

Mathematical Models and Methods for HIV Infection Dynamics

Sarah Holte
Fred Hutchinson Cancer Research Center
Seattle, WA

HIV

- HIV is a retrovirus which primarily infects CD4+ T-cells, as well as macrophages and possibly dendritic cells.
- HIV infection induces both a humoral (antibody) and cellular (CTL) immune response.
- Many HIV-infected individuals experience a long period of clinical latency (~ 10 years) before the onset of AIDS. However, mathematical models helped contribute to the discovery that the virus is extremely active during this period of time.
- Recent advances in treatment (HAART) have allowed suppression of this high level of activity which has resulted in clearly improved clinical outcomes for treated individuals.

Differential Equations

Differential equations are primarily used in situations where rates of change are of interest.

As a simple example, consider

$$\frac{dX}{dt} = aX, \quad X(0) = X_0$$

where $X(t)$ is the size of the size of some population at time t and X_0 is the initial size of the population.

This is a differential equation, and the “solution” to this equation is a function, in this case

$$X(t) = X_0 e^{at}$$

Note that the population has qualitatively different behaviour depending on whether or not $a > 0$, $a < 0$, or $a = 0$.

There are a number of types of differential equations:

- Ordinary differential equations (ODEs)
- Partial differential equations (PDEs),
- Delay differential equations (DDEs).

Current Interests

- Competition Models for two strains of HIV - with Michael Emerman
 - Recently extended to incorporate the delay between cellular infection and viral production - at the Industrial Mathematics Modeling Workshop at North Carolina State University.
- Effects of Host genotypes on immune response. Model development aimed at assessing which immune response mechanisms are affected by host genotype - with Uma Malhotra and Julie McElrath.
- Development of a mathematical model for Structured treatment Interruption and associated control theory (with HT Banks at NC State University) which could be used to assess different treatment strategies - Protocol development for a pilot study with Bob Geise and Larry Corey.
- Statistical methods for models described by ODEs, PDEs, Functional DEs.
 - Development of numerical methods required for complex non-linear random effects models that arise when means are described by differential equations - with HT Banks.
 - Modify and extend the work of HT Banks which has been used successfully in engineering literature to apply to models of interest in HIV science.

Today's Talk

- Extensions of standard linear decay models for infected cell populations using data from infected children - with Ann Melvin, Lisa Frenkel, and Jim Mullins.
- “Integrated Data Method” which can be used to estimate parameters in models with means defined by ODEs - with Peter Cornelisse, Patrick Heagerty, and Steve Self.

Mathematical models for viral decay after start of HAART

- HAART stands for Highly Active Anti-Retroviral Therapy
- Contains a Protease Inhibitor which renders free virus produced after treatment non-infectious
- Thus no new infections can occur
- However, existing infected cells continue to produce non-infectious virus, which is measured in plasma.

A Mathematical Model for Viral Decline After Treatment

- X represents a population of cells which produce large amounts of virus and die quickly, Activated CD4+ T-cells.
- Y represents a population of cells (possibly macrophages) which produce less virus but have a longer life expectancy.
- V is the population of HIV virions.

Perelson's model. Viral trajectories given by solution to the differential equations:

$$\begin{aligned}\frac{dX}{dt} &= -\delta X, & X(0) &= X_0 \\ \frac{dY}{dt} &= -\mu Y, & Y(0) &= Y_0 \\ \frac{dV}{dt} &= p_x X + p_y Y - cV, & V(0) &= V_0\end{aligned}$$

Use a numerical ODE solver and non-linear least squares regression to obtain estimates of δ , μ , p_x and p_y .

Density Dependent decay model

$$\begin{aligned}\frac{dX}{dt} &= -(\delta X^{r-1}) X, \quad X(0) = X_0 \\ \frac{dY}{dt} &= -(\mu Y^{r-1}) Y, \quad Y(0) = Y_0 \\ \frac{dV}{dt} &= p_x X + p_y Y - cV, \quad V(0) = V_0\end{aligned}$$

- If $r \neq 1$ then the per capita decay rates for infected cells which depend on the size of the decaying populations.
- If $r = 1$ this model reduces to Perelson's model.

Questions:

- Do the data support density dependent decay (Is $r > 1$)?
- Does more than one infected cell compartment contribute to viral pool (Is $p_x > 0$ and $p_y > 0$)?

We use a non-linear least squares method and numerical solutions to the system to obtain estimates of δ , μ , p_x , p_y , and r , and profile likelihood confidence intervals to test the hypothesis: $r = 1$, $p_x = 0$, and $p_y = 0$.

