Food Microbiology 78 (2019) 38–45

Contents lists available at ScienceDirect

Food Microbiology

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

Modeling growth limits of *Bacillus* spp. spores by using deep-learning algorithm

Sayuri Kuroda[a], Haruko Okuda[b], Wataru Ishida[b], Shigenobu Koseki[a,*]

[a] *Graduate School of Agricultural Science, Hokkaido University, Kita-9, Nishi-9, Kita-ku, Sapporo, 060-8589, Japan*
[b] *Nisshin Seifun Group, Inc, Fujimino, Saitama, Japan*

ARTICLE INFO

Keywords:
Bacillus
Spoilage bacteria
Growth/no growth boundary model
Logistic regression
Neural network
Deep learning

ABSTRACT

Growth/no growth boundary models for *Bacillus* spores that accounted for the effects of environmental pH, water activity (a_w), acetic acid, lactic acid, bacterial strain, and storage period were developed using conventional logistic regression and machine learning algorithms. Growth in tryptic soy broth at 317 conditions comprising nine levels of pH (4.0–6.5), six levels of a_w (0.85–1.00), six levels of acetic acid concentrations (0–0.8%), and five levels of lactic acid concentrations (0–0.8%) was examined to confirm growth limit conditions. All models developed using logistic regression, neural network, and deep learning on the basis of obtained datasets successfully described growth/no growth boundaries of three *Bacillus* species. Although the logistic regression model failed to describe growth limits under some conditions, neural network and deep learning approaches enabled to determine them in such cases. The developed models were evaluated by independent experimental data of growth in tryptic soy broth and in clam soup. The deep learning model enabled better prediction of independent data with smaller probabilistic variability values than those of the logistic regression and neural network models. The deep learning procedure can be utilized for growth boundary modeling to control bacterial growth safely and flexibly.

1. Introduction

Many growth/no growth boundary models have been published for various bacteria in order to predict bacterial growth limit conditions (Dang et al., 2010; Evans, 2004; Polese et al., 2011). Because *Bacillus* species often contaminate food ingredients and can survive in various stressful environments due to their spore-forming ability, the control of *Bacillus* spores plays an important role in quality assurance and safety of processed foods. Therefore, bacterial growth/no growth boundary conditions for *Bacillus* species were examined in many studies (Daelman et al., 2013; Lanciotti et al., 2001; Quintavalla and Parolari, 1993).

Logistic regression is generally used for the prediction of bacterial growth/no growth boundary. This modeling procedure has an advantage that any type and number of explanatory variables can be used in the model. However, logistic regression model has limited prediction flexibility because it is based on a mathematical equation. Hence, logistic regression-based prediction greatly relies on the nature of that equation. Although Mejlholm et al. (2010) reported that a highly complex model with 12 explanatory variables enabled to accurately predict growth responses in complex food, development of simpler, flexible, and accurate prediction model is needed to enable robust risk estimation in practice.

An alternative method to logistic regression model is machine learning. Several studies have reported that the neural network was superior in prediction performance to conventional regression models (Esnoz et al., 2006; Fernández-Navarro et al., 2010; Hajmeer and Basheer, 2003b, 2002; Jeyamkondan et al., 2001; Lou et al., 2001). Machine learning has a number of advantages, including an ability to implicitly detect complex relationships between explanatory and response variables based on large datasets. Machine learning usually allows a more flexible prediction than conventional regression models because machine learning does not depend on the regression equation. In addition, it is easy to add new data and new parameters to the machine learning algorithm. However, machine learning has disadvantages, such as unclear calculation procedures, requirements of large datasets, greater computational burden, and a possibility of overfitting due to data-driven prediction.

Recently, machine learning has emerged again as a promising tool because of the advancements in computing power and development of highly complex calculation methods. In particular, deep learning

---

Abbreviations: AUC, area under the ROC curve; FC, fraction correct; ROC, receiver operating characteristic; TSB, tryptic soy broth

[*] Corresponding author.

E-mail address: koseki@bpe.agr.hokudai.ac.jp (S. Koseki).

https://doi.org/10.1016/j.fm.2018.09.013
Received 25 March 2018; Received in revised form 19 September 2018; Accepted 19 September 2018

algorithm allows highly complex predictions by extending the neural network to a multilayered structure. As a result, the use of deep learning algorithms has been increasing in various fields. However, its application in predicting microbiological parameters has not yet been examined.

In the present study, we aimed to develop growth/no growth boundary models to predict a probability of growth for spores of *Bacillus* species as a function of environmental pH, water activity (a_w), organic acid concentration, bacterial strains, and incubation time by using logistic regression, neural network, and deep learning algorithm. The developed predictive models were validated by independent datasets that included experiments in culture medium and real food matrix. The accuracy and validity of the models were then compared. In particular, we focused on the possibility to utilize machine learning modeling procedures for flexible prediction of bacterial growth limits as an alternative approach to conventional logistic regression analysis.

2. Materials and methods

2.1. Bacterial strains

Cocktails of spores of ten strains of B. subtilis (NBRC13169, NBRC3936, NBRC14473, NBRC16449, NBRC100521, NBRC101592, JCM2499, NBRC13719T, NBRC12112, JCM1465), three strains of B. coagulans (NBRC12583, NBRC3557, JCM20478), and eight strains of B. megaterium (NBRC100168, NBRC100708, NBRC104465, NBRC3970, NBRC13498, NBRC15308T, NBRC12068, NBRC12108) were used. NBRC and JCM strains were obtained from the National Institute of Technology and Evaluation (NITE) Biological Resource Center (Tokyo, Japan) and RIKEN Bio Resource Center (Tsukuba, Japan), respectively.

A platinum loop was used to transfer the frozen bacterial cultures by scratching the surface of the frozen culture into plate count agar plates, and the inoculated plate was incubated at 37 °C for 5 days. Afterwards, an isolated colony of each bacterium was transferred to 5 mL of tryptic soy broth (TSB) in a sterile plastic tube, and the inoculated tube was incubated at 37 °C for 24 h. The cultures were transferred into soil extract agar (2 g/L beef extract, 3 g/L yeast extract, 10 g/L peptone, 5 g/L NaCl, 20 g/L agar, 1 g/L starch, 1 m$props_value{literPattern}/L MnSO4 solution, 250 m$props_value{literPattern}/L soil solution, 750 m$props_value{literPattern}/L pure water), and the inoculated plate was incubated at 37 °C for 10 days. After incubation, the bacterial colonies were scratched and collected by a platinum loop and suspended in 2 mL of 1/15 M phosphate buffer. Spore formation was confirmed by using phase contrast microscopy observations. Spores were collected by centrifugation (1000 × *g* for 10 min at 20 °C). The supernatant was discarded, and the spores were subsequently resuspended in 1/15 M phosphate buffer. This procedure was repeated three times and finally, the spore solutions were heated to 80 °C for 10 min to remove remaining vegetative cells. The prepared spore suspensions were stored at −80 °C and, when needed, were gently thawed and mixed with each strain. The spores were heated at 80 °C for the purpose of heat shock, and the cocktail of spore suspensions was diluted to 10^4 colony-forming units/mL using peptone water.

