Food Sci. Biotechnol. Vol. 17, No. 6, pp. 1352 ~ 1356 (2008)

RESEARCH NOTE

Food Science and Biotechnology

Application of Bootstrap Method to Primary Model of Microbial Food Quality Change

Dong Sun Lee* and Jin Pyo Park[1]

Department of Food Science and Biotechnology, Kyungnam University, Masan, Gyeongnam 631-701, Korea
[1]Department of Computer Engineering, Kyungnam University, Masan, Gyeongnam 631-701, Korea

Abstract Bootstrap method, a computer-intensive statistical technique to estimate the distribution of a statistic was applied to deal with uncertainty and variability of the experimental data in stochastic prediction modeling of microbial growth on a chill-stored food. Three different bootstrapping methods for the curve-fitting to the microbial count data were compared in determining the parameters of Baranyi and Roberts growth model: nonlinear regression to static version function with resampling residuals onto all the experimental microbial count data; static version regression onto mean counts at sampling times; dynamic version fitting of differential equations onto the bootstrapped mean counts. All the methods outputted almost same mean values of the parameters with difference in their distribution. Parameter search according to the dynamic form of differential equations resulted in the largest distribution of the model parameters but produced the confidence interval of the predicted microbial count close to those of nonlinear regression of static equation.

Keywords: microbial growth model, Baranyi and Roberts model, aerobic bacteria, nonlinear regression, confidence limit

Introduction

Quantitative modeling and estimation of microbial growth are very useful for shelf-life determination of perishable foods labile to microbiological spoilage (1-3). As a first step in microbial modeling, the microbial growth under constant conditions is described by a mathematical function, which is called as primary model: typical examples of primary model commonly used include Gompertz and logistic functions. Recently, a primary model in differential equations based on mechanistic principle of microbial growth has been proposed by Baranyi and Roberts (4) and attained wide acceptance from many researchers (5). The model can also be solved analytically and given in an integrated nonlinear equation as explicit function of time. Its nature provides with parameters of mechanistic meaning directly related to microbial growth kinetics and it can handle with the dynamic environmental condition more easily than other models.

The primary models of microbial growth are characterized by the parameters, which are usually determined by nonlinear curve fitting procedure to the experimental data. The quality of parameter estimates depends on consistency in the measurement locations, scattering of experimental data, and model fitting method (6,7). In actual practice of modeling, quantity and quality of experiment data (experimental error, data distribution, positioning in time, etc) affect the determination of the appropriate parameters in the primary model. Stochastic models with distribution and confidence interval of the parameters are applied to deal with the problem of uncertainty and variation of experimental data (8-10). Randomization techniques such as Monte Carlo method and bootstrap can be adopted for the stochastic modeling (3,11-13).

Bootstrap method is a computer-intensive statistical technique to estimate the distribution of a statistic for stochastic predicted modeling. Schaffner (14) calculated variances and distributions of microbial growth rate by bootstrapping method. Lee *et al.* (15) recently presented a method to determine the parameters of Baranyi and Roberts model by directly bootstrapping microbial count data: the bootstrapped mean microbial counts were used in searching for the parameters by numerically solving the differential equations of the model. Bootstrap method in regression may apply the model-based resampling of the residuals or the direct resampling of the cases (16-19). The different bootstrap methods for primary microbial growth model need to be compared and examined in the determining the model parameters: the parameter values themselves, their distributions and confidence intervals, and the resultant confidence bands of the microbial growth predicted by the use of parameters should be validated.

This study therefore aims to compare different bootstrap methods in estimating the parameters of a primary model of microbial growth, particularly with emphasis on testing the validity of the method proposed by Lee *et al.* (15). The effect of different methods on the confidence interval of predicted microbial growth was also investigated.

Materials and Methods

Experimental microbial growth data of chill-stored pan-fried meat patties An experimental data set of aerobic bacterial growth on Korean style pan-fried meat patties at constant temperature of 5°C was available in our laboratory for this study. The measurement of colony forming units (CFU)/g of sample was conducted on plate count agar (PCA; Difco Laboratories, Detroit, MI, USA) incubated for 72 hr at 30°C.

