Received: 18 February 2021 | Revised: 6 May 2021 | Accepted: 18 July 2021

DOI: 10.1111/jfpp.15819

ORIGINAL ARTICLE

Journal of Food Processing and Preservation | Institute of Food Science + Technology | ifst | WILEY

Development and validation of a neural network model for growth of *Salmonella* Newport from chicken on cucumber for use in risk assessment

Thomas P. Oscar

U. S. Department of Agriculture, Agricultural Research Service, Chemical Residue and Predictive Microbiology Research Unit, Center for Food Science and Technology, University of Maryland Eastern Shore, Princess Anne, Maryland, USA

Correspondence
Thomas P. Oscar, U. S. Department of Agriculture, Agricultural Research Service, Chemical Residue and Predictive Microbiology Research Unit, Room 2111, Center for Food Science and Technology, University of Maryland Eastern Shore, Princess Anne, MD, USA 21853.
Email: thomas.oscar@usda.gov

Abstract

A neural network model was developed for predicting growth of a chicken isolate of *Salmonella* Newport on cucumber portions as a function of times (0 to 8 hr) and temperatures (16 to 40°C) observed during meal preparation and serving for use in risk assessment. Model development and validation were accomplished using the test data, model performance, and model validation criteria of the Acceptable Prediction Zones (APZ) method in the Validation Software Tool (ValT). The model was considered to provide acceptable predictions when the proportion of residuals in the APZ (pAPZ) was ≥0.70. Data for model development (n = 140) and validation (n = 72) satisfied all criteria of the APZ method in ValT with pAPZ of 0.97 and 0.93, respectively. Thus, the model was successfully validated and can be used with confidence in risk assessment to predict growth of *Salmonella* Newport from chicken on cucumber during meal preparation and serving.

Novelty impact statement: The model can be used by risk assessors to help predict variability and uncertainty of consumer exposure to *Salmonella* from individual lots of chicken produced by a farm-to-table scenario that includes cross-contamination of sliced cucumbers with *Salmonella* Newport from raw chicken followed by growth of *Salmonella* Newport on cucumber for times and temperatures observed during meal preparation and serving.

1 | INTRODUCTION

Models that predict growth of pathogens on or in food are valuable tools for verifying compliance with performance standards for food safety. For example, models can be used to verify less than one-log of growth of *Clostridium perfringens* during cooling of commercially cooked red meat and poultry products as part of a Hazard Analysis and Critical Control Points (HACCP) program for food safety (Mohr et al., 2015; Smith-Simpson & Schaffner, 2005). In addition, predictive models can be used in risk assessments to help determine consumer exposure to pathogens in food produced by different farm-to-table scenarios (Gonzalez et al., 2018; Santillana Farakos et al., 2016). This is an important application of predictive models because risk assessments are used to establish new food safety regulations and practices aimed at protecting public health (Lambertini et al., 2019; Membre & Boue, 2018).

Risk assessment can also be used at the processing plant to identify unsafe food before it is shipped to consumers and causes foodborne illness (Notermans et al., 1995; Oscar, 1998). This can be done by simulating the food production chain as a series of unit operations (e.g., retail transport … meal preparation … serving) and pathogen events (e.g., growth, death, survival; Oscar, 2004; Whiting & Buchanan, 1997). Predictive models can be used to forecast how pathogen number changes within each unit operation of the food production chain (Buchanan & Whiting, 1996).

Many cases of human salmonellosis are attributed to chicken (Painter et al., 2013; Scharff, 2020). Consumers can be exposed to *Salmonella* that survive during undercooking of chicken or they can be exposed to *Salmonella* that cross-contaminate and grow on ready-to-eat (RTE) food that are prepared and served with chicken (Akil & Ahmad, 2019; Maffei et al., 2017). Models that predict growth of *Salmonella* on RTE food following cross-contamination from chicken are valuable tools for risk assessment and food safety (Jayeola et al., 2019; Scolforo et al., 2017) but are not widely available.

Salad is an RTE food that is often served with chicken and it often contains diced or sliced cucumber. Although growth of *Salmonella* on cucumber has been investigated (Elexson et al., 2011; Ha et al., 2020), insufficient data were collected for full development and validation of predictive models. In addition, there are no data or models for growth of chicken isolates of *Salmonella* on cucumber. This is important because *Salmonella* Newport from cucumbers grown on the Delmarva Peninsula, an area of intense chicken production, caused an outbreak of human salmonellosis in 2014 (Angelo et al., 2015). Thus, it is possible that isolates of *Salmonella* from chicken could cross-contaminate cucumbers in the field or during meal preparation and serving and then grow to high numbers and cause foodborne illness. Therefore, to provide a valuable tool for better assessing this risk to public health, this study was undertaken to develop and validate a predictive model for growth of a chicken isolate of *Salmonella* Newport on cucumber as a function of times (0 to 8 hr) and temperatures (16 to 40°C) observed during meal preparation and serving (da Silva et al., 2020) for use in risk assessment.

The traditional approach in predictive microbiology is to use linear and nonlinear regression methods to develop models that predict pathogen growth as a function of time, temperature, and other important independent variables (e.g., pH, water activity, atmosphere) associated with food (Buchanan & Phillips, 1990; Palumbo et al., 1991). This is a three-step process that involves: (1) primary modeling; (2) secondary modeling; and (3) tertiary modeling (Whiting, 1995). In primary modeling, pathogen growth over time for single combinations of independent variables being investigated are fitted to a regression model to obtain primary model parameters like lag time, growth rate, and maximum population density (Baranyi et al., 1993; Huang, 2014). In secondary modeling, linear and nonlinear regression models are used to predict the primary model parameters as a function of the independent variables being investigated (Gibson et al., 1988; McClure et al., 1993). Finally, in tertiary modeling, secondary models for the primary model parameters are used in the primary model to predict pathogen growth over time as a function of the ranges of independent variables investigated and usually in the form of a user-friendly computer software application (Buchanan, 1993; McClure et al., 1994). Limitations of this approach are: (1) it is time-consuming; (2) regression models are inflexible, which can result in local prediction bias; and (3) prediction errors accumulate during each step of the process. One way to improve this process is by using a global regression method that combines the three steps into one (Jia et al., 2020; Martino & Marks, 2007).

Another one-step approach that can be used to develop predictive models for pathogen growth is neural network modeling (Hajmeer et al., 1997; Najjar et al., 1997). This is a less popular method in the field of predictive microbiology. Nonetheless, it learns patterns in data using methods that simulate how the human brain processes information and is the basis for new technologies such as self-driving cars. One advantage of neural network modeling versus regression methods is their flexibility to model diverse patterns of pathogen growth without prediction bias. However, it is important to not over-train neural networks or they will not generalize or interpolate well. Thus, it is important to validate them against a properly collected independent set of data.

With the advent of commercial software programs (e.g., NeuralTools), it is now easy to develop neural network models for predictive microbiology applications (Oscar, 2015). In addition, some neural network programs are compatible with Monte Carlo simulation software programs (e.g., @Risk) that allow use of probability distributions in the model for simulating variability and uncertainty of independent variables, which is an important characteristic of predictive models used in risk assessment (Oscar, 2009). Therefore, in this study, compatible software programs (i.e., Excel, NeuralTools, and @Risk) were used to develop a neural network model that could provide stochastic predictions of *Salmonella* growth on cucumber as a function of variable and uncertain times and temperatures observed during meal preparation and serving (da Silva et al., 2020).

