Mechanistic Models of Growth and Decay

H. T. Banks
Center for Research in Scientific Computation
Center for Quantitative Sciences in Biomedicine
North Carolina State University
Raleigh, NC

Fall, 2017

1 Mechanistic Models of Growth and Decay

In the following sections, we examine and explain the characteristic features of four main types of differential equation representations that are frequently used to model growth and decay in dynamical systems of equations. The first is an exponential (also called Malthusian) growth and decay model, for which a population is assumed to grow at a rate proportional to the size of the population at any given time [3, 4, 12, 14, 20]. The second is Michaelis-Menten kinetics, which emulates enzyme mediated kinetics. Finally, we use the Gompertz and logistic (also called Verhulst-Pearl) rate laws, which both involve time dependent growth/decay rates.

1.1 Exponential Growth and Decay

Many of the mechanistic terms used in this paper were formulated using an exponential growth or decay rate. In these systems, the change in population size is directly proportional to the population size at any given time:

$$\frac{dP}{dt} = kP. \tag{1}$$

Using separation of variables, the solution to (1) is found to be

$$P(t) = P_0 e^{kt}, \tag{2}$$

where P_0 is the concentration of the population at $t = 0$ (initial population concentration). Although the exponential model is simple to implement, it poses a problem in the sense that it creates an unbounded solution because of the constant intrinsic or per capita growth/decay rates

$$\frac{\frac{dP}{dt}}{P} = k. \tag{3}$$

Such rates are unlikely to occur in nature or specifically in a population of cells. To more accurately model the conversion of one molecule to another, Michaelis-Menten kinetic rates can be used. As we shall explain below, one can also turn to *time varying* rates where the growth/decay slows as the size of the population grows/decays.

1.2 Michaelis-Menten Kinetics

In a Michaelis-Menten reaction, the rate of product concentration [P] will initially be proportional to the initial concentration of a substrate [S]. However, as the concentration of the substrate increases, the rate will lose this proportionality and approach a maximum velocity, V_{max}. The rate equation for this product formation is given by

$$\frac{d[P(t)]}{dt} = \frac{V_{max}[S]}{k_m + [S]}. \tag{4}$$

1.2.1 Michaelis-Menten Derivation

The formulae of Michaelis-Menten kinetics are restricted to approximations by the law of mass action, which assumes that the amount of product formed in a reaction is proportional to the concentration of reactants [11, 20]. The kinetics begin with the joining of an enzyme (E) and substrate (S), which bind together to form a complex (C). The enzyme then facilitates the transformation of the substrate to the product (P), from which the enzyme then dissociates:

$$E + S \Leftrightarrow C \rightarrow P.$$

If we assume E and S bind and dissociate with rates k_f and k_r, respectively, and that P is formed with rate k_{cat}, then we can set up the differential equation system:

$$\begin{aligned}
\frac{dS}{dt} &= -k_f ES + k_r C \\
\frac{dE}{dt} &= -k_f ES + k_r C + k_{cat} C \\
\frac{dC}{dt} &= k_f E\mathrm{S} - k_r ES - k_{cat} C \\
\frac{dP}{dt} &= k_{cat} C. \qquad (5)
\end{aligned}$$

Although not the original derivation of Michaelis Menten kinetics, George Briggs and John Haldane [3, 8, 20] came up with an alternate derivation in 1925 in which they assumed that the rate of change of C was negligible in comparison to the rate of change of S. Accordingly, they assumed that $\frac{dC}{dt} \approx 0$. Thus, $k_f ES - k_r C - k_{cat} C \approx 0$, which implies that

$$C = \frac{k_f}{k_r + k_{cat}} ES, \qquad (6)$$

meaning that $\frac{dP}{dt} = k_{cat}\left(\frac{k_f}{k_r + k_{cat}} ES\right)$. If we define $k_m = \frac{k_r + k_{cat}}{k_f}$, then

$$\frac{dP}{dt} = \frac{k_{cat}}{k_m} ES. \qquad (7)$$

Thus, we now have an equation for the rate of growth for the product in terms of the concentrations of the enzyme and substrate. However it would be better if we could find the rate in terms of only the concentration of the substrate. To solve for E, we may substitute k_m and (6) into the enzyme conservation law

$$E_0 = E + C = E + \frac{k_f}{k_r + k_{cat}} ES = E\left(1 + \frac{S}{k_m}\right)$$

which yields

$$E = \frac{E_0}{1 + \frac{S}{k_m}}.$$

Now that we have a formula for E, we may substitute it into (7), and find that if we define $V_{max} = k_{cat} E_0$, we obtain

$$\begin{aligned}
\frac{dP}{dt} &= \frac{S k_{cat}}{k_m} \frac{E_0}{1 + \frac{S}{k_m}} \\
&= \frac{k_{cat} S E_0}{k_m + S} = \frac{V_{max} S}{k_m + S}.
\end{aligned}$$

Under the same assumptions, we find $\frac{dS}{dt}$ to be the negative of $\frac{dP}{dt}$ using (6):

$$\begin{aligned}
\frac{dS}{dt} &= -k_f ES + k_r C = -C k_r - C k_{cat} + C k_r = -k_{cat} C \\
&= -k_{cat}\left(\frac{k_f ES}{k_r + k_{cat}}\right) = -\frac{k_{cat} ES}{k_m} \\
&= -\frac{V_{max} S}{k_m + S},
\end{aligned}$$

where k_m is the amount of substrate required for the rate to reach half of V_{max} and S is the concentration of substrate [2, 3, 20]. Note that if k_m is significantly greater than S, then this equation may approximate the exponential function, where k is substituted for $\frac{V_{max}}{k_m}$.

A plot of a Michaelis-Menten reaction is given in Figure 1. Although this may not be applicable for all of the reactions since some are not enzyme-mediated, it may work as a function with a bounded rate and requires only two parameter estimates. Alternatives such as the Gompertz and logistic (Verhulst-Pearl) laws also require only two parameters but, as we shall see next, offer the possible advantages of time-dependent growth/decay rates.

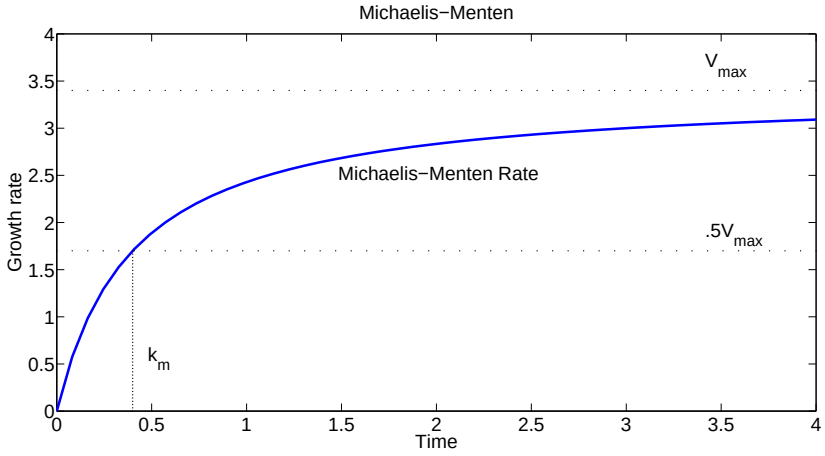


Figure 1: The rate of a Michaelis-Menten reaction with $k_m = 0.4$ and $V_{max} = 3.4$.

1.3 Logistic (Verhulst-Pearl) Growth/Decay

Another model that frequently characterizes population growth/decay is the logistic growth formulation. Under this growth assumption, the change in population size at a given time is given by:

$$\frac{dP(t)}{dt} = \alpha P(t)(1 - \frac{P(t)}{K}), \tag{8}$$

where $\alpha > 0$ is the growth/decay rate and K is the *carrying capacity* for the population.

One can readily derive the following properties for logistic models:

- The growth/decay rate (intrinsic or per capita rate) is given by

$$\frac{\dot{P}}{P} = \alpha \left[1 - \frac{P(t)}{K}\right].$$

  Note that $\frac{\dot{P}}{P}$ is positive or negative depending on whether $P < K$ or $P > K$, respectively, and it is largest/smallest near the initial population values P_0.

- The solution is given by

$$P(t) = \frac{K}{1 + (\frac{K}{P_0} - 1)e^{-\alpha t}}.$$

- The solution has a flex point at

$$\frac{K}{2}.$$

- At small values of initial population (i.e., as $P_0 \to 0^+$), we have $\dfrac{\dot{P}}{P} = \alpha\left[1 - \dfrac{P(t)}{K}\right] \to \alpha$; so initial rates remain bounded!

1.4 Gompertz Growth/Decay

The Gompertz law is again a mathematical model in which the rate of growth/decay is greatest at the start and slowest at the end. The dynamics are given by

$$\frac{dP(t)}{dt} = \alpha \ln(\frac{K}{P(t)})P(t) = \alpha[\ln(K) - \ln(P(t))]P(t), \tag{9}$$

where K is the carrying capacity (upper/lower asymptote) of the population $P(t)$ and $\alpha > 0$ is the intrinsic growth/decay rate. Either of these forms in (9) can be used in our mathematical modeling. The Gompertz law is often used to model the growth of tumors as well as for population growth, but it has described label loss rate well in previous studies [5, 21], and therefore will be a candidate model for comparison in our investigations. Characteristics for Gompertz rates can summarized by:

- The growth/decay rate (per capita) is

$$\frac{\dot{P}}{P} = \alpha[\ln K - \ln P] = \alpha \ln(\frac{K}{P}).$$

- The solution is given by

$$P(t) = Ke^{(\ln(\frac{P_0}{K})e^{(-\alpha t)})}.$$

- The solution has a flex point at

$$\frac{K}{e}.$$

- At small values of initial population (i.e., as $P_0 \to 0^+$), we have $\dfrac{\dot{P}}{P} = \alpha \ln \dfrac{K}{P} \to \infty$; so initial rates can approach unbounded values!

References

[1] Acetyl Esterase [Internet]. CPC Biotech S.r.l. 2007-2011. [cited 2012 Jun 07]. Available from: http://www.cpcbiotech.it/EN/c/d/enzyme-portfolio/enzymes/acetyl-esterase-lyophilized

[2] Peter Atkins and Julio De Paula, *Physical Chemistry* (9th edition), Freeman, New York, 2010.

[3] H.T. Banks, *Modeling and Control in the Biomedical Sciences*, Lecture Notes in Biomathematics, Vol. **6**, Heidelberg, Springer, 1975.

[4] H.T. Banks and H.T. Tran, *Mathematical and Experimental Modeling of Physical and Biological Processes*, CRC Press, Boca Raton London New York, 2009.

[5] H.T. Banks, Karyn L. Sutton, W. Clayton Thompson, Gennady Bocharov, Marie Doumic, Tim Schenkel, Jordi Argilaguet, Sandra Giest, Cristina Peligero, and Andreas Meyerhans, A new model for the estimation of cell proliferation dynamics using CFSE Data, CRSC-TR11-05, NCSU, August, 2011; *J. Immunological Methods,* **373** (2011), 143–160.

[6] H.T. Banks and W. Clayton Thompson, Mathematical models of dividing cell populations: Application to CFSE data, CRSC-TR12-10, N. C. State University, Raleigh, NC, April, 2012; *Journal on Mathematical Modelling of Natural Phenomena*, **7** (2012), 24–52. DOI: 10.1051/mmnp/20127504

[7] J.M. Berg, J.L. Tymoczko, L. Stryer, The Michaelis-Menten model accounts for the kinetic properties of many enzymes, Section 8.4 in *Biochemistry* **5th edition**, (2002). Available from: http://www.ncbi.nlm.nih.gov/books/NBK22430/

[8] George E. Briggs and John B. S. Haldane, A note on the kinetics of enzyme action, *Biochem J*, **19** (1925), 338–339.

[9] Kenneth P. Burnham and David R. Anderson, *Model Selection and Multimodal Inference* (2nd edition), Springer, New York, 2002.

[10] F. Carey, *Organic Chemistry, 8th edition*, McGraw-Hill, New York, 2011.

[11] William W. Chen, Mario Niepel and Peter K. Sorger, Classic and contemporary approaches to modeling biochemical reactions, *Genes & Development*, **24** (2010), 1861–1875.

[12] Gerda de Vries, et al., *A Course in Mathematical Biology*, SIAM Series on Mathematical Modeling and Computation, Vol. **MM12**, SIAM, Philadelphia, 2006.

[13] J. Hasenauer, D. Schittler and F. Allgöwer, Analysis and simulation of division- and label-structured population models: A new tool to analyze proliferation assays, *Bull. Math. Biol.*, **74** (2012), 2692–2732.

[14] M. Kot, *Elements of Mathematical Ecology*, Cambrodge University Press, Cambridge, UK, 2001.

[15] A. B. Lyons, Divided we stand: tracking cell proliferation with carboxyfluorescein diacetate succinimidyl ester, *Immunology and Cell Biology*, **77** (1999), 509–515.

[16] A. B. Lyons, J. Hasbold and P.D. Hodgkin, Flow cytometric analysis of cell division history using diluation of carboxyfluorescein diacetate succinimidyl ester, a stably integrated fluorescent probe, *Methods in Cell Biology*, **63** (2001), 375–398.

[17] A. B. Lyons and K. V. Doherty, Flow cytometric analysis of cell division by dye dilution, *Current Protocols in Cytometry*, (2004), 9.11.1-9.11.10.

[18] A.B. Lyons and C.R. Parish, Determination of lymphocyte division by flow cytometry, *J. Immunol. Methods*, **171** (1994), 131–137.

[19] C. Parish, Fluorescent dyes for lymphocyte migration and proliferation studies, *Immunology and Cell Biology* **77** (1999), 499–508

[20] S. I. Rubinow, *Introduction to Mathematical Biology*, J. Wiley & Sons, New York, 1975.

[21] W. Clayton Thompson, *Partial Differential Equation Modeling of Flow Cytometry Data from CFSE-based Proliferation Assays*, Ph.D. Dissertation, Dept. Mathematics, North Carolina State University, December, 2011.

[22] X. Wang, X. Duan, L. Liu, Y. Fang and Y. Tan, Carboxyfluorescein diacetate succinimidyl ester fluorescent dye for cell labeling, *Acta Biochimica et Biophysica Sinica* **37** (2005), 379–385.