We used values of V_0 , X_0 , and Y_0 obtained from data.

	r	p_x	p_y
Density Dependent Decay			
Estimate	1.435	431.5	22.7
95% CI	(1.396,1.479)	(252.2,902.7)	(10.0-49.5)
Range	1.434-1.436	28.5-3763.8	0.12-303.1
Constant Decay			
Estimate	1	253.1	60.0
95% CI	NA	(107.7,543.0)	(27.5,132.9)
Range	NA	16.7-2211.7	0.3-800.0

	δ	μ
Density Dependent Decay		
Estimate	0.00161	0.00012
95% CI	(0.00126,0.0021)	(0.00007,0.00017)
Range	0.0005-0.0041	0.00001 - 0.00035
Constant Decay		
Estimate	0.32	0.0135
95% CI	(0.21,0.43)	(0.008,0.018)
Range	0.32-0.33	0.0134-0.0135

Figure 1: Total body plasma HIV-1 RNA by months since start of treatment for two patients with fitted trajectories from the density dependent and constant decay models. Solid lines indicate fitted model trajectories from the density dependent decay model and dashed lines indicate fitted model trajectories from the constant decay model.

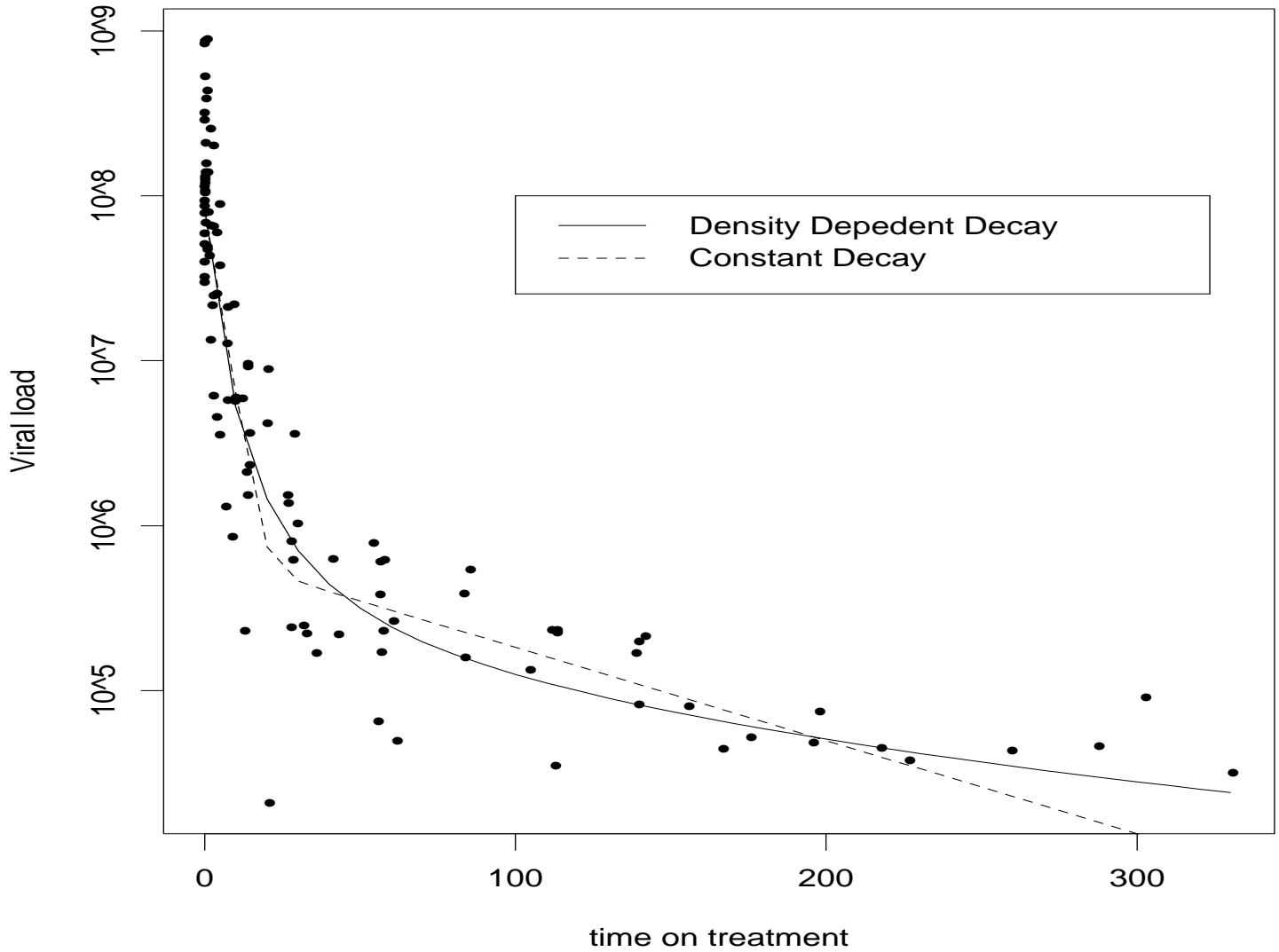
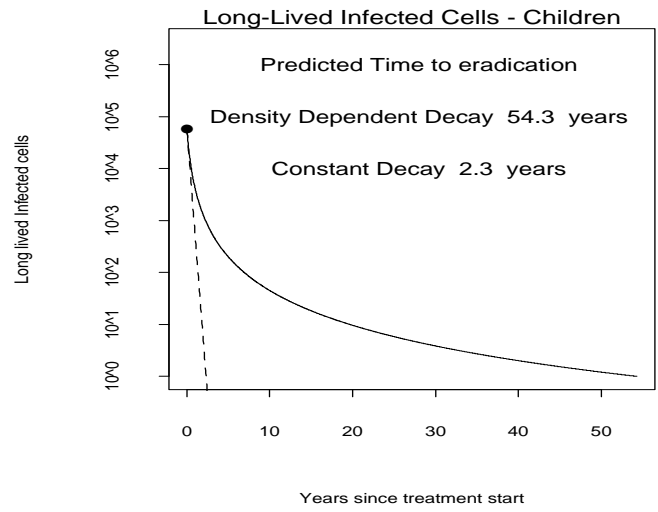
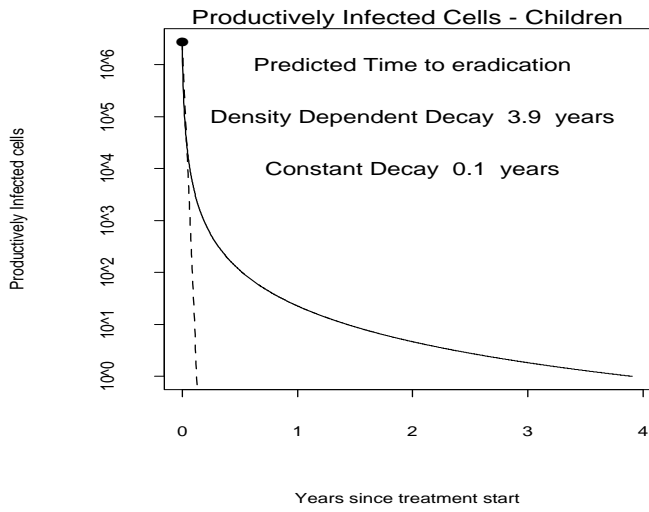


