

## USING SPREADSHEET SOFTWARE FOR PREDICTIVE MICROBIOLOGY APPLICATIONS

ROBERT L. BUCHANAN

*U.S. Department of Agriculture, ARS  
Eastern Regional Research Center  
600 E. Mermaid Lane  
Philadelphia, PA 19118*

Accepted for Publication November 12, 1990

### ABSTRACT

*Recently, multivariant models based on the use of the Gompertz function in combination with response surface analysis have been developed to predict the behavior of foodborne pathogens in response to food formulation and storage parameters, including temperature, pH, sodium chloride content, sodium nitrite concentration, and atmosphere. These models were adapted for easy use by developing a "user-friendly" application program, the Pathogen Modeling Program. This program is based on a commercially available spreadsheet program, Lotus 1-2-3™, and incorporates features such as calculation of predicted growth kinetics and time to achieve specified population densities. The current version of the software includes models for Salmonella, Shigella flexneri, Listeria monocytogenes, Staphylococcus aureus, and Aeromonas hydrophila. Microbiological modeling application software of this type appears to have great potential in relation to both developing food products with enhanced microbiological safety and teaching the multivariant nature of microbial growth in foods. The Pathogen Modeling Program is available on request.*

### INTRODUCTION

One of the concepts underlying modern food microbiology is that the growth of microorganisms in foods is controlled by the interaction of a finite number of variables associated with food composition and storage conditions (Jay 1986). Potentially, the number of variables affecting the growth kinetics of an organism is large; however, there are generally a relatively small number of factors such as storage temperature, pH,  $a_w$ , etc. that are the prime determinants of growth in food systems. This implies that quantitative knowledge of how these factors individually and interactively influence the growth of specific microorganisms

would permit one to forecast the bacteria's behavior in foods. This has been the basis for ongoing research in predictive food microbiology, including the development of mathematical and statistical models.

A number of mathematical models have been developed that predict the growth of bacteria in response to one or more of the variables associated with the formulation or storage of foods (Ratkowsky *et al.* 1983; Farber 1986; Baird-Parker and Kilsby 1987; Blankenship *et al.* 1988; Gibson *et al.* 1988; Griffiths and Phillips 1988; Chandler and McMeekin 1989; Davey 1989; Baker and Genigeorgis 1990; Buchanan and Phillips 1990). While a majority of these models provide reasonable estimates of the potential for microbial growth in food systems, there has been relatively little transfer of this technology to food microbiologists involved in nonresearch aspects of the field. The limited acceptance of predictive modeling techniques seems to be due in large part to a lack of application software that reduces to routine operations, the often complex mathematical manipulations associated with the use of the models.

Since its introduction to food science literature by Gibson *et al.* (1987), there has been increasing interest in the use of the Gompertz function (Table 1) to quantitatively describe the growth kinetics of foodborne microorganisms. Zwietering *et al.* (1990) concluded that of the models they tested the Gompertz curve was consistently the most effective means for modeling bacterial growth curves in foods, both in regard to statistical accuracy and ease of operation. The use of Gompertz parameters to describe individual growth curves has been coupled with response surface analysis techniques to successfully develop empirical models of the interactions of multiple variables on the growth of *Salmonella* (Gibson *et al.* 1988; Bratchell *et al.* 1989), *Listeria monocytogenes* (Buchanan and Phillips 1990), *Shigella flexneri* (Zaika *et al.* in preparation), *Staphylococcus aureus* (Smith *et al.* in preparation), and *Aeromonas hydrophila* (Palumbo *et al.* in press). Models developed from the combined use of the Gompertz function and response surface analysis are particularly well suited for the development of "user-friendly" applications programs. The current report describes briefly how commercially available spreadsheet software was used to develop such an application program, the "Pathogen Modeling Program" (PMP). This software is available upon request.

The current version (3.0) of PMP describes the impact of selected variables on *Salmonella* (temperature, pH, and NaCl), *L. monocytogenes* (temperature, pH, NaCl, NaNO<sub>2</sub>, and atmosphere), *S. flexneri*, *S. aureus*, and *A. hydrophila* (temperature, pH, NaCl, and NaNO<sub>2</sub>). The "heart" of the program is a series of quadratic or cubic expressions that were developed by response surface analysis of the natural logarithm (LN) transformation of the Gompertz B and M values. These values can be used in turn to derive growth kinetics values such as generation times and lag phase durations (Table 1). The decision to model

TABLE 1.  
EQUATIONS FOR GOMPERTZ FUNCTION AND DERIVED GROWTH KINETICS VALUES

GOMPERTZ'S FUNCTION:

$$L(t) = A + Ce^{-e^{-B(t-M)}}$$

where:

$L(t)$  = Log count of bacteria at time (in hours)  $t$   
[Log(cfu/ml)].