2.2. Determination of the growth/no growth boundary condition

2.2.1. Data collection for the development of growth/no growth boundary models

A condition of pre-experiment was nine levels of pH (4.0, 4.2, 4.4, 4.6, 4.8, 5.0, 5.5, 6.0, and 6.5), six levels of aw (0.85, 0.88, 0.91, 0.94, 0.96, and 1.00), six levels of acetic acid concentrations (0, 0.05, 0.1, 0.2, 0.4, and 0.8%; w/w), and five levels of lactic acid concentrations (0, 0.1, 0.2, 0.4, and 0.8%; w/w). Then a condition was narrow down to a total of 317 condition. Each condition was examined in triplicate trials. TSB was used as base medium. The value of a_w was adjusted with NaCl and then, the medium was sterilized. After that, acetic acid or lactic acid was added to the desired concentration, and pH was adjusted with 1.0 M HCl or 1.0 M NaOH. All prepared culture media were sterilized by filtration through a 0.2-µm membrane filter. An aliquot of 0.1 mL of the spore suspension was added to an airtight test tube containing 5 mL of control medium and incubated at 35 °C for 12 weeks. The presence or absence of growth in each tube was determined by evaluating medium turbidity once a week during incubation period.

2.2.2. Data collection for validation of growth/no growth boundary model

2.2.2.1. Culture medium. A condition of pre-experiment was four levels of pH (4.4, 4.8, 5.5, and 6.5), six levels of a_w (0.85, 0.88, 0.91, 0.94, 0.96, and 1.00), and two levels of acetic acid and lactic acid concentrations (0.3 and 0.6% each, w/w). Then a condition was narrow down to a total of 67 condition. Each condition was examined in triplicate trials. TSB was used as base medium. The value of a_w was adjusted by different concentrations of NaCl. Acetic acid or lactic acid was added to the desired concentration, and pH of the medium was adjusted with 1.0 M HCl or 1.0 M NaOH. All prepared culture media were sterilized by filtration through a 0.2-µm membrane filter. An aliquot of 0.1 mL of the spore suspension was added to an airtight glass test tube containing 5 mL of control medium and incubated at 35 °C for 12 weeks. The presence or absence of growth in each tube was determined by evaluating medium turbidity once a week during incubation period.

2.2.2.2. Clam soup. Clam soup (pH 5.70, a_w 0.98, Nisshin Seifun Group, Tokyo, Japan) was used as a model food substrate in this study, because the soup is similar to TSB. Data for model validation by real food matrix were collected from a total of 15 conditions comprising six levels of pH (4.2, 4.4, 4.6, 4.8, 5.5, and 6.5) and three levels of acetic acid and lactic acid concentrations (0, 0.2, and 0.4% each, w/w) in clam soup. Each condition was examined in independent triplicate trials. The solids were removed from clam soup by a sterile nonwoven fabric mesh. Acetic acid or lactic acid was added to the desired concentration, and the pH was adjusted with 1.0 M HCl or 1.0 M NaOH after removing the solid contents were removed from clam soup. Then, clam soup was sterilized at 121 °C for 15 min. An aliquot of 0.1 mL of the spore suspension was added to a sterile plastic tube containing 5 mL of control clam soup and incubated at 35 °C for 4 weeks. The presence or absence of growth in each tube was evaluated by counting the number of viable cells by the plate count method.

2.3. Model development

2.3.1. Logistic regression

On the basis of the results obtained (growth: 1; no growth: 0), logistic regression was carried out using statistical software R (ver. 3.4.1) in order to estimate the probability (*P*) of growth of *Bacillus* spp. spores as a function of pH, a_w, concentrations of acetic and lactic acids, strain, and incubation period. The equation of this model was as follows: (Eq. (1)).

$$\begin{aligned} logit(P) = {} & b_0 + b_1 pH + b_2 a_w + b_3 AcOH + b_4 lact + b_5 time + b_6 Bs \\ & + b_7 Bm + b_8 pH^2 + b_9 {a_w}^2 + b_{10} AcOH^2 + b_{11} lact^2 + b_{12} time^2 \\ & + b_{13}(pH \cdot a_w) + b_{14}(pH \cdot AcOH) + b_{15}(pH \cdot lact) + b_{16}(pH \cdot time) \\ & + b_{17}(a_w \cdot AcOH) + b_{18}(a_w \cdot lact) + b_{19}(a_w \cdot time) \\ & + b_{20}(AcOH \cdot time) + b_{21}(lact \cdot time) \end{aligned} \tag{1}$$

where logit means log (*P*/1−*P*), *AcOH* and *Lact* are concentrations of acetic and lactic acids (%, w/w), respectively. pH, a_w, *strain*, and *time* are pH, water activity, strain types (*B. subtilis*, *Bs*:1, *Bm*:0; *B. coagulans*, *Bs*:0, *Bm*:0; *B. megaterium*, *Bs*:0, *Bm*:1), and incubation period in weeks. The values of b_i ($i = 0, \ldots, 21$) are the parameters to be estimated. The model was fitted by R using linear logistic regression, according to the procedure described by Vermeulen et al. (2007). Significant parameters were selected to minimize Akaike's Information Criterion (AIC) using stepAIC function in MASS package of R program. AIC is a measure of the goodness of fit of an estimated statistical model.

2.3.2. Neural network

On the basis of the results obtained (growth: 1; no growth: 0), a feed-forward neural network with a single hidden layer was built using *nnet* package of statistical software R (ver. 3.4.1). Independent variables were pH, a_w, concentrations of acetic and lactic acids (%, w/w), strain, and incubation period in weeks. The dependent variable was growth, which could be either present (1) or absent (0). The number of units in the hidden layer was 7 and the parameter for weight decay was 0.0001. These parameters were selected by using *caret* package of statistical software R. The maximum allowable number of weights was 4,000, and the maximum number of iterations was 100. In this study, we carried out 10 iteration measurements because the prediction result of the neural network is random every time. The output of the neural network model is probability of growth. The threshold of the probability is 0.5. If the probability was less than 0.5, the condition was judged as no growth. If the probability was ≥ 0.5, the condition was judged as growth.

