

Kansas State University Libraries

New Prairie Press

Conference on Applied Statistics in Agriculture

1994 - 6th Annual Conference Proceedings

NONLINEAR REGRESSION FUNCTIONS FOR FORAGE NUTRIENT DISAPPEARANCE FROM BAGS INCUBATED IN THE RUMEN

W. J. E. Potts

B. P. Glenn

J. B. Reeves

R. A. Erdman

See next page for additional authors

Follow this and additional works at: <https://newprairiepress.org/agstatconference>



Part of the [Agriculture Commons](#), and the [Applied Statistics Commons](#)



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](#).

Recommended Citation

Potts, W. J. E.; Glenn, B. P.; Reeves, J. B.; and Erdman, R. A. (1994). "NONLINEAR REGRESSION FUNCTIONS FOR FORAGE NUTRIENT DISAPPEARANCE FROM BAGS INCUBATED IN THE RUMEN," *Conference on Applied Statistics in Agriculture*. <https://doi.org/10.4148/2475-7772.1360>

This is brought to you for free and open access by the Conferences at New Prairie Press. It has been accepted for inclusion in Conference on Applied Statistics in Agriculture by an authorized administrator of New Prairie Press. For more information, please contact cads@k-state.edu.

Author Information

W. J. E. Potts, B. P. Glenn, J. B. Reeves, and R. A. Erdman

NONLINEAR REGRESSION FUNCTIONS FOR FORAGE NUTRIENT DISAPPEARANCE FROM BAGS INCUBATED IN THE RUMEN

W. J. E. Potts, USDA/ARS and University of MD, Dept. of Animal Sciences
B. P. Glenn, USDA/ARS
J. B. Reeves III, USDA/ARS
R. A. Erdman, University of MD, Dept. of Animal Sciences

ABSTRACT

Seven nonlinear regression functions are compared for fitting rumen in situ disappearance data. The standard function is based on a simple one-compartment model. In addition, we consider a time lag modification, a two-compartment model, and functions based on underlying probability models for degradation time. The empirical suitability of the seven regression functions are assessed using two in situ experiments involving forages fed to dairy cows. A function based on the loglogistic distribution is shown to have empirical and theoretical advantages.

Key words: Compartment model, hazard function, loglogistic distribution, nonlinear regression, sigmoidal curve.

1 INTRODUCTION

Rumen in situ experiments are widely used to evaluate feed quality (Ørskov and McDonald 1979, Nocek 1988). Porous, artificial-fiber bags, containing feed, are incubated in the rumen of a fistulated animal for a range of times. The bags are deposited successively and withdrawn as a group so that each has incubated for a different time. After withdrawal, the percent disappearance is determined for a particular component of the feed, such as crude protein, neutral detergent fiber, or total dry matter. Thus, the experiment generates a time course of percent disappearance values (Figure 1).

Rumen in situ experiments usually consists of several animal \times feed combinations, with a disappearance time course measured on each. The standard strategy for analyzing this repeated measures data is univariate analysis of variance on certain nutritionally important features of the time course, such as the limiting percent disappearance (total degradability). A nonlinear curve is fit to each individual time course, and the nutritionally important features are estimated from the parameters of the regression function.

The focus of this paper is the different regression functions used to fit each disappearance time course. We compared seven regression functions:

$$\text{G1} \quad D(t) = a + b(1 - \exp(-ct)) \quad (1)$$

$$\text{G1-lag} \quad D(t) = \begin{cases} a & t \leq d \\ a + b(1 - \exp(-c(t-d))) & t > d \end{cases} \quad (2)$$

$$\text{G1G1} \quad D(t) = a + b \left(1 - \frac{c \exp(-dt) - d \exp(-ct)}{c-d} \right) \quad (3)$$

$$\text{G2} \quad D(t) = a + b(1 - \exp(-ct)(1 + ct)) \quad (4)$$

$$\text{G3} \quad D(t) = a + b \left(1 - \exp(-ct) \left(1 + ct + \frac{(ct)^2}{2} \right) \right) \quad (5)$$

$$\text{Weibull} \quad D(t) = a + b(1 - \exp(-(ct)^d)) \quad (6)$$

$$\text{Loglogistic} \quad D(t) = a + b \left(1 - \frac{1}{1 + (ct)^d} \right) \quad (7)$$

