Parameter Estimation in Complex Biological Systems

Hien Tran

Department of Mathematics
Center for Research in Scientific Computation
North Carolina State University

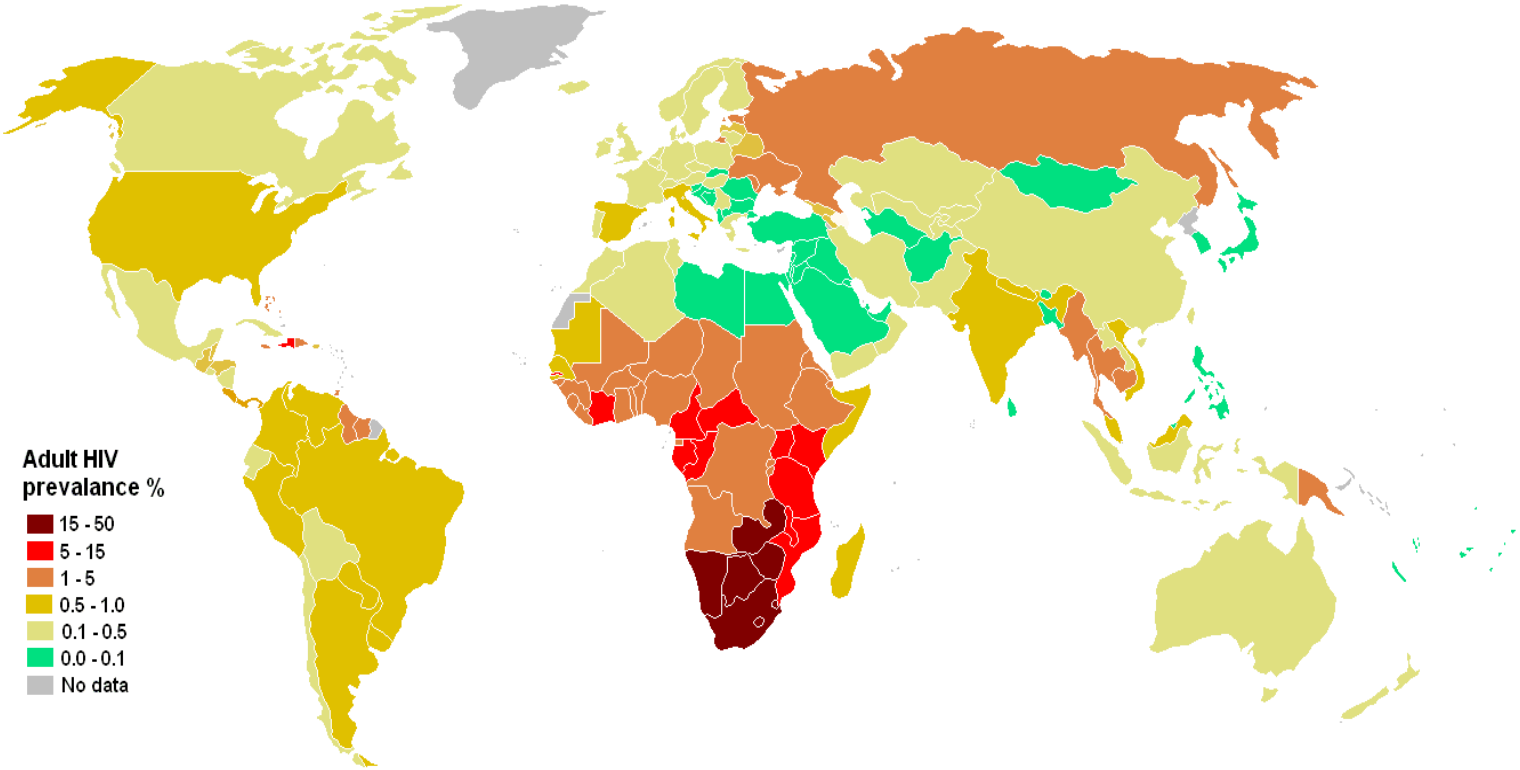
3 August 2011

(joint work with Adam Attarian and Brett Matzuka - NIAID/NSF)

NC STATE UNIVERSITY

Introduction

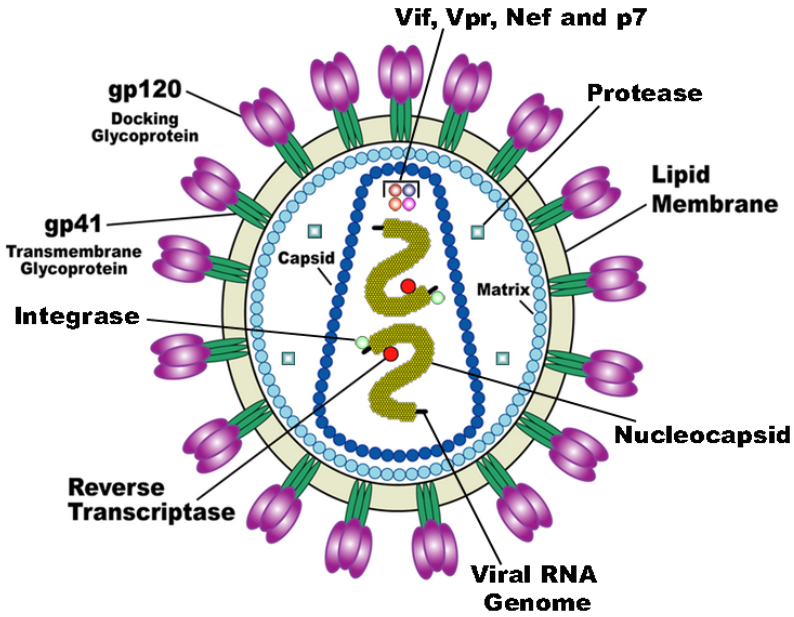
- Human Immunodeficiency Virus (HIV) affects people of all demographics, 33.3 million infected with HIV as of 2009 (WHO/UNAIDS).
- Approximately 1.8 million died in 2009 (WHO).



NC STATE UNIVERSITY

Introduction

- Target T-helper cells of the immune system.
- HIV is a retrovirus.
- Causative agent for Acquired Immune Deficiency Syndrome (AIDS).
- No known cure.