Modeling Philosophy

H.T. Banks

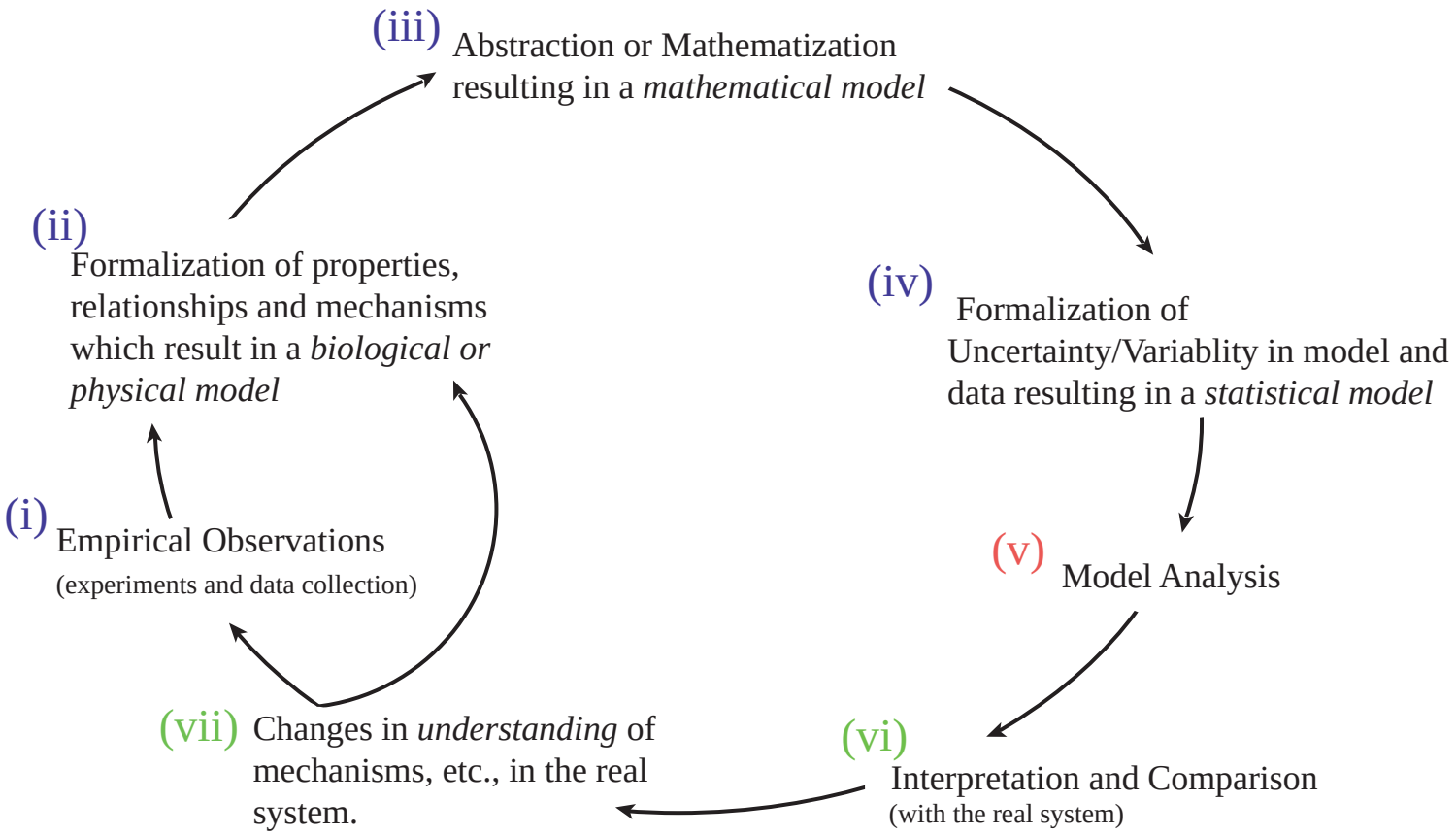
Center for Research in Scientific Computation
Center for Quantitative Sciences in Biomedicine
North Carolina State University

August 29, 2017

Modeling Philosophy

We give a brief discussion of certain philosophical notions that are important in the modeling of physical and biological systems. Modeling in our view is simply a means for providing a conceptual framework in which real systems may be investigated. The modeling process itself is (or should be) most often an iterative process: one can distinguish in it a number of rather separate steps that usually must be repeated. This iterative modeling process is schematically depicted in Figure 5. One begins with the real system under investigation and pursues the following sequence of steps:

The Iterative Modeling Process



Formation Stage: (i),(ii),(iii),(iv)	Solution Stage: (v)	Interpretation Stage: (vi), (vii)

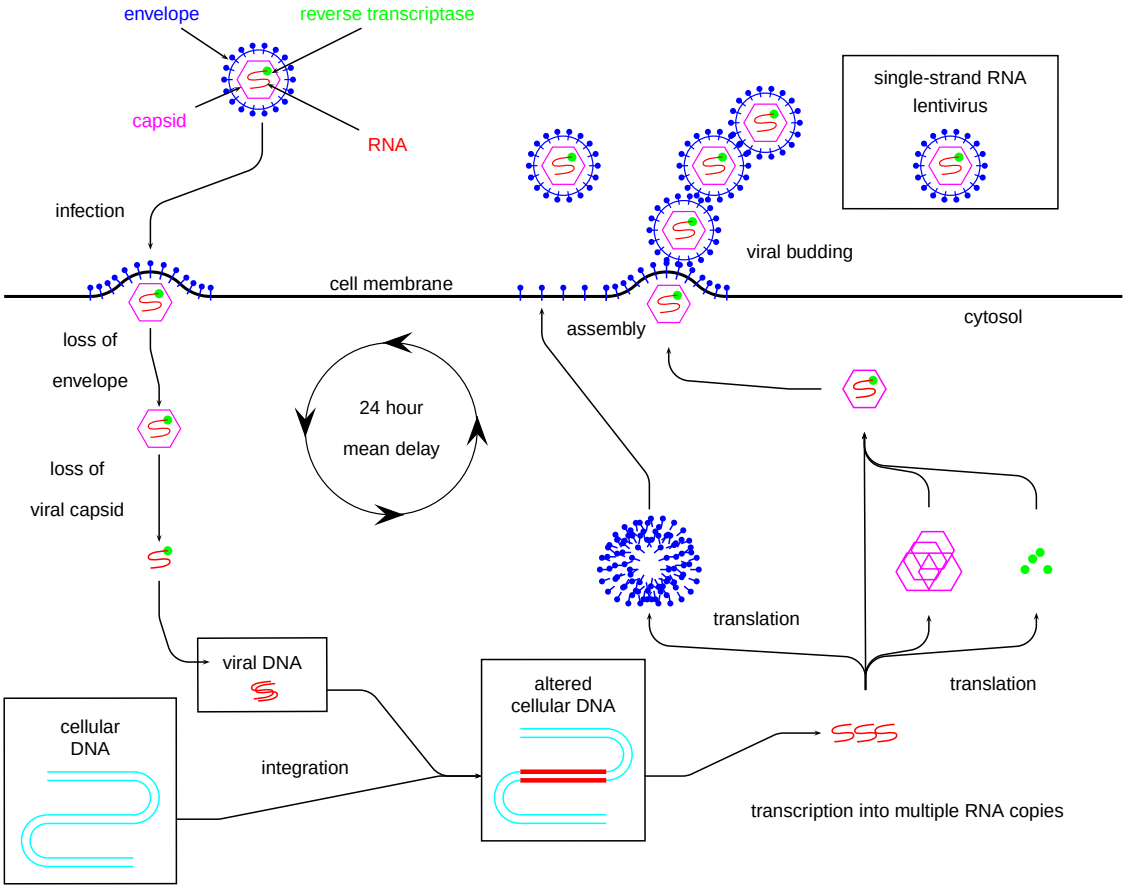


Figure 1.1: HIV Infection Pathway

15 HIV Models

Many models (elementary) in the early literature-early models were not predictive–
EXAMPLE:
The **Nowak-May model** is given by

$$\frac{dT}{dt} = s_1 - d_1 T - k_1 TV \tag{42}$$
$$\frac{dT_i}{dt} = k_1 TV - d_2 T^* \tag{43}$$
$$\frac{dV}{dt} = k_9 T^* - d_7 V \tag{44}$$

where T is the number of healthy T-cells at time t, T^* is the number of infected T-cells and V is the number of virus.

16 Introduction

Human Immunodeficiency Virus (HIV) is a retrovirus that infects T-helper cells of the immune system and is the causative agent for Acquired Immune Deficiency Syndrome (AIDS). HIV and AIDS are among the world's most serious public health concerns, affecting people of all demographics worldwide, with some regions impacted disproportionately. As of 2003, an estimated 38 million HIV-infected individuals are living worldwide, with approximately two-thirds in Africa, where 2.2 million people died from opportunistic infections related to AIDS in 2003 (UNAIDS 2004 Report on the Global HIV/AIDS Epidemic [31]). Despite many successful public health and clinical interventions since the first identification of HIV-positive patients in 1981, there remains no cure and the HIV/AIDS epidemic continues to grow.
Highly Active Antiretroviral Therapy (HAART), most commonly administered in the form of drug cocktails consisting of a protease inhibitor and at least one or more reverse transcriptase inhibitors, has been highly successful in suppressing HIV in many patients and therefore improving quality of life. However, contrary to dangerous popular myths, these drugs do not constitute a cure. While antiretroviral drugs are widely available in the United States and Western Europe, their cost and side effects may make their use challenging. In developing nations, UNAIDS estimates that only 7% of the infected population has access to HAART. Access to treatment for and education about this disease remain serious human rights issues around the world. Improved strategies are needed for efficient and appropriate use of drug therapy in both developed and underdeveloped countries.
Studies of the epidemiology of HIV and public health issues such as transmission (interhost dynamics) are important. Equally important to investigate are the effective use and improvement of antiretroviral drugs, which depend on understanding viral behavior

18 HIV Model and Inverse Problem Techniques

18.1 Model description

Many HIV models have been considered in the literature, including those surveyed in [35] and [48]. To demonstrate the potential predictive ability of such mathematical models, we employ the model developed in [26], subsequently modified in [29], and depicted in Figure 17; other models could be readily treated in our framework. The model compartments are denoted by variables T_1 (type 1 target cells, e.g., CD4$^+$ T-cells, cells/μl), T_2 (type 2 target cells, e.g., macrophages, cells/μl), V_I (infectious free virus, RNA copies/ml), V_{NI} (non-infectious free virus, RNA copies/ml), and E (cytotoxic T-lymphocytes, cells/μl). A superscript asterisk (*) denotes infected cells. The available clinical data include total CD4$^+$ T-cell count, represented by the sum $T_1 + T_1^*$, and total free virus, $V_I + V_{NI}$.

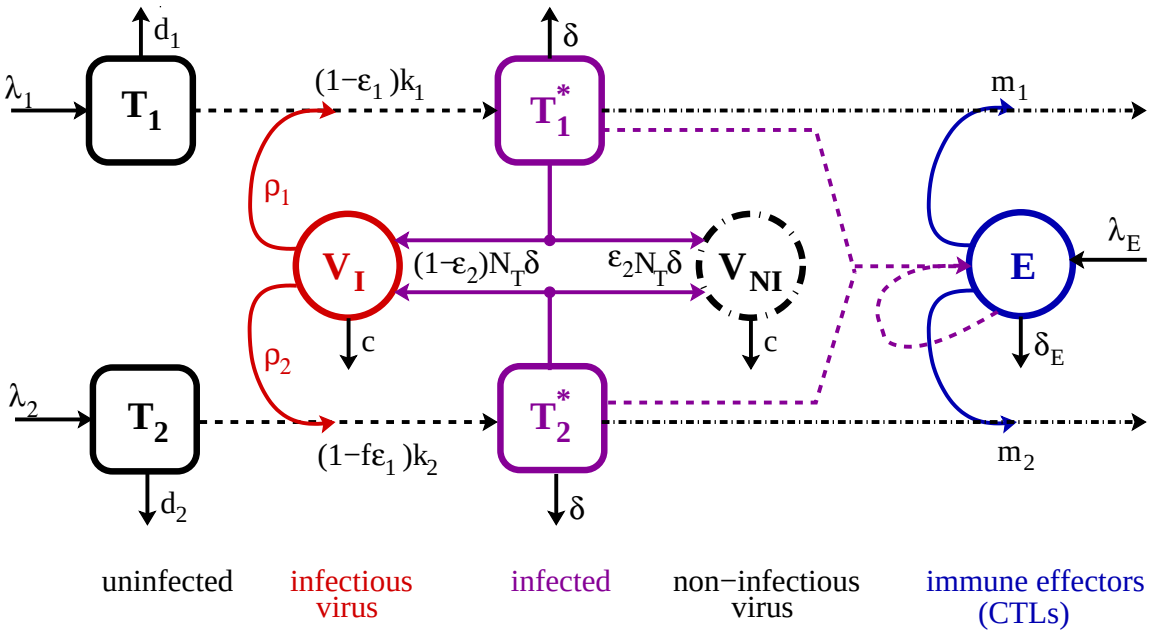


Figure 17: Schematic of compartmental HIV infection dynamics model. Only key pathways are indicated in the schematic – for further details, see the system of differential equations (45) below.

While the remaining compartments T_2, T_2^*, and E were not observed in the data used in this paper, they are important for modeling and predicting long-term longitudinal data. The presence of a secondary target cell population T_2 helps to satisfy a modeling requirement suggested by Callaway and Perelson [35] in their 2002 review paper: a reasonable model of HIV infection predicts a non-zero steady-state viral load, even in

the presence of effective drug therapy. Patients subjected to drug therapy often successfully suppress virus for a long time, potentially at undetectable levels. However, some reservoir or mechanism exists that almost invariably causes the virus to grow out to detectable levels upon removal of drug therapy. Hence one does not expect incorporation of drug therapy in the model, at a sensible efficacy, to drive the viral load to zero, but rather reduce it considerably, perhaps below the assay limits of quantification. One way to incorporate this is shown in Figure 17, where there are two co-circulating populations of target cells, potentially representing $CD4^+$ T-lymphocytes (T_1) and macrophages or other HIV-targeted cells (T_2). The two cell populations may have different activation requirements or susceptibility to drug therapy, represented by the different rate constants, thus potentially creating a non-zero, but low viral load steady state. This is crucial for modeling our long time horizon data, where patients may remain on treatment for an extended time. The differential efficacy also enables the model to exhibit reasonable sensitivity of the viral load equilibrium to treatment efficacy. For a survey of models and discussion of which exhibit reasonable sensitivity to drug efficacy, consult [35].
The documented importance of the immune system in responding to HIV infection (and especially its apparent crucial role during structured treatment interruptions) strongly motivates the inclusion of at least one model compartment representing immune response to the pathogen. We therefore include a measure E of cytotoxic T-lymphocyte (CTL) $CD8^+$ response to HIV infection. While the presently available data do not directly quantify the presence of HIV-specific CTLs, these immune responders are important for control of infected cells and may eventually be correlated to available epitope-challenge data. It is known that the immune response system is much more complicated than as represented in a single (composite) compartment denoted as CTL effectors E. Indeed, while present knowledge is incomplete, there are strong indications that a more complex modeling view of immune response involving naive and activated classes of $CD4^+$ and HIV-specific CD8+ cells as well as memory and latent reservoir classes will be important in understanding the chronic versus acute response of the immune system to HIV-1 infection [43, 46].

The corresponding compartmental ordinary differential equation (ODE) model for in-host HIV infection dynamics is given by (45). This model is essentially one suggested in Callaway–Perelson [35], but includes an immune response compartment and dynamics as suggested by Bonhoeffer, et. al. [34]. This compartment, denoted by E, represents

CTLs. The adapted system of ODEs is given by

$$\dot{T}_1 = \lambda_1 - d_1 T_1 - (1 - \bar{\epsilon}_1(t))\, k_1 V_I T_1 \tag{45a}$$

$$\dot{T}_2 = \lambda_2 - d_2 T_2 - (1 - f\bar{\epsilon}_1(t)) k_2 V_I T_2 \tag{45b}$$

$$\dot{T}_1^* = (1 - \bar{\epsilon}_1(t)) k_1 V_I T_1 - \delta T_1^* - m_1 E T_1^* \tag{45c}$$

$$\dot{T}_2^* = (1 - f\bar{\epsilon}_1(t)) k_2 V_I T_2 - \delta T_2^* - m_2 E T_2^* \tag{45d}$$

$$\begin{aligned} \dot{V}_I = {} & (1 - \bar{\epsilon}_2(t)) 10^3 N_T \delta (T_1^* + T_2^*) - c V_I \\ & - (1 - \bar{\epsilon}_1(t)) 10^3 k_1 T_1 V_I - (1 - f\bar{\epsilon}_1(t)) 10^3 k_2 T_2 V_I \end{aligned} \tag{45e}$$

$$\dot{V}_{NI} = \bar{\epsilon}_2(t) 10^3 N_T \delta (T_1^* + T_2^*) - c V_{NI} \tag{45f}$$

$$\dot{E} = \lambda_E + \frac{b_E (T_1^* + T_2^*)}{(T_1^* + T_2^*) + K_b} E - \frac{d_E (T_1^* + T_2^*)}{(T_1^* + T_2^*) + K_d} E - \delta_E E, \tag{45g}$$

together with an initial condition vector

$$(T_1(0), T_1^*(0), T_2(0), T_2^*(0), V_I(0), V_{NI}(0), E(0))^T .$$

Here the factors 10^3 are introduced to convert between microliter and milliliter scales, preserving the units from some of the published papers.

As is common in models of HIV infection, infected cells T_i^* result from encounters between uninfected target cells T_i and infectious free virus V_I in a well-mixed environment. As noted above, this model involves two co-circulating populations of target cells, perhaps representing $CD4^+$ T-lymphocytes (T_1) and macrophages (T_2). The natural infection rate k_i may differ between the two populations, which could account for suspected differences in activation rates between lymphocytes and macrophages. The treatment factor $\bar{\epsilon}_1(t)$, described further below, represents a reverse transcriptase inhibitor (RTI) that blocks new infections and is potentially more effective in population 1 (T_1, T_1^*) than in population 2 (T_2, T_2^*), where the efficacy is $f\bar{\epsilon}_1$, with $f \in [0, 1]$. The differences in infection rates and treatment efficacy help create a low, but non-zero, infected cell steady state for T_2^*, which is commensurate with the idea that macrophages may be an important source of virus after T-cell depletion. The populations of uninfected target cells T_1 and T_2 may have different source rates λ_i and natural death rates d_i.