The G1 function (1) was originally proposed by Ørskov and McDonald (1979) for in situ crude protein disappearance. The so-called *Gm* family of models, of which the G1 model is a special case, was developed for digesta passage in ruminants (Matis, Wehrly, and Ellis 1989; Pond et al 1988). Ørskov and McDonald (1979) recognized that other functions might be required to describe the degradation time course, and McDonald (1981) incorporated a time lag, resulting in the G1-lag function (2). The G1G1 function (3) is a simple extension of the *Gm* family. It was proposed for in situ residue (or equivalently, disappearance) by Van Milgen, Murphy, and Berger (1991). The Weibull function and loglogistic function (a.k.a. the Morgan-Mercer-Flodin curve) have not previously been applied to rumen in situ data (Seber and Wild 1989).

2 GRAPHS OF THE REGRESSION FUNCTIONS

The graph of each of the seven regression functions, (1)-(7), increases from an intercept at $(0, a)$ to an upper horizontal asymptote at $a+b$ (Figure 1). The parameter a is interpreted as the instantaneously degradable percentage (soluble or filterable), b is the slowly degraded percentage, and $100 - (a+b)$ is the undegradable percentage.

The chief difference in the shape of the seven curves is their ability to accommodate sigmoidal trends. The G1 curve is not sigmoidal; it increases at a constantly decreasing rate, with a maximum rate of increase at time zero.

The G1-lag curve has a flat phase between 0 and d followed by an increasing phase. The transition between the two phases is not smooth, that is, the derivative of $D(t)$ is not continuous at d .

The G1G1, G2, and G3 curves are sigmoidal, but somewhat restricted in shape. The proportion of the total increase which occurs prior to the inflection point, $(D(\text{inflection point}) - a)/b$, must be less than .264 for the G1G1 curve. For G2 and G3, the proportion exactly equals .264 and .323, respectively.

The Weibull and loglogistic curves have the flexibility to take either strictly concave-down or sigmoidal shapes depending on the parameter d . When $d \leq 1$ they are both concave down. When $d > 1$ they are both sigmoidal. The proportion of the total increase which occurs prior to the inflection point can take on a wide range of values. The proportion for the Weibull curve is $1 - \exp(1/d - 1)$ which must be less than $1 - 1/e \approx .632$. For the loglogistic curve, the proportion is $1/2(1 - 1/d)$ which must be less than .5.

3 THEORETICAL DERIVATION OF THE REGRESSION FUNCTIONS

PROBABILITY MODELS FOR DEGRADATION TIME

The expected value of the slowly degradable percentage that has degraded by time t can be derived by assuming that the feed is composed of discrete units which degrade independently. Let the unobserved random variable, T , be the time at which that an independent, discrete unit degrades. Then the expected percent degraded by t is

$$D(t) = a + b \cdot F(t) \quad (8)$$

where a is the expected value of the instantaneously degradable component, and $F(t) = \text{Prob}(T \leq t)$ is the cumulative distribution function of T . The regression functions, (1)-(7), correspond to different probability distributions for T .

The G_m family is derived from (8) by assuming that T has an Erlang distribution with scale parameter $1/c$ and shape parameter m . The Erlang distribution is a special case of the gamma distribution where the shape parameter is a positive integer. The G1 function (1) follows from assuming that T has an exponential distribution which is a special case of the Erlang (gamma) distribution where the shape parameter equals 1. Under the Erlang model, the mean degradation time is m/c .

The Weibull function (6) is derived from (8) by assuming that T has a Weibull distribution with scale parameter c and shape parameter d (Cox and Oakes 1984). The exponential distribution is also a special case of the Weibull distribution where $d = 1$. Under the Weibull model, the median degradation time is $(\ln(2))^{1/d}/c$ (Figure 1).

The loglogistic function (7) is derived from (8) by assuming that T has a loglogistic distribution (Cox and Oakes 1984), that is, $\ln(T)$ has a logistic distribution with location parameter $-\ln(c)$ and scale parameter $1/d$. Under the loglogistic model, the median degradation time is $1/c$ (Figure 1).