Figure 2: Theoretical trajectories of the total body short-lived and long-lived infected cell population by years since start of treatment.



A Model Extension which could be tested

- X represents a population of cells which produce large amounts of virus and die quickly, Activated CD4+ T-cells.
- V is the population of HIV virions.
- Z represents a population of immune systems cells (possibly CTLs) specific to HIV.

$$\begin{aligned}\frac{dX}{dt} &= kVT - \delta X^r - \alpha XZ \\ \frac{dV}{dt} &= p_x X - cV \\ \frac{dZ}{dt} &= s + \gamma XZ - \beta Z\end{aligned}$$

Hypothesis of interest:

- $r > 1$, suggests homeostasis.
- $\alpha > 0$, suggests CTL response.
- $k > 0$, suggests continued viral replication.

Mathematical Analysis of the final model

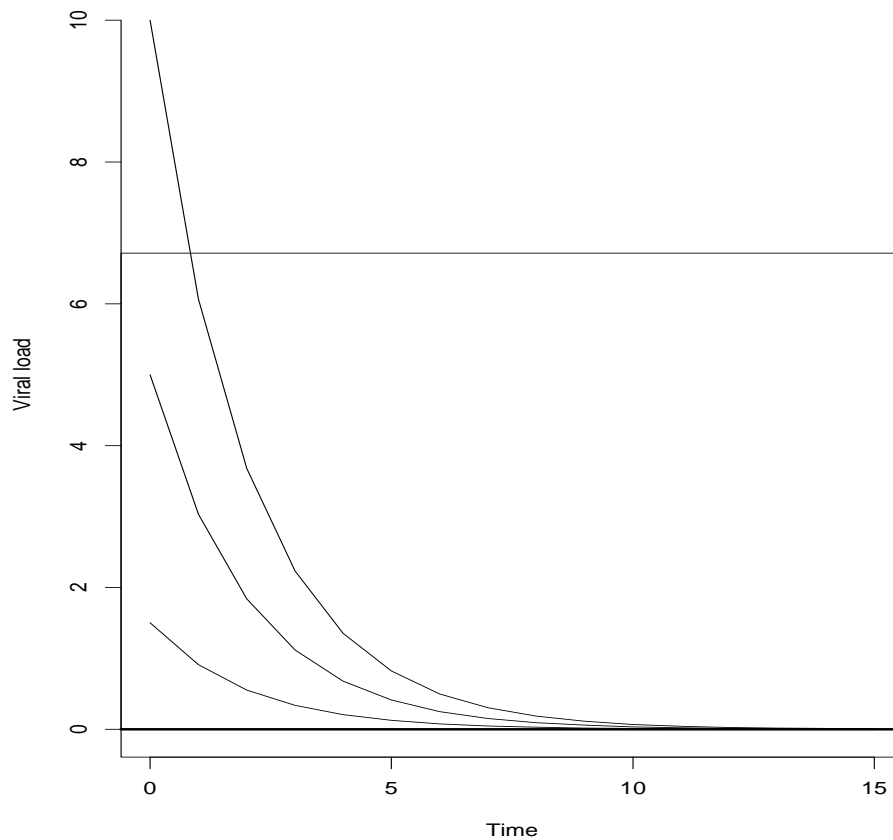
$$\begin{aligned}\frac{dX}{dt} &= \delta X - \alpha X Z \\ \frac{dV}{dt} &= pX - cV \\ \frac{dZ}{dt} &= s + \gamma X Z - \beta Z\end{aligned}$$

Equilibria of a system of differential equations are defined as those values, X , V , and Z , which satisfy the equations:

$$\begin{aligned}\frac{dX}{dt} &= 0 \\ \frac{dV}{dt} &= 0 \\ \frac{dZ}{dt} &= 0\end{aligned}$$

This model has two equilibria:

- $X = 0, V = 0, Z = \frac{s}{\beta}$
- $X = \frac{\delta \beta - \alpha s}{\delta \gamma}, V = \frac{p (\delta \beta - \alpha s)}{c \delta \gamma}, Z = \frac{\delta}{\alpha}$



Importance of stability analysis

If an equilibrium is stable, then all “nearby” trajectories converge towards it. In the previous model, the equilibrium $X = 0$, $V = 0$, $Z = \frac{s}{\beta}$ is the mathematical model of viral extinction, so we would like to know when this equilibrium is stable.

Stability analysis of the equilibria

To assess the stability of equilibrium values of the system, we assess the sign of the eigenvalues of the Jacobian of the vector field which defines the system, evaluated at the equilibrium value. Whenever the sign of all these eigenvalues is negative, the equilibrium is stable.

Fortunately, the eigenvalues associated with the first equilibrium

$$X = 0, V = 0, Z = \frac{s}{\beta}$$

are quite simple,

$$\left(-c, \delta - \frac{\alpha s}{\beta}, -\beta\right)$$

so that this equilibrium is stable whenever

$$\delta - \frac{\alpha s}{\beta} < 0 \quad \text{i.e.} \quad \delta\beta < \alpha s$$

Conclusion: The stability of the zero equilibrium for viral load depends on the the proliferation rate of infected cells, the death rate of CTL's, the rate of naive HIV specific CTL precursors into the system, and the effectiveness of CTL's in their ability to remove infected cells.