Free virus particles are produced by both types of infected cells, which we assume produce virus at the same rate (again this could be readily generalized to account for

(i) empirical observations, experiments, and data collection;

(ii) formalization of properties, relationships and mechanisms that result in a *biological or physical model* (e.g., stoichiometric relations detailing pathways, mechanisms, biochemical reactions, etc., in a metabolic pathway model; stress-strain, pressure-force relationships in mechanics and fluids);

(iii) abstraction or mathematization resulting in a *mathematical model* (e.g., algebraic and/or differential equations with constraints and initial and/or boundary conditions);

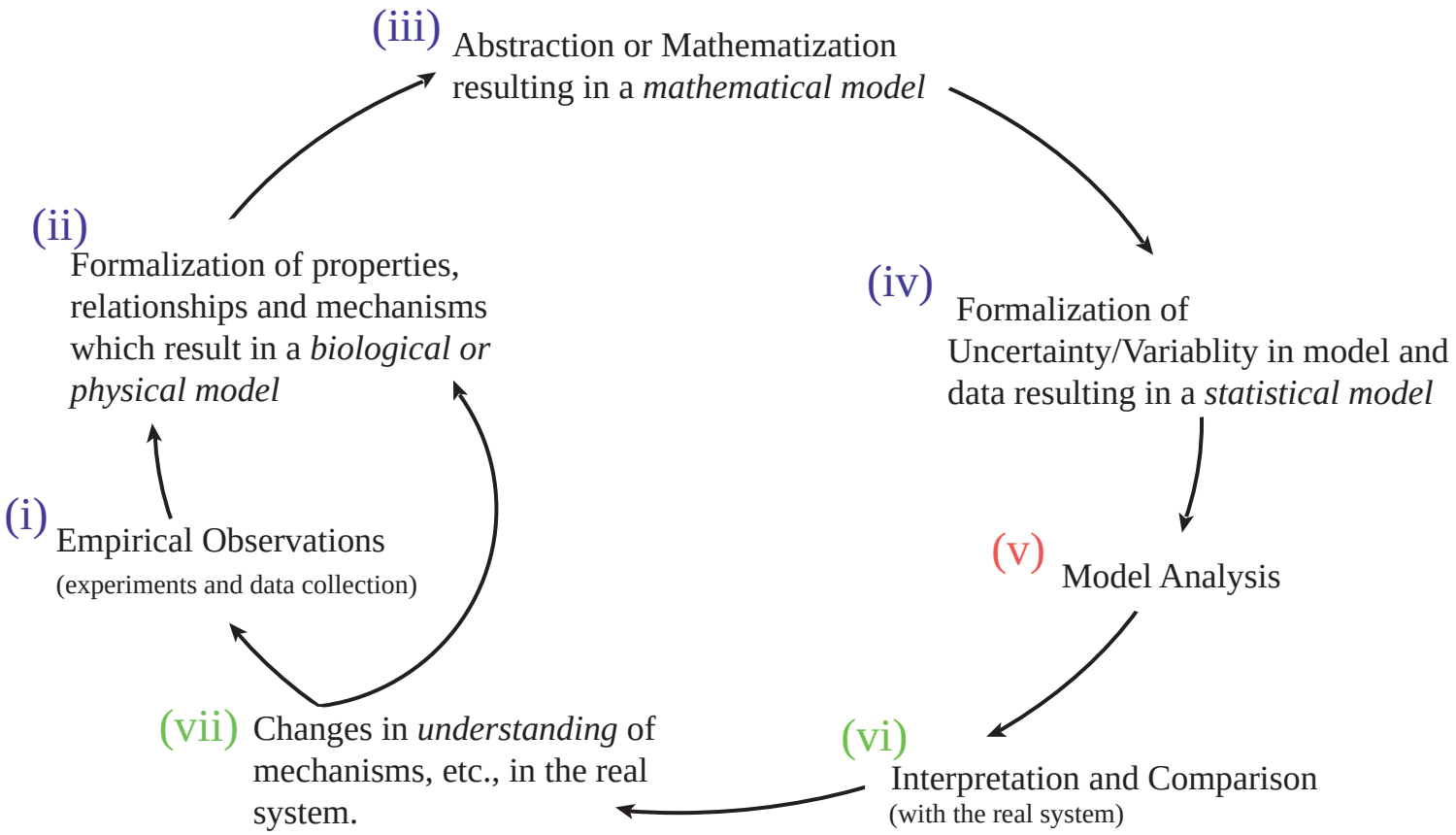
(iv) formalization of uncertainty/variability in model and data resulting in a *statistical model* (this usually involves basic assumptions about errors in modeling, observation process/measurement, etc.);

(v) model analysis that can consist of simulation studies, analytical and qualitative analysis including stability analysis, and use of mathematical techniques such as perturbation studies, parameter estimation (inverse problems) data fitting, statistical analysis;

(vi) interpretation and comparison (with the real system) of the conclusions, predictions and conjectures obtained from step (v);

(vii) changes in "understanding" of mechanisms, pathways, etc., in the real system.

The Iterative Modeling Process



Formation Stage: (i),(ii),(iii),(iv)	Solution Stage: (v)	Interpretation Stage: (vi), (vii)

Let us turn next to the **reasons frequently given for modeling**. Perhaps the one most often offered is ***simplification***: the use of models makes possible the investigation of very complex systems in a systematic manner. A second rationale is ***ease in manipulation***: investigations involving separation of subunits and hypothesis testing may often be facilitated through use of simulations in place of experimentation. The suggestive features in modeling can also help in ***formulation of hypotheses*** and in the ***design of critical experiments***. The modeling process also requires ***preciseness*** in investigation in that one must move from a **general, verbal explanation of phenomena** to a **specific, quantitative one**.

But a rationale perhaps more fundamental than any of these is that modeling leads to an ***organization*** of inquiry in that it tends

to ***polarize one's thinking and aid in posing basic questions*** concerning what one does and does not know for certain about the real system. Whatever the reasons that have been advanced to justify modeling attempts, it is sufficient perhaps to note that the primary goal must be *enlightenment*, that is, ***to gain a better understanding of the real system***, and the success or lack thereof of any modeling attempt must be appraised with this in mind.

One must recognize the various ***levels*** or ***multi-scale*** aspects of modeling in any attempt to compare or assess the validity of several models for a phenomenon. For example, consider the phenomena involved in the transmission of a nerve impulse along an axon: this process is likely to be described by the mathematician or biophysicist in terms of partial differential equations, wave phenomena, or transmission line analogies, whereas a neurophysiologist might speak in terms of local circuit analogies and changes in conductances. The cell physiologist might describe the phenomena in the context of transport properties of membranes and ion flow, while the molecular biochemist could insist that the real story lay in the theory of molecular binding.

The choice of the level (micro vs. macro) at which one models depends very much upon the training and background of the investigator. Furthermore, the perception of whether a model is a "good" one or not is also greatly influenced by this factor, and it is therefore not surprising that all of the approaches to the nerve impulse phenomena mentioned above (or indeed those for modeling any physical or biological phenomena) can be subjected to valid criticisms in any attempt to evaluate them.

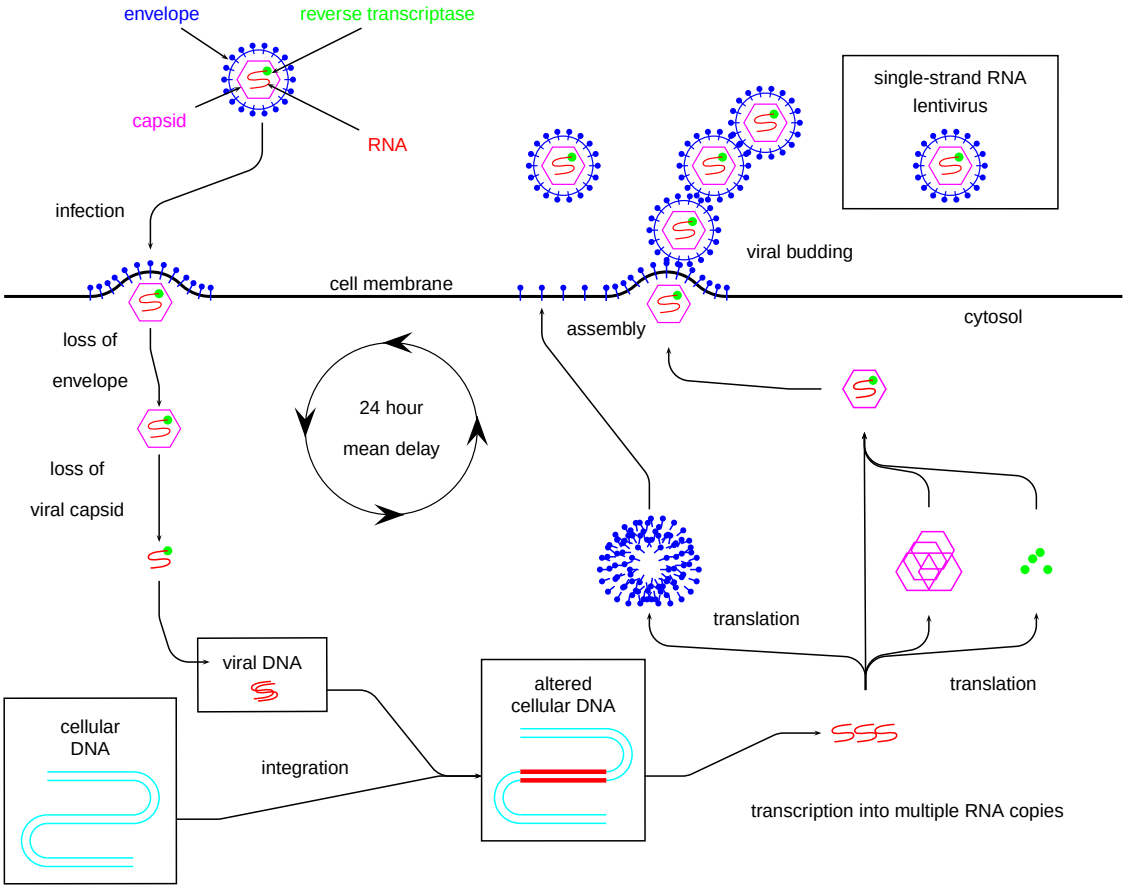


Figure 1.1: HIV Infection Pathway

10 HIV Models

The slides for this lecture are located here: http://courses.ncsu.edu/ma493/lec/003/HIV-Jan06-Cellular%20Level.pdf
These slides discuss the basis for the HIV system.

The mathematical model is a vector system

$$\dot{x}(t) = g(t, x(t), \theta),\ x(0) = x_0,\ x \in \mathbf{R}^n,\ \theta \in \mathbf{R}^p$$

To start treating the data, you have to have a statistical model

$$Y(t) = \mathcal{C}x(t, \theta) + \varepsilon_j$$

where $\mathcal{C}$ is the observation operation and ε_j is the error measurement on the jth data point.

10.1 Cellular Level

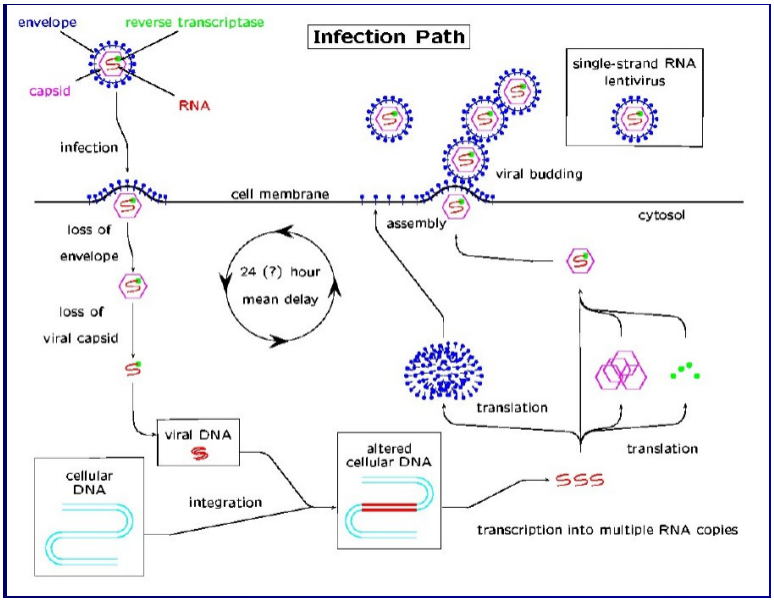


Figure 6: Overview of cellular model for HIV

This model has a probability distribution delay to account for variability across cells.

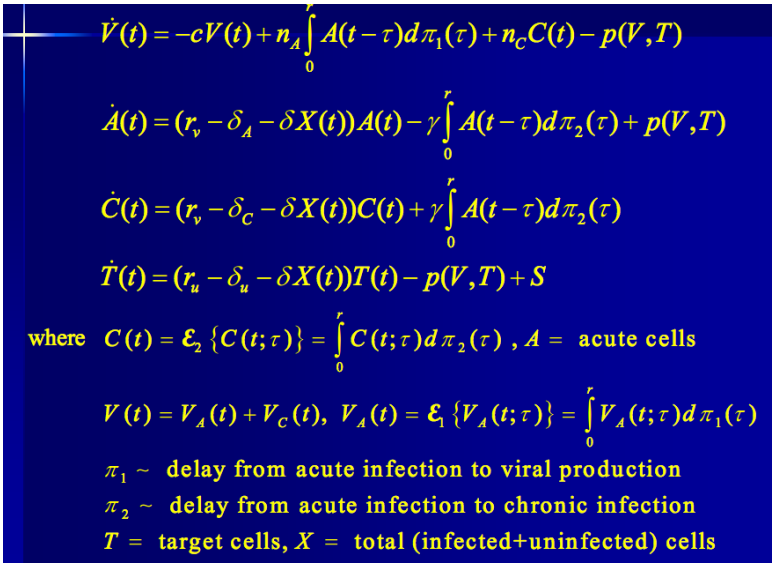
$$\frac{dV}{dt} = -cV(t) + n_1 \int_0^\infty A(t-\tau)k(\tau)d\tau + n_c C(t) - n_{vt}V(t)T(t)$$

where τ is production delay. (The rest of the variables still need to be defined...)
The system of equations describing this model is as follows: Note that the data that you get is the following $y_i^d \approx (T + A + C)(t_i) + \varepsilon_i$. Thus the observation is

$$y = \begin{bmatrix} 0 & 1 & 1 & 1 \end{bmatrix} X$$

where $X = \begin{bmatrix} V & T & A & C \end{bmatrix}'$.
Thus

$$Y(t) = \mathcal{C}X(t) + \varepsilon(t)$$



10.2 Population Level

11 Probability Review

Recall that

$$Y_i(t) = x(t_i, \theta) + \varepsilon_i$$

really just means

$$\text{“data”} = \text{“model”} + \text{“error”}$$

Notes for this lecture taken from http://courses.ncsu.edu/ma493/lec/003/Lipari_Prob_Review.pdf

- Sample Space
- Events
- Random Variables (slide 9)
- pdf and cdf (slide 10)
- Expected Value of a RV (slide 13)
- Variance (slide 14)
- Distributions (slide 21)
  - Uniform distribution
  - Normal distribution
  - Student's t distribution (slide 37)
- Correlation (slide 15) and independence (slide 17)

15 HIV Models

Many models (elementary) in the early literature-early models were not predictive–
EXAMPLE:
The **Nowak-May model** is given by

$$\frac{dT}{dt} = s_1 - d_1 T - k_1 TV \tag{42}$$
$$\frac{dT_i}{dt} = k_1 TV - d_2 T^* \tag{43}$$
$$\frac{dV}{dt} = k_9 T^* - d_7 V \tag{44}$$

where T is the number of healthy T-cells at time t, T^* is the number of infected T-cells and V is the number of virus.