COMPARTMENT MODELS

The G1-lag (2) and G1G1 (3) functions do not correspond to familiar distributions. The G1 function (1) is usually derived by modeling degradation as a one-compartment system. Let $X(t)$ be the percentage of degradable material remaining in the compartment at time t . Let $D(t)$ be the percentage of degraded material that has degraded (in the system exterior). The compartment model is represented by the system of first-order, linear differential equations

$$\begin{aligned} \frac{dX(t)}{dt} &= -cX(t) & X(0) &= b \\ \frac{dD(t)}{dt} &= cX(t) & D(0) &= a \end{aligned} \tag{9}$$

The parameter c is the transfer coefficient. The parameters b and a are the initial percentages in the system and system exterior. Solving (9) for $D(t)$ gives the G1 function. The G1-lag model is a modification of (9) to incorporate a time lag, d

$$\begin{aligned} \frac{dX(t)}{dt} &= \begin{cases} 0 & \text{if } t \leq d \\ -cX(t) & \text{otherwise} \end{cases} & X(d) &= b \\ \frac{dD(t)}{dt} &= \begin{cases} 0 & \text{if } t \leq d \\ cX(t) & \text{otherwise} \end{cases} & D(d) &= a \end{aligned}$$

The G1G1 model (Van Milgen, Murphy, and Berger 1991) has two compartments: a lag compartment and a digestion compartment. Initially, all the slowly degradable material is in the lag compartment. It transfers into the digestion compartment and then out of the system.

$$\begin{aligned} \frac{dL(t)}{dt} &= -dL(t) & L(0) &= b \\ \frac{dX(t)}{dt} &= dL(t) - cX(t) & X(0) &= 0 \\ \frac{dD(t)}{dt} &= cX(t) & D(0) &= a \end{aligned}$$

where $L(t)$ is the percentage remaining in the lag compartment, and $X(t)$ is the percentage remaining in the digestion compartment. The G2 model is a special case of the G1G1 model where $c = d$. Similarly, The G3 model can be represented as a three-compartment system with equal transfer coefficients.

THE HAZARD FUNCTION

The hazard function provides a link between the probability models and the compartment systems. The hazard function

$$h(t) = \lim_{\Delta \rightarrow 0} \frac{P(t \leq T < t + \Delta | T \geq t)}{\Delta}$$

is the probability a discrete unit degrades at time t , provided that it has survived as long as t (Cox and Oakes 1984). It measures the instantaneous proneness of an independent, discrete unit of feed to degrade.

All the regression functions, including the Weibull and loglogistic, can be derived from a one-compartment system where the transfer coefficient is a function of time. In particular, when the transfer coefficient equals the hazard function of the corresponding probability distribution

$$\begin{aligned} \frac{dX(t)}{dt} &= -h(t) X(t) & X(0) &= b \\ \frac{dD(t)}{dt} &= h(t) X(t) & D(0) &= a \end{aligned}$$

(Matis 1984).

The hazard function is a standard way of distinguishing different survival distributions. The hazard function for the exponential distribution (corresponding to the G1 function) is constant, indicating that the proneness to degrade is independent of the time a unit of material has been in the rumen (Figure 2). The hazard corresponding to the G1-lag function is discontinuous: It is initially zero

and then constant. The hazard corresponding to the G1G1 function increases to an asymptote. The hazard function for the Erlang distribution (corresponding to the G2 and G3 function) increases to an asymptote. The hazard function for the Weibull distribution can be constant, decreasing, or increasing. The hazard function for the loglogistic distribution is either decreasing or nonmonotonic—increasing to a maximum and then gradually decreasing to zero.

The constant hazard of the exponential distribution is biologically unreasonable. When feed enters the rumen, there is an early stage of reduced degradation while microbial attachment and colonization occurs. The G2 and G3 models with increasing hazards were developed and seem reasonable for *passage* of particles from the rumen, where as time increases the particle size reduces and the particles mix throughout the rumen, so that the chance of passage increases. An increasing hazard for degradation, however, might be less realistic. Forage may become increasingly lignified the longer it remains in the rumen. Consequently, after some initial stage, the proportion of the remaining feed that is more difficult to digest might increase with time. This process indicates a nonmonotonic hazard similar to that of the loglogistic distribution. A nonmonotonic hazard may be less reasonable for feedstuffs, other than forages.

The nonmonotonic hazard of the loglogistic model, in addition to providing insight into the process of degradation in the rumen, suggests features of the degradation process which may be important for summarizing feeds quality. For instance, the quantity $(d-1)^{1/d}/c$ is the time when the hazard is maximum.

