Quantitative Kinetic Models from Intravital Microcopy: A Case Study Using Hepatic Transport

Meysam Tavakoli,[†] Konstantinos Tsekouras,[‡] Richard Day,[¶] Kenneth W. Dunn,[§] and Steve Pressé[*,‖]

[†]*Dept. of Physics, Indiana University-Purdue University, Indianapolis, IN, USA*

[‡]*Dept. of Physics, Arizona State University, Tempe, AZ, USA*

[¶]*Dept. of Cellular and Integrative Physiology, Indiana University School of Medicine, Indianapolis, IN, USA*

[§]*Dept. of Medicine and Biochemistry, Indiana University School of Medicine, Indianapolis, IN, USA*

[‖]*Center for Biological Physics, Department of Physics, Arizona State University, Tempe, AZ 85287, USA and School of Molecular Sciences, Arizona State University, Tempe, AZ 85287, USA*

E-mail: spresse@asu.edu

In this supplement, we present technical details expanding upon the method presented in the main text. These include: i) additional details on the theoretical approaches used; and ii) a complete description of the inference framework developed.

APPENDIX A: DETAILED METHODS DESCRIPTION

In this section, we show the details of our adaptation of the method of cascades. The basis chosen to expand the vector, $\mathbf{x}$, in applying cascade method must have the capacity to approximate ODE solutions, as well as derivatives involved in ODEs. Ramsay and Silverman[1] discussed how to choose a basis function system that guaranteed adequate flexibility to represent the variation in the approximated function $\widehat{x}_i$ and its derivatives. In our case, we select B-splines as our basis functions because they allow appropriate control over the smoothness of the solution at specific times, including any discontinuities in the first or higher order derivatives associated with step and point changes in the inputs.[2–6]

From a mathematical viewpoint, the spline is a numeric function consisting of polynomial functions each of which describes one portion of the time trace; the time trace locations where these polynomial pieces connect are called knots.[7–9] De Boor[6] showed how to improve the spline approximation accuracy and efficiency by carefully selecting each knot.

Our approach centers on estimating parameters from two parameter classes (structural parameters and nuisance parameters) in relatively different ways by the use of a multi-criterion optimization method. In this estimation procedure, there are two optimization steps: inner and outer optimizations, in which nuisance and structural parameters are, respectively, estimated. Iteration between these inner and the outer steps continues until some convergence criterion is achieved. We thus acquire good estimates of the nuisance parameters and, more importantly, of the structural parameters, which include both the measurement parameters and kinetic rates we are interested in.

The Inner Optimization

In the inner step, the structural parameter vector, $\boldsymbol{\theta}'$, is kept fixed and we try to find an estimate, $\widehat{\mathbf{c}}$, for the basis expansion coefficients $\mathbf{c}$ (the nuisance parameters). Since we need the curve fitted to the data to also satisfy the ODEs $\frac{d\mathbf{x}(t)}{dt} = \mathbf{f}(\mathbf{x}, t|\boldsymbol{\theta})$, our estimated smoothed functions $\widehat{x}_i$ should also satisfy the ODEs, $\frac{d\widehat{\mathbf{x}}(t)}{dt} = \mathbf{f}(\widehat{\mathbf{x}}, t|\boldsymbol{\theta})$, and it is natural to express the penalty term (PEN) with respect to the differential operator vector $\mathbf{L}\{\widehat{\mathbf{x}}(t)\} = \frac{d\widehat{\mathbf{x}}(t)}{dt} - \mathbf{f}(\widehat{\mathbf{x}}, t|\boldsymbol{\theta})$

$$PEN_i(\widehat{\mathbf{x}}) = \int \left[L_i\{\widehat{x}_i(t)\}\right]^2 dt \tag{1}$$

where L_i is the i^{th} component of $\mathbf{L}\{\widehat{\mathbf{x}}(t)\} = \frac{d\widehat{\mathbf{x}}(t)}{dt} - \mathbf{f}(\widehat{\mathbf{x}}, t|\boldsymbol{\theta})$. Putting together all i components, we arrive at

$$PEN(\widehat{\mathbf{x}}) = \sum_i \lambda_i PEN_i(\widehat{\mathbf{x}}). \tag{2}$$

where the regularization factor, λ_i, allows us to differently weight each component of $\widehat{\mathbf{x}}(t)$. Thus, the fitting criterion for estimating the smooth curve is given by minimizing the penalized sum of squares

$$J_{in}(\mathbf{c}|\boldsymbol{\theta}', \boldsymbol{\lambda}, \mathbf{y}) = \sum_{i=1}^{n} |y_i - \widehat{x}_i|^2 + PEN(\widehat{\mathbf{x}}). \tag{3}$$