Proper validation of models is important because it provides users of models with confidence that predictions are reliable (Delignette-Muller et al., 1995; Ross, 1996; Walls & Scott, 1996). In addition, it helps model developers to identify prediction problems that can be repaired before models are distributed to end users and are used to make important food safety decisions. Like traditional predictive model development, model validation is a three-step process that involves evaluating model performance for: (1) dependent data or data used in model development; (2) independent data for interpolation or data not used in model development but that are within the prediction range of the model; and (3) independent data for extrapolation or data not used in model development but that are outside the prediction range of the model (Oscar, 2005).

To properly validate models, criteria are needed for test data (Oscar, 2005), model performance (Ross et al., 2000), and model validation (Oscar, 2020). These criteria help to ensure that a model is developed that provides accurate and unbiased predictions. In addition, they ensure that the model validation process is complete, accurate, unbiased, not confounded, and objective. The Acceptable Prediction Zones (APZ) method in the Validation Software Tool (ValT) is the only method in the field of predictive microbiology that has criteria for test data, model performance, and model validation (Oscar, 2020). Moreover, these criteria are statistically based and have been carefully developed over a long period of time. Consequently, models developed and validated with these criteria can be used with confidence in food safety and risk assessment applications. Thus, the APZ method in ValT was used in this study for guiding model development and for model evaluation and validation.

2 | MATERIALS AND METHODS

2.1 | Data collection

To simulate cross-contamination of cucumber with *Salmonella* from utensils used to prepare raw chicken for cooking, a chicken isolate of *Salmonella* Newport was grown for 96 hr at 22°C (meal preparation temperature) in buffered peptone water (BPW; Microbiology International, Frederick, MD) to obtain stationary phase cells (non-growing cells) for inoculation of cold (4°C) cucumber portions (0.2 g) from mesocarp with native microflora. Cucumbers obtained from a local retail store (Salisbury, MD, USA) and were the wax-coated slicing variety. Cucumber portions were inoculated with a low initial number (0.85 log) of *Salmonella* Newport because in a previous study (Oscar, 2017), it was found that the number of *Salmonella* transferred to RTE food from utensils used to prepare raw chicken for cooking was between 0 and 1 log when the raw chicken was properly stored (4°C for 6 hr) after purchase and before meal preparation.

Inoculated cucumber portions in 1.5 ml polystyrene tubes with caps were incubated at 16 to 40°C in heating and cooling blocks (ThermoStat Plus, Eppendorf, Hamburg, Germany). Heating and cooling blocks are a common piece of equipment in laboratories that do molecular biology research. They hold 1.5 ml microcentrifuge tubes, have a small footprint on the laboratory bench, and can simulate a wide range (−10 to 100°C) of temperatures encountered in the food production chain.

At designated times of storage, a sample was removed and 0.7 ml of cold (4°C) BPW was added to cover the cucumber portion and stop growth of *Salmonella*, which do not grow or die but survive at 4°C for an extended period (up to 10 days) of time (Oscar, 2011). After addition of BPW, samples were vortexed for 1 min at 3,000 rpm (Digital Disruptor Genie, Scientific Industries, Bohemia, NY) to recover *Salmonella* Newport into BPW for enumeration using a 6 replicate by 8 or 16 serial dilution (1:10) automated, whole sample enrichment, miniature most probable number (WSE-mMPN) method (Oscar, 2015). In brief, the three steps of the method were: (1) serial dilution (1:10) in BPW followed by incubation for 24 hr at 40°C; (2) transfer of 10 μl of BPW incubate to 1 ml of Rappaport Vassiliadis broth with novobiocin (RVBN) followed by incubation for 24 hr at 42°C; and (3) drop plating of 2 μl of RVBN incubate onto xylose lysine tergitol 4 (XLT4) agar followed by incubation for 24 hr at 40°C. The first two steps were conducted in 96-well, deep-well (2 ml) plates. Serial dilution, transfer, and drop plating were performed by a custom-designed robotic pipettor (SoloPlus, Hudson Robotics, Springfield, NJ). The MPN was calculated using an Excel calculator (Jarvis et al., 2010).

2.2 | Experimental designs

The experimental design for model development was a 5 × 7 full factorial of time (0, 2, 4, 6, 8 hr) and temperature (16, 20, 24, 28, 32, 36, 40°C) with four replicates per combination of independent variables. The experimental design for model validation (interpolation) was a 4 × 6 full factorial of time (1, 3, 5, 7 hr) and temperature (18, 22, 26, 30, 34, 38°C) with three replicates per combination of independent variables. All replicates were from different challenge trials.

These experimental designs were based on criteria of the APZ method in the ValT for predictive microbiology (Oscar, 2020). More specifically, even spacing of times and temperatures for model development, intermediate times and temperatures for model validation for interpolation, a minimum of four replications per combination of independent variables for model development, a minimum of two replications per combination of independent variables for model validation for interpolation, and use of the same data collection methods for model development and validation for interpolation. These criteria were developed and used to facilitate development of a model that makes accurate and unbiased predictions and to ensure that the model validation process was complete, accurate, unbiased, not confounded, and objective.

2.3 | Model development and simulation

The model was developed and simulated in Excel (Office 365; MicroSoft Corp., Redmond, WA) using NeuralTools (version 7.6, Palisade Corp., Ithaca, NY) and @Risk (version 7.6, Palisade Corp.), which are spreadsheet add-in programs. Data for model development ($n = 140$) and validation ($n = 72$) were arranged in four columns of an Excel spreadsheet: (1) tag (train or test); (2) temperature (°C; independent numerical variable); (3) time (hr; independent numerical variable); and (4) MPN (log/portion; dependent numerical variable). Data for model development were tagged "train," whereas data for model validation were tagged "test."

A multiple layer feedforward neural network with two hidden layers of two nodes each was trained using NeuralTools (Oscar, 2018). The predict function of NeuralTools was used to simulate growth of *Salmonella* Newport on cucumber as a function of time (0 to 8 hr) and temperature (16 to 40°C). In addition, to demonstrate how the model could be used in risk assessment, pert (minimum, most likely, maximum) distributions for times (0.5, 1, 6 hr) and temperatures (22, 25, 34°C) observed during meal preparation and serving (da Silva et al., 2020) were entered into the model and simulated with @Risk.

The @Risk settings for simulation of the model were Latin Hypercube sampling, Mersenne Twister, a random number generator seed of one, and 1,000 iterations. The BestFit option of @Risk and Akaike's Information Criterion were used to find the best fitting distribution for risk assessment.

2.4 | Model performance and validation

Model performance was evaluated using the APZ method in ValT (Oscar, 2020). Values for independent variables (time and temperature) and observed and predicted values (log MPN/portion) for

dependent data, and independent data for interpolation were entered in ValT, which then calculated residuals (observed – predicted) and assigned APZ values (0 to 1) for each prediction case ($n = 212$). Pivot tables were then used to count prediction cases and calculate pAPZ (proportion of residuals in the APZ) for each combination and level of independent variables and overall. These values were used to answer a series of "yes" or "no" questions for test data and model performance criteria that were organized in decision trees for dependent data and independent data for interpolation in ValT. Model validation occurred when all questions for both sets of data were answered in the affirmative ("yes").