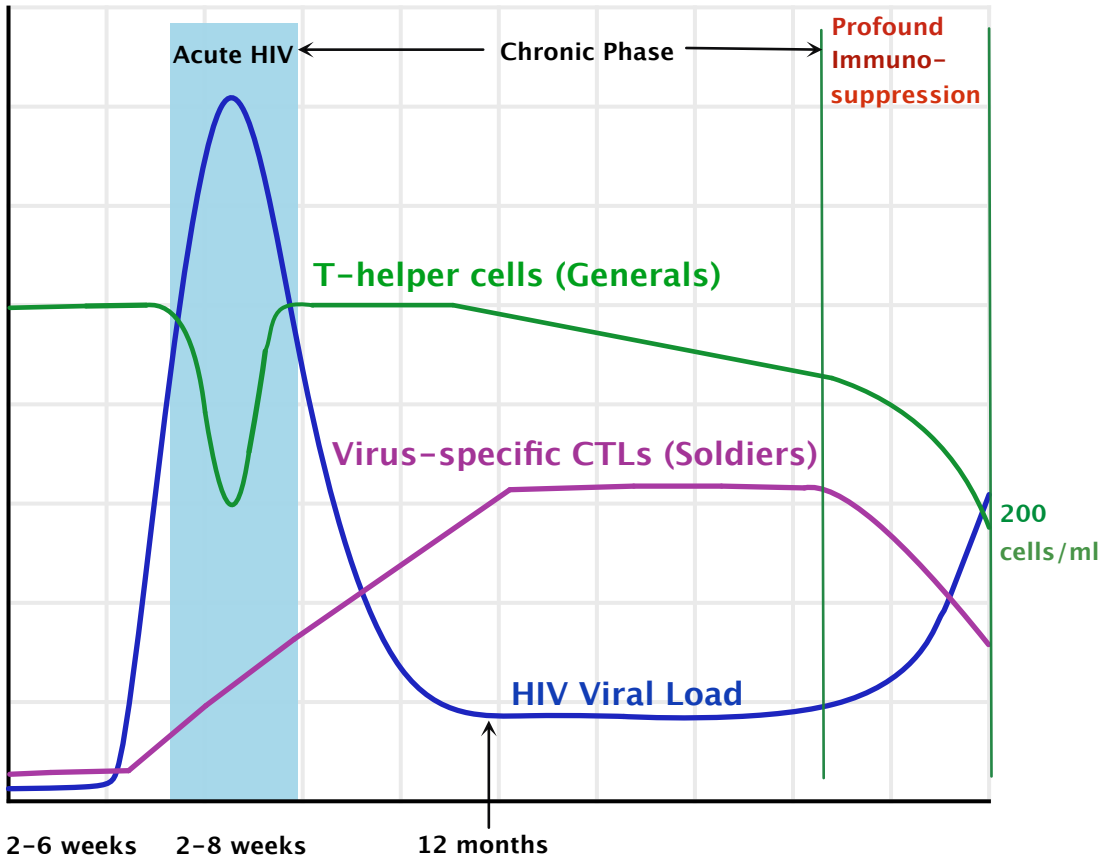


NC STATE UNIVERSITY

Introduction

HIV viral load setpoints:

- 12,000 copies/mL - Slow Progression (very slowly developing AIDS).
- 30,000 copies/mL - Moderate Progression (AIDS in 8-15 years).
- 60,000 copies/mL - Rapid Progression (AIDS in 1-4 years).



NC STATE UNIVERSITY

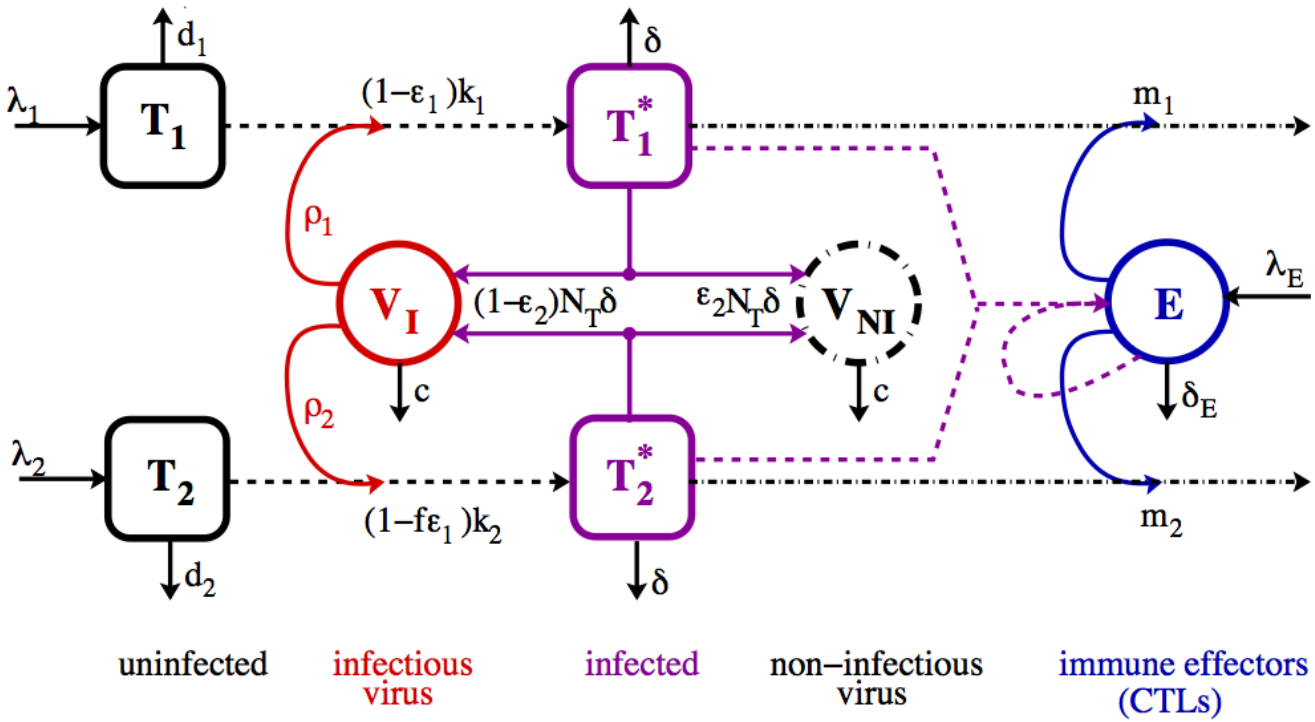
Introduction

- Most common treatment is Highly Active Antiretroviral Therapy (HAART), a combination of multiple drugs.
- The "cocktail drugs" are protease inhibitors (PI) and reverse transcriptase inhibitors (RTI).
- The best way to treat patients, and when, is still an open question.
- The use of mathematical models becoming more prevalent in treatment planning.