2.3.3. Deep learning

We examined multilayer feedforward neural networks by using *h2o* package of statistical software R (ver. 3.4.1). Independent variables were pH, a_w, concentrations of acetic and lactic acids (%, w/w), strain, and incubation period in weeks. The dependent variable was growth, which could be either present (1) or absent (0). Five hidden layers were adopted for the present study, and the number of hidden units in each hidden layer was [100, 100, 100, 100, 100]. The activation function of the node in the hidden layer was given by:

$$f(\alpha) = \frac{e^{\alpha} - e^{-\alpha}}{e^{\alpha} + e^{-\alpha}} \; f(\alpha) \in [-1.1]$$

where α denotes the weighted combination $\alpha = \sum_i w_i x_i + b$. x_i and w_i represent the firing neuron's input *P*-value and its weights, respectively. The maximum number of iterations was 200. In this study, we carried out 10 iteration measurements because the prediction result of deep learning procedure is random every time. The output of the deep learning model is probability of growth. The threshold of the probability is 0.5. If the probability was less than 0.5, the condition was judged as no growth. If the probability was ≥ 0.5, the condition was judged as growth. Deep learning was trained on a number of “training” examples and validated on another number of independent “validation” examples.

2.4. Evaluation of model performance

The predictive performances of the models were evaluated by the confusion matrix, fail-dangerous and fail-safe predictions, and receiving operator characteristic curve.

2.4.1. Confusion matrix

The method to compare the accuracies of the models was based on computing the confusion matrix. The model was classified into two classes (growth/no growth), so the output always reflected one of the four possible situations: (i) true negative (TN), when the classifier correctly identified a negative case (non-event), (ii) false positive (FP), when the classifier incorrectly identified a non-event class as an event, (iii) false negative (FN), when the classifier incorrectly identified an event case as a non-event, and (iv) true positive (TP), when the classifier correctly identified a positive case. The C-matrix was stated as:

$$C = \begin{pmatrix} N_{TN} & N_{FP} \\ N_{FN} & N_{TP} \end{pmatrix} = \begin{pmatrix} 0 \rightarrow 0 & 0 \rightarrow 1 \\ 1 \rightarrow 0 & 1 \rightarrow 1 \end{pmatrix}$$

The used performance measures were described by (Hajmeer and Basheer, 2003a): (i) fraction correct (FC), representing the proportion of cases classified correctly, (ii) false alarm (positive) rate (FAR), representing the fraction of negative cases classified as positives, (iii) probability of detection (POD), representing the fraction of true cases detected as positives by the classifier. These performance measures are expressed as:

$$FC = \frac{N_{TN} + N_{TP}}{N_{TN} + N_{FP} + N_{FN} + N_{TP}}$$

$$FAR = \frac{N_{FP}}{N_{TN} + N_{FP}}$$

$$POD = \frac{N_{TP}}{N_{FN} + N_{TP}}$$

For a perfect classifier, FC = POD = 1 and FAR = 0.

2.4.2. Fail-dangerous and fail-safe predictions

Predicted and observed growth and no growth responses were compared by calculating the percentage of all samples that were correctly predicted. Incorrect predictions were categorized as fail-dangerous (i.e., no growth predicted when growth was actually observed) or fail-safe (i.e. growth predicted when no growth was actually observed) (Mejlholm et al., 2010, 2009).

2.4.3. Area under the receiver operating characteristic curve

The receiver operating characteristic curve (ROC) was created by plotting the true positive rate against the false positive rate at various thresholds. The area under the ROC curve takes a value between 0 and 1. A value of 0.5 means that the model prediction is random, and a value of 1 means that the model predicts perfectly (Agresti, 2002). The higher value of the area under the ROC curve, the better is the prediction. In general, a value greater than 0.7 is considered acceptable, and a value greater than 0.8 is considered excellent (Hosmer and Lemeshow, 2000).

3. Results

3.1. Goodness-of-fit evaluation of the developed models

Fig. 1 shows a representative result of growth/no growth responses at 0.1% acetic acid condition for *B. subtilis* after 12 weeks of incubation. The growth/no growth boundaries were successfully estimated by logistic regression (Fig. 1A), neural network (Fig. 1B), and deep learning (Fig. 1C) models. In almost all conditions, the three models were well fitted to the growth/no growth boundary. The estimated parameters for the logistic regression model are shown in Table 1. However, the logistic regression model showed an illogical boundary in some conditions, e.g. in the case of growth/no growth responses at pH 5.5 for *B. subtilis* after 12 weeks of incubation (Fig. 2). Although the logistic regression model occasionally failed to fit adequately the growth/no growth boundary, the neural network and the deep learning models enabled better fitting in such cases (Fig. 2). As illustrated in Fig. 1, the prediction by the neural network model yielded a narrower width of probability contour lines compared with that of the logistic regression model. Furthermore, the deep learning model yielded the narrowest width of probability contour lines (it seems to be a single line) compared with those of other two models.

C-matrix coefficients and related performance indices for the models based on the trainig data and validation data are summarized in Tables 2 and 3, respectively. The models showed high accuracy, namely FC > 0.90 and AUC > 0.95, in all models. The values of the variability within replication were small for both the neural network and deep learning models (Table 2). Although the accuracies of the three models were almost comparable, the accuracy of the deep learning model was

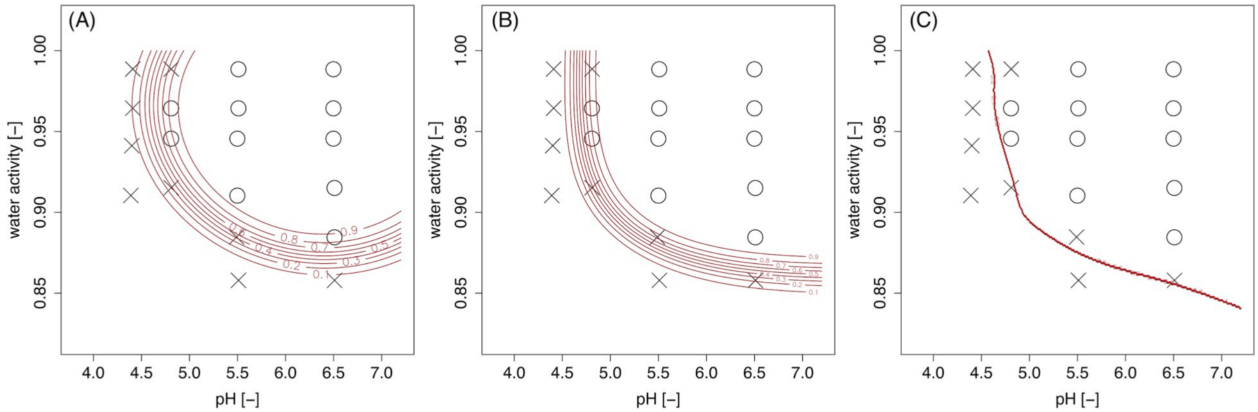

Fig. 1. Observed data for growth (○)/no growth (×) of *Bacillus subtilis* as a function of a_w and pH at 0.1% acetic acid after 12week incubation and representative growth probability contour lines obtained by using logistic regression (A), neural network (B), and deep learning (C) models.