4 FITTING THE REGRESSION FUNCTIONS TO DATA

The seven regression functions were fitted to 168 time courses. The 168 time courses came from two experiments involving dairy cows. The first experiment contributed 72 time courses for three response variables: crude protein, neutral detergent fiber, and total dry matter. The experiment involved two cows and six forages (alfalfa and orchardgrass at three stages of maturity), arranged in a crossover design. For each cow \times period \times feed combination, the percent disappearance was measured at 0, 2, 6, 12, 18, 24, 36, 48, and 72 hours of incubation.

The second experiment contributed 96 time courses for two response variables: organic matter and neutral detergent fiber. The experiment involved two cows and 12 feeds (alfalfa and switchgrass at 2 stages of maturity treated with 3 levels of sodium hydroxide) arranged in a crossover design. For each cow \times period \times feed combination, the percent disappearance was measured at 3, 6, 15, 24, 48, 72, and 96 hours of incubation.

Ordinary nonlinear least squares was used to fit each of the seven regression functions to the 168 time courses. The estimates were determined using the Gauss-Newton method (Seber and Wild 1989). Weighted least squares, where the

weights are inversely proportional to a function of $D(t)[100-D(t)]$, often is more efficient for percentage responses. But since the disappearance percentages usually ranged between 10% and 90%, the gain in efficiency was minimal.

The difference between the fitted values of the different regression functions was generally small. The loglogistic function (7) had the smallest or next to smallest mean square error (MSE) in 44% of the 168 time courses. Moreover, it had the largest or next to largest MSE in only 11% of cases. The fit of the Weibull function (6) was comparable to, but rarely better than, the loglogistic function.

The G1-lag function (2) fit nearly as well as the loglogistic. It had the smallest or next to smallest MSE in 38% of the cases and largest or next to largest in 13%. A denser placement of time points around the nonsmooth change of phase might reveal lack of fit.

The G1 function (1) had the smallest or next to smallest MSE in 35% of the time courses but the largest or next to largest in 46%. In the time courses with a weak sigmoidal trend, the parsimonious G1 function preformed well. However, the parameter d in the Weibull and loglogistic functions was estimated to be greater than 1, indicating some sigmoidal behavior, in 78% and 87% of the time courses, respectively.

The estimate of d for the Weibull and loglogistic functions indicated that the proportion of the total increase before the inflection point averaged .22. Consequently, the G2 function, where the proportion is constrained to be .264, often compared favorably with the other functions. The fit of the G1G1 curve (3) usually had a larger MSE than G2, due to the extra parameter. The proportion of increase before the inflection point was rarely as large as .32. Consequently, the G3 function (5) showed a relatively poor fit.

CONVERGENCE TROUBLE

The nonsmooth transition in the G1-lag function (2) can complicate the nonlinear least-squares estimation. The least-squares estimation is often troubled by multiple minima. The sum-of-squares surface

$$S(a, b, c, d) = \begin{cases} \sum_{t_i \leq d} (y_i - a)^2 \\ \sum_{t_i > d} (y_i - a - b(1 - \exp(-c(t_i - d))))^2 \end{cases}$$

is not smooth; the partial derivatives with respect to the parameters b , c , and d are not continuous at $d=t_i$. Geometrically, this creates sharp ridges running across the surface wherever d equals a time point. Between any two ridges a local minimum may occur, particularly between early time points. In this case, the only reliable estimation method is a grid search over values of d .

Instead of ridges, the nonsmoothness may cause a deep furrow in the sum-of-squares surface centered around a time point. In this case the least squares estimate of d equals that time point. This clearly biases the estimate when the true value of d is close to a design point. These furrows occurred on 8% of the 168 sum-of-squares surfaces.

When the least-squares estimate of d is less than or equal to the smallest time point in the data, the partial derivatives of the regression function with respect to a , b , and d are linearly dependent. This generates singularities which undermine the standard nonlinear least-squares estimation methods. This can occur when the minimum incubation time is zero and the trend is not sigmoidal. This phenomenon occurred in 8% of the time courses.