Project Goals:

- Develop patient specific mathematical model for HIV dynamics.
- Develop patient specific treatment regimen.

HIV Model



Model compartments include T_1 (type 1 target cells, e.g CD4 Th-cells), T_2 (type 2 target cells, e.g. macrophages), V_I and V_{NI} (infectious and non-infectious virus), and E (cytotoxic T-Lymphocytes). An asterisk denotes infected cells.

NC STATE UNIVERSITY

HIV Model

The associated differential equations are

$$\dot{T}_1 = \lambda_1 - d_1 T_1 - (1-\varepsilon_1) k_1 V_I T_1$$
$$\dot{T}_2 = \lambda_2 - d_2 T_2 - (1-f\varepsilon_1) k_2 V_I T_2$$
$$\dot{T}_1^* = (1-\varepsilon_1) k_1 V_I T_1 - \delta T_1^* - m_1 T_1^* E$$
$$\dot{T}_2^* = (1-f\varepsilon_1) k_2 V_I T_2 - \delta T_2^* - m_2 T_2^* E$$
$$\dot{V}_I = (1-\varepsilon_2) N_T \delta (T_1^* + T_2^*) - \left(c + (1-\varepsilon_1)\rho_1 k_1 T_1 + (1-f\varepsilon_1)\rho_2 k_2 T_2\right) V_I$$
$$\dot{V}_{NI} = \varepsilon_2 N_T \delta (T_1^* + T_2^*) - c V_{NI}$$
$$\dot{E} = \lambda_E + b_E \frac{T_1^* + T_2^*}{T_1^* + T_2^* + K_b} E - d_E \frac{T_1^* + T_2^*}{T_1^* + T_2^* + K_d} E - \delta_E E.$$

$\Rightarrow$ 7 states, 20 parameters, and two control inputs $\varepsilon_1(t) = \varepsilon_1 u(t)$, $\varepsilon_2(t) = \varepsilon_2 u(t)$.

NC STATE UNIVERSITY

HIV Model

Let x denote the vector of solutions to the ODE model. That is,

$$x(t) = \left[T_1(t), T_2(t), T_1^*(t), T_2^*, V_I(t), V_{NI}, E(t)\right].$$

The model can then be written as

$$\frac{dx}{dt} = f(t, x; q), \qquad x \in \mathbb{R}^n, q \in \mathbb{R}^m,$$

where q is a vector of model parameters. The observations of the model are given in terms of the states

$$\begin{aligned} y_1(t) &= T_1(t) + T_1^*(t) \\ y_2(t) &= V_I(t) + V_{NI}(t), \end{aligned}$$

corresponding to $CD4^+$ count and the viral load, respectively.

Parameter Estimation

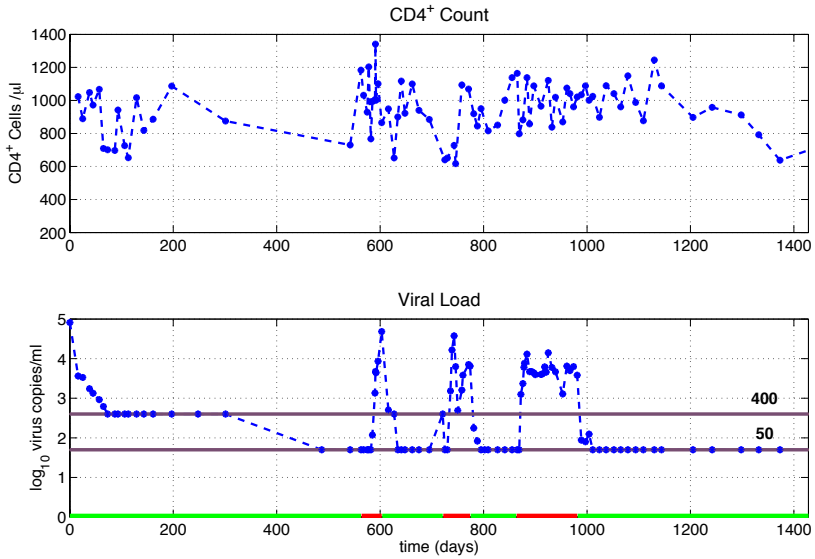
We wish to use the HIV model to describe patient specific clinical data and make patient specific predictions, so it first must be calibrated to patient data by estimating the model parameters. However, due to model structure and possible lack of measurements, some parameters may not be identifiable. In addition,

- Parameters may have a very weak effect on the measured outputs (sensitivity).
- Effect of certain parameters on the measured outputs may be nearly linearly dependent (statistically correlated).

Estimation of weak and/or nearly linearly dependent effects can lead to degradation in the predictive capability of the model.

Parameter Estimation

Clinical study from 1996 to 2004 of over 100 adults with acute HIV-1 infection, administered by Mass General Hospital.



- Data are unique in that the study members were identified soon after initial infection, capturing acute viral dynamics.
- The viral load is *censored* at either 400 copies/ml or 50 copies/ml depending on the sensitivity of the assay.
- Statistical methods needed to estimate truth data, e.g. actual viral load below the censor level.

NC STATE UNIVERSITY

Parameter Estimation

After identifying which parameters can be best suited for estimation from censored data, *how to estimate these parameter*?

$\Rightarrow$ The classical approach for parameter estimation is to minimize the residuals. This is done by defining a cost function as the difference in the squares of the model and the data, and subsequently using an optimization technique (Levenberg-Marquardt, Conjugate Gradient, Nelder-Mead, etc.) to find the minimum.

$$J(q) = \sum_{i=1}^{N} \gamma_i \left[y_i^{data} - y_i^{model}(q) \right]^2$$

where $y^{data} - y_i^{model}(q)$ is the residual.