*Corresponding author: Tel: +82-55-249-2687; Fax: +82-55-249-2995
E-mail: dongsun@kyungnam.ac.kr
Received March 4, 2008; Revised May 24, 2008;
Accepted May 24, 2008

Parameter estimation for the microbial growth model First, model-based resampling of the residuals was applied to the regression of the microbial growth data according to the static version of the Baranyi and Roberts microbial growth model:

$$\mathrm{LogN} = \log N_o + \frac{\mu_{max}}{\ln 10} \cdot A - \frac{1}{\ln 10} \cdot \ln\left(1 + \frac{e^{\mu_{max} A} - 1}{10^{(\log N_{max} - \log N_o)}}\right) \quad (1)$$

where A is defined by $A = t + \frac{1}{\mu_{max}} \cdot \ln\left(\frac{e^{-\mu_{max} t} + 1/(e^{t_{lag}\mu_{max}} - 1)}{1 + 1/(e^{t_{lag}\mu_{max}} - 1)}\right)$,

μ_{max} is the maximum specific growth rate (1/day), N is the microbial count in CFU/g at time t, N_o and N_{max} are the initial and maximum cell densities in CFU/g, respectively, and t_{lag} is lag time (day).

Equation 1 is a form reparameterizing q_o to t_{lag} from the original equation given by Baranyi and Roberts (4) according to Eq. 2.

$$q_o = \frac{1}{e^{t_{lag}\mu_{max}} - 1} \quad (2)$$

where q_o is the initial state for the physiological state of the cell population.

This formulation makes the nonlinear regression for parameter estimation more stable (7,11).

The bootstrapping procedure starts with curve fitting of the static version Eq. 1 to the experimental data based on the criteria of minimizing sum of error squares. The nonlinear regression subroutine Rnlin from IMSL Library (Visual Numerics, Houston, TX, USA) based on a modified Levenberg-Marquardt method was used to obtain a set of parameters (t_{lag}, log N_o, μ_{max}, and log N_{max}). Next a set of residuals from the best-fit microbial count estimate of regression curve was centered to the mean error and then used to establish samples for resampling. The centered residuals were then resampled randomly with replacement and added to the best-fit value to form a new set of microbial count data vs. time, which was subsequently used again for another nonlinear regression; the resampling and regression were repeated 1,000 times to obtain 1,000 sets of parameter estimates. The regression procedure was conducted separately for a set consisting of all the experimental data, and also for a mean microbial count data set.

As another way of bootstrapped curve-fitting, the dynamic version model of Baranyi and Roberts (4) in differential equations of logarithmic form (Eq. 3 and 4) was applied to estimating sets of the parameters (log q_o, log N_o, μ_{max}, and log N_{max}) following the method of Lee *et al.* (15).

$$\frac{d(\log q)}{dt} = \frac{\mu_{max}}{\ln 10} \quad (3)$$

$$\frac{d(\log N)}{dt} = \frac{\mu_{max}}{\ln 10}\left(\frac{1}{1 + 10^{-\log q}}\right)\left(1 - 10^{(\log N - \log N_{max})}\right) \quad (4)$$

where q is the physiological state of the cell population at time t. In solving the differential equations, q and N at initial time should be supplied as q_o and N_o, respectively (refer to Eq. 1 and 2).

Each experimental set consisting of mean microbial counts during storage was constructed by resampling individual plate counts and then supplied for the parameter estimation, which was done by minimizing the sum of squares of errors between the experimental data and those simulated by solving differential Eq. 3 and 4 for a guessed parameter set. The minimization algorithm of Box's complex method (20) written in Fortran code was used for the parameter search. Resampling and parameter estimation were repeated 1,000 times to obtain the distribution of log q_o, log N_o, μ_{max}, and log N_{max}.

The parameter, t_{lag} of Eq. 1 obtained from static version regression was converted to q_o following Eq. 2 to compare the different methods in the estimation.

Results and Discussion

Parameter distribution Figure 1 presents the microbial count data with mean values at experimental measurement times for the product stored at 5°C. The aerobic bacterial count change possessed typical pattern of microbial growth on food, consisting of lag, exponential growth, and stationary phases. Each sampling time measurement has variation in the microbial count and the growth pattern suffers from some fluctuation along with time, which is the commonly observed phenomenon in the destructive microbial quality measurement of the stored food products.