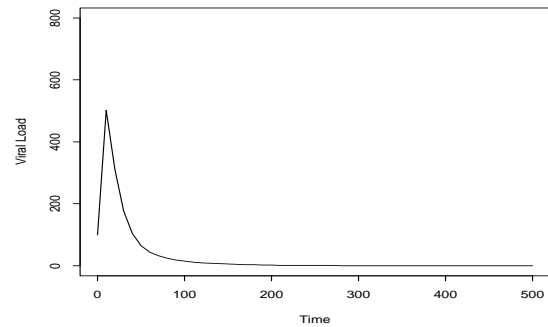
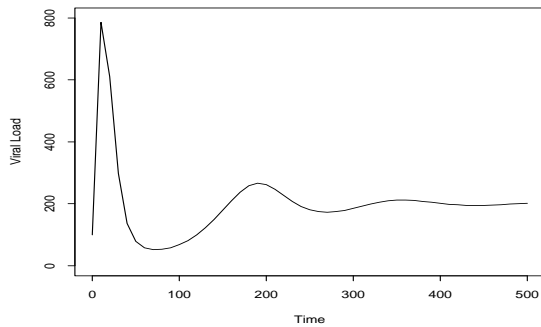
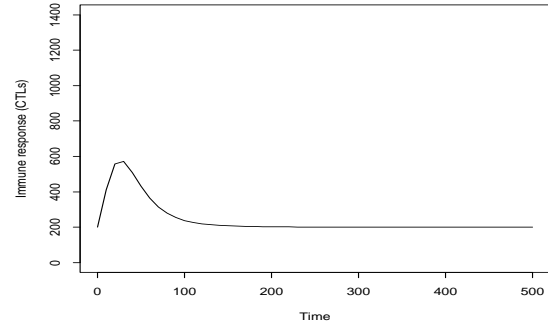
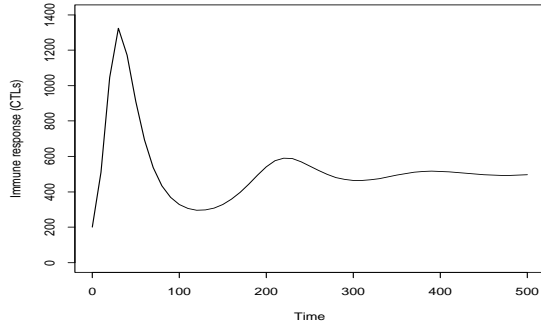
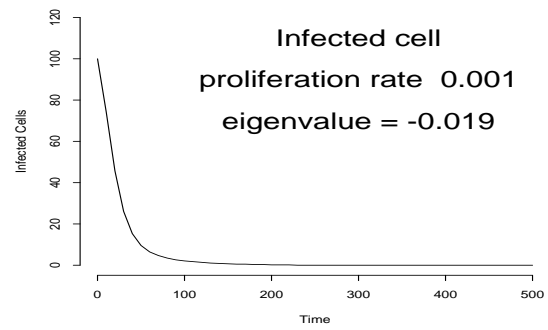
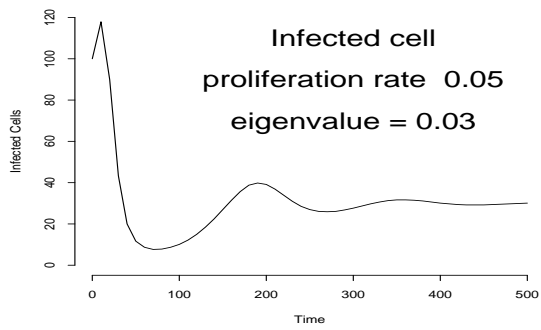
Eigenvalues for equilibria values are rarely this simple

The eigenvalues for the other equilibrium of this system

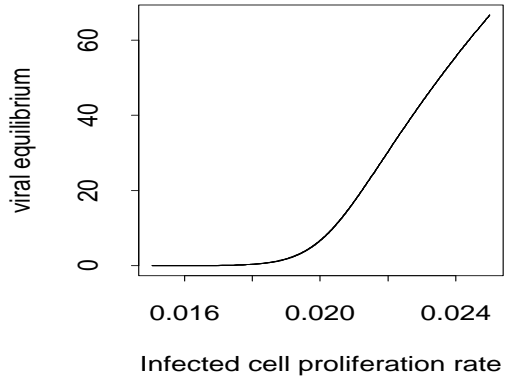
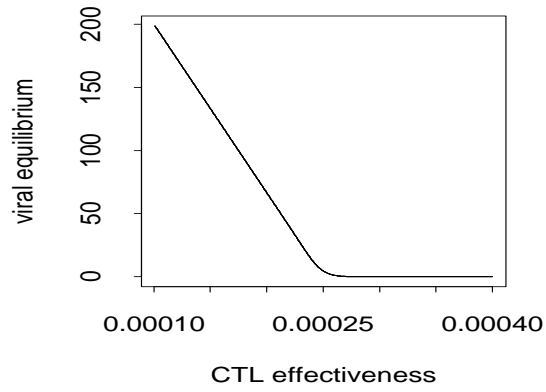
$$X = \frac{\delta \beta - \alpha s}{\delta \gamma}, V = \frac{p (\delta \beta - \alpha s)}{c \delta \gamma}, Z = \frac{\delta}{\alpha}$$

are

$$\left(-c, \frac{\alpha s \pm \sqrt{4\delta^2 \alpha s + \alpha^2 s^2 - 4\delta^2 \beta}}{2\delta}\right)$$



Example of the effect on model based viral load by changing a single parameter



Example of a bifurcation in stability of zero equilibrium for viral load

Conclusions

- Our estimates support density-dependent decay in comparison to constant decay
- Mechanism not investigated - Slowing decay rate could be results of some other time-dependent mechanism
- Further modeling and estimation which includes immune response could shed more light.
- If immune response wanes as antigen declines, therapeutic vaccines designed to stimulate the immune system to higher levels could expedite eradication of antigen

A method for estimating parameters in models with means defined by ordinary differential equations