The four-parameter G1G1 function (3) was unstable and often degenerated into simpler three-parameter functions. The iterative estimation sequence diverged in almost half of the time courses. In 9% of the time courses, where the trend was not strongly sigmoidal, the residual sum of squares did not have a global minimum. In these cases, the least-squares estimate of d was infinite. This phenomenon induces the G1 function

$$\lim_{d \rightarrow \infty} (D_{G1G1}(t)) = a + b(1 - \exp(-ct)) .$$

In 36% of the time courses, the parameters c and d of the G1G1 function approached the same value. The partial derivatives of the G1G1 regression function with respect to c and d are linearly dependent when $d \rightarrow c$. This generates singularities which undermine the standard nonlinear least-squares estimation methods. This phenomenon induces the G2 function

$$\lim_{d \rightarrow c} (D_{G1G1}(t)) = a + b(1 - \exp(-ct) (1 + ct)) .$$

Van Milgen, Murphy, and Berger (1991) found this degeneration of the G1G1 function in 16% of the time courses they studied.

5 EFFECTIVE DEGRADABILITY

One important feature of the rumen in situ disappearance data is the total percentage of potentially degradable material in the feed. Degradation, however, is overestimated in the in situ experiment, because, under normal conditions, some of the bag contents would have passed out of the rumen before they had a chance to degrade. The effective degradability, is the limiting degradable percentage, corrected for passage. Ørskov and McDonald (1979) calculated effective degradability as

$$D(0) + \int_0^{\infty} \frac{dD(t)}{dt} \exp(-kt) dt \quad (10)$$

where $\exp(-kt)$ is the fraction of rumen contents which entered the rumen at time zero and had not yet passed at time t . (The parameter, k , is either determined from a separate passage experiment, or taken from the literature.) The integrand is the degradation rate corrected for passage. The effective degradability (10) can be evaluated for the first five of the seven regression functions:

| | |
|--------|--------------------------------|
| G1 | $a + \frac{bc}{c+k}$ |
| G1-lag | $a + \frac{bc}{c+k} \exp(-kd)$ |
| G1G1 | $a + \frac{bcd}{(c+k)(d+k)}$ |
| G2 | $a + \frac{bc^2}{(c+k)^2}$ |
| G3 | $a + \frac{bc^3}{(c+k)^3}$ |

The integral in (10) can not be explicitly solved for the Weibull and loglogistic regression functions. The integrals can be evaluated to any degree of precision using numerical quadrature. Alternatively, closed form approximations can be constructed. The following approximations are the result of changing variables from t to $D(t)$ and then applying Simpson's rule on four equally spaced intervals between a and $a+b$.

$$\text{Weibull} \quad a + \frac{b}{12} \left(1 + 4 \exp\left(-\left(\ln\left(\frac{4}{3}\right)\right)^{1/d} \frac{k}{c}\right) + 2 \exp\left(-\left(\ln(2)\right)^{1/d} \frac{k}{c}\right) + 4 \exp\left(-\left(\ln(4)\right)^{1/d} \frac{k}{c}\right) \right)$$

$$\text{Loglogistic} \quad a + \frac{b}{12} \left(1 + 4 \exp\left(-\left(\frac{1}{3}\right)^{1/d} \frac{k}{c}\right) + 2 \exp\left(-\frac{k}{c}\right) + 4 \exp\left(-3^{1/d} \frac{k}{c}\right) \right)$$

The quality of these approximations was evaluated using the SAS/IML numerical integration function QUAD (SAS 1991). The approximations of effective degradability were within one percentage point of the exact value in 82% and 74% of the cases, respectively.

6 SUMMARY

The loglogistic function has empirical and theoretical advantages over the other six regression functions. It has a flexible shape, gave the best overall fit to the experimental data, can be derived from a sensible underlying degradation model, and suggests new meaningful parameters. However, effective degradability must be calculated numerically or approximated.

The standard G1 function provides an erratic fit and implies an unrealistic constant hazard. The G1-lag modification improves the empirical fit at the expense of an unrealistic and ill-behaved regression function. The G1G1 function makes the time lag smooth, but is unstable and often induces the simpler G2 model. The *Gm* family of models were originally derived from an underlying probability model. The Weibull and loglogistic models are extensions of this probability model using different distributions.

It is not clear that one regression function is satisfactory for modeling in situ degradation for all feed types. The probability models for degradation time provide a means for developing new functions corresponding to sensible hazard rates.