When the data in Fig. 1 were used for parameter estimations of static version Eq. 1 and dynamic version Eq. 3 and 4, three methods applied produced nearly same mean values of the parameters (Table 1 and Fig. 2). Poschet *et al.* (7) also reported that differences in data frequency and way

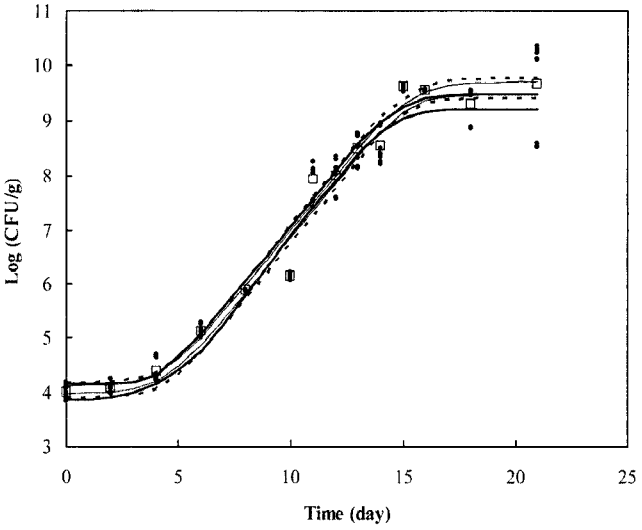

Fig. 1. Aerobic bacterial growth of the pan-fried meat patties during storage at 5°C. Nine individual plate counts were measured for each experimental sampling time except for initial measurement with 15 replicates. ●: Individual plate count; □: mean count; —: 95% confidence limit based on static version (Eq. 1) with all the experimental data; □: 95% confidence limit based on static version (Eq. 1) with average experimental data; —: 95% confidence limit based on dynamic version (Eq. 3 and 4).

Table 1. Microbial growth model parameters determined by different bootstrapping methods

Method	Parameters of Baranyi and Roberts model[1)] (Eq. 1-4)			
	Log q_o	Log N_o	μ_{max}	Log N_{max}
Nonlinear regression to static Eq. 1 with all the experimental data	-2.233±0.030	4.050±0.063	1.176±0.014	9.593±0.103
Nonlinear regression to static Eq. 1 with averaged experimental data	-2.233±0.100	4.054±0.129	1.200±0.033	9.600±0.154
Fitting with dynamic differential Eq. 3 and 4	-2.239±0.581	3.993±0.160	1.204±0.092	9.347±0.147

[1)]Value is given as mean±2 SD.

of Monte Carlo sampling did not affect the mean value of the model parameters of the microbial growth. Even with very slight difference among the methods, the mean parameter values were less different between two types of regressions to static Eq. 1 than between any of static equation regressions and dynamic differential equation fitting. This phenomenon is reasoned to have resulted from the assumptions adopted in the different versions: the static version equation regressions are based on the assumption that errors with homogeneous variance are resampled randomly, while fitting with dynamic equation directly bootstraps the experimental counts to obtain the mean count data and does not make any assumption on error variance (19).

Figure 2 shows the distributions of microbial model parameters determined from different resampling methods. Most of parameter distributions are symmetrical around the average values and overlaid with probability density functions of the normal distribution.

Nonlinear regression to static Eq. 1 with all the plate count data resulted in the narrowest ranges of parameters relatively closer to those from adopting, in the same equation, mean plate count on each storage time than to those of fitting to the dynamic Eq. 3 and 4 (Table 1 and Fig. 2). In nonlinear regression, generally both fitting to averaged data and fitting to individual replicates (same in number/measurement time) are known to give the same parameters but with possible difference in their standard errors or confidence intervals (21). Even though statistically it is suggested to treat every replicate as a separate data point in the regression analysis, averaging the dependent variable values of each measurement is widely used for