In the general case, if we have d differential equations and r measurements at any given time, the fitting criterion can be generalized to

$$J_{in}(\mathbf{c}|\boldsymbol{\theta}', \boldsymbol{\lambda}, \mathbf{y}) = \sum_{j=1}^{r} \omega_j \sum_{i=1}^{n} |y_j(t_{ij}) - \widehat{x}_j(t_{ij})|^2 + \sum_{k=1}^{d} \lambda_k PEN_k = \sum_{j=1}^{r} SSE_j + \sum_{k=1}^{d} \lambda_k PEN_k \tag{4}$$

where y_j is intended to indicate the vectors of measured values at the time t_{ij}. The norm notation $|.|^2$ is used here to represent a sum of squared error (SSE) measures of fit and ω_j is the normalization factor. We note that in our case all observables are fluorescence intensities and thus normalizing weights are not needed, unlike the general case where observables with different units would require use of such weights.

The second term in Eq. (3) expresses the penalty term when assessing the fidelity of $\widehat{x}_j$ to the ODE system. Small λ_k reduce the importance of smoothing (i.e., the second term in Eq. (3)). As is common in regularization problems, the value of λ_k is normally adjusted based on the expected noise level of the data.[10] In fact, we expect λ_k to be directly proportional to the constant variance of the data assumed.

By minimizing the criterion $J_{in}(\mathbf{c}|\boldsymbol{\theta}', \boldsymbol{\lambda}, \mathbf{y})$ in Eq. (3), we find the estimate $\widehat{\mathbf{c}}$ of the nuisance coefficient vector $\mathbf{c}$. In other words, $\widehat{\mathbf{c}}$ can be treated as a function of the ODE parameter vector $\boldsymbol{\theta}'$. When our ODEs are linear, $\widehat{\mathbf{c}}$ is explicit and when the ODEs are nonlinear, the examples of nonlinear ODEs using the method of cascades have previously been dealt with in Ref.[11]

The Outer Optimization

In the outer optimization step, $J_{out}(\boldsymbol{\theta}'|\boldsymbol{\lambda}, \mathbf{y})$, is optimized with respect to the structural parameters, $\boldsymbol{\theta}'$, alone. The dependency of J_{out} on $\boldsymbol{\theta}'$ has two parts: directly, and implicitly through the involvement of $\widehat{\mathbf{c}}(\boldsymbol{\theta}', \boldsymbol{\lambda})$ in describing the fit $\widehat{\mathbf{x}}$. Since $\widehat{\mathbf{c}}(\boldsymbol{\theta}', \boldsymbol{\lambda})$ is regularized already, $J_{out}(\boldsymbol{\theta}'|\boldsymbol{\lambda}, \mathbf{y})$ does not need further regularization and is a straightforward measure of fit.

The structural parameter vector $\boldsymbol{\theta}'$ values are varied in the outer optimization, and are required to minimize the relation

$$J_{out}(\boldsymbol{\theta}'|\boldsymbol{\lambda}, \mathbf{y}) = \sum_{j=1}^{r} \omega_j \sum_{i=1}^{n} |y_j(t_{ij}) - \widehat{x}_j(t_{ij}|\boldsymbol{\theta}')|^2 \tag{5}$$

which calculates only the squared distance between the data and the smoothed curve (a simple χ^2 metric). We use the notation $\widehat{x}_j(t_i|\boldsymbol{\theta}')$ to emphasize that the fit to variable j at this stage is actually a function of $\boldsymbol{\theta}'$, since each time $\widehat{x}_j(t_i|\boldsymbol{\theta}')$ is re-estimated, $\boldsymbol{\theta}'$ changes. For the optimization of Eq. (3) involving a non-linear penalty term, we use the Newton-

Raphson method and, within Newton-Raphson, we compute gradients and Hessian matrices analytically. For more information see.[11–15]

APPENDIX B: IDENTIFIABILITY

Parameter Identifiability

We ask here if it is possible to find a unique solution for each of the unknown parameters of our ODE model, from data collected from experiments.[16,17]