$A$  = Asymptotic log count of bacteria as time decreases indefinitely (i.e., initial level of bacteria)  
[Log(cfu/ml)].

$C$  = Asymptotic amount of growth that occurs as  $t$  increases indefinitely (i.e., number of log cycles of growth) [Log(cfu/ml)].

$M$  = Time at which the absolute growth rate is maximal [hr].

$B$  = Relative growth rate at  $M$ . [(Log(cfu/ml))/hr]

DERIVED GROWTH KINETICS EQUATIONS:

$$\text{Exponential Growth Rate (EGR)} = BC/e \quad [(\text{Log}(\text{cfu/ml}))/\text{hr}]$$

$$\text{Generation Time (GT)} = (\text{Log}(2))/BC \quad [\text{hr}]$$

$$\text{Lag Phase Duration (LPD)} = M - (1/B) \quad [\text{hr}]$$

$$\text{Maximum Population Density (MPD)} = A + C \quad [\text{Log}(\text{cfu/ml})]$$

the  $B$  and  $M$  terms instead of derived kinetics values was based in large part on the usefulness of the models in applications software, particularly in relation to generating predicted growth curves and calculating the time it would take the bacteria to achieve a specific population density. Models for the Gompertz  $A$  and  $C$  terms were not employed based on laboratory data that indicated that (1) over a broad range, inoculum size did not affect the growth kinetics of the microorganisms, and (2) except for combinations of extreme growth conditions, if a species initiated growth it typically achieved a characteristic maximum population density (MPD) between  $10^8$ – $10^{10}$  cfu/g (Gibson *et al.* 1988; Buchanan and Phillips 1990). The impact of these assumptions on the development of the application software will be discussed later.

The decision to employ a commercially available spreadsheet program for automating the use of the models was based on a number of factors such as the highly visual nature of spreadsheets, their array of useful mathematical functions, ready accessibility of good graphics, and simplicity of developing menu-driven option selection protocols. Using commercially available software proved much simpler and more effective than attempting to develop a program "from scratch," and provided a template that potential users were likely to already have a high degree of familiarity. The specific selection of Lotus 1-2-3 (Trademark of the Lotus Development Corp.) was based on several pragmatic reasons including the wide distribution of this commercially available program and the author's familiarity with its macro command language. However, there is no inherent reason why other commercially available spreadsheet software with similar capabilities could not be used to generate similar application programs.

### STRUCTURE OF PATHOGEN MODELING PROGRAM

The application program was designed around a "multiple file" architecture that employs a "master menu" file through which the files for individual pathogens are accessed (Fig. 1). A primary reason for using a multiple file structure