Table 1
Statistical parameters obtained for logistic regression model for growth limits of *Bacillus* spores.

Factors	Coefficient.	Std. error	*p*-value
Intercept	−1095	53.97	0.000
pH	41.48	4.904	0.000
a_w	2007	92.20	0.000
AcOH	−4.415	10.52	0.675
lact	−83.31	14.66	0.000
time	1.695	0.353	0.000
Bs	−0.156	0.084	0.064
Bm	−0.808	0.086	0.000
pH^2	−2.778	0.147	0.000
$a_w{}^2$	−1013	40.93	0.000
$AcOH^2$	7.819	0.862	0.000
$lact^2$	−6.693	1.066	0.000
$time^2$	−0.017	0.003	0.000
pH × a_w	−6.751	3.906	0.084
pH × *AcOH*	7.720	0.568	0.000
pH × *lact*	8.350	0.844	0.000
pH × *time*	−0.066	0.018	0.000
a_w × *AcOH*	−59.64	8.796	0.000
a_w × *lact*	42.95	11.73	0.000
a_w × *time*	−1.133	0.294	0.000
AcOH × *time*	0.230	0.049	0.000
lact × *time*	–	–	NS*

NS*: not significant.

slightly higher than those of the other two models.

3.2. Model validation

3.2.1. Culture medium

Fig. 3 shows a representative result of growth/no growth responses at 0.3% of acetic acid for *B. subtilis* after 12 weeks of incubation in TSB with predictions of the boundary condition estimated by the logistic regression (Fig. 3A), neural network (Fig. 3B), and deep learning (Fig. 3C) models. Although all models successfully predicted growth boundaries against independently observed data, the machine learning procedures (neural network and deep learning models) allowed narrower and sharper predictions than that afforded by the logistic regression model.

Table 4 shows the comparison of predictive performances of the three classifiers using C-matrix and performance indices for the independent validation data of growth in TSB. All models were highly accurate with FC > 0.90 and AUC > 0.95. Although the accuracies were slightly lower when compared with the predictions for the model development data, the predictions for the independent validation data were nonetheless sufficiently accurate. Although the accuracies of all three models were almost comparable, the accuracy of the deep learning model was slightly higher than those of the other two models. In particular, deep learning model exhibited lower fail-dangerous and higher fail-safe ratios than those afforded by the logistic regression and the neural network models.

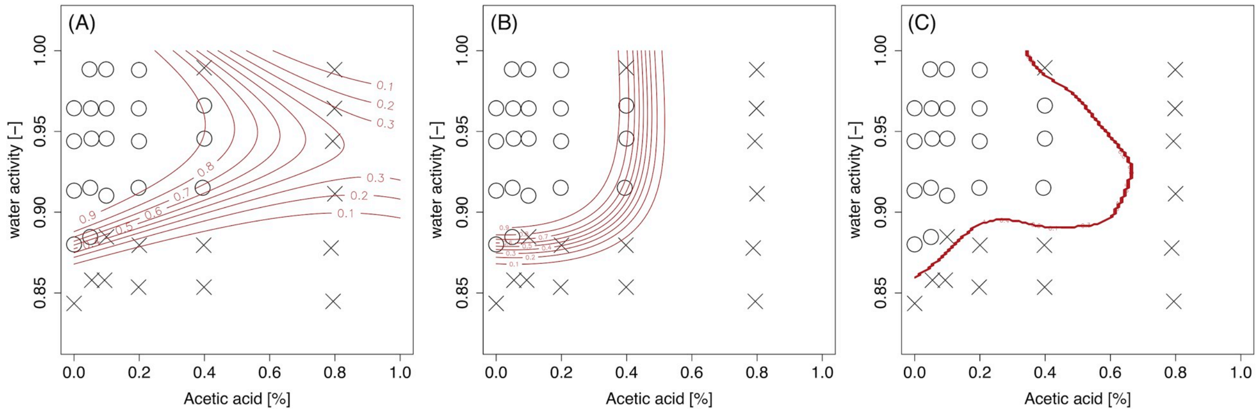

Fig. 2. Observed data for growth (○)/no growth (×) of *Bacillus subtilis* as a function of a_w and acetic acid concentration at pH 5.5 after 12week incubation and representative growth probability contour lines obtained by using logistic regression (A), neural network (B), and deep learning (C) models.

Table 2

Comparison of predictive performance of three classifiers using C-matrix and performance measures based on training data. The output of the models are probability of growth. The threshold of the probability is 0.5.

Modeling approach		[a]*n*	[b]C-matrix				Accuracy indices					
			TN	FP	FN	TP	[c]FC	[d]FAR	[e]POD	Fail-dangerous (%)	Fail-safe (%)	[f]AUC
Logistic regression		14340	7301	477	529	6033	0.930	0.061	0.919	3.7	3.3	0.980
Neural network	[g]Replicate											
	1	14340	7313	465	505	6057	0.932	0.060	0.923	3.5	3.2	0.981
	2	14340	7291	487	518	6044	0.930	0.063	0.921	3.6	3.4	0.978
	3	14340	7296	482	543	6019	0.929	0.062	0.917	3.8	3.4	0.979
	4	14340	7284	494	516	6046	0.930	0.064	0.921	3.6	3.4	0.978
	5	14340	7355	423	463	6099	0.938	0.054	0.929	3.2	2.9	0.982
	6	14340	7276	502	542	6020	0.927	0.065	0.917	3.8	3.5	0.974
	7	14340	7408	370	527	6035	0.937	0.048	0.920	3.7	2.6	0.981
	8	14340	7298	480	547	6015	0.928	0.062	0.917	3.8	3.3	0.975
	9	14340	7293	485	435	6127	0.936	0.062	0.934	3.0	3.4	0.972
	10	14340	7297	481	458	6104	0.935	0.062	0.930	3.2	3.4	0.980
	Mean						0.932	0.060	0.923	3.5	3.3	0.978
Deep learning	Replicate											
	1	14340	7406	372	307	6255	0.953	0.048	0.953	2.1	2.6	0.991
	2	14340	7280	498	481	6081	0.932	0.064	0.927	3.4	3.5	0.981
	3	14340	7364	414	487	6075	0.937	0.053	0.926	3.4	2.9	0.986
	4	14340	7470	308	414	6148	0.950	0.040	0.937	2.9	2.1	0.990
	5	14340	7529	249	445	6117	0.952	0.032	0.932	3.1	1.7	0.987
	6	14340	7298	480	592	5970	0.925	0.062	0.910	4.1	3.3	0.975
	7	14340	7410	368	504	6058	0.939	0.047	0.923	3.5	2.6	0.981
	8	14340	7513	265	422	6140	0.952	0.034	0.936	2.9	1.8	0.986
	9	14340	7364	414	361	6201	0.946	0.053	0.945	2.5	2.9	0.985
	10	14340	7493	285	391	6171	0.953	0.037	0.940	2.7	2.0	0.985
	Mean						0.944	0.047	0.933	3.1	2.5	0.985