Model performance was considered acceptable (i.e., acceptable prediction accuracy and bias) when pAPZ was ≥0.7 and there were no local prediction problems. A local prediction problem occurred when pAPZ was <0.7 for three consecutive combinations of independent variables or when pAPZ was <0.7 for one level of an independent variable.

The pAPZ or average APZ value for individual levels and combinations of independent variables and overall were calculated using four APZ: (1) 0 to −1 log/portion (fail-safe and fully acceptable); (2) < −1 to > −2 log/portion (fail-safe and partly acceptable); (3) 0 to 0.5 log/portion (fail-dangerous and fully acceptable); and (4) >0.5 and <1 log/portion (fail-dangerous and partly acceptable). Residuals in the fully acceptable APZ were assigned an APZ value of one, residuals in partly acceptable APZ were assigned APZ values from >0 to <1 depending on their linear distance from the corresponding fully acceptable APZ, and residuals outside the APZ were assigned an APZ value of zero.

3 | RESULTS

3.1 | *Salmonella* growth

Representative MPN data used for model development and validation are shown in Figure 1 along with predicted growth curves. These data show that *Salmonella* Newport growth on cucumber increased as a function of time and temperature. The data are shown to provide a representative sample for visual evaluation of data quality and repeatability among independent challenge trials. A more formal and complete analysis of data quality is provided below.

3.2 | Model predictions

Figure 2 shows the neural network model that was developed in Excel and was simulated with NeuralTools and @Risk. The model was used to predict growth of *Salmonella* Newport on cucumber for times and temperatures used and not used in model development but that were within ranges of time (0 to 8 hr) and temperature (16 to 40°C) used for model development. For example, in Figure 2, the model predicted growth of *Salmonella* Newport on cucumber from 0 to 8 hr at 25°C, a temperature that was not investigated but that was within the range of temperatures investigated and modeled.

The model was also used to make stochastic predictions for use in risk assessment. For example, for the variable time and temperature scenario shown in Figure 2, the model predicted that *Salmonella* Newport growth on cucumber would range from 0 to 1.2 with an average of a 0.05 log increase per portion. The distribution that fitted best to these results was a Kumaraswamy. This distribution can be used in a risk assessment model to simulate growth of *Salmonella* Newport on cucumber after cross-contamination from utensils used to prepare raw chicken for cooking.

3.3 | Model performance and validation

3.3.1 | Dependent data

Figure 3 shows representative plots of residuals for data used in model development and validation. These plots provide a visual assessment of model performance and were used to look for unacceptable systematic prediction bias, which was not observed. They also show distribution of residuals among fully acceptable and partly acceptable APZ. A more formal and complete evaluation of model performance is provided next.

Data used for model development met all criteria for test data of the APZ method as indicated by answers of "yes" to questions 1 to 4 in the decision tree for dependent data in ValT (Table 1). This indicated that these data could be used with confidence to provide a complete, accurate, unbiased, and nonconfounded evaluation of model performance for data used in model development.

Overall pAPZ for dependent data ($n = 140$) was 0.97, whereas pAPZ ranged from 0.95 to 1 for time, from 0.91 to 1 for temperature, from 0.74 to 1 for combinations of time and temperature, and the maximum number of consecutive pAPZ <0.70 was zero (Table 1). Consequently, answers to questions 5 to 7 for model performance criteria of the APZ method in the decision tree for dependent data in ValT were "yes" (Table 1). This indicated that the model did not have any global or local prediction problems for dependent data. Thus, the model was validated for dependent data because it satisfied all criteria for test data and model performance of the APZ method in ValT.

3.3.2 | Independent data for interpolation

Data used to evaluate the model for interpolation met all criteria for test data of the APZ method as indicated by answers of "yes" to questions 1 to 6 in the decision tree for interpolation in ValT (Table 2). This indicated that these data could be used with confidence to provide a complete, accurate, unbiased, and nonconfounded evaluation of model performance for interpolation.

Journal of Food Processing and Preservation Institute of Food Science + Technology ifst WILEY

