented by a polynomial of the form:

$$y = c_1 + c_2 x_1 + c_3 x_2 + c_4 x_1^2 + c_5 x_2^2 + c_6 x_1 x_2$$

where y is the parameter of the system being modelled, and $c_1, c_2, \ldots, c_6$ are coefficients to be determined, using a least squares optimisation procedure. Fernández and Peck [5] describe the use of a tri-quadratic response surface to model the natural logarithm of the time to the first observation of growth (y) in the five replicates performed for each heat

treatment regime at each incubation temperature. The coefficients obtained for the second order polynomial were

$$
\begin{aligned}
\log_e y \quad = \quad & 0.0037I^2 - 0.0056H^2 - 0.00000015t^2 \\
- \quad & 0.000014tI + 0.0002Ht + 0.0005HI \\
- \quad & 0.2748I + 1.0054H - 0.0119t \\
- \quad & 39.14
\end{aligned}
$$

where H is the heating temperature, t is the heating time and I is the incubation temperature. While a model of this type clearly provides a useful tool in ensuring food safety, there remain areas where improvements may be made:

- The experimental data are not fully utilised; only the time to first growth for each heat treatment regime, at each incubation temperature, is used to form the model.

- It is difficult to place a statistical interpretation of the output of the model. As the model is not based on an estimate of the underlying distribution of the data, it is difficult to quantify the risk of growth earlier than predicted by the model.

- Growth is observed in the experimental data earlier, in some cases, than is predicted by the model.

The food manufacturer requires a model that provides an indication that a given heat treatment is sufficient to ensure the safety of the consumer, over the expected range of storage conditions. A model of the distribution of the time to growth, based on all of the experimental observations, would allow the *probability* of growth and therefore toxin production to be estimated, for a given heat treatment and incubation temperature. This would clearly provide a more useful resource in risk analysis, as it would then be possible to establish the conditions where growth will not occur, at a well defined confidence level. The remaining sections of this paper describe the use of artificial neural networks to construct a model of this type.

4. Modelling of Time to Growth Data Using an Artificial Neural Network

It has been shown that the outputs of classifiers based on the multi-layer perceptron can be regarded as estimates of Bayesian *a posteriori* probabilities, given a suitable combination of output coding method, activation function and error metric (Baum and Wilczec [1], Bridle [3], Richard and Lippmann [13], Hampshire and Pearlmutter [6]). This section describes the construction of a model, based on this property of artificial neural networks, of the probability of growth from spores, and hence toxin production, of non-proteolytic *C. botulinum* following heat treatment and subsequent incubation at chill temperatures.
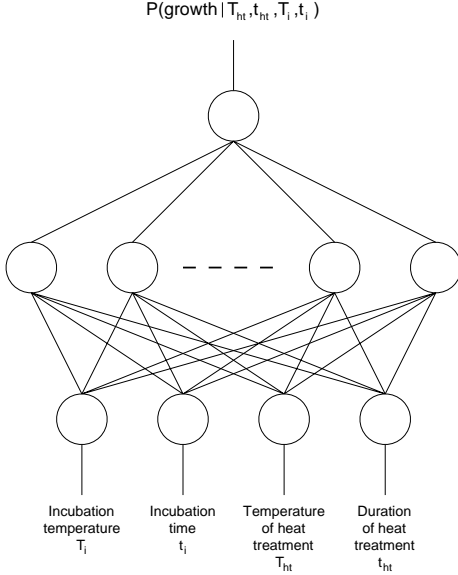
$P(\text{growth} \mid T_{ht}, t_{ht}, T_i, t_i)$

Incubation temperature T_i

Incubation time t_i

Temperature of heat treatment T_{ht}

Duration of heat treatment t_{ht}

Figure 1. Architecture of the proposed model of time to growth data based on a multi-layer perceptron.

Figure 1 shows a schematic diagram of the proposed model of the dataset described in section 2, based on a multi-layer perceptron. The input layer consists of four neurons, representing the temperature (T_{ht}) and duration (t_{ht}) of the heat treatment, the incubation temperature (T_i) and incubation time (t_i). Each input is first standardised (*i.e.* linearly rescaled) to have a zero mean and unit variance, as this simplifies weight initialisation (Bishop [2]) and can have a beneficial effect on training time (Sarle [16]). For initial experiments a hidden layer of sixteen neurons, with symmetric logistic sigmoidal activation functions (Stornetta and Huberman [18]), was deemed to be appropriate, given the size and apparent complexity of the dataset. The output layer consists of a single neuron, with an assymetric logistic activation function, trained to estimate the *a posteriori* probability of growth, using the backpropagation algorithm (Rumelhart *et al.* [14]).