By using a method detailed in Ref.,[18] we first find unidentifiable parameters and fix their value based on additional pieces of information. For pedagogical reasons, before explaining how we applied this method to your full hepatic transport model, we demonstrate the method on a simple two states system with two rates and a measurement model provided by Eq. (5) from the main text. Fig. (1) in the main text shows a schematic of this system. This system is described by the Eqs. (6) and (7) with parameter vector $\boldsymbol{\theta}' = [k_+, k_-, \alpha, \beta]$.

$$\begin{cases} \dfrac{dX_1}{dt} = -k_+ X_1 + k_- X_2 \\ \\ \dfrac{dX_2}{dt} = k_+ X_1 - k_- X_2 \end{cases} \tag{6}$$

with measurements

$$\begin{pmatrix} y_{X_1} \\ y_{X_2} \end{pmatrix} = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix} \tag{7}$$

We first Laplace transform Eq. (6) which yields

$$\begin{cases} s\widehat{X_1}(s) - X_1(0) = -k_+ \widehat{X_1}(s) + k_- \widehat{X_2}(s) \\ \\ s\widehat{X_2}(s) - X_2(0) = k_+ \widehat{X_1}(s) - k_- \widehat{X_2}(s) \end{cases} \tag{8}$$

$$\Rightarrow \begin{cases} \widehat{X_1}(s) = \dfrac{k_- \widehat{X_2}(s)}{s + k_+} + \dfrac{X_1(0)}{s + k_+} \\ \\ \widehat{X_2}(s) = \dfrac{k_+ \widehat{X_1}(s)}{s + k_-} + \dfrac{X_2(0)}{s + k_-} \end{cases} \tag{9}$$

From Eq. (9) we can isolate $\widehat{X_1}(s)$. From this, we can write $\widehat{y}_{X_1}(s)$ as follows

$$\begin{aligned} &\Rightarrow \widehat{X_1}(s) = \frac{k_- X_2(0)}{s^2 + (k_+ + k_-)s} + \frac{(s + k_-)X_1(0)}{s^2 + (k_+ + k_-)s} \\ &\Rightarrow \widehat{y}_{X_1}(s) = \frac{\alpha k_- X_2(0)}{s^2 + (k_+ + k_-)s} + \frac{\alpha(s + k_-)X_1(0)}{s^2 + (k_+ + k_-)s}. \end{aligned} \tag{10}$$

Insofar as identifiability is concerned, Eq. (10) captures the model structure with unknown parameter vector $\{\alpha, k_+, k_-\}$. Now, from Eq. (10), we create an analogous model structure form for the hypothetical true parameter values of the same form

$$\widehat{y}_{X_1}(s) = \frac{\alpha^* k_-^* X_2(0)}{s^2 + (k_+^* + k_-^*)s} + \frac{\alpha^*(s + k_-^*)X_1(0)}{s^2 + (k_+^* + k_-^*)s} \tag{11}$$

where the * sign shows the true or known value of the parameter. By comparing Eqs. (10) and (11) we find

$$\Rightarrow \begin{cases} \alpha k_- = \alpha^* k_-^* \\ \\ k_+ + k_- = k_+^* + k_-^*. \end{cases} \tag{12}$$

Likewise, for the second species, $\widehat{X_2}(s)$, we similarly have

$$\begin{aligned} &\Rightarrow \widehat{X_2}(s) = \frac{k_+ k_- X_2(0)}{(s + k_-)(s^2 + (k_+ + k_-)s)} + \frac{k_+ X_1(0)}{s^2 + (k_+ + k_-)s} + \frac{k_- X_2(0)}{(s + k_-)} \\ &\Rightarrow \widehat{y}_{X_2}(s) = \frac{\beta k_- k_+ X_2(0)}{(s + k_-)(s^2 + (k_+ + k_-)s)} + \frac{\beta k_+ X_1(0)}{s^2 + (k_+ + k_-)s} + \frac{\beta k_- X_2(0)}{(s + k_-)}. \end{aligned} \tag{13}$$

Similarly, we find

$$\Rightarrow \begin{cases} \beta k_- k_+ = \beta^* k_-^* k_+^* \\ \\ \beta k_+ = \beta^* k_+^* \\ \\ \beta k_- = \beta^* k_-^* \\ \\ k_+ + k_- = k_+^* + k_-^*. \end{cases} \tag{14}$$

By using either Eqs. (12) or (14) we find that in order to uniquely express all the starred quantities in terms of unstarred quantities we need to pre-specify the value of either α or β.