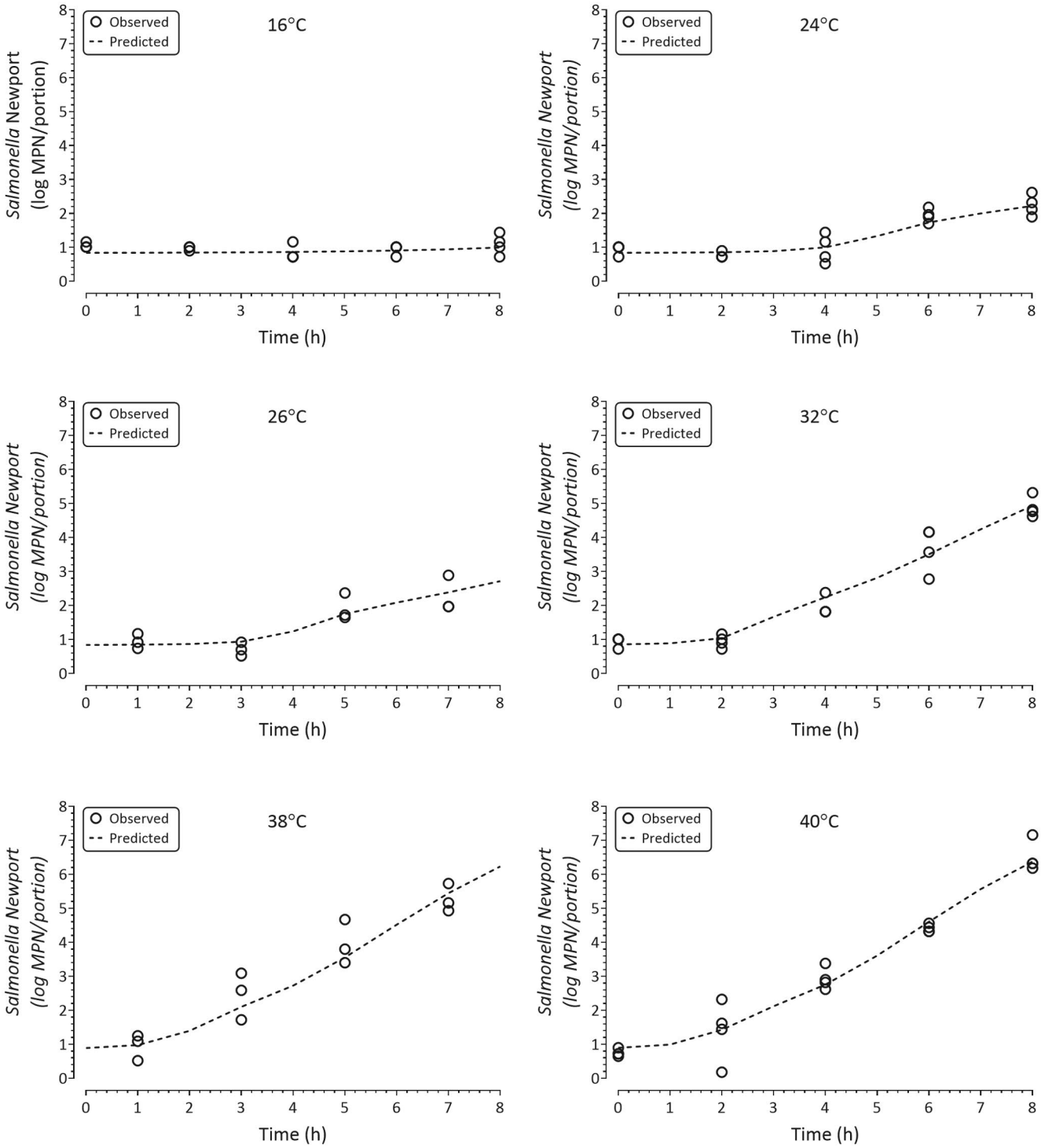

FIGURE 1 Observed (symbols) and predicted (lines) growth of a chicken isolate of *Salmonella* Newport on cucumber as a function of time and temperature for representative data used in model development (16, 24, 32, 40°C) and validation (26, 38°C)

Overall pAPZ for independent data for interpolation (n = 72) was 0.93, whereas pAPZ ranged from 0.9 to 1 for time, from 0.83 to 1 for temperature, from 0.67 to 1 for combinations of time and temperature, and the maximum number of consecutive pAPZ <0.70 was two (Table 2). Consequently, answers to questions 7 to 9 for model performance criteria of the APZ method in the decision tree for interpolation in ValT were "yes" (Table 2). This indicated that the model did not have any global or local prediction problems for the independent data for interpolation. Thus, the model was successfully validated for interpolation because it satisfied all test data and model performance criteria of the APZ method in ValT.

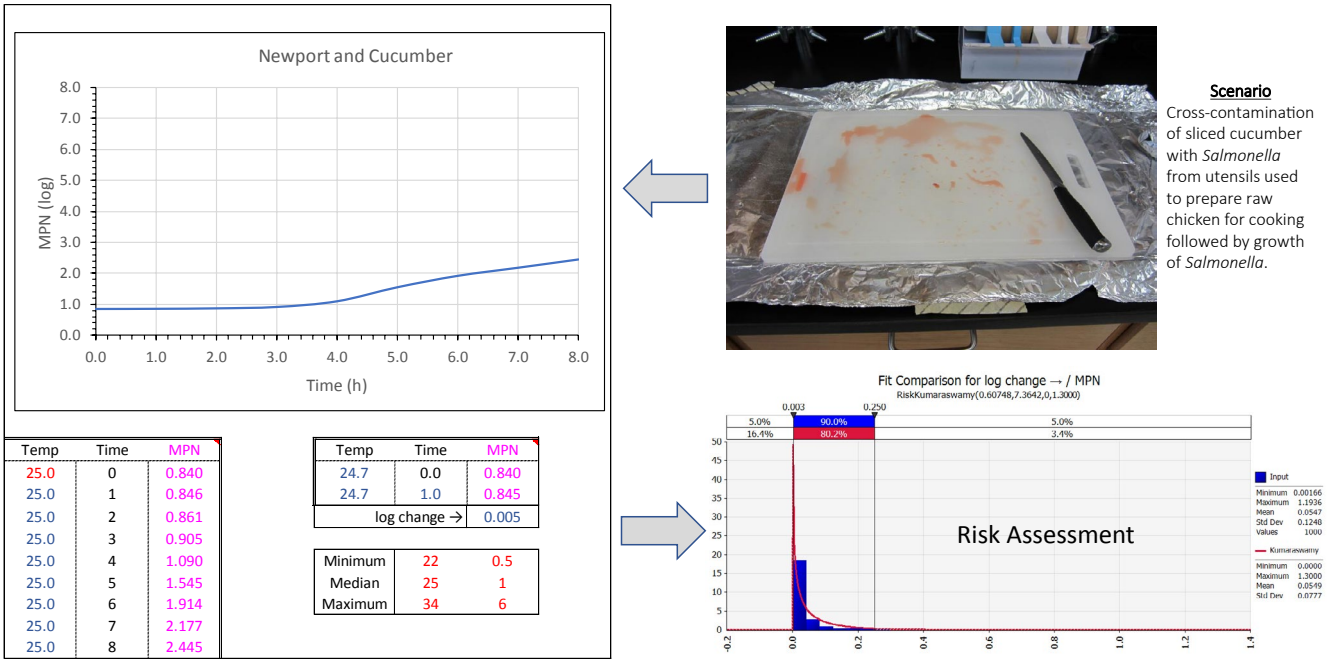

FIGURE 2 Neural network model for simulating and predicting growth of a chicken isolate of *Salmonella* Newport on cucumber after simulated cross-contamination from utensils used to prepare raw chicken for cooking. An output of the model is a probability distribution for growth (log increase) of *Salmonella* Newport on cucumber held for variable (pert: minimum, most likely, maximum) times (0.5, 1, 6 hr) and temperatures (22, 25, 34°C) observed during meal preparation and serving for use in risk assessment

4 | DISCUSSION

Results of this study indicated that after simulated cross-contamination from utensils used to prepare raw chicken for cooking, *Salmonella* Newport (0.85 log) grew on cucumber portions (0.2 g) with native microflora at times (0 to 8 hr) and temperatures (16 to 40°C) observed during meal preparation and serving (da Silva et al., 2020). These results agree with other studies that investigated growth of *Salmonella* on cucumber slices held at room temperatures. For example, Elexson et al. (2011) investigated growth of *Salmonella* Enteritidis (1 or 3 log) on cucumber slices (10 g) added to egg sandwiches and then held at refrigeration (4°C) or room temperature (not specified) for 0 to 6 hr. They observed time-dependent growth (0 to 4 log/10 g) at both inoculum sizes and temperatures during 6 hr of storage. Although *Salmonella* Enteritidis is often isolated from chicken meat and eggs (Borges et al., 2017; Jackson et al., 2013) and thus, is a good serotype for model development, data in the study of Elexson et al. (2011) were insufficient (only two temperatures) to develop and validate a predictive model for growth of *Salmonella* on cucumber for use in risk assessment.

In another study, Bardsley et al. (2019) investigated growth of a mixture of outbreak isolates of *Salmonella* serotypes Michigan, Enteritidis, Meunchen, Newport, and Saintpaul (3 log) on cucumber slices (10 g) incubated at 23°C for 0 to 8 hr. After 5 and 8 hr of storage at 23°C, 1.5 and 2.9 log of growth were observed, respectively. In comparison, in this study, *Salmonella* Newport growth on cucumber portions after 5 and 8 hr at 23°C was predicted to be 0.3 and 1.2 log, respectively. No firm conclusions can be made about why growth was predicted to be less in this study than in the study of Bardsley et al. (2019) because data on which these comparisons are based were collected with different methods (*Salmonella* serotypes, inoculum size, previous history, enumeration method). Nonetheless, both studies show that *Salmonella* can grow on cucumber stored for times and a temperature (i.e., 23°C) observed during meal preparation and serving (da Silva et al., 2020).