[a] *n*: numbers of tested condition.

[b] C-matrix: confusion matrix, TN; true negative, FP; false positive, FN; false negative, TP; true positive.

[c] FC: fraction correct.

[d] FAR: false alarm rate.

[e] POD: probability of detection.

[f] AUC: area under the ROC curve.

[g] Replicate: numbers of calculation.

3.2.2. *Clam soup*

Fig. 4 shows a representative result of growth/no growth responses at 0.3% acetic acid for *B. subtilis* after 4 weeks of incubation in clam soup with the predictions of the boundaries estimated by the logistic regression (Fig. 4A), neural network (Fig. 4B), and deep learning (Fig. 4C) models. Compared with the validation results of the experiments in culture medium, the accuracies of prediction for clam soup were slightly lower in all models. Although there were no significant differences among the three models tested, the deep learning procedure illustrates a prediction with narrower growth boundaries than with the other two models. The deep learning model accurately predicted growth probability. However, the accuracy of the boundary prediction by the neural network model was lower when compared with the prediction results of the experiments in culture medium.

Table 4 shows the comparison of predictive performances of the three classifiers using C-matrix and performance indices for the independent validation data of growth in clam soup. All models showed high accuracy with FC > 0.80 and AUC > 0.90. The accuracies of the three models were comparable, and for all models, the ratio of fail-safe was higher than that of fail-dangerous. Although the accuracies were overall lower when compared with the prediction results obtained in the experiments with culture medium, nonetheless the predictions for clam soup were acceptably accurate.

4. Discussion

This study successfully demonstrated the application of three modeling techniques for determining the growth/no growth boundary of *Bacillus* species as a function of six environmental factors. The machine learning (neural network and deep learning) procedures exhibited equal or better performance than that of the conventional logistic regression method. All models developed in the present study showed highly accurate predictions with AUC values over 0.95. Hosmer and Lemeshow (2000) noted that an AUC value greater than 0.7 is considered acceptable, and the value greater than 0.8 is considered excellent. All three models tested enabled an accurate description of the actual growth/no growth boundary of *Bacillus* species.

The predictions of machine learning models showed more flexibility than those of the logistic regression model. Logistic regression failed to approximate accurately the growth boundary under some conditions (Fig. 2A). In contrast, machine learning always described the growth boundary accurately under the same conditions (Fig. 2B and C). To improve the accuracy of the logistic regression model, square-root type model would be applicable (Vermeulen et al., 2007). However, we focused on simply modeling based on the observed data as much as possible to compare different modeling procedures in the present study. Valero et al. (2007) obtained a similar result to that of the present study, having shown that machine learning was more flexible than logistic regression. In particular, Valero et al. (2007) reported that growth probability of the neural network model was more sharply defined by small changes in the levels of the factors than that of the logistic regression model. Therefore, machine learning outperformed logistic regression in terms of classification flexibility without increasing the number of coefficients of the model. However, in turn, the flexibility of machine learning might cause overfitting due to data-driven prediction.

We compared the three modeling approaches in terms of the probability contour line accuracy. The deep learning model yielded the narrowest width of probability contour lines, which seemed to be one

Table 3
Comparison of predictive performance of three classifiers using C-matrix and performance measures based on validation data of tryptic soy broth. The output of the models are probability of growth. The threshold of the probability is 0.5.

Modeling approach		[a]*n*	[b]C-matrix				Accuracy indices					
			TN	FP	FN	TP	[c]FC	[d]FAR	[e]POD	Fail-dangerous (%)	Fail-safe (%)	[f]AUC
Logistic regression		3015	1185	26	255	1549	0.907	0.021	0.859	8.5	0.9	0.982
Neural network	[g]Replicate											
	1	3015	1180	31	263	1541	0.902	0.026	0.854	8.7	1.0	0.974
	2	3015	1173	38	231	1573	0.911	0.031	0.872	7.7	1.3	0.972
	3	3015	1194	17	261	1543	0.908	0.014	0.855	8.7	0.6	0.970
	4	3015	1200	11	258	1546	0.911	0.009	0.857	8.6	0.4	0.973
	5	3015	1174	37	266	1538	0.900	0.031	0.853	8.8	1.2	0.967
	6	3015	1189	22	257	1547	0.907	0.018	0.858	8.5	0.7	0.967
	7	3015	1200	11	270	1534	0.907	0.009	0.850	9.0	0.4	0.963
	8	3015	1188	23	258	1546	0.907	0.019	0.857	8.6	0.8	0.952
	9	3015	1172	39	205	1599	0.919	0.032	0.886	6.8	1.3	0.941
	10	3015	1184	27	262	1542	0.904	0.022	0.855	8.7	0.9	0.975
	Mean						0.908	0.021	0.860	8.4	0.8	0.965
Deep learning	Replicate											
	1	3015	1093	118	140	1664	0.914	0.097	0.922	4.6	3.9	0.969
	2	3015	1052	159	97	1707	0.915	0.131	0.946	3.2	5.3	0.973
	3	3015	1088	123	148	1656	0.910	0.102	0.918	4.9	4.1	0.957
	4	3015	1126	85	155	1649	0.920	0.070	0.914	5.1	2.8	0.972
	5	3015	1110	101	165	1639	0.912	0.083	0.909	5.5	3.3	0.962
	6	3015	1142	69	157	1647	0.925	0.057	0.913	5.2	2.3	0.974
	7	3015	1106	105	124	1680	0.924	0.087	0.931	4.1	3.5	0.974
	8	3015	1112	99	148	1656	0.918	0.082	0.918	4.9	3.3	0.973
	9	3015	1143	68	195	1609	0.913	0.056	0.892	6.5	2.3	0.959
	10	3015	1123	88	203	1601	0.903	0.073	0.887	6.7	2.9	0.965
	Mean						0.916	0.084	0.915	5.1	3.4	0.968

[a] *n*: numbers of tested condition.
[b] C-matrix: confusion matrix, TN; true negative, FP; false positive, FN; false negative, TP; true positive.
[c] FC: fraction correct.
[d] FAR: false alarm rate.
[e] POD: probability of detection.
[f] AUC: area under the ROC curve.
[g] Replicate: numbers of calculation.