16 Introduction

Human Immunodeficiency Virus (HIV) is a retrovirus that infects T-helper cells of the immune system and is the causative agent for Acquired Immune Deficiency Syndrome (AIDS). HIV and AIDS are among the world's most serious public health concerns, affecting people of all demographics worldwide, with some regions impacted disproportionately. As of 2003, an estimated 38 million HIV-infected individuals are living worldwide, with approximately two-thirds in Africa, where 2.2 million people died from opportunistic infections related to AIDS in 2003 (UNAIDS 2004 Report on the Global HIV/AIDS Epidemic [31]). Despite many successful public health and clinical interventions since the first identification of HIV-positive patients in 1981, there remains no cure and the HIV/AIDS epidemic continues to grow.
Highly Active Antiretroviral Therapy (HAART), most commonly administered in the form of drug cocktails consisting of a protease inhibitor and at least one or more reverse transcriptase inhibitors, has been highly successful in suppressing HIV in many patients and therefore improving quality of life. However, contrary to dangerous popular myths, these drugs do not constitute a cure. While antiretroviral drugs are widely available in the United States and Western Europe, their cost and side effects may make their use challenging. In developing nations, UNAIDS estimates that only 7% of the infected population has access to HAART. Access to treatment for and education about this disease remain serious human rights issues around the world. Improved strategies are needed for efficient and appropriate use of drug therapy in both developed and underdeveloped countries.
Studies of the epidemiology of HIV and public health issues such as transmission (interhost dynamics) are important. Equally important to investigate are the effective use and improvement of antiretroviral drugs, which depend on understanding viral behavior

18 HIV Model and Inverse Problem Techniques

18.1 Model description

Many HIV models have been considered in the literature, including those surveyed in [35] and [48]. To demonstrate the potential predictive ability of such mathematical models, we employ the model developed in [26], subsequently modified in [29], and depicted in Figure 17; other models could be readily treated in our framework. The model compartments are denoted by variables T_1 (type 1 target cells, e.g., CD4$^+$ T-cells, cells/μl), T_2 (type 2 target cells, e.g., macrophages, cells/μl), V_I (infectious free virus, RNA copies/ml), V_{NI} (non-infectious free virus, RNA copies/ml), and E (cytotoxic T-lymphocytes, cells/μl). A superscript asterisk (*) denotes infected cells. The available clinical data include total CD4$^+$ T-cell count, represented by the sum $T_1 + T_1^*$, and total free virus, $V_I + V_{NI}$.

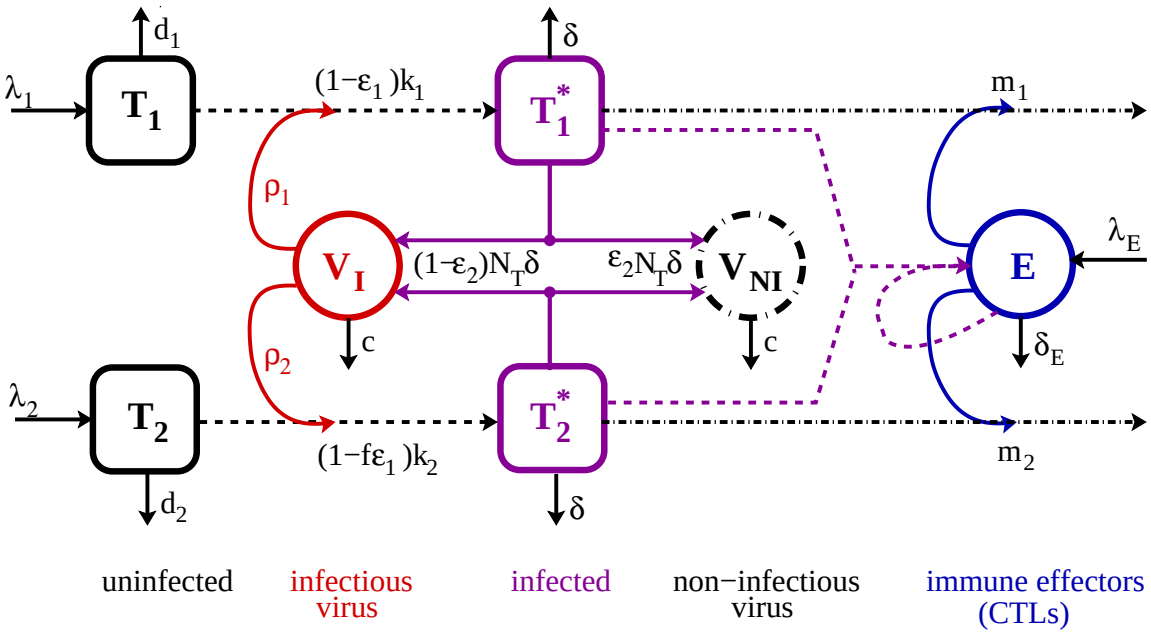


Figure 17: Schematic of compartmental HIV infection dynamics model. Only key pathways are indicated in the schematic – for further details, see the system of differential equations (45) below.

While the remaining compartments T_2, T_2^*, and E were not observed in the data used in this paper, they are important for modeling and predicting long-term longitudinal data. The presence of a secondary target cell population T_2 helps to satisfy a modeling requirement suggested by Callaway and Perelson [35] in their 2002 review paper: a reasonable model of HIV infection predicts a non-zero steady-state viral load, even in

CTLs. The adapted system of ODEs is given by

$$\dot{T}_1 = \lambda_1 - d_1 T_1 - (1 - \bar{\epsilon}_1(t))\, k_1 V_I T_1 \tag{45a}$$

$$\dot{T}_2 = \lambda_2 - d_2 T_2 - (1 - f\bar{\epsilon}_1(t)) k_2 V_I T_2 \tag{45b}$$

$$\dot{T}_1^* = (1 - \bar{\epsilon}_1(t)) k_1 V_I T_1 - \delta T_1^* - m_1 E T_1^* \tag{45c}$$

$$\dot{T}_2^* = (1 - f\bar{\epsilon}_1(t)) k_2 V_I T_2 - \delta T_2^* - m_2 E T_2^* \tag{45d}$$

$$\begin{aligned} \dot{V}_I = {} & (1 - \bar{\epsilon}_2(t)) 10^3 N_T \delta (T_1^* + T_2^*) - c V_I \\ & - (1 - \bar{\epsilon}_1(t)) 10^3 k_1 T_1 V_I - (1 - f\bar{\epsilon}_1(t)) 10^3 k_2 T_2 V_I \end{aligned} \tag{45e}$$

$$\dot{V}_{NI} = \bar{\epsilon}_2(t) 10^3 N_T \delta (T_1^* + T_2^*) - c V_{NI} \tag{45f}$$

$$\dot{E} = \lambda_E + \frac{b_E (T_1^* + T_2^*)}{(T_1^* + T_2^*) + K_b} E - \frac{d_E (T_1^* + T_2^*)}{(T_1^* + T_2^*) + K_d} E - \delta_E E, \tag{45g}$$

together with an initial condition vector

$$(T_1(0), T_1^*(0), T_2(0), T_2^*(0), V_I(0), V_{NI}(0), E(0))^T .$$

Here the factors 10^3 are introduced to convert between microliter and milliliter scales, preserving the units from some of the published papers.

As is common in models of HIV infection, infected cells T_i^* result from encounters between uninfected target cells T_i and infectious free virus V_I in a well-mixed environment. As noted above, this model involves two co-circulating populations of target cells, perhaps representing $CD4^+$ T-lymphocytes (T_1) and macrophages (T_2). The natural infection rate k_i may differ between the two populations, which could account for suspected differences in activation rates between lymphocytes and macrophages. The treatment factor $\bar{\epsilon}_1(t)$, described further below, represents a reverse transcriptase inhibitor (RTI) that blocks new infections and is potentially more effective in population 1 (T_1, T_1^*) than in population 2 (T_2, T_2^*), where the efficacy is $f\bar{\epsilon}_1$, with $f \in [0, 1]$. The differences in infection rates and treatment efficacy help create a low, but non-zero, infected cell steady state for T_2^*, which is commensurate with the idea that macrophages may be an important source of virus after T-cell depletion. The populations of uninfected target cells T_1 and T_2 may have different source rates λ_i and natural death rates d_i.

Free virus particles are produced by both types of infected cells, which we assume produce virus at the same rate (again this could be readily generalized to account for

Common difficulties and limitations:

(a) Availability and accuracy of data;

(b) Analysis of the mathematical model;

(c) Use of local representations that are invalid for the overall system;

(d) Obsession with the solution stage;

(e) Assumption that the "model" is the real system;

(f) Communication in interdisciplinary efforts.

Finally, we turn to the question of how one appraises a specific modeling attempt. There are a number of criteria that one might use. Among those proposed by various authors are the suggestions that a good model should:

- fit data accurately;
- be theoretically consistent with the real system;
- have parameters with physical meaning that can be measured independently of each other;
- prove useful in prediction;
- not so much explain or predict, but organize and economize thinking;
- pose new empirical questions and help answer them through the iterative process;

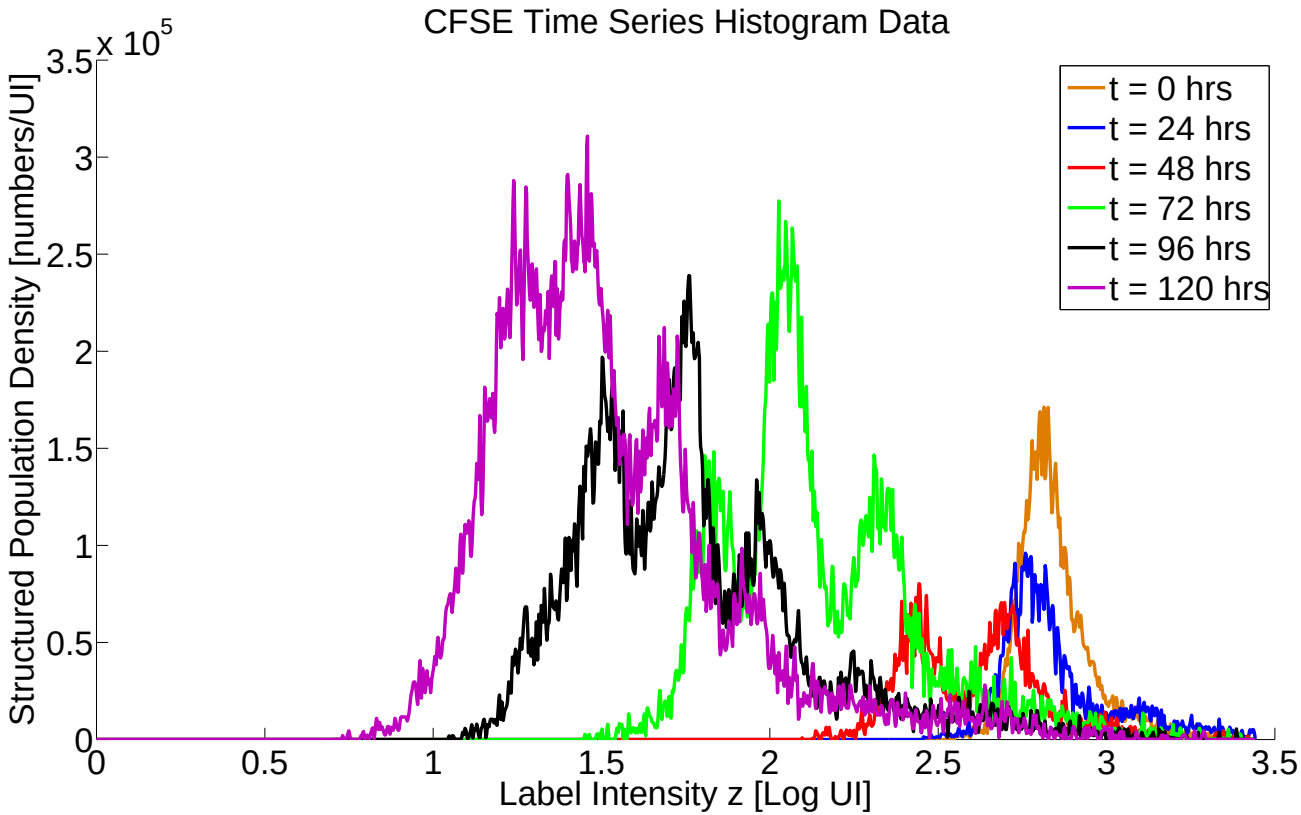
- help us understand the phenomena it represents and think comfortably about them;
- point to inadequacies in some way of available data.

It is clear, though, that for a modeling investigation to be deemed a success, it must have **enhanced our overall knowledge and understanding of the phenomena in question.** As one of our students (having been attacked by other students for some rather unorthodox and, at the time, unsupported hypothesis about mechanisms) noted in defending his efforts, "We learn little indeed if the models we build never stretch our understanding, but only tell us what we already feel is safely known." We remind the reader of the often quoted truth "all models are incorrect, but some are more useful than others".

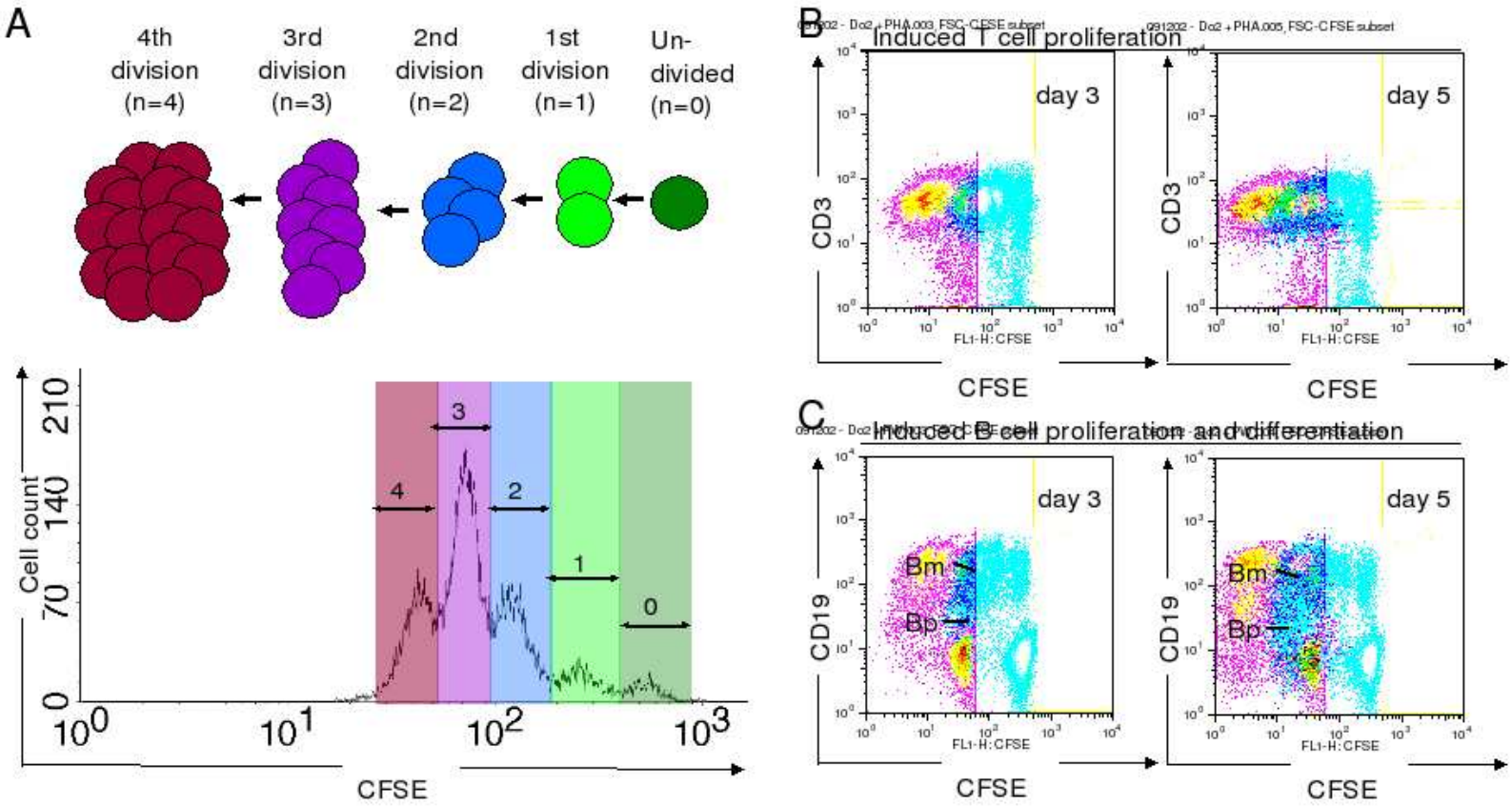
MODELS USUALLY BASED ON CONSERVATION LAWS

- **Force and Momentum Balance (mechanics, physical models)**
- **Mass Balance (biological, chemical)**
- **Energy Balance (thermal)**