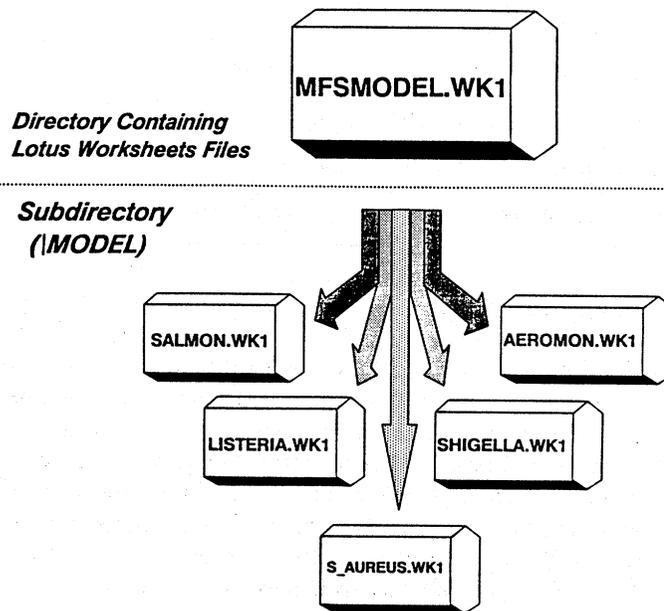


FIG. 1. MULTI-FILE STRUCTURE OF PATHOGEN MODELING PROGRAM 3.0

was that the speed of processing is enhanced greatly if the size of any single worksheet file is minimized, particularly when employing a 8088-based micro-computer without a numeric co-processor. Further, a multiple file approach greatly simplifies the inclusion of additional pathogens as suitable mathematical models become available. The primary function of the "master menu" file is to provide the user with a menu of the program options (i.e., available pathogen models) by automating the Lotus/*File-Retrieve* command sequence using a {MENUBRANCH} macro. (Details of the various macros used in the Pathogen Modeling Program are not included in the current report but are available upon request or can be obtained from a copy of the program.) The program was organized further by isolating the files for the various pathogens in a separate subdirectory, with only the master menu being present in the directory normally containing the user's worksheet files. This was done to help prevent the files containing the models from being accessed inadvertently during other uses of Lotus 1-2-3.

The basic structure of the individual files for the five pathogens is similar, with the file for *L. monocytogenes* being somewhat more complex due to the inclusion of models (Buchanan and Phillips 1990) for both aerobic and anaerobic growth. It will be used for further discussion. The structure of this worksheet file is diagrammatically depicted in Fig. 2. Upon retrieval, an autoexec macro (/O) is used to automatically initiate a sequence of subroutines wherein the user selects, via two data entry boxes (Fig. 2, B), the conditions to be evaluated by the model. The first data entry box asks for an assumed level of contamination and a level of concern. The second asks for assumed values for the storage and formulation variables considered by the models. Included in the data entry blocks are the limits of permissible values. As with any response surface model of this nature, prediction of the microorganisms' behavior in response to the interaction of the various variables should not be made beyond the limits for which experimental data were generated and analysed. For *L. monocytogenes*, the models' limits were temperature: 5°–37°C, pH: 4.5–7.5, NaCl: 0.5–5.0%, and NaNO<sub>2</sub>: 0–200 µg/mL. Upon answering the final query, the assumed values are substituted into the mathematical models (Fig. 2, C & D) which are solved to generate predicted B and M values for aerobic and anaerobic cultures. The user is then presented with the primary menu (Fig. 3), that was again developed using a {MENU-BRANCH} macro.

The potential selections associated with the primary menu fall into two categories, "results" and "options." The latter includes [MICROBE], [FOOD], [ATMOSPHERE], and [SPECIES/EXIT]. Each option has an underlying macro which returns the user to the appropriate data entry box, switches between the aerobic and anaerobic models, or re-initiates the master menu file. The "results" options employ a series of macros to move the user around the worksheet to where the results of the various calculations are displayed. The [GOMPERTZ]