REFERENCES

- Cox, D. R. and Oakes, D. (1984). *Analysis of Survival Data*. Chapman and Hall.
- Matis, J. H. (1984). A Generalized Approach to Compartmental Modeling Based on Retention Time Distributions. In *Modeling Ruminant Digestion and Metabolism, Proceedings of the Second International Workshop UC Davis*, editors: Baldwin, R. L. and Bywater, A. C.
- Matis, J. H., Wehrly, T. E., and Ellis, W. C. (1989). Some Generalized Stochastic Compartment Models for Digesta Flow. *Biometrics* 45, 703-720.
- McDonald, I. (1981). A Revised Model for the Estimation of Protein Degradability in the Rumen. *J. agric. Sci. Camb.* 96, 251-252.

- Nocek, J. E. (1988). In Situ and Other Methods to Estimate Ruminal Protein and Digestibility: A Review. *J Dairy Sci.* 71, 2051-2069.
- Ørskov, E. R. and McDonald, I. (1979). The Estimation of Protein Degradability in the Rumen from Incubation Measurements Weighted according to Passage. *J. agric. Sci. Camb.* 92, 499-503.
- Pond, K. R., Ellis, W. C., Matis, J. H., Ferreiro, H. M., & Sutton, J. D. (1988). Compartment Models for Estimating Attributes of Digesta Flow in Cattle. *British Journal of Nutrition* 60, 571-595.
- SAS Technical Report P-220 (1991). *SAS/IML Software: Changes and Enhancements, Release 6.07* 27-34. SAS Institute inc.
- Seber, G. A. F. and Wild, C. J. (1989). *Nonlinear Regression*. Wiley.
- Van Milgen, J., Murphy, M. R., and Berger, L. L. (1991). A Compartmental Model to Analyze Ruminal Digestion. *J. Dairy Sci.* 74, 2515-2529.

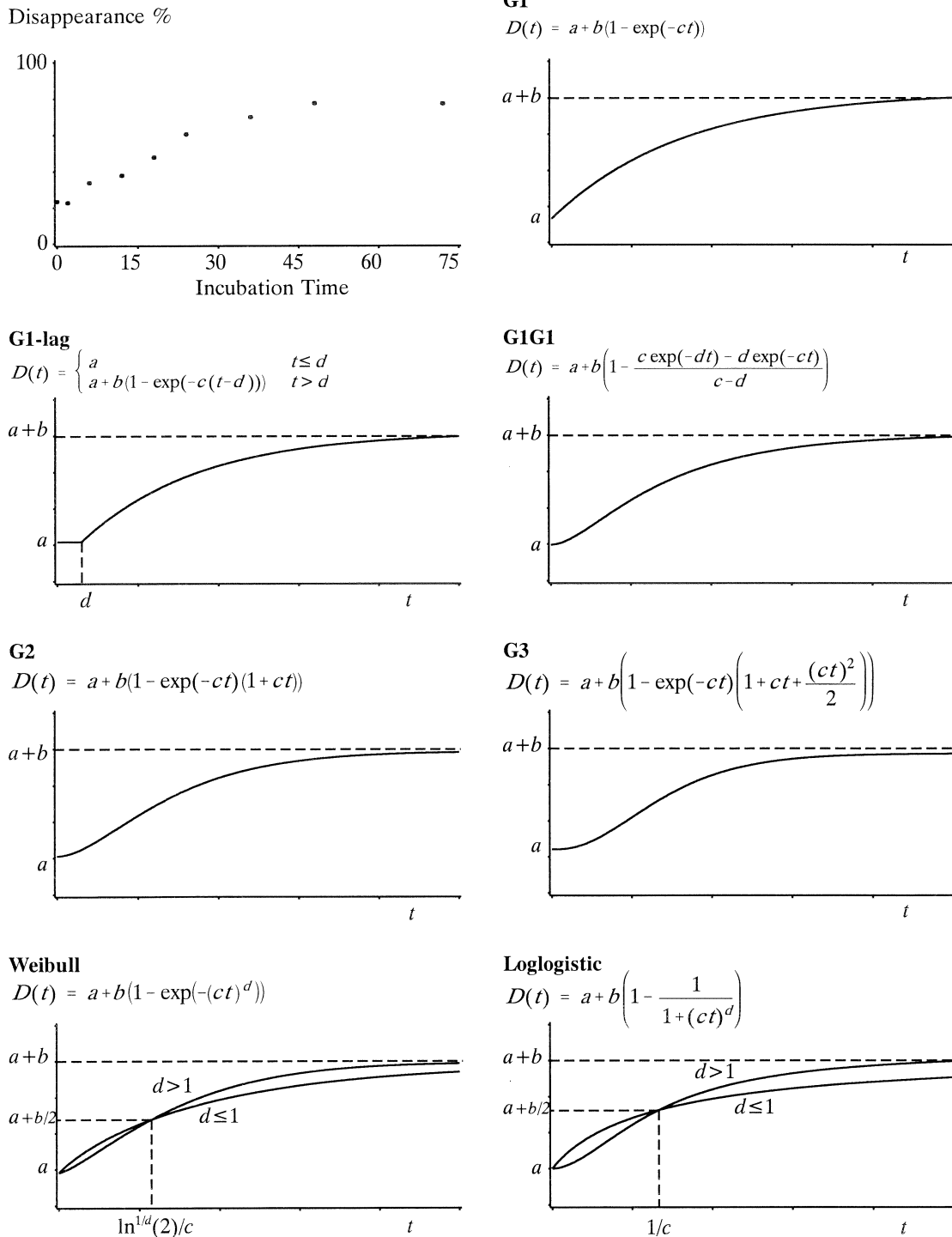
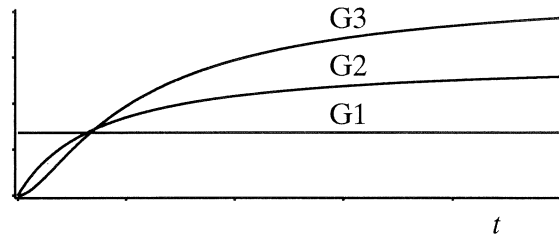