The interpretation of the output of a neural network as an *a posteriori* probability holds for a wide range of error functions, including the standard squared error and cross entropy (Hinton [7]) metrics. The cross entropy, C, is given by:

$$C = -\sum_{j,p} t_{j,p} \log_2(o_{j,p}) + (1 - t_{j,c}) \log_2(1 - o_{j,p})$$

where $t_{j,c}$ is the target output for unit j for pattern p and $o_{j,p}$ is the output of the network for pattern p. As we are principally interested in defining the conditions under which growth is unlikely, the cross-entropy error metric was

considered most appropriate for this application, because it places a greater emphasis on errors for probabilities close to zero.

Before training, the weights were initialised to small random values between -0.3 and 0.3. A simple early stopping procedure was employed to prevent over-fitting of the training data (Sarle [15]). As only a limited amount of training data was available, a ten-fold cross-validation scheme was then used to estimate the true accuracy of the model (Stone [17]).

5. Results Obtained Using a Neural Model of Time to Growth Data

Figure 2 shows a graph of the average root mean square error against cycles trained, for the ten neural time to growth models generated by the cross-validation procedure, over both the training and validation data. The models are seen to converge slowly, with little indication of overfitting.
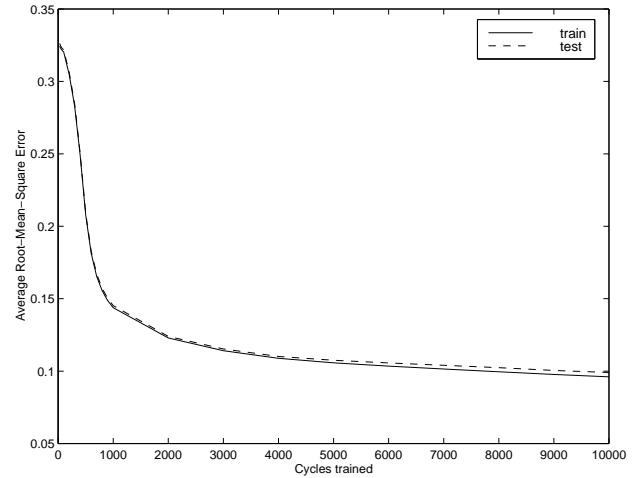

Figure 2. Graph of the average root mean square error against cycles trained, for ten neural time to growth models, over both the training and validation datasets.

Figure 3 shows a contour plot of a cumulative density function (C.D.F.), generated by the neural time-to-growth model, of the probability of growth against incubation temperature and incubation time, following heat treatment at a temperature of 75°C for 463 minutes. The C.D.F. seems well in agreement with the experimental data obtained for this heat treatment schedule, denoted by circles in figure 3. In general, C.D.F.s generated by the neural network are smooth, monotonic functions, and appear to be subjectively reasonable interpolations of the experimental data.
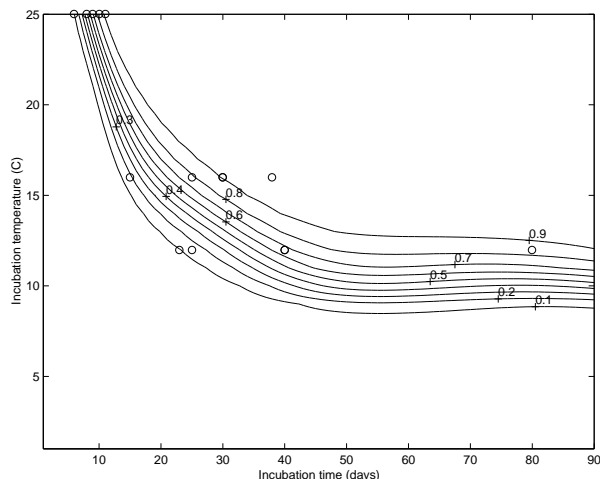
Figure 3. Contour plot of the cumulative density function for growth of non-proteolytic strains of *C. botulinum*, in a meat-based medium containing lysozyme, following heat treatment at 75°C for 463 minutes and subsequent incubation at 5–25°C. The circles denote experimental observations of growth (no growth was observed for incubation temperatures of 5 and 8°C for this heat treatment).