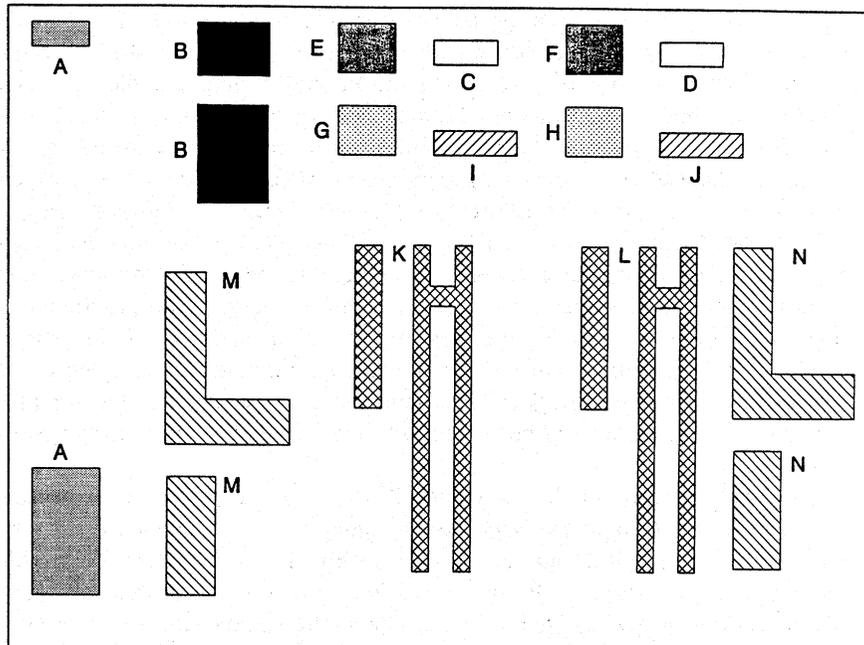


FIG. 2. DIAGRAMATIC REPRESENTATION OF THE COMPONENTS OF LISTERIA.WK1 WORKSHEET FILE

- A. Message Boxes
- B. Data Entry Boxes
- C. Mathematical Model—Aerobic
- D. Mathematical Model—Anaerobic
- E. Gompertz Values—Aerobic
- F. Gompertz Values—Anaerobic
- G. Growth Kinetics Values—Aerobic
- H. Growth Kinetics Values—Anaerobic
- I. Time to Level of Concern—Aerobic
- J. Time to Level of Concern—Anaerobic
- K. Calculations for Growth Curve—Aerobic
- L. Calculations for Growth Curve—Anaerobic
- M. Macros—Aerobic
- N. Macros—Anaerobic

option displays the four Gompertz parameters (Fig. 2, E and F), with B and M being derived from the model, A being the initial level assumed by the user, and C being calculated based on the A value and the assumed maximum population density (MPD). The [KINETICS] options displays the calculated or assumed growth kinetics values (Fig. 2, G and H) most often used by food

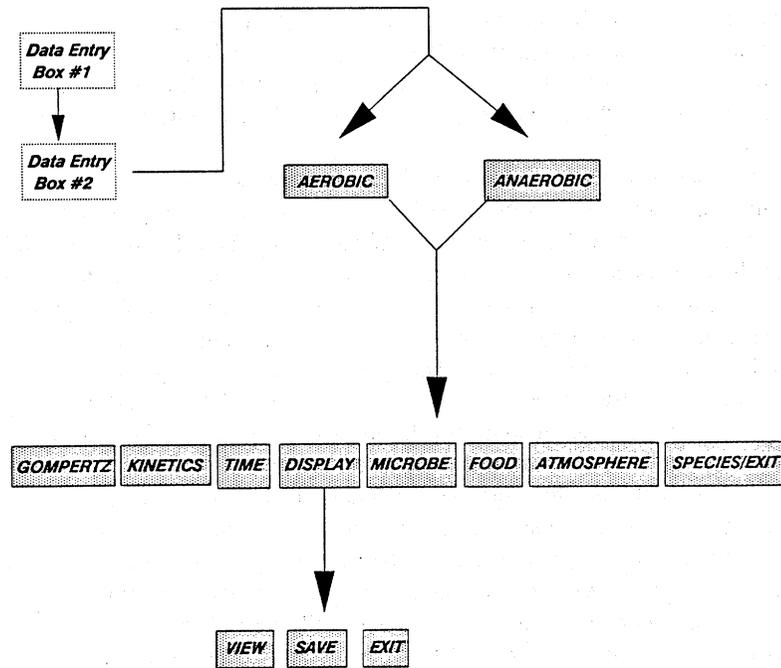


FIG. 3. SCHEMATIC OF MENU SELECTIONS FOR LISTERIA.WK1 FILE

microbiologists including exponential growth rate (EGR), generation time (GT), lag phase duration (LPD), and MPD. The [TIME] option automatically calculates the estimated time that it should take to go from the initial level of the pathogen to the specified level of concern, via a rearrangement of the Gompertz equation such that it is solved for  $t$ .

Selection of the primary menu's [DISPLAY] option initiates another series of macros that generates the predicted growth curve for the selected conditions. In the underlying routines (Fig. 2, K and L), the Gompertz equation is solved in 5 h time increments, which was selected on the basis of providing sufficient detail without unnecessarily increasing the file size. However, this time interval can be altered with a minor revision of the macro. The total time being considered for the graph can be specified by the user. This capability was developed by automating the use of the Lotus/*Data-Fill* command sequence, which is also the means by which the 5 h increment was established. A limit of 600 h was selected since few of the experiments underlying the models were run for a longer period. However, if extended data were available, the limit could be altered readily by making minor changes to the worksheet. Once the duration

of the growth curve has been entered by the user, the underlying macro solves the equation, draws the figure using the graphics capabilities of the spreadsheet software, and displays a secondary menu (Fig. 3) that further automates the selection options to view the graph, save the graph as a .PIC file, or exit to the primary menu.

A number of options within Lotus' macro command language were used to enhance the speed of the worksheet files. For example, the pathogen files were designed so that calculations are not performed automatically after the addition of each value entered in the data entry boxes. Instead, they are performed only after a command for recalculation is issued. When working with large spreadsheets containing a large number of formulas, this significantly increases the speed of operation, particularly when employing data entry boxes (Fig. 2, B). Only after all entries have been entered are the calculations performed by incorporating a {calc} command into the macro. Similarly, the speed of the applications were enhanced through the use of the macro commands, {windowoff} and {windowon}, that turn off and on, respectively, the sequential display of Lotus menus as a macro steps through series of worksheet functions.