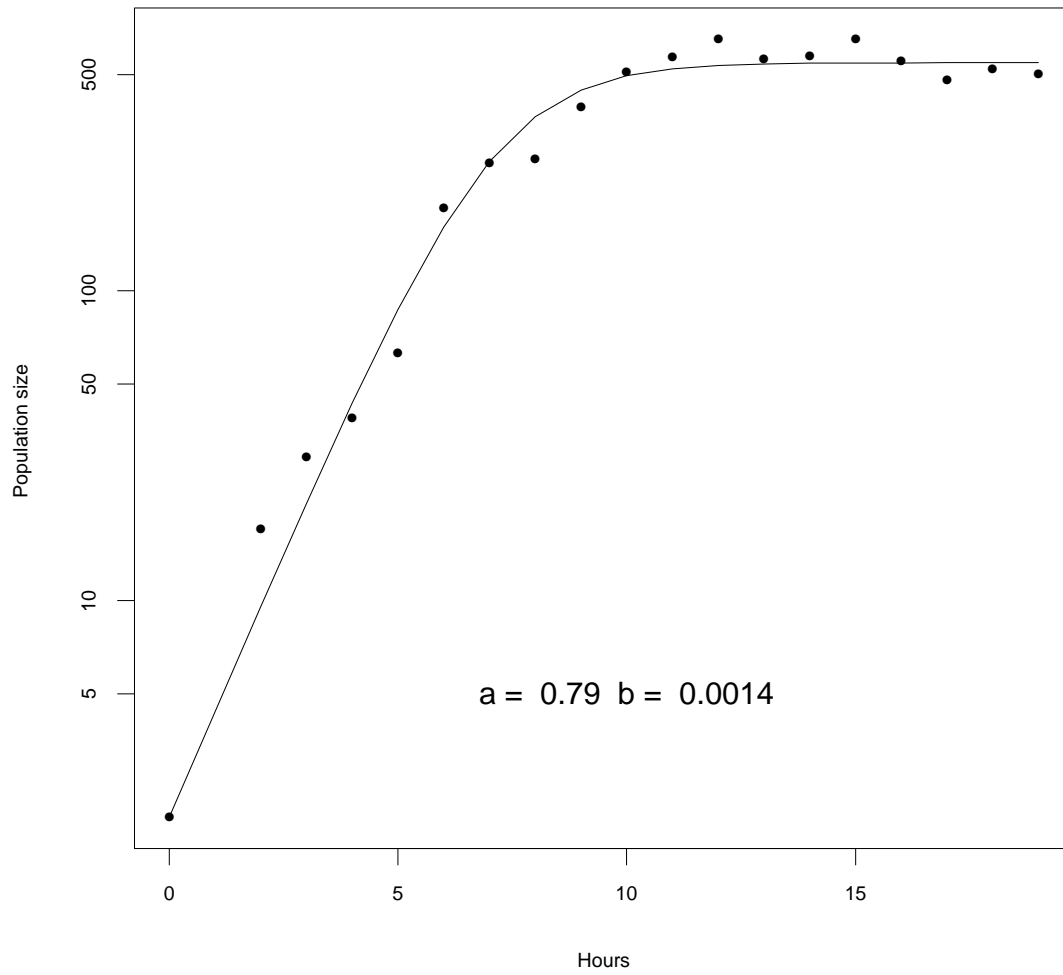
Data on Growth Curves of *paramecium aurelium* - Gause 1934
(From StatLib and described in Diggle - Time Series: A Biostatistical Introduction)

The population growth model - non-linear ODE

$$\frac{dX}{dt} = aX - bX^2, \quad X(0) = X_0$$

where a is the per-capita birth rate and $\frac{a}{b}$ is the carrying capacity of the population