6. Conclusions and Further Work

The accuracy with which *a posteriori* probabilities are estimated by the outputs of multi-layer perceptron classifiers depends on two principal factors [13]:

- **Complexity of the neural network** - the multi-layer perceptron must be sufficiently complex to represent the form of the statistical distribution of the data. Further experimentation is required to determine the optimal network size for this dataset.

- **Adequacy of the data set** - the dataset must be representative of the underlying statistical distribution. It is not clear at the current time that the existing data set is sufficiently large to be able to estimate the *a posteriori* probability of growth accurately. A larger dataset would allow research to be conducted into the amount of data required to form statistical models of this type.

Modelling the growth, survival and death of populations of various microbial organisms is a common task in food safety research. There is scope, therefore to apply profitably the methods described in this paper to a number of related datasets in this field.

References

[1] E. B. Baum and F. Wilczec. Supervised learning of probability distributions by neural networks. In D. Anderson, editor, *Neural Information Processing Systems*, pages 52–61. American Institute of Physics, 1988.

[2] C. M. Bishop. *Neural Networks for Pattern Recognition*. Clarendom Press, Oxford, 1995.

[3] J. S. Bridle. Probabilistic interpretation of feedforward classification network outputs, with relation to statistical pattern recognition. In F. Fogelman Soulié and J. Hérault, editors, *Neurocomputing*, pages 227–236. Springer-Verlag, 1990.

[4] F. Carlin and M. W. Peck. Growth and toxin production by non-proteolytic and proteolytic *Clostridium botulinum* in cooked vegetables. *Letters in Applied Microbiology*, 20:152–156, 1995.

[5] P. S. Fernández and M. W. Peck. A predictive model that describes the effect of prolonged heating at 70°–90°C and incubation at refrigeration temperatures on growth and toxigenesis by non-proteolytic *Clostridium botulinum* in the presence of lysozyme. submitted to *Applied and Environmental Microbiology*, 1997.

[6] J. B. Hampshire and B. Pearlmutter. Equivalence proofs for multi-layer perceptron classifiers and the Bayesian discriminant function. In D. S. Touretzky, J. L. Elman, T. J. Sejnowski, and G. E. Hinton, editors, *Proceedings of the 1990 Connectionist Models Summer School*, pages 159–172. Morgan-Kaufmann, 1990.

[7] G. E. Hinton. Connectionist learning procedures. *Artificial Intelligence*, 40:185–234, 1989.

[8] B. M. Lund and S. H. W. Notermans. Potential hazards associated with REPFEDs. In Clostridium botulinum*: ecology and control in foods*, pages 279–301. Marcel Dekker, New York, 1992.

[9] B. M. Lund and M. W. Peck. Heat resistance and recovery of spores of non-proteolytic *Clostridium botulinum* in relation to refrigerated, processed foods with extended shelf-life. *Applied Bacteriology*, 76:115–128, 1994.

[10] M. W. Peck. *Clostridium botulinum* and the safety of refrigerated processed foods of extended durability. *Trends in Food Science and Technology*, 8:186–192, 1997.

[11] M. W. Peck, B. M. Lund, D. A. Fairbairn, A. S. Kassperson, and P. C. Undeland. Effect of heat treatment on survival of, and growth from, spores of non-proteolytic *Clostridium botulinum* at refrigeration temperatures. *Applied and Environmental Microbiology*, 61:1780–1785, May 1995.

[12] M. D. Potter, J. Meng, and P. Kimsey. An ELISA for detection of botulinal toxin types A, B and E in innoculated food samples. *Journal of Food Protection*, 56:856–861, 1993.

[13] M. D. Richard and R. P. Lippmann. Neural network classifiers estimate bayesian *a posteriori* probabilities. *Neural Computation*, 3:461–483, 1991.

[14] D. E. Rummelhart, J. L. McClelland, and the PDP Research Group. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, volume 1: Foundations. The MIT Press, 1986.

[15] W. S. Sarle. Stopped training and other remedies for overfitting. In *Proceedings of the 27th Symposium on the Interface*, 1995.

[16] W. S. Sarle (ed.). Neural network FAQ part 1 of 7: *Introduction*. periodic posting to the Usenet newsgroup comp.ai.neural-nets, also available via the WWW (URL: ftp://ftp.sas.com/pub/neural/FAQ.html), 1997.

[17] M. Stone. Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society*, B 36(1):111–147, 1974.

[18] W. S. Stornetta and B. A. Huberman. An improved three-layer backpropagation algorithm. In *Proceedings of the I.E.E.E. International Conference on Neural Networks*, pages 637–644, San Diego, California, U.S.A., 1987.