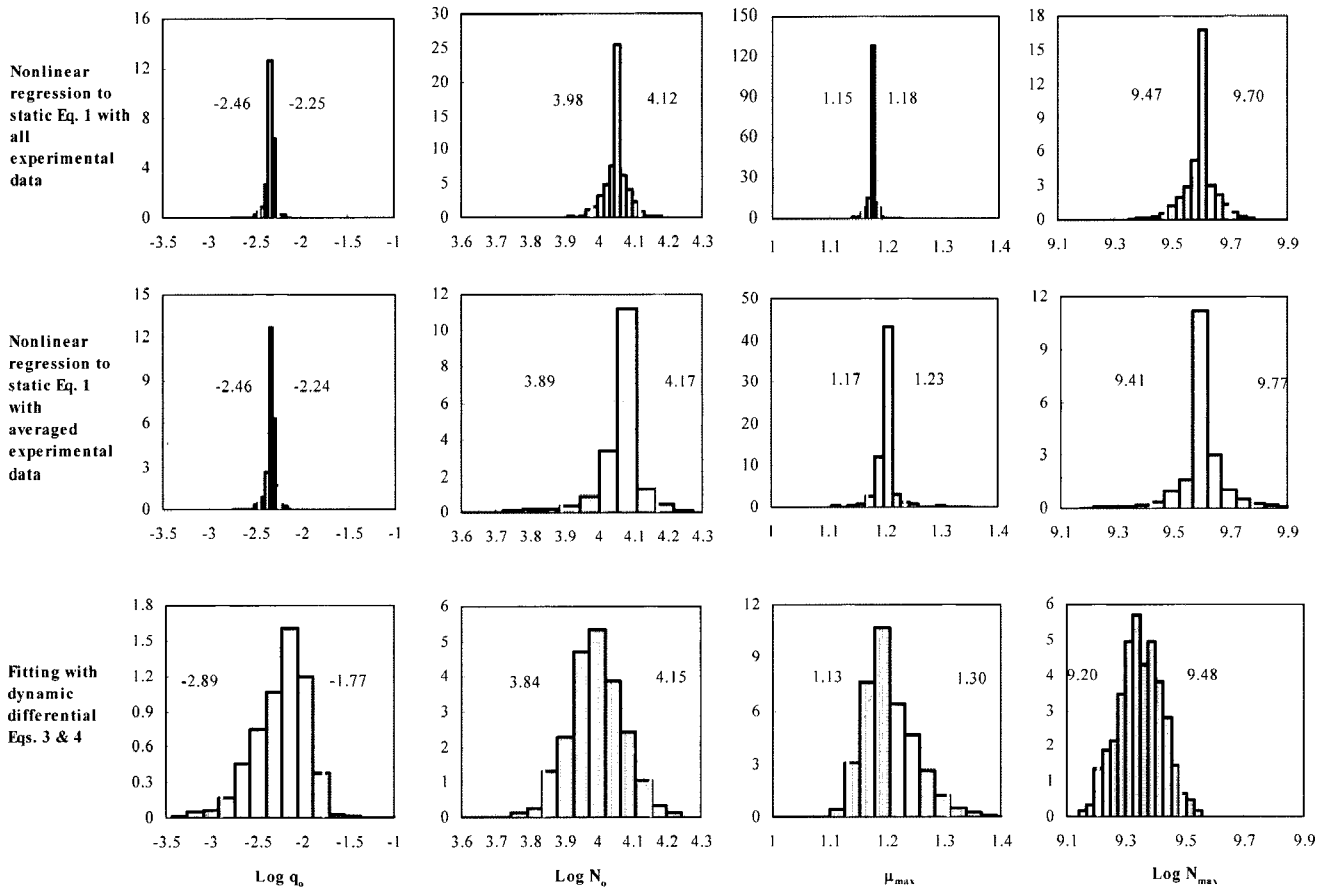

Fig. 2. Comparison of model parameter distributions obtained by different bootstrap curve-fitting techniques. The 2 vertical lines in the graphs depict 95% bootstrap confidence intervals. For the distribution function parameters, refer to Table 1.

practical kinetic analysis of food quality changes (22-25). Hietala *et al.* (25) applied the bootstrapped nonlinear regression to the averaged data for estimating the errors of the parameter estimates of yeast growth under cadmium toxicity. Treating individual data points as separate data points in this study had the effects of increasing data points for the regression to the microbial growth model (Eq. 1) and reducing the confidence limits. Poschet *et al.* (7) also observed that increased data points reduced the standard deviation of the parameter distribution of Baranyi and Roberts model; particularly intense replicate sampling during certain growth phases such as lag phase and transition zone could reduce the uncertainty of the model parameters.