We can now apply the same procedure to our full hepatic transport model described by Eqs. (15), (16) and Eq. (17).

$$\begin{cases} \dfrac{dS}{dt} = -k_{S\to H} S + k_{H\to S} H \\ \\ \dfrac{dH}{dt} = k_{S\to H} S - (k_{H\to S} + k_{H\to C}) H - k_{HT} H \\ \\ \dfrac{dC}{dt} = k_{H\to C} H - k_C C \end{cases} \tag{15}$$

and

$$\begin{cases} \dfrac{dS'}{dt} = -k'_{S\to H} S' + k'_{H\to S} H' \\ \\ \dfrac{dH'}{dt} = k'_{S\to H} S' - (k'_{H\to S} + k_{H\to C}) H' + k_{HT} H \\ \\ \dfrac{dC'}{dt} = k'_{H\to C} H' - k'_C C'. \end{cases} \tag{16}$$

With the measurement model

$$\begin{pmatrix} y_S \\ y_H \\ y_C \end{pmatrix} = \begin{pmatrix} \alpha & 0 & 0 & \alpha' & 0 & 0 \\ 0 & \beta & 0 & 0 & \beta' & 0 \\ 0 & 0 & \gamma & 0 & 0 & \gamma' \end{pmatrix} \begin{pmatrix} S \\ H \\ C \\ S' \\ H' \\ C' \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}. \tag{17}$$

Laplace transforming Eqs. (15) and (16) gives

$$\begin{cases} s\widehat{S}(s) - S(0) = -k_+\widehat{S}(s) + k_-\widehat{H}(s) \\ \\ s\widehat{H}(s) - H(0) = k_+\widehat{S}(s) - (k_- + k_2 + k_{HT})\widehat{H}(s) \\ \\ s\widehat{C}(s) - C(0) = k_2\widehat{H}(s) - k_{3l}\widehat{C}(s) \end{cases} \tag{18}$$

$$\begin{cases} s\widehat{S'}(s) - S'(0) = -k'_+\widehat{S'}(s) + k'_-\widehat{H'}(s) \\ \\ s\widehat{H'}(s) - H'(0) = k'_+\widehat{S'}(s) - (k'_- + k'_2)\widehat{H'}(s) + k_{HT}\widehat{H}(s) \\ \\ s\widehat{C'}(s) - C'(0) = k'_2\widehat{H'}(s) - k'_{3l}\widehat{C'}(s). \end{cases} \tag{19}$$

By re-arranging Eqs. (18) and (19), we can determine $\widehat{S}(s)$, $\widehat{H}(s)$, and $\widehat{C}(s)$ (for fluorescein)

and their primed counter-parts (for glucuronidated fluorescein). Thus, we have

$$\Rightarrow \begin{cases} \widehat{S}(s) = \dfrac{k_- \widehat{H}(s)}{s + k_+} + \dfrac{S(0)}{s + k_+} \\ \\ \widehat{H}(s) = \dfrac{k_+ \widehat{S}(s)}{s + k_- + k_2 + k_{HT}} + \dfrac{H(0)}{s + k_- + k_2 + k_{HT}} \\ \\ \widehat{C}(s) = \dfrac{k_2 \widehat{H}(s)}{s + k_{3l}} + \dfrac{C(0)}{s + k_{3l}}. \end{cases} \tag{20}$$

Likewise, for the primed counter-parts, we have

$$\Rightarrow \begin{cases} \widehat{S'}(s) = \dfrac{k'_- \widehat{H'}(s)}{s + k'_+} + \dfrac{S'(0)}{s + k'_+} \\ \\ \widehat{H'}(s) = \dfrac{k'_+ \widehat{S'}(s)}{s + k'_- + k'_2} + \dfrac{k_{HT} \widehat{H}(s)}{s + k'_- + k'_2} + \dfrac{H'(0)}{s + k'_- + k'_2} \\ \\ \widehat{C'}(s) = \dfrac{k'_2 \widehat{H'}(s)}{s + k'_{3l}} + \dfrac{C'(0)}{s + k'_{3l}}. \end{cases} \tag{21}$$

By isolating $\widehat{H}(s)$ and $\widehat{H'}(s)$ from Eqs. (20) and (21), we find

$$\widehat{S}(s) = \frac{k_- H(0)}{s^2 + (k_+ + k_- + k_2 + k_{HT})s + k_+ k_2 + k_+ k_{HT}} + \frac{S(0)(s + k_- + k_2 + k_{HT})}{s^2 + (k_+ + k_- + k_2 + k_{HT})s + k_+ k_2 + k_+ k_{HT}} \tag{22}$$

$$\widehat{S'}(s) = \frac{k_{HT} k'_- \widehat{H}(s)}{s^2 + (k'_+ + k'_- + k'_2)s + k'_+ k'_2} + \frac{k'_- H(0)}{s^2 + (k'_+ + k'_- + k'_2)s + k'_+ k'_2} + \frac{S'(0)(s + k'_- + k'_2)}{s^2 + (k'_+ + k'_- + k'_2)s + k'_+ k'_2}. \tag{23}$$