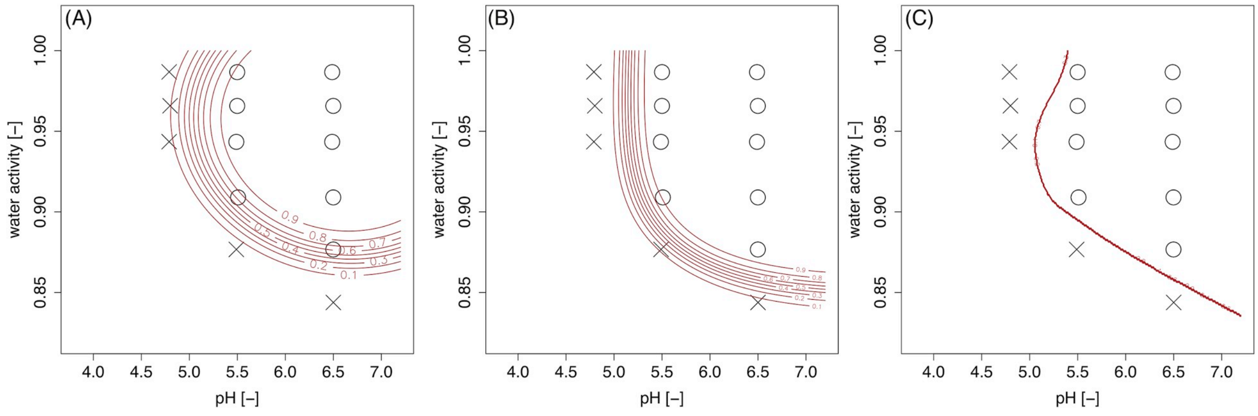

Fig. 3. Observed data for growth (◯)/no growth (×) of *Bacillus subtilis* as a function of a_w and pH at 0.3% acetic acid after 12week incubation in tryptic soy broth with the predictions of growth probability estimated by logistic regression (A), neural network (B), and deep learning (C) models.

line due to considerable overlapping compared with the widths calculated by other two approaches. The width of probability contour lines derived from the neural network model was slightly narrower than that of the logistic regression model. Whereas the logistic regression model predicted that the shift from growth to no growth condition is gradual, the deep learning model predicted that shift to be abrupt. Hajmeer and Basheer (2003b) reported a similar result as their hybrid artificial neural network built by the Bayesian approach yielded the narrower width of probability contour lines compared with that afforded by the nonlinear logistic regression model. In addition, Le Marc et al. (2005), Presser et al. (1998), and Tienungoon et al. (2000) reported that the shift from growth to no growth was rapid and sometimes, it was even close to the limit of pH and water activity measurements. These reports (Le Marc et al., 2005; Presser et al., 1998; Tienungoon et al., 2000) are therefore in agreement with the performance of our deep learning model, which predicted that the shift from growth to no growth is rapid. In addition, since the deep learning model might tend to overfit to the observed data, there are still rooms for improving for the deep learning algorithm by tuning the detail of calculation procedure in the future study. Furthermore, because the number of repetitions of one

Table 4
Comparison of predictive performance of three classifiers using C-matrix and performance measures based on validation data of clam soup. The output of the models are probability of growth. The threshold of the probability is 0.5.

Modeling approach		[a]*n*	[b]C-matrix				Accuracy indices					
			TN	FP	FN	TP	[c]FC	[d]FAR	[e]POD	Fail-dangerous (%)	Fail-safe (%)	[f]AUC
Logistic regression		405	101	45	22	237	0.835	0.308	0.915	5.4	11.1	0.925
Neural network	[g]Replicate											
	1	405	116	30	16	243	0.886	0.205	0.938	4.0	7.4	0.948
	2	405	92	54	19	240	0.820	0.370	0.927	4.7	13.3	0.909
	3	405	107	39	31	228	0.827	0.267	0.880	7.7	9.6	0.920
	4	405	113	33	37	222	0.827	0.226	0.857	9.1	8.1	0.918
	5	405	119	27	16	243	0.894	0.185	0.938	4.0	6.7	0.931
	6	405	92	54	43	216	0.760	0.370	0.834	10.6	13.3	0.903
	7	405	119	27	16	243	0.894	0.185	0.938	4.0	6.7	0.949
	8	405	104	42	16	243	0.857	0.288	0.938	4.0	10.4	0.926
	9	405	101	45	16	243	0.849	0.308	0.938	4.0	11.1	0.931
	10	405	119	27	16	243	0.894	0.185	0.938	4.0	6.7	0.938
	Mean						0.851	0.259	0.913	5.6	9.3	0.927
Deep learning	[g]Replicate											
	1	405	107	39	16	243	0.864	0.267	0.938	4.0	9.6	0.891
	2	405	110	36	16	243	0.872	0.247	0.938	4.0	8.9	0.908
	3	405	110	36	16	243	0.872	0.247	0.938	4.0	8.9	0.904
	4	405	116	30	16	243	0.886	0.205	0.938	4.0	7.4	0.924
	5	405	107	39	16	243	0.864	0.267	0.938	4.0	9.6	0.874
	6	405	110	36	16	243	0.872	0.247	0.938	4.0	8.9	0.923
	7	405	90	56	15	244	0.825	0.384	0.942	3.7	13.8	0.868
	8	405	104	42	16	243	0.857	0.288	0.938	4.0	10.4	0.890
	9	405	110	36	16	243	0.872	0.247	0.938	4.0	8.9	0.897
	10	405	111	35	15	244	0.877	0.240	0.942	3.7	8.6	0.919
	Mean						0.866	0.264	0.939	3.9	9.5	0.900

[a] *n*: numbers of tested condition.
[b] C-matrix: confusion matrix, TN; true negative, FP; false positive, FN; false negative, TP; true positive.
[c] FC: fraction correct.
[d] FAR: false alarm rate.
[e] POD: probability of detection.
[f] AUC: area under the ROC curve.
[g] Replicate: numbers of calculation.

condition was only three times in this study, it will be necessary to carefully judge the accuracy of the probability estimation. Consequently, further studies are needed in order to clarify the accuracy of the prediction of growth probability.