Non-linear least squares fit of log transformed data using numerical ODE solver



An alternate method of fitting - Integrated data

The corresponding log transformed differential equation is

$$\frac{d(\log(X))}{dt} = a - bX, \quad \log(X(0)) = \log(X_0).$$

which can be written as the integral equation

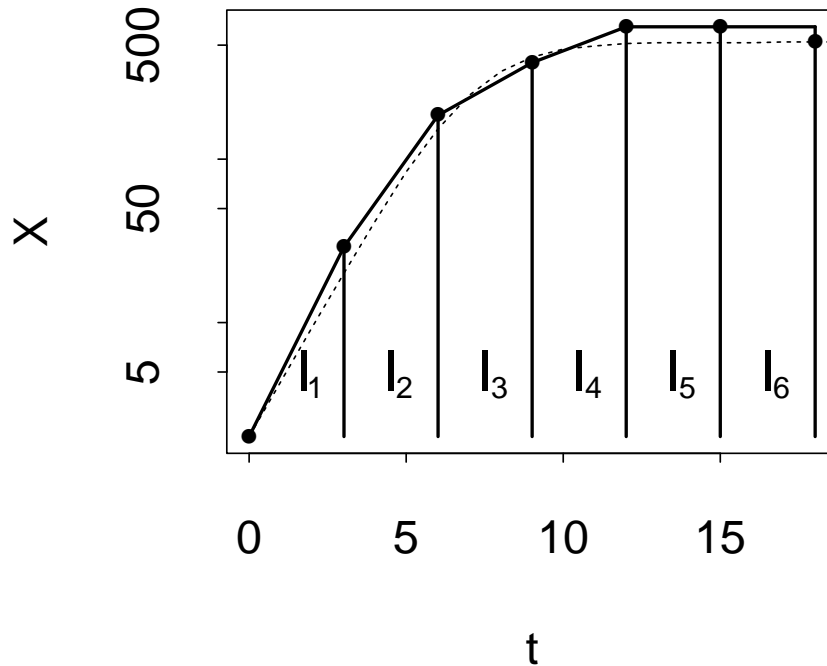
$$\log(X(t)) - \log(X_0) = a \int_0^t ds - b \int_0^t X ds$$

Response = Covariates

Define the linear regression problem

$$\mathbf{y} \sim a\mathbf{t} + b\mathbf{z} + \epsilon_{\mathbf{y}}, \quad \epsilon_{\mathbf{y}} \sim N(0, \sigma^2).$$

- $\mathbf{y} = \log(\mathbf{x}) - \log(x_0)$
- $\mathbf{t} = \int_0^t ds$
- \mathbf{z} approximates $\int_0^t X ds$



Defining covariates for the regression model

$$z_0 = 0$$

$$z_i = \frac{(x_i + x_{i-1})(t_i - t_{i-1})}{2} * e^{-\sigma^2} + z_{i-1}, \quad i = 1, \dots, n.$$

where $\mathbf{X} = (x_0, x_1, \dots, x_n)$ is the observed data.