Now by using the first measurement equation, $y_S = \alpha S + \alpha' S'$, by comparing Eq. (23) with the model structure for known parameters (same as we did for Eq. (11)) the we obtain

$$\Rightarrow \begin{cases} \alpha k_- = \alpha^* k_-^* \\ k_+ + k_- + k_2 + k_{HT} = k_+^* + k_-^* + k_2^* + k_{HT}^* \\ k_+ + k_- = k_+^* + k_-^* \\ \alpha' k_{HT} k_2' = \alpha'^* k_{HT}^* k_2'^* \\ \alpha'^* k_-'^* = \alpha'^* k_-'^* \\ k_+' + k_-' = k_+'^* + k_-'^*. \end{cases} \tag{24}$$

Therefore, according to Eq. (24), we find that we need to pre-specify the values of α, α' and k_{HT} to uniquely determine the other parameters.

References

(1) Ramsay, J. O. Functional data analysis. *Encyclopedia of Statistical Sciences* **2004**, *4*, 37–111.

(2) Wahba, G. *Spline Models for Observational Data*; SIAM, 1990.

(3) Wahba, G. Bayesian Confidence Intervals for the Cross-validated Smoothing Spline. *J. R. Stat. Soc. Series B Stat. Methodol.* **1983**, 133–150.

(4) Friedman, J. H. Multivariate Adaptive Regression Splines. *Ann. Stat.* **1991**, 1–67.

(5) Wand, M. P. A Comparison of Regression Spline Smoothing Procedures. *Comput. Stat.* **2000**, *15*, 443–462.

(6) De Boor, C.; De Boor, C.; Mathématicien, E.-U.; De Boor, C.; De Boor, C. *A Practical Guide to Splines*; Springer-Verlag New York, 1978.

(7) Schumaker, L. *Spline Functions: Basic Theory*; Cambridge University Press, 2007.

(8) Bartels, R. H.; Beatty, J. C.; Barsky, B. A. *An Introduction to Splines for Use in Computer Graphics and Geometric Modeling*; Morgan Kaufmann, 1995.

(9) Judd, K. L. *Numerical Methods in Economics*; MIT press, 1998.

(10) Welch, G.; Bishop, G. An Introduction to the Kalman Filter. *University of North Carolina at Chapel Hill: ACM, Inc.* **1995**,

(11) Ramsay, J. O.; Hooker, G.; Campbell, D.; Cao, J. Parameter Estimation for Differential Equations: A Generalized Smoothing Approach. *J. R. Stat. Soc. Series B Stat. Methodol.* **2007**, *69*, 741–796.

(12) Cao, J.; Ramsay, J. O. Parameter Cascades and Profiling in Functional Data Analysis. *Comput. Stat.* **2007**, *22*, 335–351.

(13) Cao, J.; Fussmann, G. F.; Ramsay, J. O. Estimating a Predator-Prey Dynamical Model with the Parameter Cascades Method. *Biometrics* **2008**, *64*, 959–967.

(14) Cao, J.; Ramsay, J. O. Generalized Profiling Estimation for Global and Adaptive Penalized Spline Smoothing. *Comput. Stat. Data Anal.* **2009**, *53*, 2550–2562.

(15) Cao, J.; Zhao, H. Estimating Dynamic Models for Gene Regulation Networks. *Bioinformatics* **2008**, *24*, 1619–1624.

(16) Wu, H.; Zhu, H.; Miao, H.; Perelson, A. S. Parameter Identifiability and Estimation of HIV/AIDS Dynamic Models. *Bull. Math. Biol.* **2008**, *70*, 785–799.

(17) Miao, H.; Xia, X.; Perelson, A. S.; Wu, H. On Identifiability of Nonlinear ODE Models and Applications in Viral Dynamics. *SIAM Rev.* **2011**, *53*, 3–39.

(18) Walter, E.; Pronzato, L.; Norton, J. *Identification of Parametric Models from Experimental Data*; Springer Berlin, 1997.