Other Lotus options were used to enhance the appearance and security of the worksheet files. For example, all the macros were "hidden" to reduce unneeded clutter on the worksheet displays. Likewise, all areas of the worksheets except cells requiring input on the part of the user were protected so that one could not inadvertently alter the worksheet structure.

### MAKING CALCULATIONS INDEPENDENT OF INOCULUM SIZE

The two experimentally based assumptions introduced earlier had a significant impact on the methods that were employed for the using the Gompertz equation to determine EGR's/GT's, to calculate the time for the microorganism to achieve a specified population density, and to "draw" predicted growth curves. It required that a modification of the Gompertz equation be employed so that derived growth kinetics values were unaffected by the initial inoculum size. Without modification, altering the Gompertz A term would alter the subsequent calculation of EGR's/GT's. This was remedied using the first assumption that EGR's and LPD's were not affected by inoculum size along with the second assumption that MPD was independent of the variables (i.e., a constant). Mathematically, the two assumptions can be expressed as:

$$\text{EGR} = BC/e = k_1 \quad (1)$$

$$\text{LPD} = M - (1/B) = k_2 \quad (2)$$

$$\text{MPD} = A + C = k_3 \quad (3)$$

Calculation of EGR includes a C term which would be dependent on the initial inoculum level, A. However, since the C term was independent of the experimental variables, this parameter could be estimated using the grand mean of all experimental data for C (designated  $C_0$ ). This assumption is similar to the one that the MPD could be estimated using the grand mean of all experimentally determined MPD values. The  $C_0$  and MPD values for the various pathogens are summarized in Table 2. One can then substitute  $C_0$  into Eq. (1) thereby allowing calculation of EGR (and GT) independent of the inoculum level.

$$\text{EGR} = BC_0/e = k_1$$

The three constants can then be used to generate growth curves and growth calculations that have LPD's and EGR's/GT's that are independent of inoculum size by first rearranging Eq. (1)–(3):

$$B = (\text{EGR})e/C = k_1e/C \quad (4)$$

$$M = (\text{LPD}) + (1/B) = k_2 + (1/B) \quad (5)$$

$$C = \text{MPD} - A = k_3 - A \quad (6)$$

Eq. (4) is then substituted into Eq. (5):

$$M = k_2 + C/k_1e \quad (7)$$

The derived expressions for B (4), C (6), and M (7) terms were then substituted into Gompertz equation:

$$L(t) = A + (k_3 - A)e^{-e^{-[k_1e/(k_3 - A)][t - (k_2 + (k_3 - A)/k_1e)]}} \quad (8)$$

In this manner, the curve is defined on the basis of an assumed A value along with a series of constants derived from a combination the initial determination of the B and M value via the model along with terms derived directly from the experimental data. Eq. (8) was used for performing calculations associated with determination of the predicted growth curves and “times to a specific population,” allowing them to be generated for any initial population (i.e., A) without alterations in predicted LPD's or EGR/GT's.

TABLE 2.  
GRAND MEAN OF EXPERIMENTALLY DERIVED VALUES FOR GOMPERTZ A AND C  
TERMS AND MPD (MAXIMUM POPULATION DENSITY)

	A			C			MPD		
	MEAN	VAR	SEM	MEAN	VAR	SEM	MEAN	VAR	SEM
<i>Salmonella</i>	3.25	NA	NA	5.98	NA	NA	9.23	NA	NA
<i>Listeria monocytogenes</i>									
<i>Aerobic</i>	3.73	0.26	0.01	5.77	1.24	0.07	9.50	1.01	0.05
<i>Anaerobic</i>	3.69	0.08	0.01	5.65	0.58	0.04	9.34	0.49	0.03
<i>Shigella flexneri</i>	2.93	0.26	0.01	6.20	0.95	0.05	9.13	0.80	0.04
<i>Aeromonas hydrophila</i>	3.31	0.28	0.02	6.57	1.28	0.11	9.88	1.22	0.11
<i>Staphylococcus aureus</i>	3.02	0.83	0.07	5.57	1.33	0.11	8.58	0.96	0.08

All values in units of Log(cfu/ml).