CFSE Data Set



Data Overview

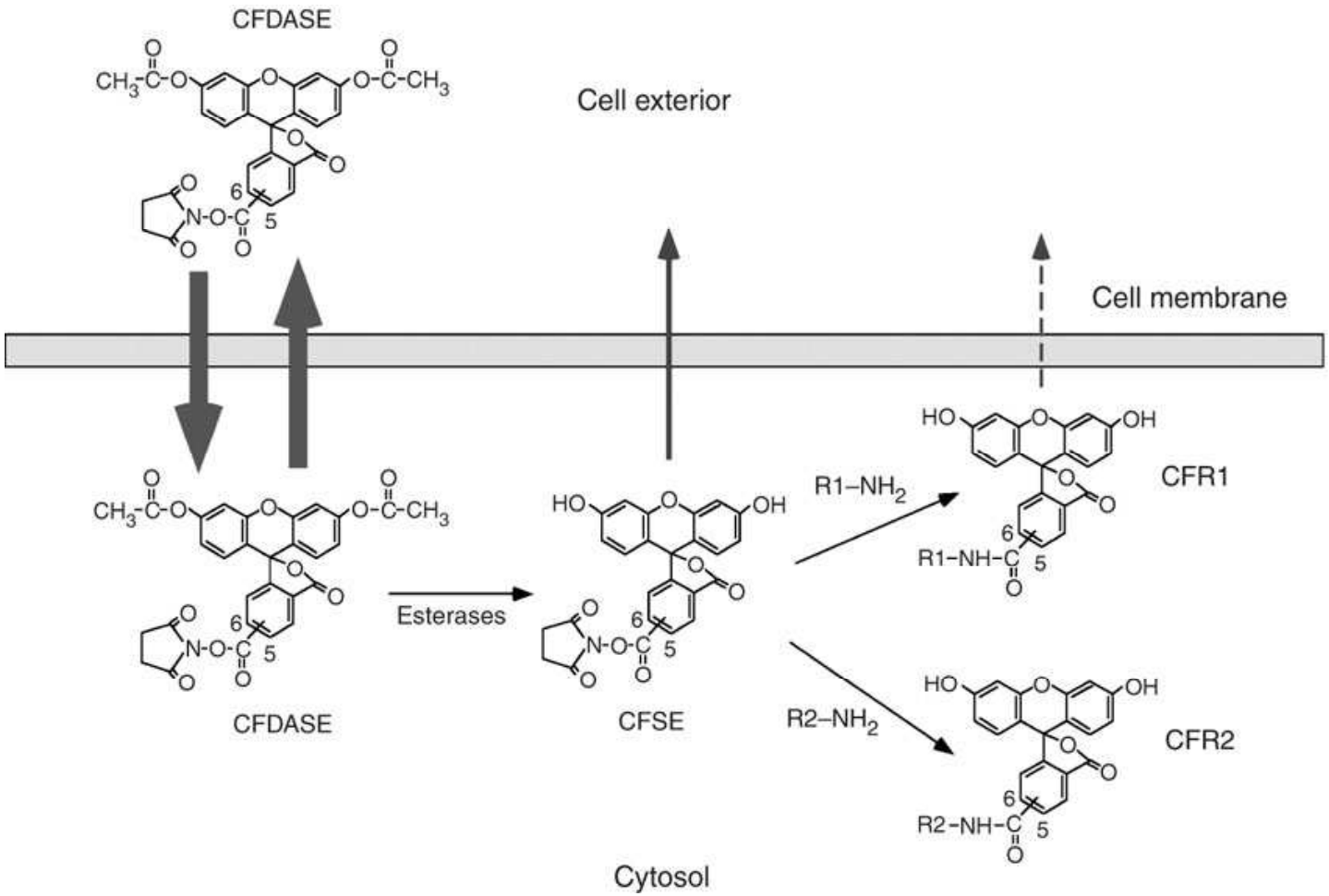


(A. Meyerhans)

CFSE Labeling (Lyons and Parish, 1994)

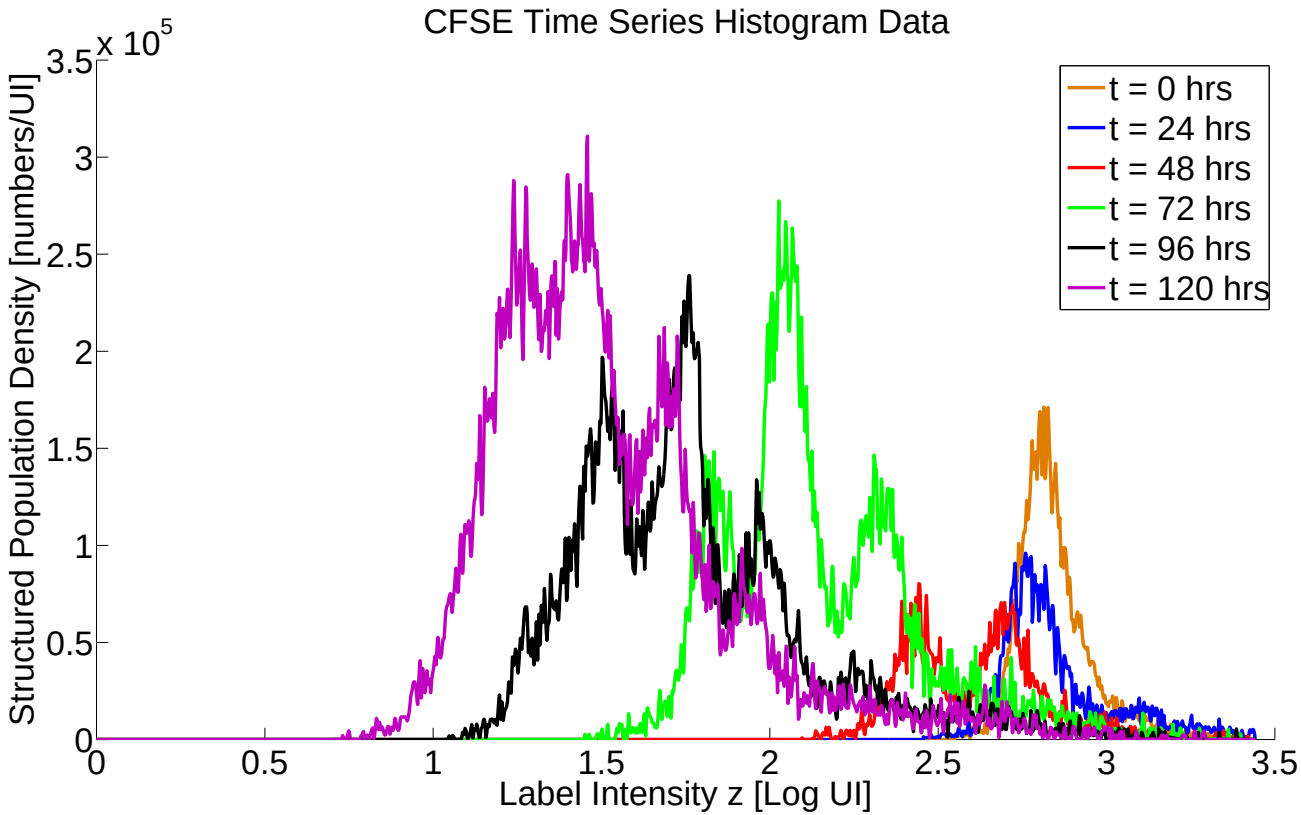
- Cells cultured with CFDA-SE (carboxyfluorescein diacetate succinimidyl ester) then washed
- CFDA-SE becomes protein-bound and fluorescent CFSE (the fluorescent dye carboxyfluorescein succinimidyl ester)
- Dye split between daughter cells at division
- Dye naturally turns over/degrades (very slowly)
- Fluorescence Intensity (FI) of CFSE measured via flow cytometry
- FI linear with dye concentration $\Rightarrow$ FI $\propto$ mass
- Several advantages over other dyes/techniques

CFSE Labeling (Lyons and Parish, 1994)



(C. Parish, Fluorescent dyes for lymphocyte migration and proliferation studies, *Immunology and Cell Biol.* **77** (1999), 499–508.)

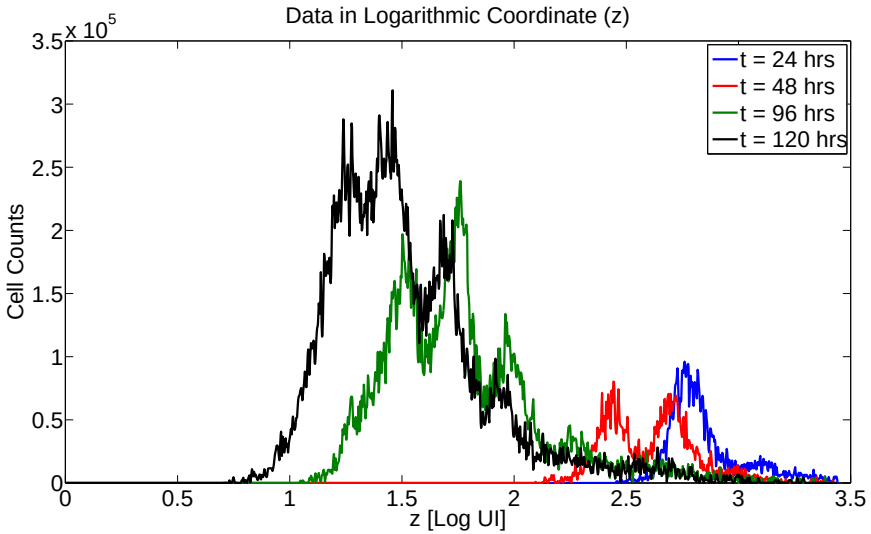
CFSE Data Set



Label-Structured Model (cont'd)

This model must account for (Luzyanina et al., 2007):

- Slow decay of CFSE FI over time
- Dilution of CFSE as cells divide
- Asynchronous division times



Goals of Modeling

- Cellular ‘Dynamic Responsiveness’
- Link cell counts with proliferation/death rates
  - Population doubling time
  - Cell viability
  - Biological descriptors (cell cycle time, etc.)
- Uncertainty Quantification...
  - ... in the experimental procedure
  - ... for estimated rates/etc
- Analyze cell differentiation and division-linked changes
- Investigate immunospecific extracellular signaling pathways
- Comparison among donors/cell types/disease progression

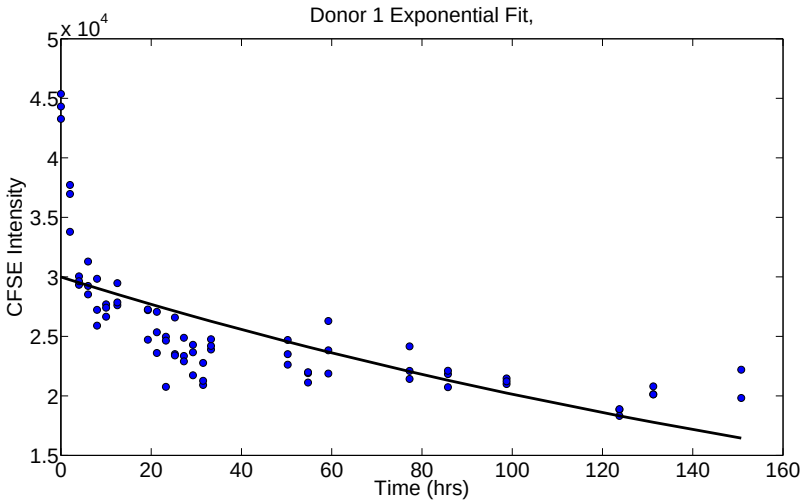
Initial PDE Model

- Structured density $n(t,x)$ (cells/UI)
- (Exponential) Proliferation rate $\alpha(x)$
- (Exponential) Death rate $\beta(x)$

$$\frac{\partial n}{\partial t}(t,x) + \frac{\partial [v(x)n(t,x)]}{\partial x} = -(\alpha(x)+\beta(x))n(t,x) + \chi_{[x_{\min}, x_{\max}/\gamma]} 2\gamma\alpha(\gamma x)n(t,x)$$

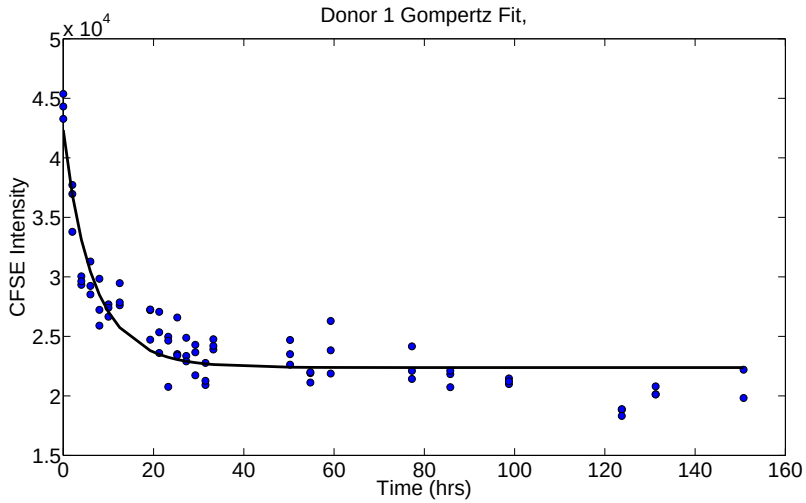
- Assume exponential loss for natural decay of label, $v(x) = \frac{dx}{dt} = -cx$
- Assume even partitioning of CFSE upon division
  $\gamma \equiv$ mother-to-daughter CFSE ratio

‘Biphasic Decay’



$$\frac{dx}{dt} = v(x) = c(x - x_a)$$

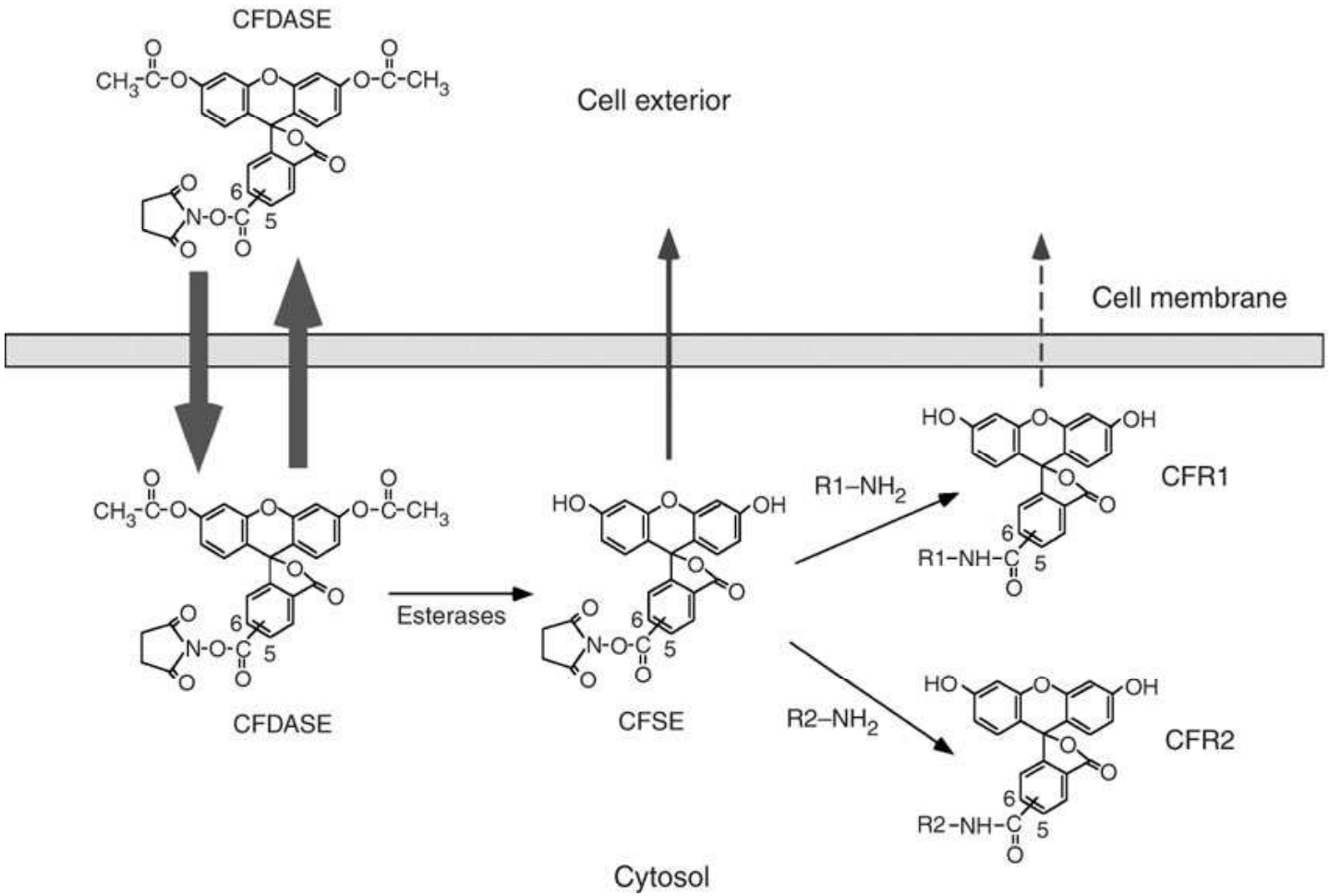
Exponential



$$\frac{dx}{dt} = v(t,x) = c(x - x_a)e^{-kt}$$

Gompertz

CFSE Labeling (Lyons and Parish, 1994)



(C. Parish, Fluorescent dyes for lymphocyte migration and proliferation studies, *Immunology and Cell Biol.* **77** (1999), 499–508.)