$\Rightarrow$ **Kalman filter** (on-line estimation)

Sensitivity Analysis

Consider a mathematical model

$$\frac{dx}{dt}(t) = f(t, x(t); q)$$

with observation process

$$y(t, q) = h(t, x(t; q), q)$$

The sensitivity of outputs y_i with respect to parameters q_j is defined by

$$\frac{dy_i}{dq_j}$$

Using the chain rule for differentiation,

$$\frac{dy}{dq} = \frac{\partial h}{\partial x}\frac{dx}{dq} + \frac{\partial h}{\partial q}$$

$$\frac{d}{dt}\frac{dx}{dq} = \frac{\partial f}{\partial x}\frac{dx}{dq} + \frac{\partial f}{\partial q}$$

NC STATE UNIVERSITY

Sensitivity Identifiability

This notion is defined in terms of the output sensitivity functions with respect to the parameters, that is, $dy_i/dq_j(q^*)$ (local concept)
Define the sensitivity matrix function

$$\mathbf{S}(\mathbf{q},\mathbf{t}) \equiv \begin{bmatrix} s_{1,1}(\mathbf{q},t_1) & \cdots & s_{1,m}(\mathbf{q},t_1) \\ s_{1,1}(\mathbf{q},t_2) & \cdots & s_{1,m}(\mathbf{q},t_2) \\ \vdots & \ddots & \vdots \\ s_{1,1}(\mathbf{q},t_n) & \cdots & s_{1,m}(\mathbf{q},t_n) \\ s_{2,1}(\mathbf{q},t_1) & \cdots & s_{2,m}(\mathbf{q},t_1) \\ \vdots & & \vdots \\ s_{k,1}(\mathbf{q},t_n) & \cdots & s_{k,m}(\mathbf{q},t_n) \end{bmatrix}, \quad s_{i,j}(\mathbf{q},t_\ell) = \frac{dy_i(t_\ell)}{dq_j}$$

Now, let $\Delta q = q - q^*$ denote a small perturbation from q^*. This gives rise to a small perturbation in the output $\Delta y = y(t,q) - y(t,q^*)$. By the chain rule, we obtain the following (approximate) relationship

$$\Delta y \approx S \Delta q$$

NC STATE UNIVERSITY

Sensitivity Identifiability

The model is called sensitivity identifiable if

$$\Delta y \approx S \Delta q$$

can be solved uniquely (in the local sense) for Δq. This is the case if and only if $\text{rank}(S) = m$ (m is the number of parameter) or $\det(S^\top S) \neq 0$.

$\Rightarrow$ Need to compute the numerical rank k of $S \Rightarrow k$ identifiable parameters.

$$\text{rank}(\mathbf{S}, \varepsilon) = \max \left\{ i \left| \frac{|\sigma_i|}{|\sigma_1|} > \varepsilon \|\mathbf{S}\| m \right. \right\},$$

for the ordered singular values σ_i. Here, $\varepsilon = \sqrt{10^{-16}} = 10^{-8}$.

Subset Selection

Now that we know we can estimate k_j parameters for patient j, which k parameters? We consider the distance between subspaces for each subset ρ.

- Consider all $\binom{n}{k}$ subsets of parameters.
- Compute the subspace spanned by the k most significant eigenvectors of $S^\top S$. Call this W_1.
- Define W_2^i to be the subspace spanned by the elementary basis generated by $\rho_i, i \in [1, 2, \ldots, \binom{n}{k}]$
- Compute the minimum distance between W_1 and W_2^i, $\min_i \text{dist}(W_1, W_2^i)$.
- The *best identifiable* subset of parameters is then ρ_{min}, and is different for each patient.

For patient 1, we have $k = 9$ and $\rho = [d_1\ k_1\ k_2\ m_1\ m_2\ N_T\ \delta_E\ \varepsilon_1]$.

NC STATE UNIVERSITY

Inverse Problem

For each patient j, the goal is to fit the ODE model to data by minimizing

$$q^{*j} = \arg\min J(q) = \sum_{s=1}^{2} \frac{1}{N_s^j} \sum_{i=1}^{N_s^j} \left| z_s(t_s^{ij}; q) - y_s^{ij} \right|^2.$$

Minimizing J corresponds to a maximum likelihood estimation of q assuming the measurements y_s^{ij} are normally distributed,

$$y_s^{ij} \sim \mathcal{N}(z_s(t^i; q^0), \sigma_s^2), \quad s = 1, 2,$$

for some true parameter q^0 and variance σ_s^2. For viral load in the absence of censoring, the log-likelihood function is

$$\mathcal{L}(q, \sigma_2) = -\frac{N}{2} \log 2\pi - N \log \sigma_2 - \sum_{i=1}^{N} \frac{(y_2^i - x_2^i)}{2\sigma_2^2}.$$

NC STATE UNIVERSITY

Censored Data

Censored data violate the assumption of full normality, so we assume it is distributed according to a *truncated* normal distribution. Define the standard normal pdf and cdf,

$$\phi(\xi) = \frac{1}{\sqrt{2\pi}} \exp(\xi^2/2), \quad \Phi(\xi) = \int_{-\infty}^{\xi} \phi(s) ds,$$

then the log-likelihood function for all data, based on a truncated normal distribution, is given by

$$\mathcal{L}(q, \sigma_2) = \sum_{i=1}^{N} \left(\chi^i \left[\log \phi \left(\frac{y_2^i - z_2^i}{\sigma_2} \right) - \log \sigma_2 \right] + (1 - \chi^i) \left[\log \Phi \left(\frac{L^i - z_2^i}{\sigma_2} \right) \right] \right),$$

where $\chi = 1$ for uncensored data and $\chi = 0$ for censored data. We wish to maximize $\mathcal{L}$ for q, σ_2. This is accomplished through the Expectation Maximization (EM) algorithm.