In another study, Ha et al. (2020) isolated three strains of *Salmonella* from cucumber and then inoculated them as a mixture (3.6 to 5.2 log) onto cucumber slices (25 g) by dipping. The inoculated cucumber slices were stored in pairs at 10°C (0 to 96 hr), 20°C (0 to 48 hr), 25°C (0 to 48 hr), and 30°C (0 to 48 hr) and then *Salmonella* growth over time of storage was determined by viable counts (CFU). The CFU data were fitted to a primary growth model (Baranyi) and then secondary models (polynomial) were developed that predicted lag time and growth rate (primary model parameters) as a function of temperature (10 to 30°C).

Growth of *Salmonella* on the cucumber slices was observed to increase as a function of time and temperature and at 8 hr of storage was predicted to be 0.4 log at 20°C, 1.5 log at 25°C, and 2.1 log at 30°C (Ha et al., 2020). In comparison, in this study, growth of *Salmonella* Newport on cucumber at 8 hr of storage was predicted to be 0.6 log at 20°C, 1.6 log at 25°C, and 3.3 log at 30°C. However, these comparisons are confounded by differences in data collection (*Salmonella* type, inoculum size, previous history, enumeration method) and modeling methods (regression vs. neural network) and thus, no firm conclusions can be made about why growth was predicted to be higher in this study than in the study of Ha et al. (2020).

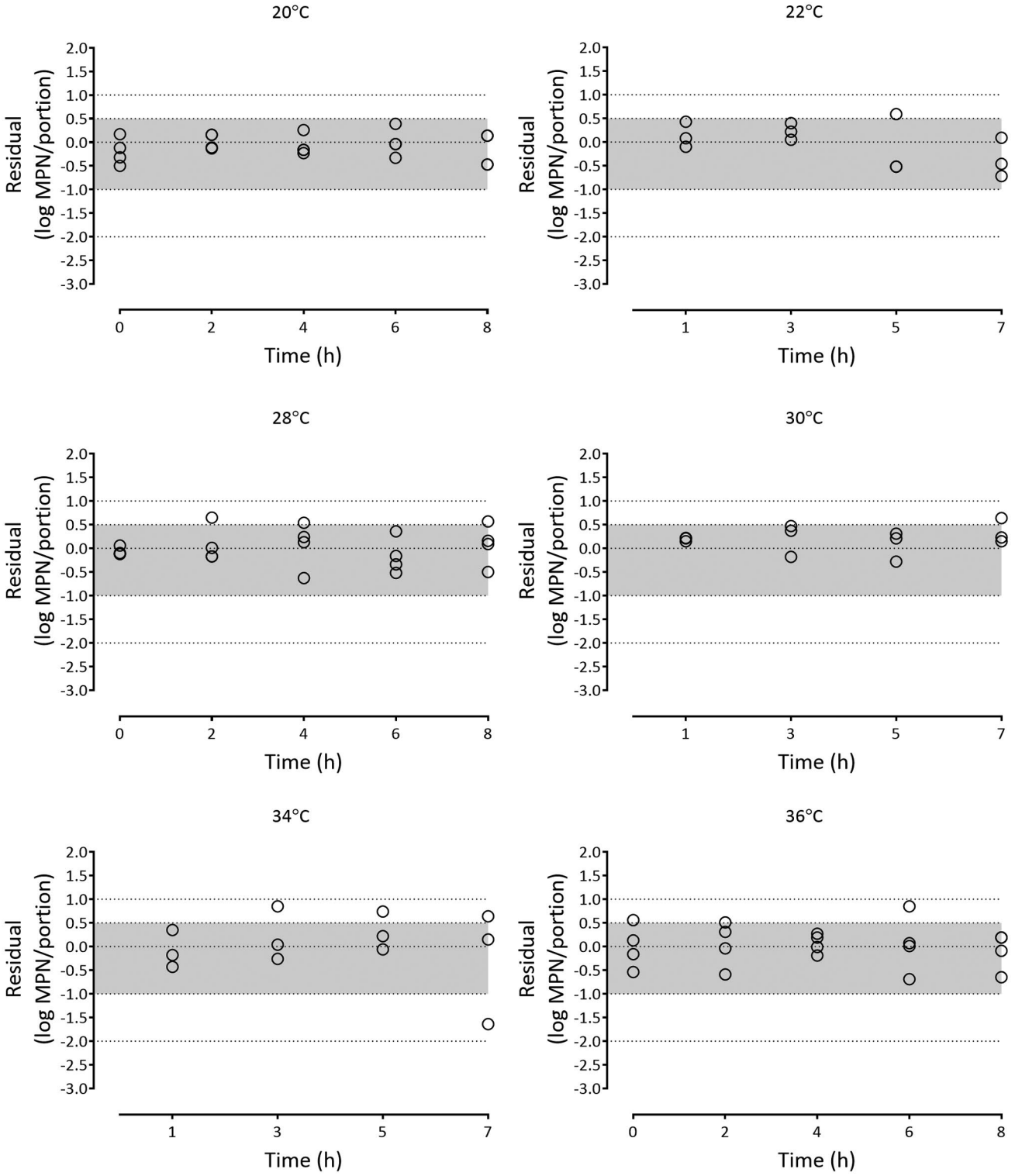

FIGURE 3 Representative residual (observed – predicted) plots for growth of a chicken isolate of *Salmonella* Newport on cucumber as a function of time and temperature for data used in model development (20, 28, 36°C) or validation (22, 30, 34°C). Residuals in the gray shaded area were fully acceptable, residuals in the non-shaded areas with dashed borders were partly acceptable, and residuals outside these areas were unacceptable

Nonetheless, both studies indicate that *Salmonella* can grow on cucumber at a range of times and temperatures encountered during meal preparation and serving (da Silva et al., 2020).

For data used in model development, Ha et al. (2020) reported coefficients of determination (R^2) of 0.996 for the secondary model for lag time and 0.971 for the secondary model for growth rate.

TABLE 1 Acceptable prediction zones analysis for data used in model development

Question	Answer	Decision tree for dependent data				
1	yes	Were the data used to develop the model?				
2	yes	Were the independent variables evenly spaced?				
3	yes	Was there a minimum of four prediction cases per combination of independent variables?				
4	yes	Did all combinations of independent variables have the same number of prediction cases?				
5	yes	Was the overall pAPZ ≥0.70?				
6	yes	Was pAPZ for all individual levels of independent variables ≥0.70?				
7	yes	Was a single pAPZ ≥0.70 for every three consecutive combinations of the independent variables?				
8	yes	Was the model validated for dependent data?				
pAPZ	Time					
Temp	0	2	4	6	8	Average
16	1.00	1.00	1.00	1.00	1.00	1.00
20	1.00	1.00	1.00	1.00	1.00	1.00
24	1.00	1.00	1.00	1.00	1.00	1.00
28	1.00	0.93	0.98	1.00	0.97	0.98
32	1.00	1.00	1.00	0.85	1.00	0.97
36	0.97	1.00	1.00	0.83	1.00	0.96
40	1.00	0.74	0.94	1.00	0.86	0.91
Average	1.00	0.95	0.99	0.95	0.97	0.97

Abbreviations: pAPZ, proportion of residuals in the acceptable prediction zones; Temp, temperature in °C; Time, time in hours.

