Aggregate Data and Inverse Problems

H. T. Banks

Center for Research in Scientific Computation

Center for Quantitative Sciences in Biomedicine

North Carolina State University

May 22, 2011

Populations with Aggregate Data and Uncertainty

- Consider approximation methods in estimation or inverse problems–quantity of interest is a probability distribution
- assume we have parameter ($q \in Q$) dependent system with model responses $x(t, q)$ describing population of interest
- For data or observations, we are given a set of values $\{y_l\}$ for the expected values

$$\mathcal{E}[x_l(q)|P] = \int_Q x_l(q) dP(q)$$

for model $x_l(q) = x(t_l, q)$ wrt unknown probability distribution P describing distribution of parameters q over population

- Use data to choose from a given family $\mathcal{P}(Q)$ the distribution P^* that gives best fit of underlying model to data
- Formulate ordinary least squares (OLS) problem–not essential–could equally well use a WLS, MLE, etc., approach
- Seek to minimize

$$J(P) = \sum_l |\mathcal{E}[x_l(q)|P] - y_l|^2$$

  over $P \in \mathcal{P}(Q)$

- Even for simple dynamics for x_l is an infinite dimensional optimization problem–need approximations that lead to computationally tractable schemes
- That is, it is useful to formulate methods to yield finite dimensional sets $\mathcal{P}^M(Q)$ over which to minimize $J(P)$

- Of course, we wish to choose these methods so that "$\mathcal{P}^M(Q) \to \mathcal{P}(Q)$" in some sense
- In this case we shall use *Prohorov metric* [BBPP, Bi] of weak star convergence of measures to assure the desired approximation results

General theoretical framework is given in [BBPP] with specific results on the approximations we use here given in [BB, BP]. Briefly, ideas for the underlying theory are as follows:

- One argues *continuity of* $P \to J(P)$ on $\mathcal{P}(Q)$ with the Prohorov metric
- If Q is compact then $\mathcal{P}(Q)$ is a *complete metric space*–also *compact*–when taken with Prohorov metric
- Approximation families $\mathcal{P}^M(Q)$ are chosen so that elements $P^M \in \mathcal{P}^M(Q)$ can be found to approximate elements $P \in \mathcal{P}(Q)$ in Prohorov metric
- *Well-posedness* (existence, continuous dependence of estimates on data, etc.) obtained along with *feasible computational methods*

The data $\{y_l\}$ available (which, in general, will involve longitudinal or time evolution data) determines the nature of the problem.

- **Type I**: The most classical problem (which we shall refer to as a *Type I* problem) is one in which individual longitudinal data is available for each member in the population. In this case there is a wide statistical literature (in the context of hierarchical modeling, mixing distributions, mixed or random effects, mixture models, etc.)
  [BS, DGa1, DGa2, DG1, DG2, L1, L2, LL, Ma, SRM, S1, S2]
  which provides theory and methodology for estimating not only individual parameters but also population level parameters and allows one to investigate both intra-individual and inter-individual variability in the population and data.

- **Type II**: In what we shall refer to as *Type II* problems one has only *aggregate* or population level longitudinal data available. This is common in marine, insect, etc., *catch and release* experiments [BK] where one samples at different times from the same population but cannot be guaranteed of observing the same set of individuals at each sample time. This type of data is also typical in experiments where the organism or population member being studied is sacrificed in the process of making a single observation (e.g., certain physiologically based pharmacokinetic (PBPK) modeling [BPo, E, Po] and whole organism transport models [BK]). In this case one may still have dynamic (i.e., time course) models for individuals, but no individual data is available.

- **Type III**: Finally, the third class of problems which we shall refer to as *Type III* problems involves dynamics which depend explicitly on the probability distribution P itself. In this case one only has dynamics (*aggregate dynamics*) for the expected value

$$\bar{x}(t) = \int_Q x(t,q)dP(q)$$

of the state variable. No dynamics are available for individual trajectories $x(t,q)$ for a given $q \in Q$. Such problems arise in *viscoelasticity* and *electromagnetics* as well as biology (the HIV cellular models of Banks, Bortz and Holte [BBH]) see also [BBPP, BG1, BG2, BP, G].

- While the approximations we discuss below are applicable to all three types of problems, we shall illustrate the computational results in the context of *size-structured marine populations (mosquitofish, shrimp)* and *PBPK problems (TCE)* where the inverse problems are of Type II.
- Finally, we note that in the problems considered here, one *can not sample directly from the probability distribution* being estimated and this again is somewhat different from the usual case treated in some of the statistical literature, e.g., see [Wahba1, Wahba2] and the references cited therein.

Example 1: The Growth Rate Distribution Model and Inverse Problem in Marine Populations

- Motivating application: estimation of growth rate distributions for size-structured mosquitofish and shrimp populations.
- Mosquitofish used in place of pesticides to control mosquito populations in rice fields–Marine biologists desire to correctly predict growth and decline of mosquitofish populatio–in order to determine the optimal densities of mosquitofish to use to control mosquito populations–a mathematical model that accurately describes the mosquitofish population would be beneficial in this application, as well as in other problems involving population dynamics and age/size-structured data.