NC STATE UNIVERSITY

EM Algorithm

EM iteratively updates q, σ_2, until the maximum is achieved. The data are updated according to

$$\tilde{y}^{i(k)} = \chi^i y_2^i + (1 - \chi^i) \left[z_2^{i(k)}(q^k) - \sigma_2^{(k)} \frac{\phi(\zeta^{i(k)})}{\Phi(\zeta^{i(k)})} \right], \quad \zeta^{i(k)} = \frac{L - z_2^{i(k)}}{\sigma_2^k}.$$

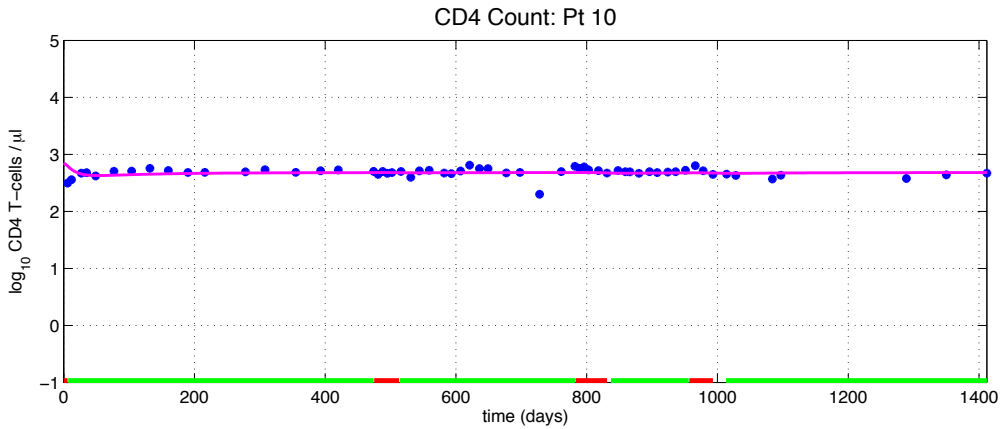
The parameters and variance estimates are then updated by performing the weighted least squares minimization of

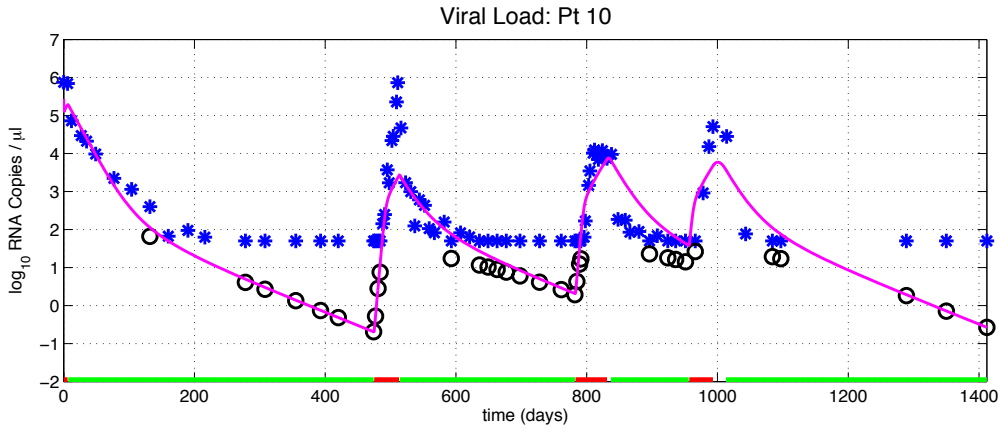
$$q^{k+1} = \arg\min \left[\frac{1}{(\sigma_1^k)^2} \sum_{i=1}^{N_1} \left((y_1^i - z_1(t_1^i; q) \right)^2 + \frac{1}{(\sigma_2^k)^2} \sum_{i=1}^{N_2} \left((\tilde{y}_i^2 - z_2(t_2^i; q) \right)^2 \right].$$

Algorithm terminates on small relative changes of q, σ_1, σ_2.

Model Validation

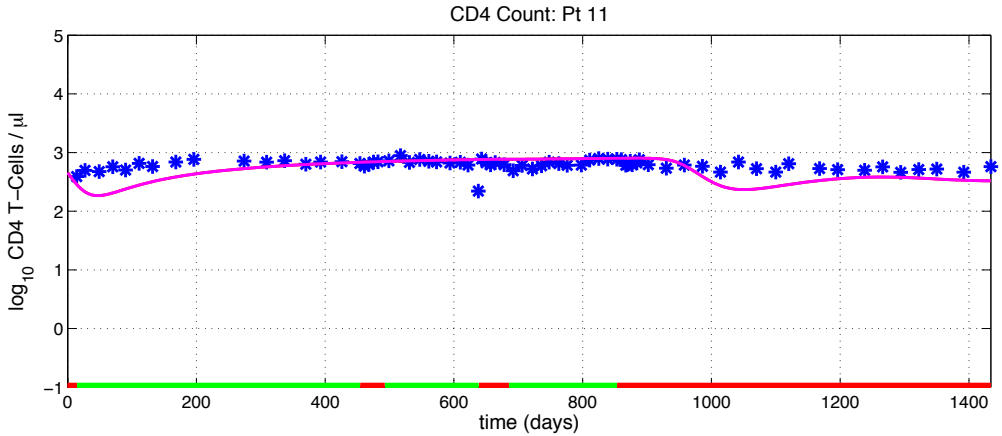
Identifiability analysis is performed on all 14 patients, resulting in 14, subsets of identifiable parameters. The non-identifiable subset $\bar{p}$ is fixed to the result of a global optimization process.