Given the random variable Y_j^d of observations, the ordinary least squares technique can be used as an estimator for $\vec{\theta_0}$ by minimizing

$$J(\vec{\theta}) = \sum_{j=1}^{n} |Y_j^d - f(t_j, \vec{\theta})|^2 \tag{12}$$

where $f(t_j, \vec{\theta})$ is the model solution at time t_j with parameter $\vec{\theta} = \big(k_1, k_2, \cdots, k_n\big)$. Thus, we have an estimator, $\vec{\theta}_{OLS}$, which we are using in hopes of finding an accurate estimate to fit to our data.

$$\vec{\theta}_{OLS} = \arg\min \sum_{j=1}^{n} |Y_j^d - f(t_j, \vec{\theta})|^2 \tag{13}$$

The data set shown in Figure 1, $\{y_j\}_{j=1}^n$, is then assumed to be a realization of the random variable (11). We can implement OLS in (13) to find an estimate, $\hat{\theta}_{OLS}$, which best fits our mathematical model to this data set, i.e.,

$$\hat{\theta}_{OLS} = \arg\min \sum_{j=1}^{n} |y_j - f(t_j, \vec{\theta})|^2. \tag{14}$$

4.2 Biological Model 1

Due in part to the presence of two acetate esters in its structure, CFDA-SE has a high lipophilicity which allows it to passively diffuse across cell membranes [19], suggesting that both an inflow rate and an outflow rate should be accounted for. However, the data set we are examining is the result of a particular procedure where, after initial exposure to CFDA-SE, the cell culture was flushed with water, eliminating any excess label [21]. Thus, the only source of CFDA-SE inflow would be re-entering CFDA-SE, which the data suggests is insignificant. Therefore, we considered it reasonable to assume that there is no continuing flow of CFDA-SE into the cell, but rather that all CFDA-SE is present inside the cell at the start of the procedure, as depicted in Figure 3.

Once inside the cell, CFDA-SE reacts with intracellular esterases, resulting in the formation of the highly fluorescent carboxyfluorescein succinimidyl ester (CFSE). At this point, our knowledge of the reaction is limited. Nonetheless, we know a great deal about the structure of the product, CFSE. The structure of CFSE lacks the two acetate esters present in CFDA-SE [19, 21, 22]. This absence decreases the lipophilicity of CFSE and renders it less membrane permeable. As before, this knowledge necessitates both inflow and outflow rates, but the data suggests that the mass of re-entering CFSE is insignificant in comparison to the mass of CFSE leaving the cell. Therefore, only the outflow of CFSE from the cell is assumed in our model. Additionally, the succinimidyl ester present in the structure of CFDA-SE is also present in CFSE. This succinimidyl moiety of CFSE is highly reactive with amino groups and can covalently couple 5-6-carboxyfluorescein (CF) to intracellular molecules [19]. Our knowledge of this reaction and its products is currently limited, but we are aware that two types of coupling can occur, yielding two types of products. One type of coupling occurs when CF is bound to a type of intracellular molecule, which we arbitrarily call R1-NH_2 that results in the conjugate CF-R1 which is unstable and quickly exits the cell or is degraded. The second type of coupling occurs when CF is bound to a type of long lived intracellular molecule, which we arbitrarily call R2-NH_2 that results in the conjugate CF-R2 that is stable and essentially membrane impermeable, maintaining a fluorescent label [19]. A initial schematic of this process within a cell is depicted in Figure 3 and is designated as *Biological Model 1.* A list of related state variables and kinetic parameters for this model and the subsequent models used below are given in Table 2.

Parameter	Description
k_1	Rate of decay of CFDA-SE
k_2	Rate of conversion of CFDA-SE to CFSE
v_2	V_{max} for Michaelis-Menten reaction between CFDA-SE and CFSE
K_2	k_m for Michaelis-Menten reaction between CFDA-SE and CFSE
k_3	Rate of decay of CFSE
k_4	Rate of uniform conversion of CFSE to CF-R1 and CF-R2
f	Fraction of CFSE that converts to CF-R1
$k_{4,1}$	Rate of conversion of CFSE to CF-R1
$v_{4,1}$	V_{max} for Michaelis-Menten reaction between CFSE and CF-R1
$K_{4,1}$	k_m for Michaeli-Menten reaction between CFSE and CF-R1
$k_{4,2}$	Rate of conversion of CFSE to CF-R2
$v_{4,2}$	V_{max} for Michaelis-Menten reaction between CFSE and CF-R2
$K_{4,2}$	k_m for Michaelis-Menten reaction between CFSE and CF-R2
k_5	Rate of decay of CF-R1
k_6	Rate of decay of CF-R2
k_7	Rate of conversion of CFSE to general CF-R term
v_7	V_{max} for Michaelis-Menten reaction between CFSE and general CF-R term
K_7	k_m for Michaelis-Menten reaction between CFSE and general CF-R term
k_8	Rate of decay of general CF-R term
x_0	Initial fluorescence
c	Determines how much initial fluorescence CF-R1 & CF-R2 receive in Model 5.1
α	Rate of decay in Model 6.2 and 6.3
K	Carrying capacity (lower asymptote) for total fluorescence in Model 6.2 and 6.3
Component	**Description**
x_1	CFDA-SE
x_2	CFSE
x_3	CF-R1
x_4	CF-R2
x_5	general CF-R term (CF-R1 + CF-R2)

Table 2: List of variables used in the paper.

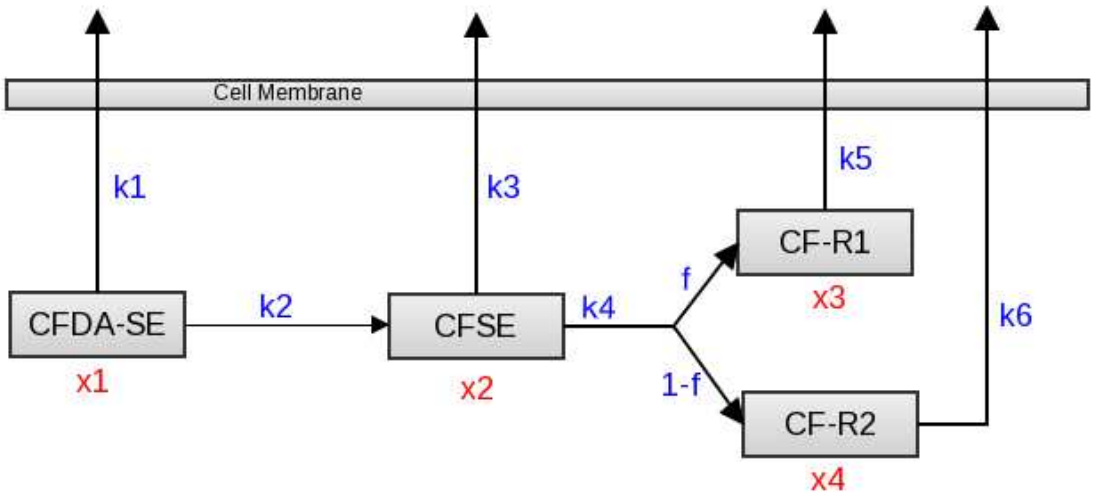


Figure 3: A schematic of *Biological Model 1*, representing the serial dilution process of the intracellular dye within a cell.

4.2.1 Model 1.1: Exponential Decay

Based on the law of mass action, the following system of equations was created to model the conversion and decay of CFDA-SE in a cell:

$$\frac{d\vec{x}}{dt} = \begin{cases} \frac{dx_1}{dt} = -k_1 x_1 - k_2 x_1 \\ \frac{dx_2}{dt} = k_2 x_1 - k_3 x_2 - k_4 x_4 \\ \frac{dx_3}{dt} = k_4 f x_2 - k_5 x_3 \\ \frac{dx_4}{dt} = k_4 (1-f) x_2 - k_6 x_4. \end{cases} \tag{15}$$

A list of variables used and their meanings can be found in Table 2. Continuing to implement the modeling process in [4], we were able to calculate the $\vec{\theta}$ that produced the lowest $J(\vec{\theta})$. This was evaluated by comparing the data to our model, which is the solution $\vec{x}$ (the masses of CFDA-SE, CFSE, CF-R1, and CF-R2) multiplied by the observation matrix (1 1 1 1). This first model, based on simple laws of mass action in the kinetics, is plotted against the data versus time in Figure 4.

4.3 Biological Model 2

In the first mathematical model, Model 1.1, the rates at which CF-R1 and CF-R2 are formed from CFSE were assumed to be the same. However, based on our general knowledge of chemical reactions and energy diagrams, we questioned the appropriateness of this assumption. Typically, the rate determining step of a reaction is the step in which the most unstable state is being formed. Applying this notion to the conversion of CFSE, the reaction forming the unstable conjugate would occur much more slowly than the reaction forming the stable conjugate. Therefore, we concluded that the conversion of CFSE to CF-R1 and CF-R2 should be characterized by two different rates instead of one rate in two different proportions.

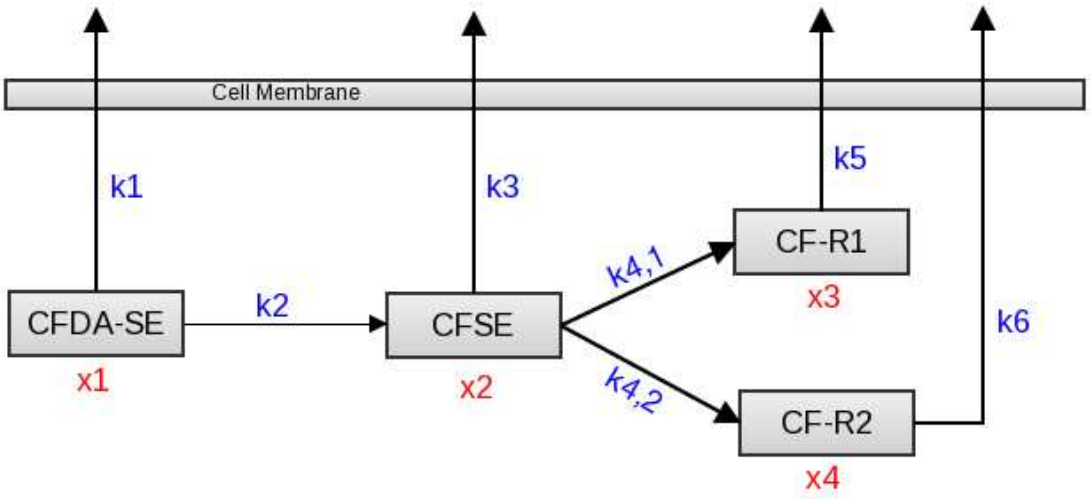


Figure 5: A schematic of *Biological Model 2*, with independent reaction rates from CFSE to CF-R1 and CF-R2.

4.3.1 Model 2.1: Exponential Decay

To follow this new biological model, the f parameter was removed, and the reactions of CFSE to CF-R1 and CF-R2 were treated as two completely independent reactions. This change resulted in the system of equations

$$\frac{d\vec{x}}{dt} = \begin{cases} \dfrac{dx_1}{dt} = -k_1 x_1 - k_2 x_1 \\ \dfrac{dx_2}{dt} = k_2 x_1 - k_3 x_2 - k_{4,1} x_2 - k_{4,2} x_2 \\ \dfrac{dx_3}{dt} = k_{4,1} x_2 - k_5 x_3 \\ \dfrac{dx_4}{dt} = k_{4,2} x_2 - k_6 x_4, \end{cases} \tag{16}$$

where $k_{4,1}$ and $k_{4,2}$ represent the rates of conversion to CF-R1 and CF-R2, respectively, as given in Table 2. The plots for Model 2.1 versus the data set are presented in Figure 6.

4.3.2 Model 2.2: Michaelis-Menten Kinetics for the First Conversion

Exponential models, though simple to implement, represent models with possibly unbounded solutions, which one does not see in typical biological systems. Thus, to more accurately estimate various conversion rates, we decided

4.4 Biological Model 3

Upon further consideration, we generated the hypothesis that CFDA-SE is converted to CFSE through the catalyzed hydrolysis of its acetyl esters by acetylesterase. It is generally accepted that intracellular esterases are responsible for the conversion of CFDA-SE to CFSE [19, 22]. These esterases are hydrolase enzymes that cleave the acetyl esters present in CFDA-SE into their parent carboxylic acid, acetate, and an alcohol [10]. The particular esterase which specializes in removing acetyl groups is called acetylesterase [1].

Applying the mechanisms of hydrolysis to CFDA-SE produces the exact structure of CFSE that is described in the literature and Section 4.2 above. Under basic conditions, a reaction known as saponification [10] takes place (see Figure 9). First, the hydroxide, functioning as a nucleophile, attacks the electrophilic C found in the double bond of the ester, breaking the π bond and forming a tetrahedral intermediate. This intermediate then collapses to form a carboxylic acid when the alkoxide leaving group is kicked off and the π bond is re-formed. The alkoxide that was previously lost then functions as a base, quickly deprotonating the carboxylic acid and forming the final products: the parent carboxylic acid (acetate) and an alcohol.

Under acidic conditions, the reverse of Fischer esterification [10] takes place (see Figure 10). To begin with, the ester must be activated since the nucleophile present is weak and the electrophile present is poor. To do this, the oxygen of the carbonyl ester is protonated to make it more electrophilic. The molecule then encounters water, which functions as a nucleophile, attacking the electrophilic C found in the double bond of the ester, breaking the π bond and forming a tetrahedral intermediate. Another water molecule then deprotonates the oxygen that came from the water molecule, neutralizing its charge. As with the base hydrolysis, the alkoxide needs to leave, but in this case, it is not a good enough leaving group. Therefore, it must first be protonated. Following protonation, the electrons from the adjacent oxygen help push it off, reforming the π bond and creating an alcohol. Yet another water molecule then deprotonates the oxonium ion forming the final products: the parent carboxylic acid (acetate), a regenerated acid catalyst (hydronium), and an alcohol (from the previous step). Such correlation leads us to conclude that the catalyzed hydrolysis of its acetyl esters by acetylesterase (via either mechanism above) is a reasonable explanation for the reactants and products observed in the biological process.

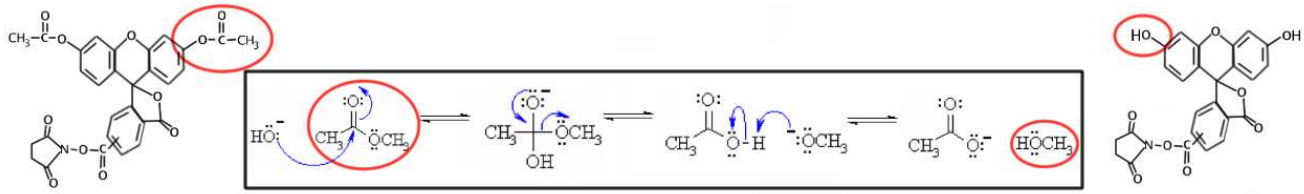


Figure 9: The above shows the mechanism by which the acetyl esters present in CFDA-SE are cleaved to form CFSE under basic conditions. This base hydrolysis of esters is known as saponification [10].

This knowledge raised further questions about the importance of the diffusion of CFDA-SE out of the cell, leading to the creation of a third biological model. Although we were comfortable with the assumption that there is no significant inflow of CFDA-SE into cells, previous models showed an efficient transfer of CFDA-SE to CFSE. This is consistent with the notion that catalyzed reactions occur quickly, so in the next model (denoted as *Biological Model 3* and depicted in Figure 11), all CFDA-SE was assumed to have already been converted to CFSE at $t = 0$ without any CF-R1 or CF-R2 yet present. .