The covariates \mathbf{z} are defined as $z_0 = 0$ and for $i = 1, \dots, n$,

$$z_i = \sum_{j=1}^i I_j$$

Integrated data fit of logistic growth model

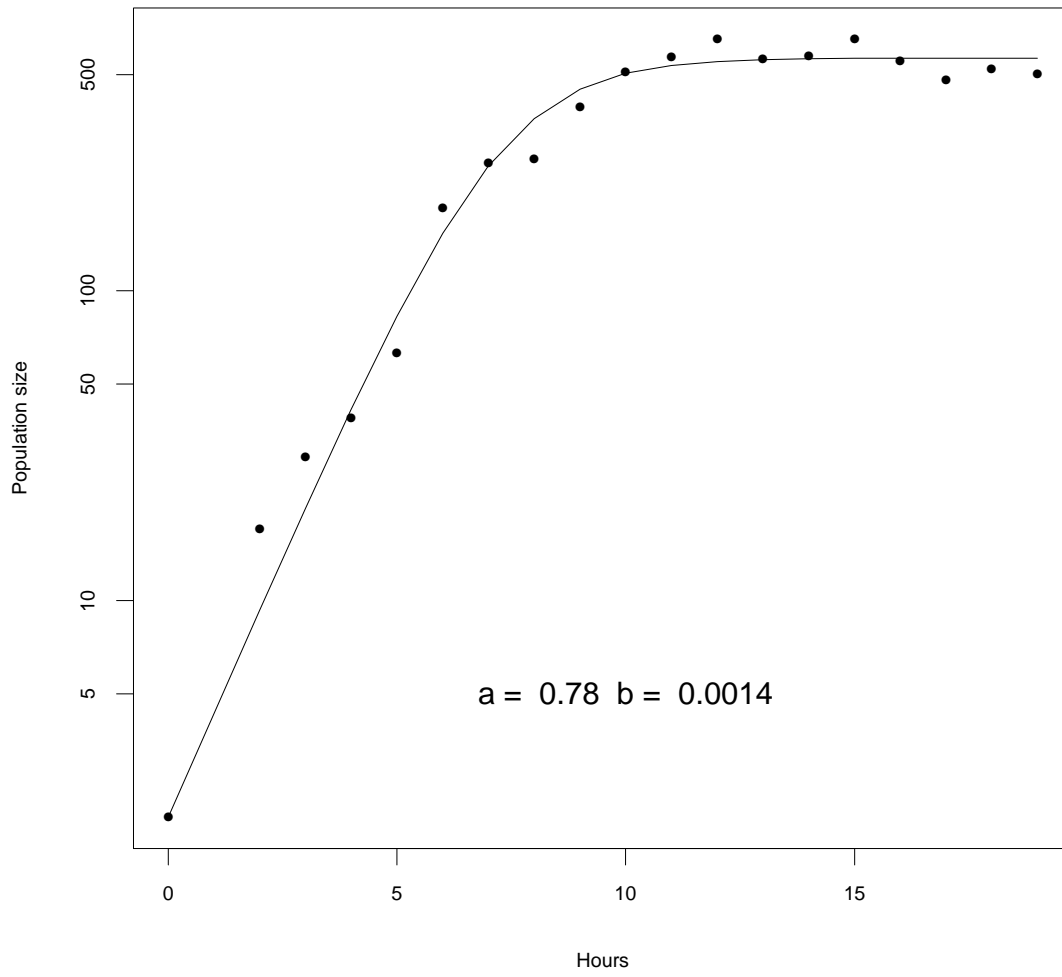


Table 1: Results of NLS and ID methods for estimation of parameters in models for growth curves of colonies of paramecium aurelium.

	<i>a</i>		<i>b</i>	
	NLS	ID	NLS	ID
Data Set 1	0.79 (0.75 , 0.86)	0.78 (0.73 , 0.84)	0.0014 (0.0012 , 0.0017)	0.0014 (0.0011 , 0.0015)
Data Set 2	0.84 (0.80 , 0.89)	0.80 (0.76 , 0.87)	0.0017 (0.0014 , 0.0019)	0.0016 (0.0013 , 0.0017)
Data Set 3	0.89 (0.86 , 0.94)	0.87 (0.83 , 0.91)	0.0016 (0.0015 , 0.0017)	0.0015 (0.0013 , 0.0016)
Data Set 4	0.84 (0.81 , 0.89)	0.82 (0.78 , 0.87)	0.0016 (0.0014 , 0.0017)	0.0015 (0.0013 , 0.0016)

Table 2: Simulation Results for Parameters in Logistic Growth Model using Nonlinear Least-squares (NLS) and Integrated Data (ID) Methods.

	a=0.8		b=0.0015	
	NLS	ID	NLS	ID
	N=1000	N=1000	N=1000	N=1000
Mean	0.800	0.800	0.0015	0.0015
Standard Error	0.017	0.022	0.00009	0.00010
95% Coverage				
Parametric Bootstrap	0.950	0.943	0.947	0.931
Non-Parametric Bootstrap		0.943		0.892

Simulation Experiment

To compare the two methods for accuracy and efficiency we performed a simulation experiment, simulating data from the model

$$\log\{y(\mathbf{t})\} = \log\{\mu(\mathbf{t})\} + \epsilon, \quad \epsilon \sim i.i.d N(0, 0.25^2)$$

where $\mu(\mathbf{t})$ is the solution of equation (1) with parameters set as follows: $a = 0.8$, $b = 0.0015$, and $y_0 = 2$. We chose these parameters (as well as 0.25 for the standard error) based on the fits to the four data sets we evaluated. We replicated the experiment 1000 times.

Table 5: Simulation Results

	$E(a)$	$E(b)$	σ_a	σ_b	ρ	σ_y
True Value	0.800	0.00150	0.050	0.00010	0.100	0.150
Mean	0.798	0.00151	0.053	0.00011	0.121	0.163
Standard Error	0.0054	0.000020	0.0059	0.000014	0.1243	0.003

Extention to Random Effects

To assess the possibility of extending the method to one where the model parameters themselves are viewed as random variable (hierarchical or random effects model), we performed a simulation experiment, simulating data \mathbf{y} under the following assumptions:

- We sampled 30 colonies at \mathbf{t} sampling times; 50 observations equally spaced between 0 and 20.
- for each colony, we simulated data as $\log\{y(\mathbf{t})\} = \log\{\mu(\mathbf{t})\} + \epsilon$, $\epsilon \sim i.i.d N(0, 0.15^2)$ where $\mu(\mathbf{t})$ is the solution of equation (1) with parameters simulated as follows:
 - $a \sim N(0.8, 0.05)$
 - $b \sim N(0.0015, 0.0001)$
 - $\text{corr}(a,b) = 0.10$

The experiment was replicated 3000 times.

Conclusions

- Integrated Data method is a viable alternative to nonlinear least-squares with an analytic or numerical solution to a system of ordinary differential equations.
- Can be used to obtain starting values for nonlinear least-squares estimation if desired.
- Requires observations for all “states” of the system.
- For a single “state” system requires estimate of the variance whereas nonlinear least-squares does not.
- Simulations suggest that it may be possible to extend to nonlinear random effects estimation.