VAR = Variance

SEM = Standard Error of Mean

NA = Not available

## POTENTIAL ENHANCEMENTS

The initial development of the PMP concentrated on establishing a general format that provided the user a tool that required minimal training, and that could be adapted readily for use with the mathematical models for various pathogens. The program was written using version 2.01 of Lotus 1-2-3, which is the release that was (and likely remains) available to most individuals. However, this restricted the program to considering the pathogens on an individual basis only, since each of the microorganisms had to be segregated into independent worksheet files to manage the size (and thus the speed) of worksheets. The newer releases of Lotus 1-2-3 (and similar programs) include a number of features that are being explored in relation to enhancing the effectiveness of the PMP in regard to assisting the user. Many of the improvements being considered are based on suggestions received from individuals using the software. For example, the ability to link multiple worksheets appears to have potential for greatly enhancing the software through the development of consolidated "results files" wherein the results from the individual files for the various microorganisms could be imported and displayed. This option should allow development of capabilities, such as a consolidated graph, that simultaneously displays the predicted growth curves for all the pathogens considered by PMP. These and other improvements are actively being considered for inclusion in future PMP releases.

## REFERENCES

- BAIRD-PARKER, A. C. and KILSBY, D. C. 1987. Principles of predictive food microbiology. J. Appl. Bacteriol. Symp. Supplement. 43S-49S.
- BAKER, D. A., and GENIGEORGIS, C. 1990. Predicting the safe storage of fresh fish under modified atmospheres with respect to *Clostridium botulinum* toxigenesis by modeling length of the lag phase of growth. J. Food Protection 53, 131-140.
- BLANKENSHIP, L. C., CRAVEN, S. E., LEFFLER, R. G. and CUSTER, C. 1988. Growth of *Clostridium perfringens* in cooked chili during cooling. Appl. Environ. Microbiol. 54, 1104-1108.
- BRATCHELL, N., GIBSON, A. M., TRUMAN, M., KELLY, T. M. and ROBERTS, T. A. Predicting microbial growth: The consequences of quantity of data. Internat. J. Food Microbiol. 8, 47-58.
- BUCHANAN, R. L. and 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*. J. Food Protection 53, 370-376, 381.

- CHANDLER, R. E. and MCMEEKIN, T. A. 1989. Modeling the growth response of *Staphylococcus xylosum* to changes in temperature and glycerol concentration/water activity. *J. Appl. Bacteriol.* 66, 543–548.
- DAVEY, K. R. 1989. A predictive model for combined temperature and water activity on microbial growth during the growth phase. *J. Appl. Bacteriol.* 67, 483–488.
- FARBER, J. M. 1986. Predictive modeling of food deterioration and safety. In *Foodborne Microorganisms and Their Toxins: Developing Methodologies*. (M. D. Pierson and N. J. Stern, eds.) pp. 57–90. Marcell Dekker, New York.
- GIBSON, A. M., BRATCHELL, N. and ROBERTS, T. A. 1987. The effect of sodium chloride and temperature on the rate and extent of growth of *Clostridium botulinum* type A in pasteurized pork slurry. *J. Appl. Bacteriol.* 62, 479–490.
- GIBSON, A. M., BRATCHELL, N. and ROBERTS, T. A. 1988. Predicting microbial growth: Growth responses of salmonellae in a laboratory medium as affected by pH, sodium chloride, and storage temperature. *Int. J. Food Microbiol.* 6, 155–178.
- GRIFFRITH, M. W. and PHILLIPS, J. D. 1988. Prediction of the shelf-life of pasteurized milk at different storage temperatures. *J. Appl. Bacteriol.* 65, 269–278.
- JAY, J. M. 1986. *Modern Food Microbiology*, 3rd ed. pp. 33–60. Van Nostrand Reinhold Co. New York.
- PALUMBO, S. A., WILLIAMS, A. C., BUCHANAN, R. L. and PHILLIPS, J. G. 1991. Model for the aerobic growth of *Aeromonas hydrophila* K144. *J. Food Protection*. (In Press)
- RATKOWSKY, D. A., LOWRY, R. K., MCMEEKIN, T. A., STOKES, A. N. and CHANDLER, R. E. 1983. Model for bacterial culture growth rate throughout the entire biokinetic temperature range. *J. Bacteriol.* 154, 1222–1226.
- ZWIETERING, M. H., JONGENBURGER, I. ROMBOUTS, F. M. and VAN't RIET, K. 1990. Modeling of the bacterial growth curve. *Appl. Environ. Microbiol.* 56, 1875–1881.