Prediction accuracies of the developed models were validated by using independent experimental data of growth in TSB. Validation results indicated that predictions afforded by all three modeling approaches demonstrated sufficient accuracies. In addition, the results of model predictions were in concordance with ComBase (ID: B261_164–166) and other published works (Lanciotti et al., 2001; Marshall et al., 1971). For example, Lanciotti et al. (2001) reported that growth/no growth boundary condition of *B. cereus* in terms of pH was between 5.5 and 4.75 at 32.5 °C with a_w 0.960 and ethanol concentration 1.5% (w/w). Marshall et al. (1971) reported that growth limit of *B. subtilis* at 30 °C was $a_w = 0.90$ in the case of a_w adjustment with NaCl.

Results of the validation experiments for clam soup indicated that the growth boundary models developed in the present study correctly predicted growth conditions in clam soup with over 80% accuracy. Because the a_w level of the clam soup could not be changed due to product nature, and the predictive models developed in the present

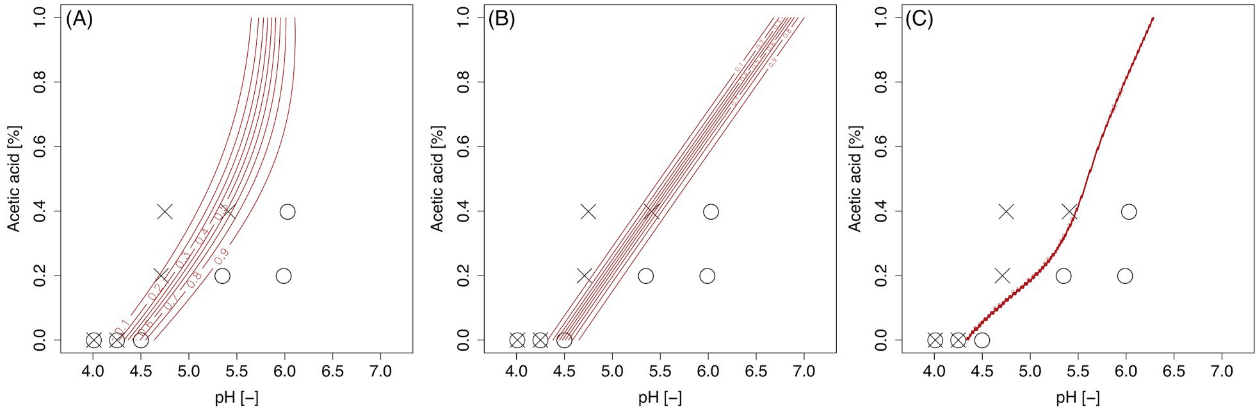

Fig. 4. Observed data for growth (○)/no growth (×) of *Bacillus subtilis* as a function of acetic acid and pH at a_w 0.98 after 4week incubation in clam soup with the predictions of growth probability estimated by logistic regression (A), neural network (B) and deep learning (C) models.

study include the extent of the storage period, a 4-week storage duration was set as an example for validation. Performances of our predictive models for this food matrix were comparable with the logistic regression model predictions of growth/no growth, such as 87.7% for *Listeria monocytogenes* in cheese (Bolton et al., 1999) and with the neural network model prediction of growth/no growth of 83% for *L. monocytogenes* in processed and ready-to-eat foods (Mejlholm et al., 2010). As shown in Fig. 4 (c), even though the deep learning model could not accurately predict in some cases such as lower pH conditions, it might be due to limitation of the model performance and/or uncertainty of experimental observed results. To improve the model accuracy, further inputs possibly affecting bacterial growth such as nutrition and/or spices should be included in the future work.

All the models developed in the present study provided high fail-safe predictions for clam soup indicating that growth was predicted in some cases when no growth actually occurred. The fail-safe predictions are known to be biased toward safe predictions. For example, McKellar et al. (2002) sought to predict the interface between survival and death of *Escherichia coli* O157:H7 in a mayonnaise model system using logistic regression. Although the results of validation experiments indicated that the survival model correctly predicted survival in mayonnaise in most cases with one exception, the failure was considered to be fail-safe. McKellar et al. (2002) suggested that the presence of additional antimicrobial agents in mayonnaise (e.g., lysozyme in whole eggs) may explain the increased microbial sensitivity observed in the validation data relative to that in model development data that were based on *E. coli* growth in TSB. Thus, the results of McKellar et al. (2002) suggested that the predictions in food using logistic regression were safer than the predictions about growth in TSB. The results reported by McKellar et al. (2002) were consistent with the results of the present study. In addition, the deep learning procedure generated fewer false negative cases (i.e., cases when no growth was predicted, but it was observed) than other models. Therefore, deep learning was the best model in the present study that predicted safely the properties of bacteria in clam soup. This fact is important because deep learning would be possibly applied to predict in the near future and in other foods bacterial growth or no-growth conditions in accumulating experimental datasets.

In conclusion, growth/no growth boundary models for *Bacillus* spores using logistic regression, neural network, and deep learning were successfully developed. The predictive models based on neural network and deep learning procedures enabled flexible predictions for independent conditions in culture media. In addition, the deep learning model allowed a safe prediction of the growth boundary of *Bacillus* spores in clam soup. In future, the deep learning algorithm will be able to predict more accurately and flexibly by accumulating and learning new experimental datasets and environmental parameters. Accordingly, machine learning approaches will help food processing establishments to efficiently design minimally processed high quality food whilst ensuring microbiological safety.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fm.2018.09.013.

References

Agresti, A., 2002. Categorical Data Analysis. Wiley, New York.

Bolton, L.F., protection, J.F.J.O.F., 1999. Defining the growth/no-growth interface for *Listeria monocytogenes* in Mexican-style cheese based on salt, pH, and moisture content. J. Food Protect. 62, 601–609. https://doi.org/10.4315/0362-028X-62.6.601.

ComBase Modelling Toolbox. ID: B261_164–166. http://ifrsvwwwdev.ifrn.bbsrc.ac.uk/CombasePMP/GP/(accessed 1 February, 2018).

Daelman, J., Vermeulen, A., Willemyns, T., Ongenaert, R., Jacxsens, L., Uyttendaele, M., Devlieghere, F., 2013. Growth/no growth models for heat-treated psychrotrophic *Bacillus cereus* spores under cold storage. Int. J. Food Microbiol. 161, 7–15. https://doi.org/10.1016/j.ijfoodmicro.2012.11.017.