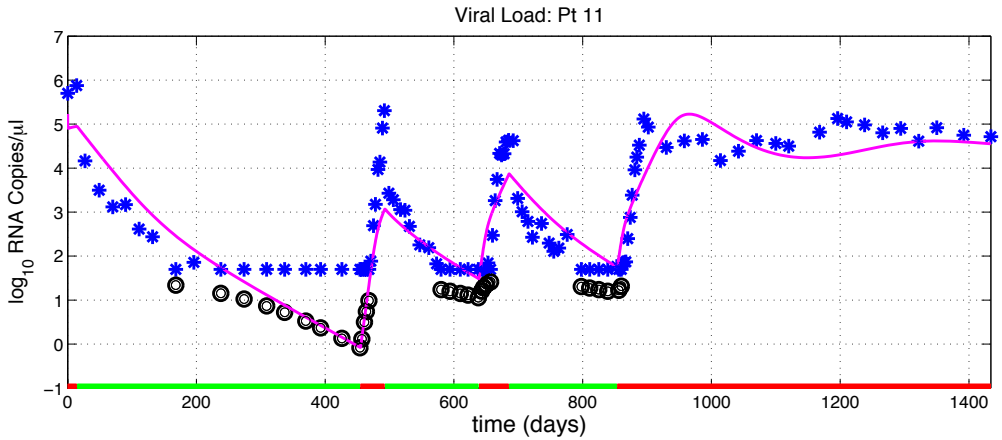




NC STATE UNIVERSITY

Model Validation





NC STATE UNIVERSITY

Kalman Filter (On-Line Estimation)

Dual estimation problems consist of estimating both the states, x_k, and the parameters, θ_k, given noisy data, y_k.

Joint Filtering

$$\begin{aligned} \dot{x} &= f(t, x; q) \\ \dot{q} &= 0 \end{aligned}$$

- Increase the number of states (large number of parameters)
- Errors propagate from the state into the parameter (which subsequently propagate back into the state)

  $\Rightarrow$ Can lead to inaccurate results or divergence of the filter

NC STATE UNIVERSITY

Kalman Filter (On-Line Estimation)

Dual Filtering

Idea: Running two filters concurrently

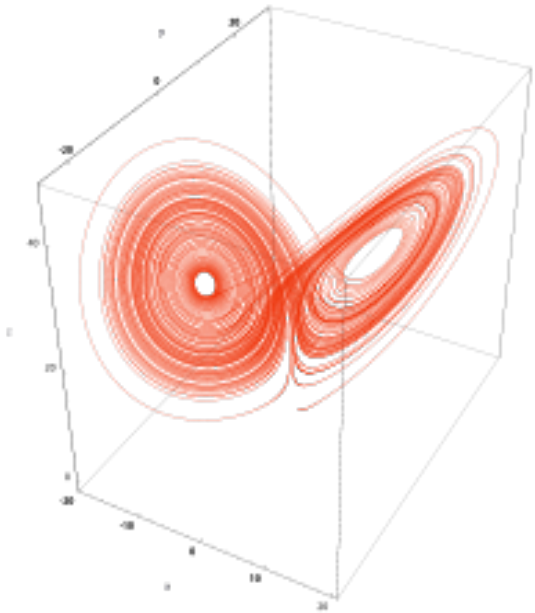
- State Filter estimates the state using the current parameter estimate.
- Parameter Filter estimates the parameters using the current state estimate.
  - Do not increase the number of states for estimation.
  - The error of each filter is still passed, but it is better handled due to each filter updated its covariance.

Estimation Examples

Lorenz Equations

Edward N. Lorenz (a meteorologist and mathematician from MIT) in 1963 derived the following simplified three-dimensional system for convection motion of fluid cells:

$$\frac{dx}{dt} = \sigma(y - x)$$
$$\frac{dy}{dt} = x(\rho - z) - y$$
$$\frac{dz}{dt} = xy - \beta z$$



NC STATE UNIVERSITY

For our test case, we generated simulated data using the following initial conditions and parameters (parameter values change at $t = 10$):

$$\begin{pmatrix} x_0 \\ y_0 \\ x_0 \end{pmatrix} = \begin{pmatrix} 0.9 \\ 1 \\ 1.1 \end{pmatrix}, \quad \begin{pmatrix} \sigma \\ \rho \\ \beta \end{pmatrix} = \begin{pmatrix} 10; 7 \\ 28; 21 \\ \frac{8}{3}; 1 \end{pmatrix}, \quad w_k = 0.1, \quad \Delta t = 0.01$$

NC STATE UNIVERSITY

Filtered Results Using Simulated Data

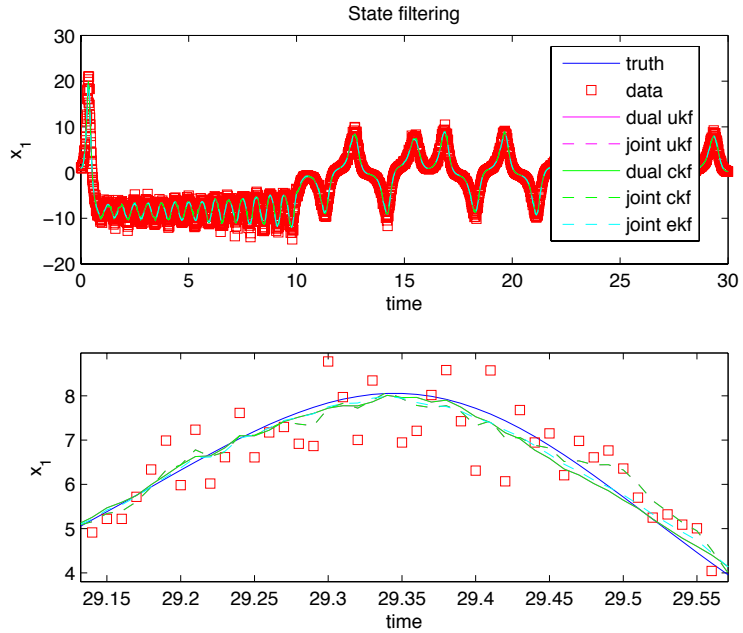


Figure: Plot of state estimates using joint and dual filters for x_1.

	truth	j. ukf	j. ckf	j. ekf	d. ukf	d. ckf
SST	0	3.0251	3.0251	1.1403	0.7987	0.8095
SSP	0	0.1182	0.11784	0.855	0.031416	0.031383
σ (1 std)	10; 7	7.22 (3.04)	7.22 (3.27)	6.91 (0.91)	6.93 (1.02)	6.93 (1.02)
ρ (1 std)	28; 21	21.23 (2.86)	21.23 (3.14)	20.63 (0.81)	20.83 (0.47)	20.83 (0.47)
β (1 std)	$\frac{8}{3}$; 1	0.87 (0.94)	0.87 (1.14)	0.84 (0.37)	1.00 (0.10)	1.00 (0.10)