Predicted microbial growth From 1,000 sets of microbial growth model parameters for each bootstrapping method, 1,000 growth curves can be obtained by substituting them into Eq. 1 or Eq. 3 and 4. The 95% confidence limit for any storage time was obtained from 2.5 and 97.5% percentiles of the estimated microbial counts, which were also overlaid onto Fig. 1. Three methods applied generally gave almost identical microbial growth patterns and similar count estimates with comparable but a little different confidence bands. The two kinds of static version regression showed the similar steady confidence bands with a little narrower one for the case using all the experimental data points. The dynamic version gave the time variant confidence interval generally close to that of the static version regression with average count on each sampling time, except for the later stationary period of growth showing some non-overlapped band which comes from small difference in the estimated log N_{max}. The bootstrapped nonlinear regression to static model equation using all the experimental replicate data reduced the uncertainty of the model parameters as discussed above, and their reduced distribution in turn contributed a little to narrowing down of the confidence band in the predicted microbial count values. It has been demonstrated by Labuza and Kamman (22) that the way of kinetic analysis in the same experimental data set of chemical quality change can affect appreciably the confidence limits of the dependent variable; consideration of practical limitation and usefulness is suggested to come along with statistical analysis.

Statistically the confidence bands in Fig. 1 mean the area having a 95% chance of containing the true best-fit curve (21), and are often understood as confidence intervals of the estimated means (26). Assuming homogenous variance and supplying large number data as individual points in regression of static version Eq. 1 would have reduced the range of true regression curve obtainable and provided relatively high assurance on the fitted model. Supply of mean counts by direct resampling from experimental data to the dynamic version Eq. 3 and 4 would have worked to give the dispersed band and weaker assurance due to assuming none on the sample or error distributions. In this context, the experimental and curve-fitting errors have been taken more appropriately into consideration by the dynamic version bootstrapping. The method is also noted to give confidence range similar to those of static version regressions as shown in Fig. 1. However, the dynamic version bootstrapping method required more delicate algorithm and long procedure of calculation, which is a drawback of this method.

This study reports the behavior of different bootstrapped curve-fitting approaches to describe the microbial quality change with time. The uncertainty and variability of the microbial data would have been handled in different degrees depending on the method used. Theoretical and practical implications of the different methods may probably be elucidated by further deeper statistical work and experimental confirmation. Some intuitional analysis and trial may be helpful for practical applications such as shelf-life estimation as discussed by Labuza and Kamman (22).

Conclusively, three different curve-fitting methods using bootstrapping to deal with uncertainty and variability of the experimental data in microbial growth were compared in the Baranyi and Roberts model parameters determined. All the methods gave almost same mean values of the parameters but with some difference in their confidence intervals. The dynamic version fitting of differential equations gave confidence intervals in the predicted microbial count comparable to those from nonlinear regression of static equation, suggesting its validity in estimating the parameters of the primary model.

Acknowledgments

This work was supported by the Korea Science and Engineering Foundation (Project # R01-2005-000-10235-0). The authors appreciate So Jung Kim for providing the experimental data set for this work.

References

1. McMeekin TA, Olley J, Ratkowsky DA, Ross T. Predictive microbiology: Towards the interface and beyond. Int. J. Food Microbiol. 73: 395-407 (2002)
2. Baranyi J, Pin C. Modeling microbiological safety. pp. 383-401. In: Food Process Modeling. Tijskens LMM, Hertog MLATM, Nicolai BM (eds). Woodhead Publishing, Cambridge, UK (2001)
3. Lee DS, Hwang K-J, Seo I, Park JP, Paik H-D. Estimation of shelf life distribution of seasoned soybean sprouts using the probability of *Bacillus cereus* contamination and growth. Food Sci. Biotechnol. 15: 773-777 (2006)
4. Baranyi J, Roberts TA. A dynamic approach to predicting bacterial growth in food. Int. J. Food Microbiol. 23: 277-294 (1994)
5. McKellar RC, Lu X. Primary models. pp. 21-62. In: Modeling Microbial Responses in Food. McKellar RC, Lu X (eds). CRC Press, Boca Raton, FL, USA (2004)
6. Baranyi J, Roberts TA. Mathematics of predicted food microbiology. Int. J. Food Microbiol. 26: 199-218 (1995)
7. Poschet F, Bernaerts K, Geeraerd AH, Scheerlinck N, Nicolai BM, Van Impe JF. Sensitivity analysis of microbial growth parameter distributions with respect to data quality and quantity by using Monte Carlo analysis. Math. Comput. Simulat. 65: 231-243 (2004)
8. Park JP, Kim HW, Lee DS, Paik H-D. Seasonal and market group variation in the microbiological quality of seasoned soybean sprouts. Food Sci. Biotechnol. 16: 325-328 (2007)
9. Park JP, Lee DS, Paik H-D. Distribution functions describing the microbiological contamination of seasoned soybean sprouts. Food Sci. Biotechnol. 17: 659-663 (2008)
10. Park JP, Lee DS. Analysis of temperature effects on microbial growth parameters and estimation of food shelf life with confidence band. J. Food Sci. Nutr. 13: 104-111 (2008)
11. Poschet F, Geeraerd AH, Scheerlinck N, Nicolai BM, Van Impe JF. Monte Carlo analysis as a tool to incorporate variation on experimental data in predictive microbiology. Food Microbiol. 20:

285-295 (2003)
12. Moreau Y, Couvert O, Thuault D. Estimation of the confidence band of bacterial growth simulation. The Sym'Previus approach. Acta Horticulturae 674: 415-420 (2005)
13. Manly BFJ. Randomization, Bootstrap, and Monte Carlo Methods in Biology. 3rd ed. Chapman & Hall/CRC, Boca Raton, FL, USA. pp. 169-202 (2007)
14. Schaffner DW. Application of a statistical bootstrapping technique to calculate growth rate variance for modeling psychrotrophic pathogen growth. Int. J. Food Microbiol. 24: 309-314 (1994)
15. Lee DS, Hwang K-J, An DS, Park JP, Lee HJ. Model on the microbial quality change of seasoned soybean sprouts for on-line shelf life prediction. Int. J. Food Microbiol. 18: 285-293 (2007)
16. Efron B, Tibshirani RJ. An Introduction to the Bootstrap. Chapman & Hall/CRC, Boca Raton, FL, USA. pp. 11-16 (1993)
17. Yafune A, Ishiguro M. Bootstrap approach for constructing confidence intervals for population pharmacokinetic parameters. I: A use of bootstrap standard error. Stat. Med. 18: 581-599 (1999)
18. Hass CN, Rose JB, Gerba CP. Quantitative Microbial Risk Assessment. John Wiley & Sons, New York, NY, USA. pp. 278-304 (1999)
19. Davison AC, Hinkley DV. Bootstrap Methods and Their Application. Cambridge University Press, Cambridge, UK. pp. 261-266 (1997)
20. Saguy I. Optimization methods and applications. pp. 264-320. In: Computer-aided Techniques in Food Technology. Saguy I (ed). Marcel Dekker, New York, NY, USA (1983)
21. Motulsky H, Christopoulos A. Fitting Models to Biological Data Using Linear and Nonlinear Regression. Oxford University Press, New York, NY, USA. pp. 29-137 (2004)
22. Labuza TP, Kamman JF. Reaction kinetics and accelerated tests simulations as a function of temperature. pp. 71-115. In: Computer-Aided Techniques in Food Technology. Saguy I (ed). Marcel Dekker, New York, NY, USA (1983)
23. Corbo MR, Del Nobile MA, Sinigaglia M. A novel approach for calculating shelf life of minimally processed vegetables. Int. J. Food Microbiol. 106: 69-73 (2006)
24. Koseki S, Isobe S. Prediction of pathogen growth on iceberg lettuce under real temperature history during distribution from farm to table. Int. J. Food Microbiol. 104: 239-248 (2005)
25. Hietala KA, Lynch ML, Allshouse JC, Johns CJ, Roane TM. A mathematical model of *Saccharomyces cerevisiae* growth in response to cadmium toxicity. J. Basic Microb. 46: 196-202 (2006)
26. Nauta MJ. Uncertainty and variability in predictive models of microorganisms in food. pp. 44-66. In: Modeling Microorganisms in Food. Brul S, Van Gerwen S, Zwietering M (eds). Woodhead Publishing, Cambridge, UK (2007)