Dang, T.D.T., Mertens, L., Vermeulen, A., Geeraerd, A.H., Van Impe, J.F., Debevere, J., Devlieghere, F., 2010. Modelling the growth/no growth boundary of *Zygosaccharomyces bailii* in acidic conditions: a contribution to the alternative method to preserve foods without using chemical preservatives. Int. J. Food Microbiol. 137, 1–12. https://doi.org/10.1016/j.ijfoodmicro.2009.10.020.

Esnoz, A., Periago, P.M., Conesa, R., Palop, A., 2006. Application of artificial neural networks to describe the combined effect of pH and NaCl on the heat resistance of *Bacillus stearothermophilus*. Int. J. Food Microbiol. 106, 153–158. https://doi.org/10.1016/j.ijfoodmicro.2005.06.016.

Evans, D., 2004. Use of survival analysis and Classification and Regression Trees to model the growth/no growth boundary of spoilage yeasts as affected by alcohol, pH, sucrose, sorbate and temperature. Int. J. Food Microbiol. 92, 55–67. https://doi.org/10.1016/j.ijfoodmicro.2003.07.008.

Fernández-Navarro, F., Valero, A., Hervás-Martínez, C., Gutiérrez, P.A., García-Gimeno, R.M., Zurera-Cosano, G., 2010. Development of a multi-classification neural network model to determine the microbial growth/no growth interface. Int. J. Food Microbiol. 141, 203–212. https://doi.org/10.1016/j.ijfoodmicro.2010.05.013.

Hajmeer, M., Basheer, I., 2003a. Comparison of logistic regression and neural network-based classifiers for bacterial growth. Food Microbiol. 20, 43–55. https://doi.org/10.1016/S0740-0020(02)00104-1.

Hajmeer, M., Basheer, I., 2002. A probabilistic neural network approach for modeling and classification of bacterial growth/no-growth data. J. Microbiol. Meth. 51, 217–226. https://doi.org/10.1016/S0167-7012(02)00080-5.

Hajmeer, M.N., Basheer, I.A., 2003b. A hybrid Bayesian–neural network approach for probabilistic modeling of bacterial growth/no-growth interface. Int. J. Food Microbiol. 82, 233–243. https://doi.org/10.1016/S0168-1605(02)00308-2.

Hosmer, D.W., Lemeshow, S., 2000. Applied Logistic Regression, second ed. Wiley, New York.

Jeyamkondan, S., Jayas, D.S., Holley, R.A., 2001. Microbial growth modelling with artificial neural networks. Int. J. Food Microbiol. 64, 343–354. https://doi.org/10.1016/S0168-1605(00)00483-9.

Lanciotti, R., Sinigaglia, M., Gardini, F., Vannini, L., Guerzoni, M.E., 2001. Growth/no growth interfaces of *Bacillus cereus, Staphylococcus aureus* and *Salmonella enteritidis* in model systems based on water activity, pH, temperature and ethanol concentration. Food Microbiol. 18, 659–668. https://doi.org/10.1006/fmic.2001.0429.

Le Marc, Y., Pin, C., Baranyi, J., 2005. Methods to determine the growth domain in a multidimensional environmental space. Int. J. Food Microbiol. 100, 3–12. https://doi.org/10.1016/j.ijfoodmicro.2004.10.003.

Lou, W., Nakai, S., 2001. Application of artificial neural networks for predicting the thermal inactivation of bacteria: a combined effect of temperature, pH and water activity. Food Res. Int. 34, 573–579. https://doi.org/10.1016/S0963-9969(01)00074-6.

Marshall, B.J., Ohye, D.F., Christian, J.H.B., 1971. Tolerance of bacteria to high concentrations of NaCl and glycerol in the growth medium. Appl. Microbiol. 21, 363–364.

McKellar, R.C., Lu, X., Delaquis, P.J., 2002. A probability model describing the interface between survival and death of *Escherichia coli* O157:H7 in a mayonnaise model system. Food Microbiol. 19, 235–247. https://doi.org/10.1006/fmic.2001.0449.

Mejlholm, O., Gunvig, A., Borggaard, C., Blom-Hanssen, J., Mellefont, L., Ross, T., Leroi, F., Else, T., Visser, D., Dalgaard, P., 2010. Predicting growth rates and growth boundary of *Listeria monocytogenes* — an international validation study with focus on processed and ready-to-eat meat and seafood. Int. J. Food Microbiol. 141, 137–150. https://doi.org/10.1016/j.ijfoodmicro.2010.04.026.

Mejlholm, O., protection, P.D.J.O.F., 2009. 2009. Development and validation of an extensive growth and growth boundary model for *Listeria monocytogenes* in lightly preserved and ready-to-eat shrimp. J. Food Protect. 72, 2132–2143. https://doi.org/10.4315/0362-028X-72.10.2132.

Polese, P., Del Torre, M., Spaziani, M., Stecchini, M.L., 2011. A simplified approach for modelling the bacterial growth/no growth boundary. Food Microbiol. 28, 384–391. https://doi.org/10.1016/j.fm.2010.09.011.

Presser, K.A., Ross, T., Ratkowsky, D.A., 1998. Modelling the growth limits (growth/no growth interface) of *Escherichia coli* as a function of temperature, pH, lactic Acid Concentration, and water activity. Appl. Environ. Microbiol. 64, 1773–1779.

Quintavalla, S., Parolari, G., 1993. Effects of temperature, a_w and pH on the growth of *Bacillus cells* and spores: a response surface methodology study. Int. J. Food Microbiol. 19, 207–216. https://doi.org/10.1016/0168-1605(93)90078-U.

Tienungoon, S., Ratkowsky, D.A., McMeekin, T.A., Ross, T., 2000. Growth limits of Listeria monocytogenes as a function of temperature, pH, NaCl, and lactic acid. Appl. Environ. Microbiol. 66, 4979–4987. https://doi.org/10.1128/AEM.66.11.4979-4987.2000.

Valero, A., Hervas, C., Garciagimeno, R., Zurera, G., 2007. Product unit neural network models for predicting the growth limits of *Listeria monocytogenes*. Food Microbiol. 24, 452–464. https://doi.org/10.1016/j.fm.2006.10.002.

Vermeulen, A., Devlieghere, F., Bernaerts, K., Van Impe, J., Debevere, J., 2007. Growth/no growth models describing the influence of pH, lactic and acetic acid on lactic acid bacteria developed to determine the stability of acidified sauces. Int. J. Food Microbiol. 119, 258–269. https://doi.org/10.1016/j.ijfoodmicro.2007.08.003.