NC STATE UNIVERSITY

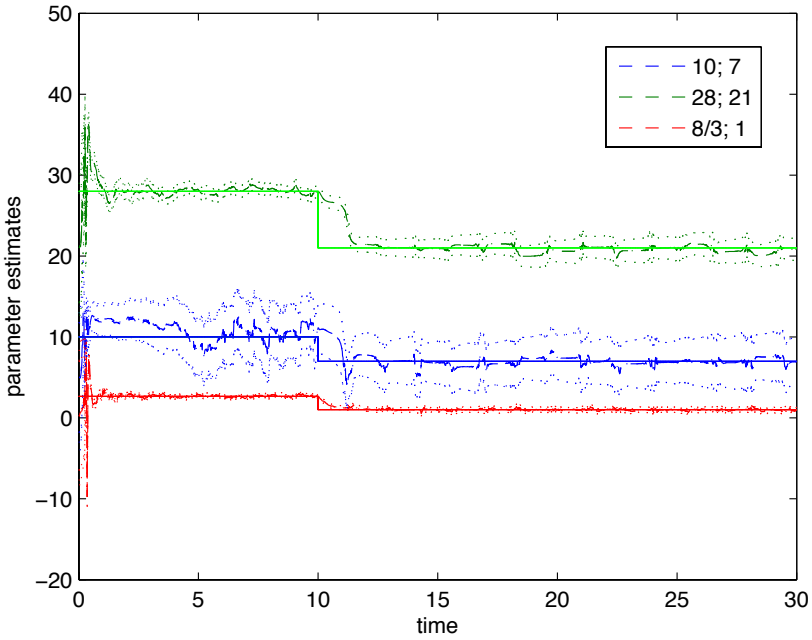


Figure: Plot of parameter estimates for the dual UKF and dual CKF.

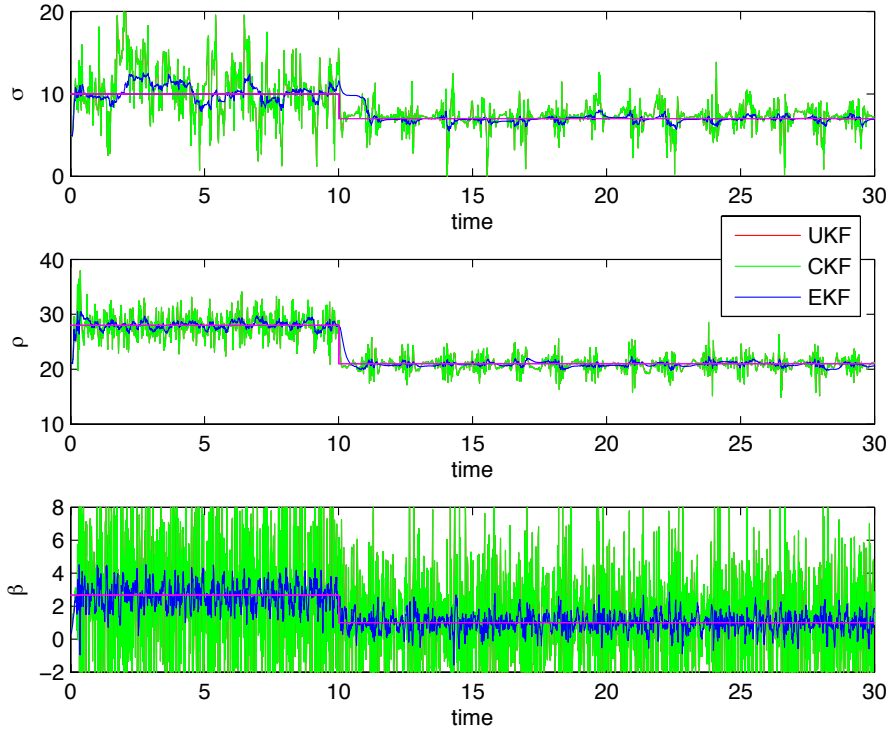
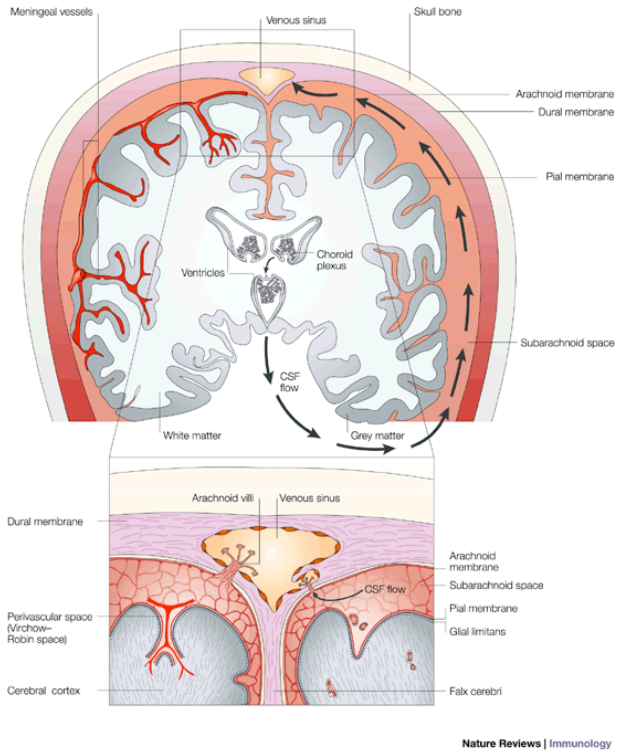


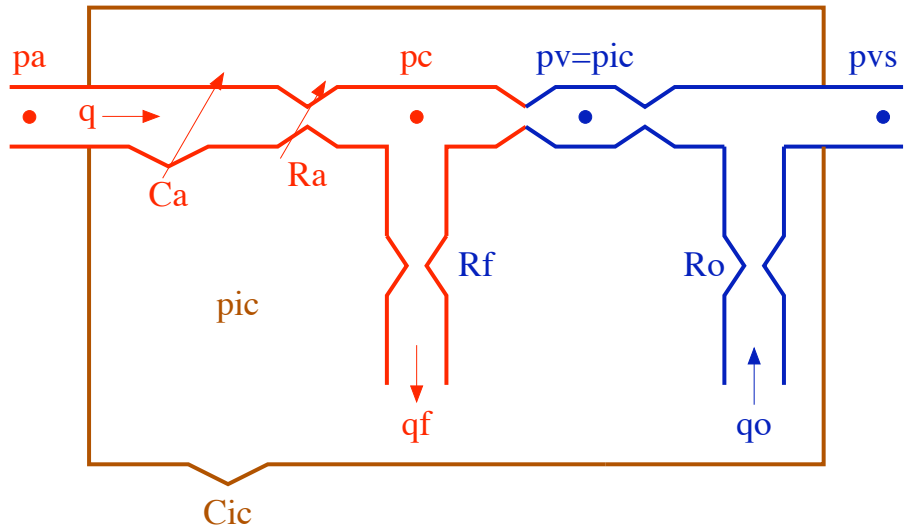
Figure: Plot of parameter estimates for the joint EKF, joint UKF and joint CKF.