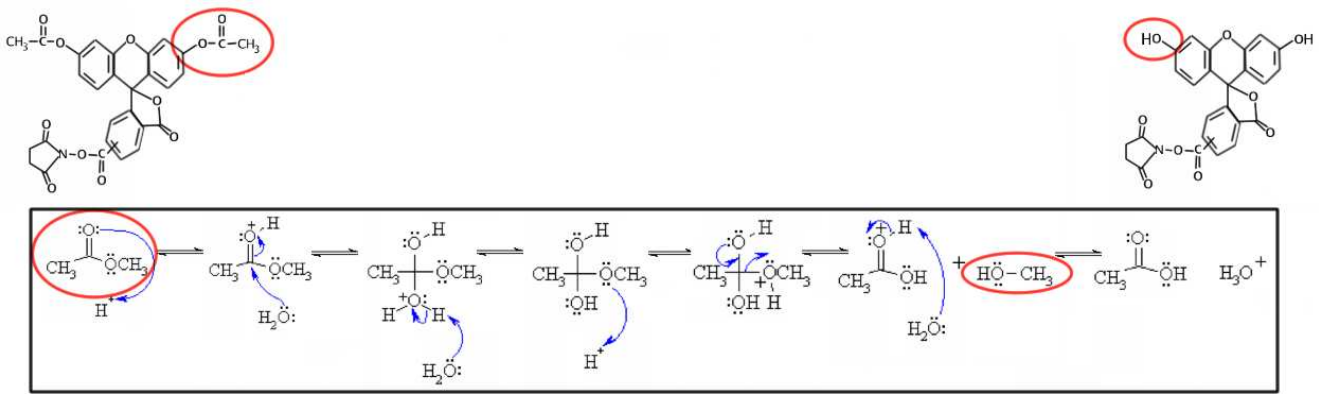


Figure 10: The above shows the mechanism by which the acetyl esters present in CFDA-SE are cleaved to form CFSE under acidic conditions. This acid catalyzed hydrolysis of esters is known to be the reverse of Fischer esterification [10].

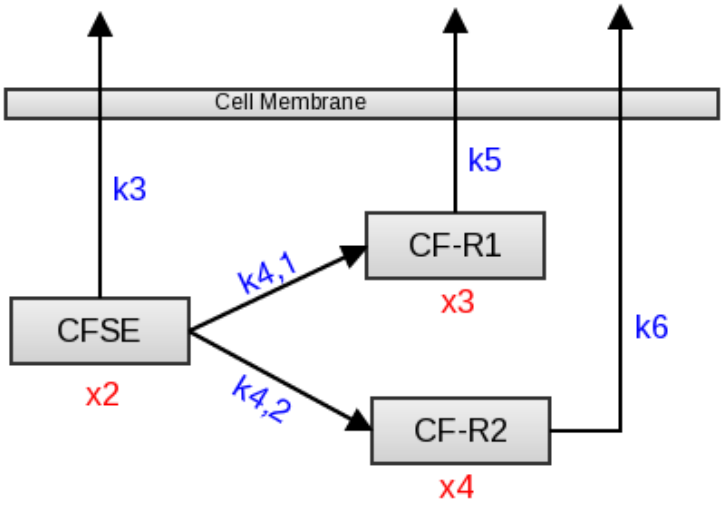


Figure 11: A schematic of *Biological Model 3*, assuming all CFDA-SE has already been converted to CFSE at $t = 0$.

4.4.1 Model 3.1: Exponential decay

Based on the assumption that CFDA-SE is immediately converted to CFSE, our original exponential system was altered to only depend on five rate parameters instead of seven, yielding the system of equations (19). The variables are defined in Table 2, and depicted in the schematic for Model 3 (Figure 11). Plots of the corresponding fits-to-data are given in Figure 12.

$$\frac{d\vec{x}}{dt} = \begin{cases} \dfrac{dx_2}{dt} = -k_3 x_2 - k_{4,1} x_2 - k_{4,2} x_2 \\ \dfrac{dx_3}{dt} = k_{4,1} x_2 - k_5 x_3 \\ \dfrac{dx_4}{dt} = k_{4,2} x_2 - k_6 x_4. \end{cases} \tag{19}$$

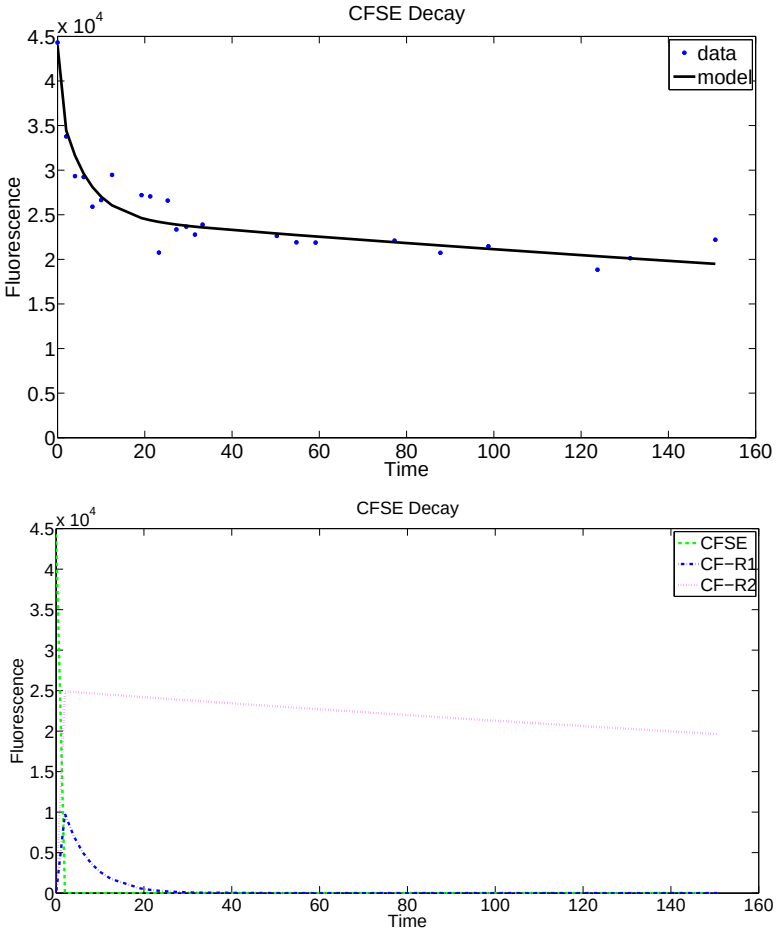


Figure 12: A plot of Model 3.1 and the first data set from Donor 1 using exponential decay is depicted in the top figure, assuming all of the CFDA-SE is immediately converted to CFSE and little dye leaks out of the cell. The individual components of the system are plotted in the bottom graph.

4.4.2 Model 3.2: Michaelis-Menten Kinetics

We again consider Biological Model 3 but with Michaelis-Menten kinetics. After applying Michaelis-Menten kinetics to all of the conversion rates and leaving the rates of decay as exponential rates, the system of differential equations (20) was created. The resulting plots versus the data along with component compartments are given in Figure 13.

$$\frac{d\vec{x}}{dt} = \begin{cases} \dfrac{dx_2}{dt} = -k_3 x_2 - \dfrac{v_{4,1} x_2}{K_{4,1} + x_2} - \dfrac{v_{4,2} x_2}{K_{4,2} + x_2} \\ \dfrac{dx_3}{dt} = \dfrac{v_{4,1} x_2}{K_{4,1} + x_2} - k_5 x_3 \\ \dfrac{dx_4}{dt} = \dfrac{v_{4,2} x_2}{K_{4,2} + x_2} - k_6 x_4. \end{cases} \tag{20}$$

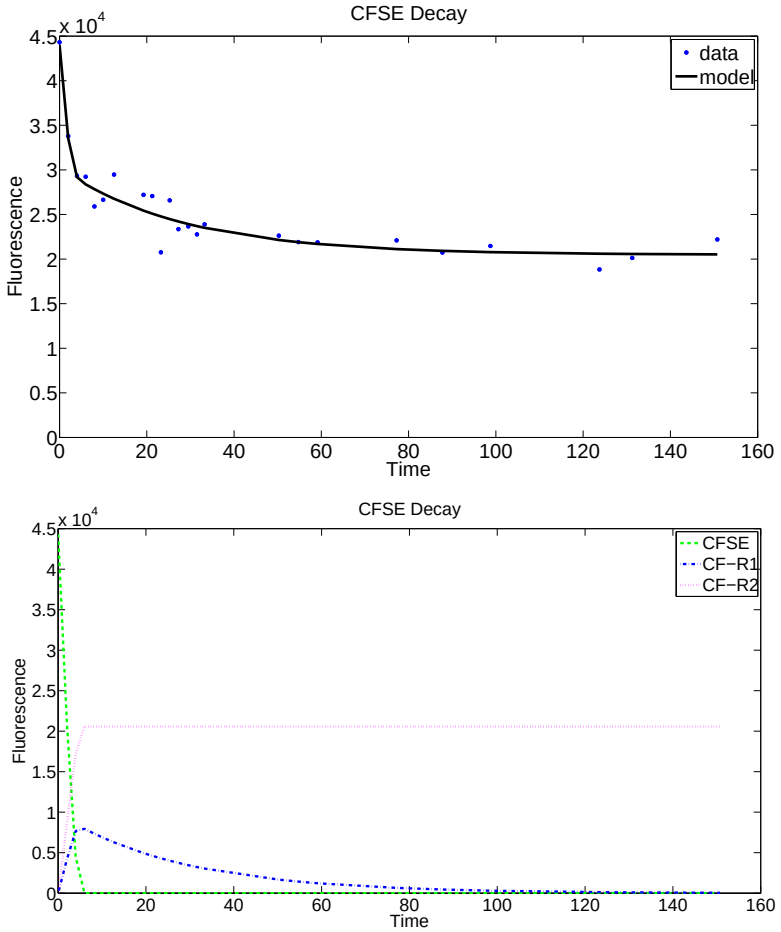


Figure 13: A plot of the Model 3.2 and the first data set from Donor 1, using Michaelis-Menten kinetics for the conversion of CFSE to CF-R1 and CF-R2 is given in the top graph. The individual components are plotted in the bottom graph.

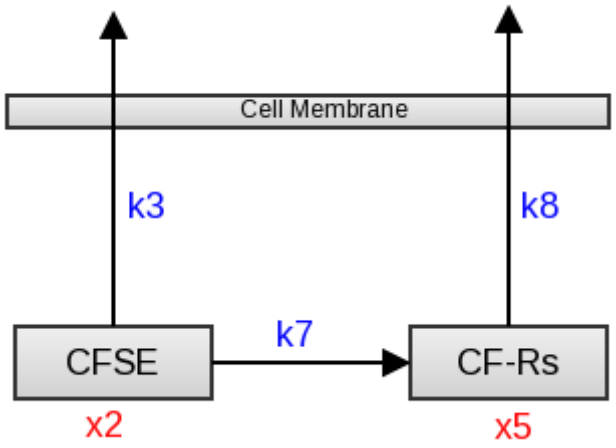


Figure 14: A schematic of *Biological Model 4*, in which the two CF-R1 and CF-R2 components combined as a single component.

4.5 Biological Model 4

Encouraged by our success in simplifying our model from four to three components, we decided to test this simplicity even further with several two-component systems. The first such system combines both of the CF-R1 and CF-R2 terms together as one single component. A schematic, depicted as *Biological Model 4*, for this model is shown in Figure 14.

4.5.1 Model 4.1: Exponential Decay

In the first mathematical model for Biological Model 4, we assume an exponential conversion rate between each term. This is given by the differential equation (21). The results of the inverse problem calculations for Model 4.1 is plotted against the third data set from Donor 1, and is given in Figure 15.

$$\frac{d\vec{x}}{dt} = \begin{cases} \dfrac{dx_2}{dt} = -k_3 x_2 - k_7 x_2 \\ \dfrac{dx_5}{dt} = k_7 x_2 - k_8 x_5. \end{cases} \tag{21}$$

4.5.2 Model 4.2: Michaelis-Menten Kinetics

In this mathematical model, we assume a Michaelis-Menten conversion rate between CFSE and the general CF-R term, composed of both CF-R1 and CF-R2. Although there is no real biological basis for using a Michaelis-Menten reaction here since there is no enzyme involved in this step, it still provides a nice bounded solution model. The differential equation system created by using Michaelis-Menten kinetics for the conversion rates and exponential decay for the decay rates is given in (22). This model was fit to the third data set from Donor 1, with results given in Figure 16.

$$\frac{d\vec{x}}{dt} = \begin{cases} \dfrac{dx_2}{dt} = -k_3 x_2 - \dfrac{v_7 x_2}{K_7 + x_2} \\ \dfrac{dx_5}{dt} = \dfrac{v_7 x_2}{K_7 + x_2} - k_8 x_5. \end{cases} \tag{22}$$

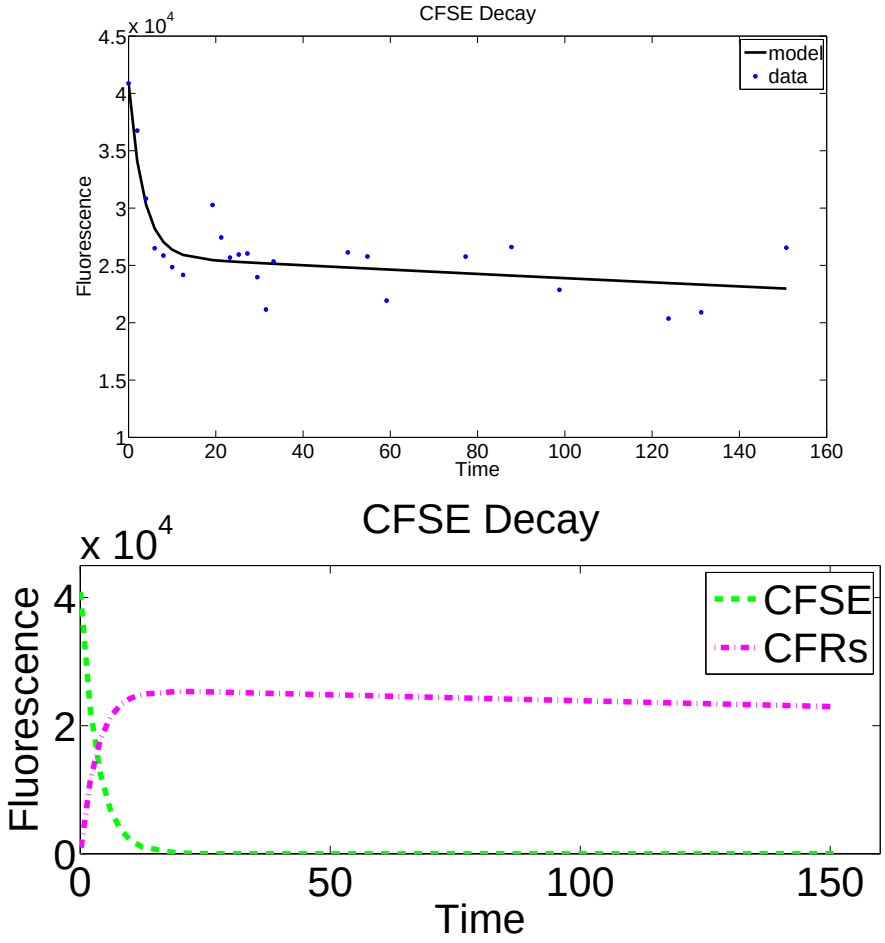


Figure 15: Top: Model 4.1 fit to the first data set from Donor 2; Bottom: the components of the system.