TABLE 2 Acceptable prediction zones analysis for data used in model validation for interpolation

Question	Answer	Decision tree for interpolation				
1	yes	Was the model validated for dependent data?				
2	yes	Were the data independent?				
3	yes	Were the data collected using the same methods as dependent data?				
4	yes	Were the independent variables at values intermediate to those used in model development?				
5	yes	Was there a minimum of two prediction cases per combination of independent variables?				
6	yes	Did all combinations of independent variables have the same number of prediction cases?				
7	yes	Was the overall pAPZ ≥0.70?				
8	yes	Was pAPZ for all individual levels of independent variables ≥0.70?				
9	yes	Was a single pAPZ ≥0.70 for every three consecutive combinations of the independent variables?				
10	yes	Was the model validated for interpolation?				
pAPZ	Time					
Temp	1	3	5	7	Average	
18	1.00	1.00	1.00	1.00	1.00	
22	1.00	1.00	0.94	1.00	0.99	
26	1.00	1.00	0.92	1.00	0.98	
30	1.00	1.00	1.00	0.91	0.98	
34	1.00	0.77	0.84	0.69	0.83	
38	1.00	0.67	0.67	1.00	0.84	
Average	1.00	0.91	0.90	0.93	0.93	

Abbreviations: pAPZ, proportion of residuals in the acceptable prediction zones; Temp, temperature in °C; Time, time in hours.

Although the models had high goodness-of-fit ($R^2 > 0.97$), they were not evaluated for their ability to predict data for lag time and growth rate that were not used in model development. In other words, they were not validated for interpolation. In contrast, in this study, the neural network model for growth of *Salmonella* Newport on cucumber was validated for interpolation using the test data, model performance, and model validation criteria of the APZ method in ValT (Oscar, 2020).

In contrast to this study, Ha et al. (2020) evaluated their model for extrapolation to a fluctuating temperature condition. They reported a root mean squared error (RMSE) of 0.367 log and concluded that the model provided good predictions of *Salmonella* growth on cucumber. However, this conclusion was based on one scenario for fluctuating temperature and a limited amount of CFU data ($n = 14$). In contrast, in this study, model validation for interpolation against independent data was based on six scenarios and a large amount of MPN data ($n = 72$).

Although the current model was not evaluated for extrapolation to a fluctuating temperature condition, data used in model development and validation were collected under changing temperature conditions. In this study, cucumber portions were cold (4°C) when inoculated with *Salmonella* Newport. They were then incubated at higher temperatures (16 to 40°C) for 0 to 8 hr. This resulted in collection of data for *Salmonella* growth over time under changing temperature conditions as cucumber portions warmed from 4°C to the test temperatures. In fact, these storage trial scenarios were designed to simulate transition from the previous unit operation (i.e., refrigerated storage of cucumbers) and pathogen event (cross-contamination of utensils) in the food production chain to the unit operation (i.e., meal preparation and serving) and pathogen event (i.e., growth) being simulated for use in risk assessment.

Proper development and validation of models for foodborne pathogens is important because it provides model users with confidence that predictions are reliable (Ross, 1996; Ross et al., 2000). In addition, it helps modelers identify prediction problems that can be repaired before models are shared with end users. In this study, the neural network model for growth of *Salmonella* Newport on cucumber was carefully developed and validated using the test data, model performance, and model validation criteria of the APZ method in ValT (Oscar, 2020). In contrast, the model of Ha et al. (2020) for

growth of *Salmonella* on cucumber was developed and validated without using criteria for test data, model performance, and model validation. In addition, it was not evaluated for local prediction problems like the current model was.

Results of model validation in this study indicated that the neural network model for growth of *Salmonella* Newport on cucumber provided highly reliable predictions (pAPZ > 0.92) and thus, does not need to be repaired before distribution to end users. However, that does not mean that the current model cannot be improved. In fact, the current model can be improved by seeing how broadly it can be applied to other independent variables such as other inoculum sizes, other serotypes, other previous histories, and other food matrices.

More specifically, the current model can be improved using the test data, model performance, and model validation criteria for extrapolation in the APZ method of ValT (Oscar, 2020). These criteria require that the independent data for extrapolation to a new independent variable are collected using the same experimental design and methods as used to collect data for model development except for the new independent variable being evaluated. In addition to providing a proper evaluation for model extrapolation to the new independent variable, this approach makes it possible to expand the current model to include the new independent variable if it fails validation for extrapolation.

In a previous study (Oscar, 2018), a neural network model for growth of *Salmonella* Newport on Roma tomato portions was evaluated for extrapolation to ten other serotypes of *Salmonella* using the APZ method in ValT. Most (7 of 10) of them had similar growth kinetics as *Salmonella* Newport. However, three serotypes failed the validation for extrapolation. Thus, the neural network model for growth of *Salmonella* Newport on Roma tomato portions can be improved by expanding it to include these serotypes.

5 | CONCLUSIONS

A neural network model for growth of a chicken isolate of *Salmonella* Newport on cucumber portions incubated for 0 to 8 hr at 16 to 40°C was successfully developed and validated using the test data, model performance, and model validation criteria of the APZ method in ValT. Thus, the model can be used with confidence to fill an important data and modeling gap in risk assessments for *Salmonella* and chicken. Namely, lack of data and models for predicting growth of *Salmonella* on cucumber (RTE food) after cross-contamination from utensils used to prepare raw chicken for cooking.

ACKNOWLEDGMENTS

The author appreciates the outstanding technical assistance of Bethany Boleratz of the USDA, Agricultural Research Service (ARS), and undergraduate student researchers Florence Oyebanje, Nakya Robinson, and Anthony McCray of the University of Maryland Eastern Shore. This project was funded by CRIS-8072-42000-079-00D of the USDA, ARS. Mention of trade names or commercial products is solely for providing specific information and does not imply recommendation or endorsement by the USDA, which is an equal opportunity provider and employer.

CONFLICT OF INTEREST

The author declares that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

AUTHOR CONTRIBUTIONS

Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Software; Supervision; Validation; Visualization; Writing-original draft; Writing-review & editing: **Thomas Patrick Oscar**

DATA AVAILABILITY STATEMENT

Data will be archived on the Poultry FARM website (www.ars.usda.gov/nea/errc/PoultryFARM) and in ComBase (https://portal.errc.ars.usda.gov/ComBase.aspx).

ORCID

Thomas P. Oscar https://orcid.org/0000-0001-6253-1286

REFERENCES

Akil, L., & Ahmad, H. A. (2019). Quantitative risk assessment model of human salmonellosis resulting from consumption of broiler chicken. *Diseases*, *7*, 19. https://doi.org/10.3390/diseases7010019

Angelo, K. M., Chu, A., Anand, M., Nguyen, T. A., Bottichio, L., Wise, M., Williams, I., Seelman, S., Bell, R., Fatica, M., Lance, S., Baldwin, D., Shannon, K., Lee, H., Trees, E., Strain, E., & Gieraltowski, L. (2015). Outbreak of Salmonella Newport infections linked to cucumbers–United States, 2014. *MMWR Morbidity and Mortality Weekly Report*, *64*, 144–147.