Figure 1. In situ disappearance data and graphs of the regression functions.

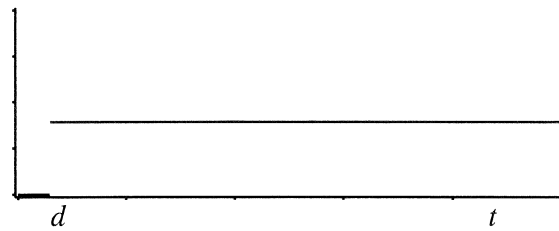
G_m

$$h(t) = \frac{c^m t^{m-1}}{(m-1)! \sum_{j=0}^{m-1} \frac{(ct)^j}{j!}}$$



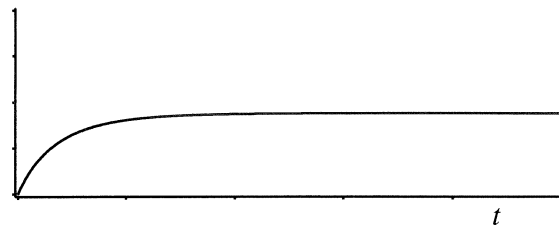
G₁-lag

$$h(t) = \begin{cases} 0 & t \leq d \\ c & t > d \end{cases}$$



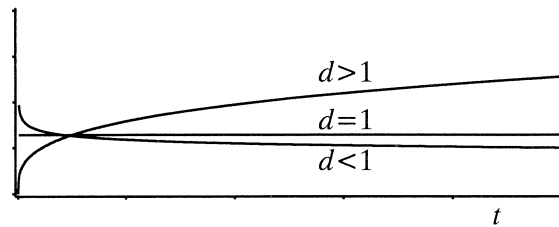
G₁G₁

$$h(t) = \frac{cd(e^{-dt} - e^{-ct})}{ce^{-dt} - de^{-ct}}$$



Weibull

$$h(t) = dc^d t^{d-1}$$



Loglogistic

$$h(t) = \frac{dc^d t^{d-1}}{1 + (ct)^d}$$

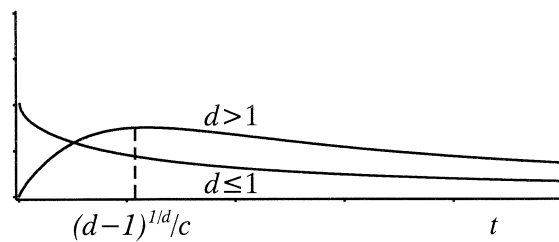


Figure 2. Hazard functions corresponding to the regression models.