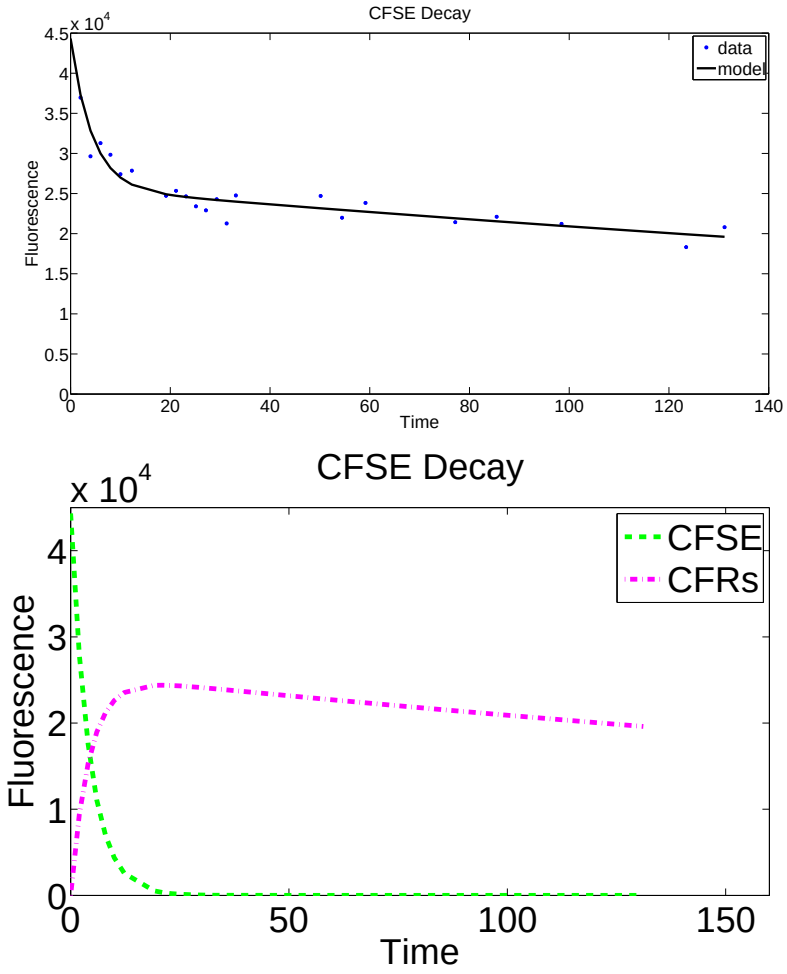


Figure 16: Model 4.2 fit to the third data set from Donor 1 in the top graph, with the components of the system depicted in the bottom graph.

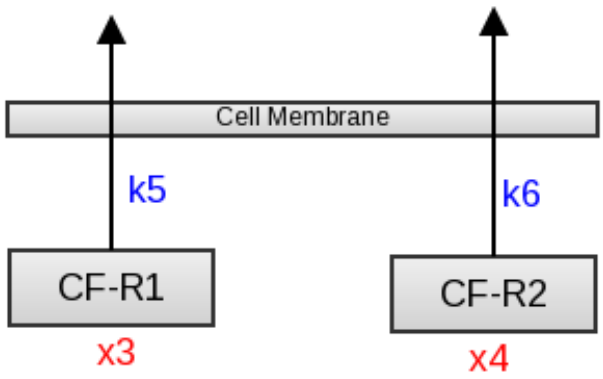


Figure 17: Schematic of *Biological Model 5*, which only accounts for the fluorescence given by CF-R1 and CF-R2.

4.6 Biological Model 5

We next considered a second two-component system, where we solely assume that fluorescence is composed of the CF-R1 and CF-R2 terms, which are constantly decaying. This *Biological Model 5* is depicted in Figure 17.

4.6.1 Model 5.1

There is only one mathematical equation for this model, and the differential equation can actually be solved analytically very easily. Accordingly, the equation for the fluorescence is given by

$$x(t) = cx_0e^{-k_5t} + (1-c)x_0e^{-k_6t} \tag{23}$$

where the c term, $0 \leq c \leq 1$, allows both components to receive some of the initial fluorescence, denoted by x_0. A plot of this model fitted against the third data set from Donor 1 is given in Figure 18.

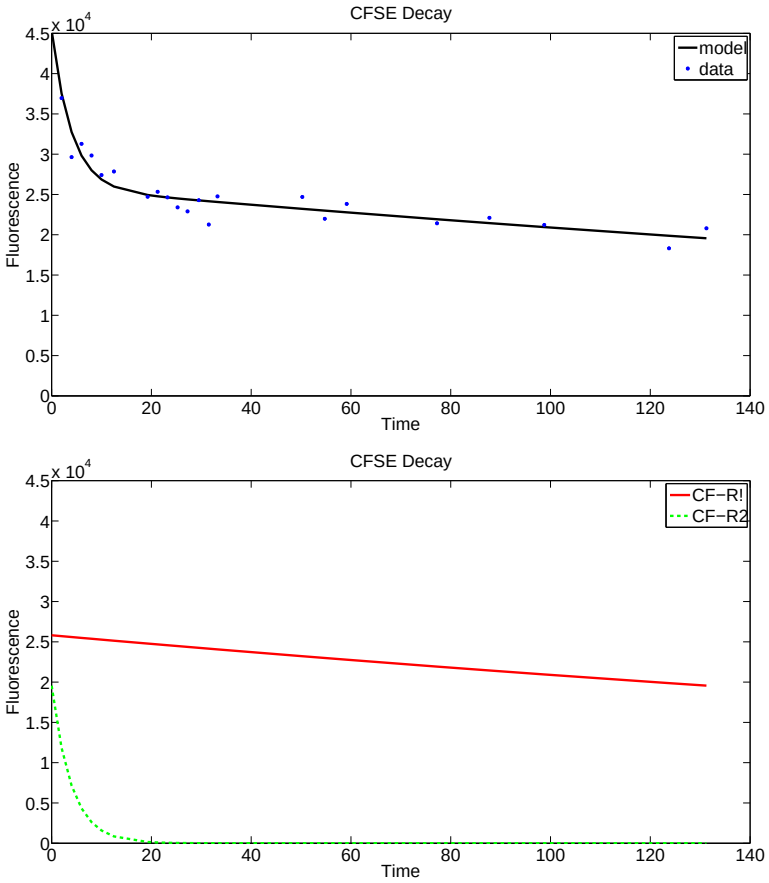


Figure 18: Model 5.1 fit to the third data set from Donor 1 in the top graph, with the components plotted in the bottom graph.

4.7 Biological Model 6

In order to further test simplification in our models, we decided to carry out the inverse problem for the data sets using models with only one differential equation instead of an entire system. In our previous models, we were able to plot each of the components of CFSE and make predictions about how the mass of each component changes as the dye is converted from one form to another. The models in this section do not allow us to calculate the mass of each component, but they do tell us the rate of change in the total fluorescence in the cell.

4.7.1 Model 6.1: Exponential Decay

Using the exponential decay model, we created the differential equation with only one rate parameter given by

$$\frac{dx}{dt} = rx. \tag{24}$$

Initial attempts at fitting the data with a fixed initial condition of the first data point resulted in exponential curves that cut through the middle of the data. We found that adding an additional parameter for estimating the initial condition allowed the exponential curve to fit the data much more smoothly, resulting in a smaller cost function but still not a convincing fit-to-data. This fit can be seen in Figure 19 against two different data sets.

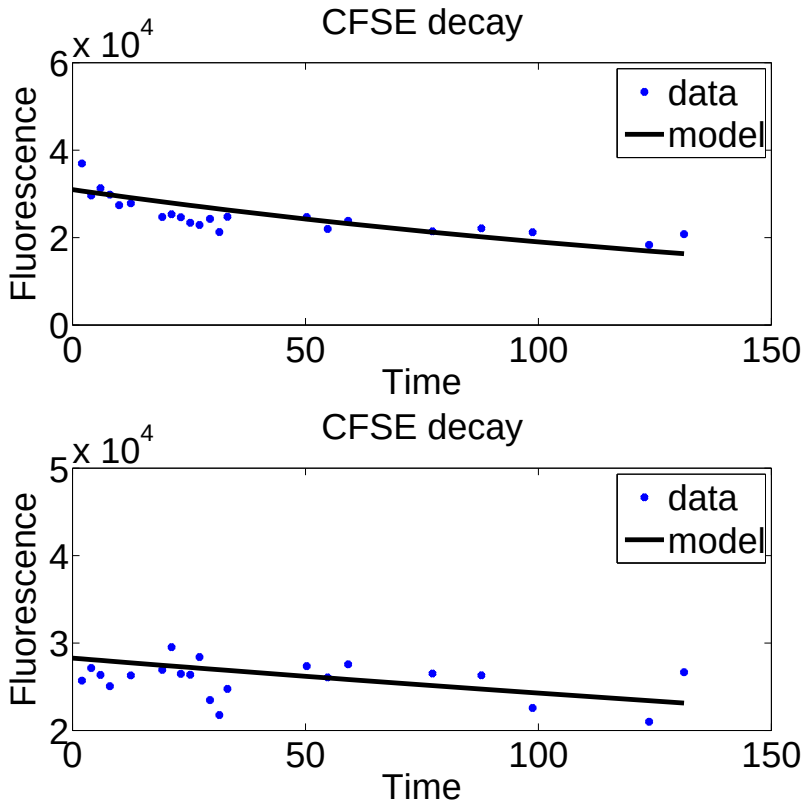


Figure 19: Model 6.1 fit to the third data set from Donor 1 (top) and the third data set from Donor 2 (bottom), using a once component exponential decay.

4.7.2 Model 6.2: Logistic Decay

Using the logistic decay model, we created a differential equation for decay with two parameters

$$\frac{dx}{dt} = \alpha x(1 - \frac{x}{K}), \tag{25}$$

where α is the decay rate and K is the carrying capacity of the population, as given in Table 2. A plot for Model 6.2 versus the data is given in Figure 20.

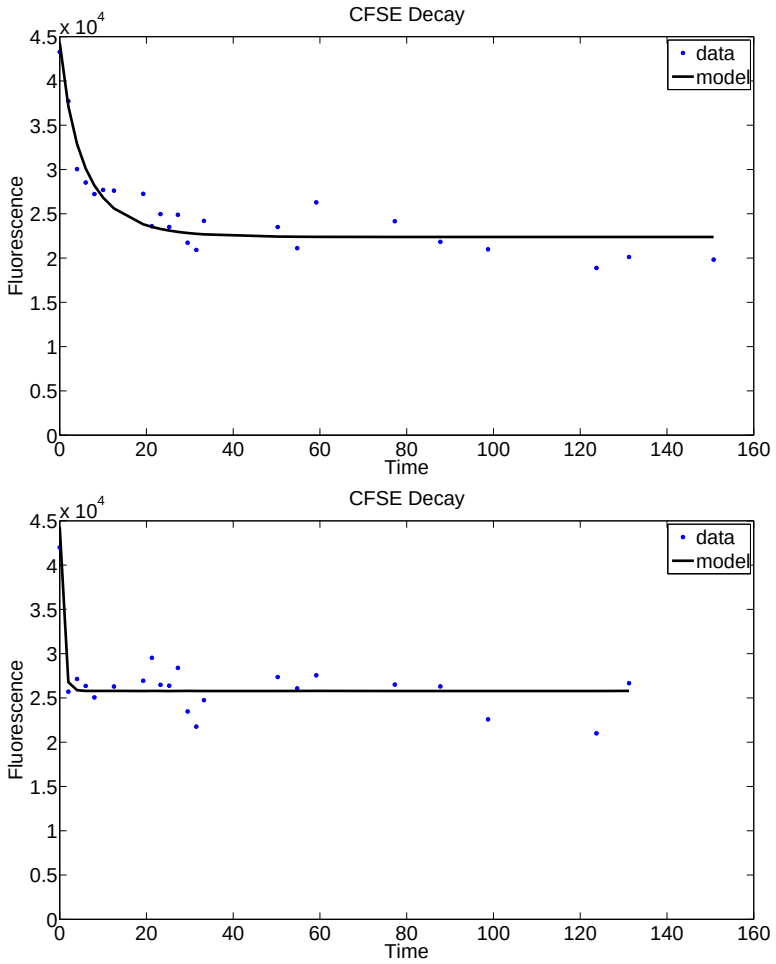


Figure 20: Model 6.2 fit to the second data set from Donor 1 (top) and the third data set from Donor 2 (bottom), using logistic decay and one component.

4.7.3 Model 6.3: Gompertz Decay

We next used the Gompertz model to describe fluorescence decay, which has the same two parameters as Model 6.2. A plot of Model 6.3 against the data is given in Figure 21 with the equation given by

$$\frac{dx}{dt} = \alpha x \log \frac{K}{x}. \tag{26}$$

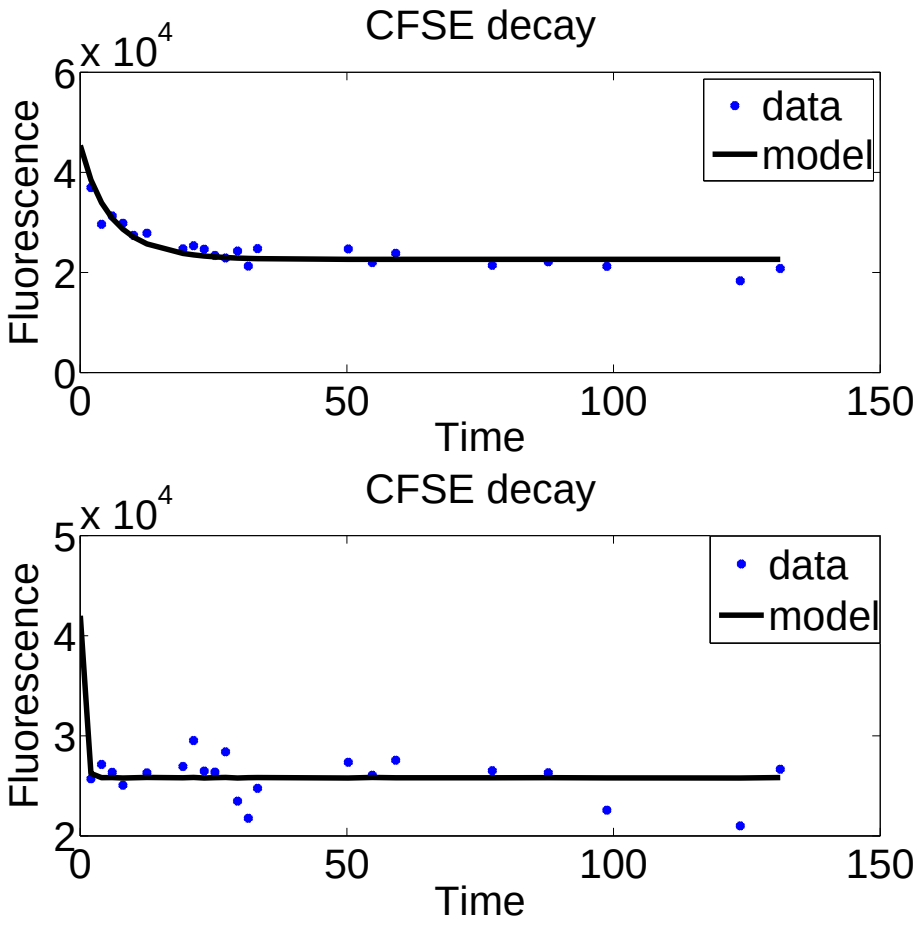


Figure 21: Model 6.3 fit to the third data set from Donor 1 (top) and the third data set from Donor 2 (bottom), using the Gompertz rate and one component.

References

[1] H.T. Banks, Frederique Charles, Marie Doumic, Karyn L. Sutton, and W. Clayton Thompson, Label structured cell proliferation models, *Appl. Math. Letters* **23** (2010), 1412–1415; doi:10.1016/j.aml.2010.07.009

[2] H. T. Banks, K. L. Sutton, W. C. Thompson, G. Bocharov, D. Roose, T. Schenkel, and A. Meyerhans, Estimation of cell proliferation dynamics using CFSE data, CRSC-TR09-17, August, 2009; *Bull. Math. Biology*, **73** (2011), 116–150.

[3] H. T. Banks, K. L. Sutton, W. C. Thompson, G. Bocharov, M. Doumic, T. Schenkel, J. Argilaguet, S. Giest, C. Peligero and A. Meyerhans, A new model for the estimation of cell proliferation dynamics using CFSE data, CRSC-TR11-05, March, 2011; *J. Immunological Methods*, DOI:10.1016/j.jim.2011.08.014.

[4] H. T. Banks and H. T. Tran, *Mathematical and Experimental Modeling of Physical and Biological Processes*, CRC Press, Boca Raton, FL, 2009.

[5] A.V. Gett and P.D. Hodgkin, A cellular calculus for signal integration by T cells, *Nature Immunology* **1** (2000), 239–244.

[6] C. Parish, Fluorescent dyes for lymphocyte migration and proliferation studies, *Immunology and Cell Biol.* **77** (1999), 499–508.

[7] B. Quah, H. Warren and C. Parish, Monitoring lymphocyte proliferation in vitro and in vivo with the intracellular fluorescent dye carboxyfluorescein diacetate succinimidyl ester, *Nature Protocols* **2**:9 (2007), 2049–2056.