Baranyi, J., Roberts, T. A., & McClure, P. (1993). A non-autonomous differential equation to model bacterial growth. *Food Microbiology*, *10*, 43–59. https://doi.org/10.1006/fmic.1993.1005

Bardsley, C. A., Truitt, L. N., Pfuntner, R. C., Danyluk, M. D., Rideout, S. L., & Strawn, L. K. (2019). Growth and survival of *Listeria monocytogenes* and *Salmonella* on whole and sliced cucumbers. *Journal of Food Protection*, *82*, 301–309. https://doi.org/10.4315/0362-028X.JFP-18-341

Borges, K. A., Furian, T. Q., de Souza, S. N., Tondo, E. C., Streck, A. F., Salle, C. T., de Souza Moraes, H. L., & do Nascimento, V. P. (2017). Spread of a major clone of *Salmonella enterica* serotype Enteritidis in poultry and in salmonellosis outbreaks in Southern Brazil. *Journal of Food Protection*, *80*, 158–163. https://doi.org/10.4315/0362-028X.JFP-16-299

Buchanan, R. L. (1993). Developing and distributing user-friendly application software. *Journal of Industrial Microbiology*, *12*, 251–255. https://doi.org/10.1007/BF01584198

Buchanan, R. L., & Phillips, J. G. (1990). Response surface model for predicting the effects of temperature, pH, sodium chloride content, sodium nitrite concentration and atmosphere on the growth of *Listeria monocytogenes*. *Journal of Food Protection*, *53*, 370–376. https://doi.org/10.4315/0362-028X-53.5.370

Buchanan, R. L., & Whiting, R. C. (1996). Risk assessment and predictive microbiology. *Journal of Food Protection*, Supplement, 31–36. https://doi.org/10.4315/0362-028X-59.13.31

da Silva, D. C., Lopes, S. M., de Aquino, N. S. M., Elias, S. O., Duda, H. A., & Tondo, E. C. (2020). Mathematical modelling and validation of *Salmonella enterica* growth in sushi exposed to different time-temperature scenarios found in real sushi establishments. *Food*

Research International, *136*, 109609. https://doi.org/10.1016/j.foodres.2020.109609

Delignette-Muller, M. L., Rosso, L., & Flandrois, J. P. (1995). Accuracy of microbial growth predictions with square root and polynomial models. *International Journal of Food Microbiology*, *27*, 139–146. https://doi.org/10.1016/0168-1605(94)00158-3

Elexson, N., Tuan Zainazor, T. C., Tunung, R., Ubong, A., & Son, R. (2011). Time course study on the growth of *Salmonella* Enteritidis on raw vegetables used in sandwiches. *International Food Research Journal*, *18*, 1505–1508.

Gibson, A. M., Bratchell, N., & Roberts, T. A. (1988). Predicting microbial growth: Growth responses of salmonellae in a laboratory medium as affected by pH, sodium chloride and storage temperature. *International Journal of Food Microbiology*, *6*, 155–178. https://doi.org/10.1016/0168-1605(88)90051-7

Gonzalez, S. C., Possas, A., Carrasco, E., Valero, A., Bolivar, A., Posada-Izquierdo, G. D., Garcia-Gimeno, R. M., Zurera, G., & Perez-Rodriguez, F. (2018). 'MicroHibro': A software tool for predictive microbiology and microbial risk assessment in foods. *International Journal of Food Microbiology*, *290*, 226–236. https://doi.org/10.1016/j.ijfoodmicro.2018.10.007

Ha, J., Park, E., Kim, J.-S., Lee, S., Kim, S., Lee, J., Choi, Y., Yoon, Y., Oh, H., Kim, Y., Lee, Y., Seo, Y., & Kang, J. (2020). Prevalence of *Salmonella* in cucumbers, antibiotic and acid resistances and description of the kinetic behavior with dynamic model during storage. *Journal of Food Safety*, *40*, e12760.

Hajmeer, M. N., Basheer, I. A., & Najjar, Y. M. (1997). Computational neural networks for predictive microbiology II. Application to microbial growth. *International Journal of Food Microbiology*, *34*, 51–66. https://doi.org/10.1016/S0168-1605(96)01169-5

Huang, L. (2014). IPMP 2013–a comprehensive data analysis tool for predictive microbiology. *International Journal of Food Microbiology*, *171*, 100–107. https://doi.org/10.1016/j.ijfoodmicro.2013.11.019

Jackson, B. R., Griffin, P. M., Cole, D., Walsh, K. A., & Chai, S. J. (2013). Outbreak-associated *Salmonella enterica* serotypes and food commodities, United States, 1998–2008. *Emerging and Infectious Disease*, *19*, 1239–1244. https://doi.org/10.3201/eid1908.121511

Jarvis, B., Wilrich, C., & Wilrich, P. T. (2010). Reconsideration of the derivation of Most Probable Numbers, their standard deviations, confidence bounds and rarity values. *Journal of Applied Microbiology*, *109*, 1660–1667. https://doi.org/10.1111/j.1365-2672.2010.04792.x

Jayeola, V., Jeong, S., Almenar, E., Marks, B. P., Vorst, K. L., Brown, J. W., & Ryser, E. T. (2019). Predicting the growth of *Listeria monocytogenes* and *Salmonella* Typhimurium in diced celery, onions, and tomatoes during simulated commercial transport, retail storage, and display. *Journal of Food Protection*, *82*, 287–300. https://doi.org/10.4315/0362-028X.JFP-18-277

Jia, Z., Peng, Y., Yan, X., Zhang, Z., Fang, T., & Li, C. (2020). One-step kinetic analysis of competitive growth of *Salmonella* spp. and background flora in ground chicken. *Food Control*, *117*, 107103.

Lambertini, E., Ruzante, J., Chew, R., Apodaca, V. L., & Kowalcyk, B. B. (2019). The public health impact of different microbiological criteria approaches for *Salmonella* in chicken parts. *Microbial Risk Analysis*, *12*, 44–59. https://doi.org/10.1016/j.mran.2019.06.002

Maffei, D. F., Sant'Ana, A. S., Franco, B. D. G. M., & Schaffner, D. W. (2017). Quantitative assessment of the impact of cross-contamination during the washing step of ready-to-eat leafy greens on the risk of illness caused by *Salmonella*. *Food Research International*, *92*, 106–112. https://doi.org/10.1016/j.foodres.2016.12.014

Martino, K. G., & Marks, B. P. (2007). Comparing uncertainty resulting from two-step and global regression procedures applied to microbial growth models. *Journal of Food Protection*, *70*, 2811–2818. https://doi.org/10.4315/0362-028X-70.12.2811

McClure, P. J., Baranyi, J., Boogard, E., Kelly, T. M., & Roberts, T. A. (1993). A predictive model for the combined effect of pH, sodium chloride and storage temperature on the growth of *Brochothrix thermosphacta*. *International Journal of Food Microbiology*, *19*, 161–178. https://doi.org/10.1016/0168-1605(93)90074-Q