Estimation Examples

A Cardiovascular Regulation Model

Reference: M. Ursino and C.A. Lodi, *Journal of Applied Physiology*, 81:1256-1269, 1997.





NC STATE UNIVERSITY

$$\frac{dp_{ic}}{dt} = \frac{k_E p_{ic}}{1 + C_a k_E p_{ic}} \left[C_a \frac{dp_a}{dt} + \frac{dC_a}{dt}(p_a - p_{ic}) + \frac{p_c - p_{ic}}{R_f} - \frac{p_{ic} - p_{vs}}{R_0} \right]$$

$$R_a = \frac{k_R}{V_a^2}, \quad V_a = C_a(p_a - p_{ic}), \quad p_c = \frac{p_a R_{pv} + p_{ic} R_a}{R_{pv} + R_a}$$

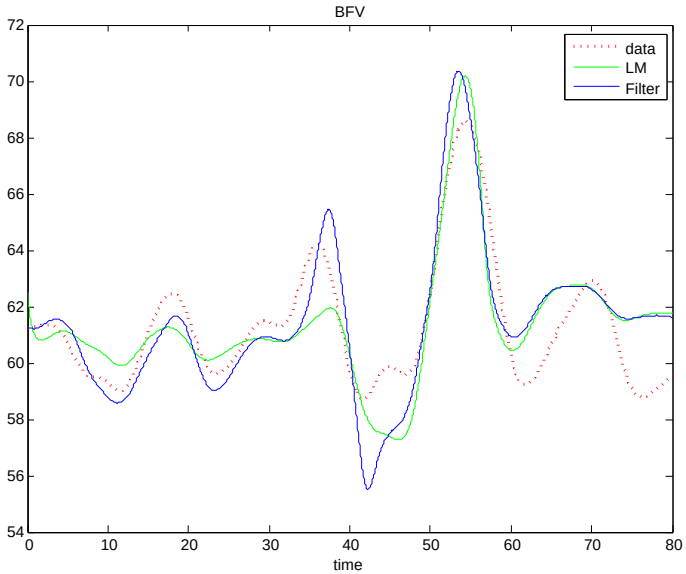
$$\frac{dC_a}{dt} = \frac{1}{\tau}\left[-C_a + \sigma(x)\right]$$

$$\sigma(Gx) = \frac{(C_{an} + \Delta C_a/2) + (C_{an} - \Delta C_a/2)e^{Gx/k_\sigma}}{1 + e^{Gx/k_\sigma}}$$

$$\hat{v}(t_i) = \frac{1}{A_c R_{pv}} \left[\frac{(p_a - p_{ic})^3}{p_a^2 - 2p_a p_{ic} + p_{ic} + \frac{k_R}{C_a^2 R_{pv}}} \right]$$

Filtered Results Using Clinical Data

For the cardiovascular model, sensitivity and subset selection reveal that only 5 parameters are most identifiable and sensitive (locally) (out of 13 parameters).



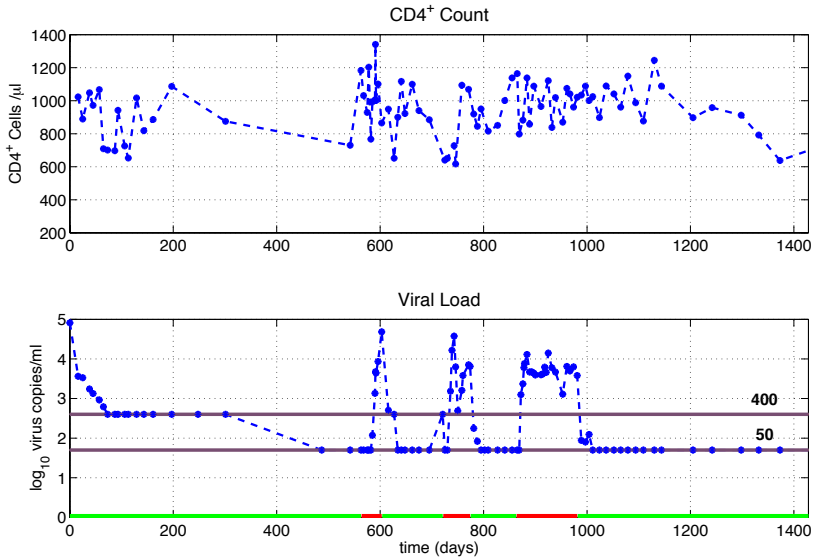
	NLS	Filter
SSR	3.578e3	3.8024e3
Ke (1 std)	.117	0.13 (0.335)
τ (1 std)	10.05	12.28 (0.334)
G (1 std)	2.14	1.89 (0.334)
C_{an} (1 std)	0.176	0.19 (0.334)
kR (1 std)	4.12e4	4.84e4 (0.334)

Figure: Plot of a patient blood flow velocity (model versus clinical data).

Clinical data: Dr. Vera Novak, Director, Syncope and Falls in the Elderly laboratory at Beth Israel Deaconess Medical Center and Harvard Medical School.

NC STATE UNIVERSITY

Conclusions



$$E\{x(k)|Z_{nc}(k), Z_c(k) \in \mathcal{Z}\}$$

Reference: B. Ibarz-Gabardos, *A Kalman filter with censored data*, IEEE, 2005.

NC STATE UNIVERSITY