McClure, P. J., Blackburn, C. W., Cole, M. B., Curtis, P. S., Jones, J. E., Legan, J. D., Ogden, I. D., Peck, M. W., Roberts, T. A., Sutherland, J. P., & Walker, S. J. (1994). Modelling the growth, survival and death of microorganisms in foods: The UK Food Micromodel approach. *International Journal of Food Microbiology*, *23*, 265–275. https://doi.org/10.1016/0168-1605(94)90156-2

Membre, J. M., & Boue, G. (2018). Quantitative microbiological risk assessment in food industry: Theory and practical application. *Food Research International*, *106*, 1132–1139. https://doi.org/10.1016/j.foodres.2017.11.025

Mohr, T. B., Juneja, V. K., Thippareddi, H. H., Schaffner, D. W., Bronstein, P. A., Silverman, M., & Cook, L. V. J. (2015). Assessing the performance of *Clostridium perfringens* cooling models for cooked, uncured meat and poultry products. *Journal of Food Protection*, *78*, 1512–1526. https://doi.org/10.4315/0362-028X.JFP-15-015

Najjar, Y. M., Basheer, I. A., & Hajmeer, M. N. (1997). Computational neural networks for predictive microbiology: I. *Methodology. International Journal of Food Microbiology*, *34*, 27–49. https://doi.org/10.1016/S0168-1605(96)01168-3

Notermans, S., Gallhoff, G., Zwietering, M. H., & Mead, G. C. (1995). The HACCP concept: Specification of criteria using quantitative risk assessment. *Food Microbiology*, *12*, 81–90. https://doi.org/10.1016/S0740-0020(95)80082-4

Oscar, T. P. (1998). The development of a risk assessment model for use in the poultry industry. *Journal of Food Safety*, *18*, 371–381. https://doi.org/10.1111/j.1745-4565.1998.tb00227.x

Oscar, T. P. (2004). A quantitative risk assessment model for *Salmonella* and whole chickens. *International Journal of Food Microbiology*, *93*, 231–247. https://doi.org/10.1016/j.ijfoodmicro.2003.12.002

Oscar, T. P. (2005). Validation of lag time and growth rate models for *Salmonella* Typhimurium: Acceptable prediction zone method. *Journal of Food Science*, *70*, M129–M137.

Oscar, T. P. (2009). General regression neural network and Monte Carlo simulation model for survival and growth of *Salmonella* on raw chicken skin as a function of serotype, temperature, and time for use in risk assessment. *Journal of Food Protection*, *72*, 2078–2087. https://doi.org/10.4315/0362-028X-72.10.2078

Oscar, T. P. (2011). Development and validation of a predictive microbiology model for survival and growth of *Salmonella* on chicken stored at 4 to 12°C. *Journal of Food Protection*, *74*, 279–284. https://doi.org/10.4315/0362-028X.JFP-10-314

Oscar, T. P. (2015). Neural network model for survival and growth of *Salmonella enterica* serotype 8,20:-:z_6 in ground chicken thigh meat during cold storage: Extrapolation to other serotypes. *Journal of Food Protection*, *78*, 1819–1827. https://doi.org/10.4315/0362-028X.JFP-15-093

Oscar, T. P. (2017). Risk of salmonellosis from chicken parts prepared from whole chickens sold in flow pack wrappers and subjected to temperature abuse. *Journal of Food Protection*, *80*, 1496–1505. https://doi.org/10.4315/0362-028X.JFP-17-097

Oscar, T. P. (2018). Development and validation of a neural network model for predicting growth of *Salmonella* Newport on diced Roma tomatoes during simulated salad preparation and serving: Extrapolation to other serotypes. *International Journal of Food Science and Technology*, *53*, 1789–1801. https://doi.org/10.1111/ijfs.13767

Oscar, T. P. (2020). Validation software tool (ValT) for predictive microbiology based on the acceptable prediction zones method. *International Journal of Food Science and Technology*, *55*, 2802–2812. https://doi.org/10.1111/ijfs.14534

Painter, J. A., Hoekstra, R. M., Ayers, T., Tauxe, R. V., Braden, C. R., Angulo, F. J., & Griffin, P. M. (2013). Attribution of foodborne illnesses, hospitalizations, and deaths to food commodities by using outbreak data,

United States, 1998–2008. *Emerging Infectious Diseases*, *19*, 407–415. https://doi.org/10.3201/eid1903.111866

Palumbo, S. A., Williams, A. C., Buchanan, R. L., & Phillips, J. G. (1991). Model for the aerobic growth of *Aeromonas hydrophila* K144. *Journal of Food Protection*, *54*, 429–435. https://doi.org/10.4315/0362-028X-54.6.429

Ross, T. (1996). Indices for performance evaluation of predictive models in food microbiology. *Journal of Applied Bacteriology*, *81*, 501–508. https://doi.org/10.1111/j.1365-2672.1996.tb03539.x

Ross, T., Dalgaard, P., & Tienungoon, S. (2000). Predictive modeling of the growth and survival of *Listeria* in fishery products. *International Journal of Food Microbiology*, *62*, 231–245.

Santillana Farakos, S. M., Pouillot, R., Anderson, N., Johnson, R., Son, I., & Van Doren, J. (2016). Modeling the survival kinetics of *Salmonella* in tree nuts for use in risk assessment. *International Journal of Food Microbiology*, *227*, 41–50. https://doi.org/10.1016/j.ijfoodmicro.2016.03.014

Scharff, R. L. (2020). Food attribution and economic cost estimates for meat- and poultry-related illnesses. *Journal of Food Protection*, *83*, 959–967. https://doi.org/10.4315/JFP-19-548

Scolforo, C., Bairros, J. V., Rezende, A. C. B., Silva, B. S., Alves, R. B. T., Costa, D. S., Andrade, N. J., Sant'Ana, A. S., & Pena, W. E. L. (2017). Modeling the fate of *Listeria monocytogenes* and *Salmonella enterica* in the pulp and on the outer rind of Canary melons (*Cucumis melo* (Indorus group)). *LWT-Food Science and Technology*, *77*, 290–297. https://doi.org/10.1016/j.lwt.2016.11.059

Smith-Simpson, S., & Schaffner, D. W. (2005). Development of a model to predict growth of *Clostridium perfringens* in cooked beef during cooling. *Journal of Food Protection*, *68*, 336–341. https://doi.org/10.4315/0362-028X-68.2.336

Walls, I., & Scott, V. N. (1996). Validation of predictive mathematical models describing the growth of *Escherichia coli* O157:H7 in raw ground beef. *Journal of Food Protection*, *59*, 1331–1335. https://doi.org/10.4315/0362-028X-59.12.1331

Whiting, R. C. (1995). Microbial modeling in foods. *Critical Reviews in Food Science and Nutrition*, *32*, 467–494. https://doi.org/10.1080/10408399509527711

Whiting, R. C., & Buchanan, R. L. (1997). Development of a quantitative risk assessment model for *Salmonella enteritidis* in pasteurized liquid eggs. *International Journal of Food Microbiology*, *36*, 111–125. https://doi.org/10.1016/S0168-1605(97)01262-2

How to cite this article: Oscar, T. P. (2021). Development and validation of a neural network model for growth of *Salmonella* Newport from chicken on cucumber for use in risk assessment. *Journal of Food Processing and Preservation*, 45, e15819. https://doi.org/10.1111/jfpp.15819