- Based on data collected from rice fields, a reasonable mathematical model would have to predict two key features that are exhibited in the data: *dispersion* and *bifurcation* (i.e., a unimodal density becomes a bimodal density) of the population density over time [BBKW, BF, BFPZ].
- *Growth rate distribution (GRD) model*, developed in [BBKW] and [BF], captures both of these features in its solutions.
- Model is a modification of the *Sinko-Streifer model* (used for modeling age/size-structured populations) which takes into account that individuals have different characteristics or behaviors.

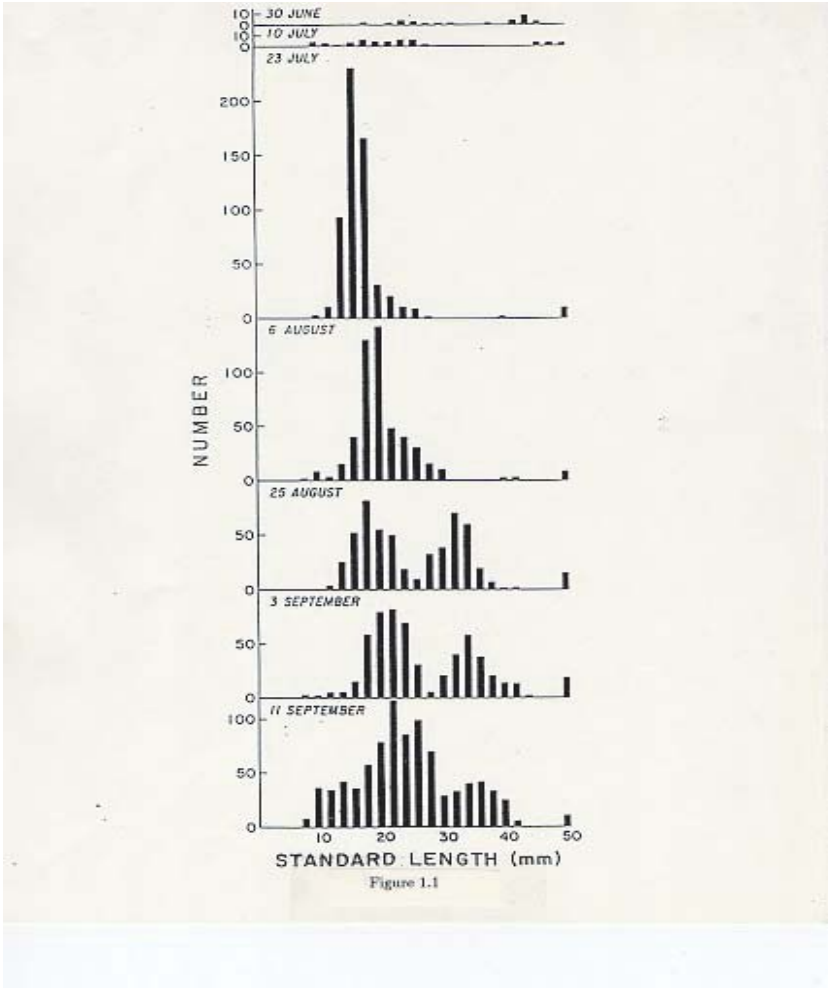


Figure 1: Mosquitofish data.

- *Sinko-Streifer model (SS)* for size-structured mosquitofish populations given by

$$\begin{aligned} \frac{\partial v}{\partial t} + \frac{\partial}{\partial x}(gv) &= -\mu v, \quad x_0 < x < x_1, \quad t > 0 \qquad (1) \\ v(0,x) &= \Phi(x) \\ g(t,x_0)v(t,x_0) &= \int_{x_0}^{x_1} K(t,\xi)v(t,\xi)\partial\xi \\ g(t,x_1) &= 0. \end{aligned}$$

Here $v(t,x)$ represents size (given in numbers per unit length) or population density, where t represents time and x represents length of mosquitofish–growth rate of individual mosquitofish given by $g(t,x)$, where

$$\frac{dx}{dt} = g(t,x) \qquad (2)$$

for each individual (all mqf of given size have same growth rate)

- In SS $\mu(t,x)$ represents *mortality rate* of mosquitofish–function $\Phi(x)$ represents initial size density of the population, while K represents the *fecundity kernel.* The boundary condition at $x = x_0$ is *recruitment*, or *birth rate*, while the boundary condition at $x = x_1 = x_{max}$ ensures the maximum size of the mosquitofish is x_1. The SS model cannot be used as is to model the mosquitofish population because it *does not predict dispersion or bifurcation* of the population in time under biologically reasonable assumptions [BBKW, BF].

- By modifying the SS model so that the individual growth rates of the mosquitofish *vary across the population* (instead of being the same for all individuals in the population), one obtains a model, known as the *growth rate distribution (GRD) model*–does in fact exhibit both dispersal in time and development of a bimodal density from a unimodal density (see [BF, BFPZ]).
- In (GRD) model, population density $u(t, x; P)$, discussed in [BBKW] and developed in [BF], is actually given by

$$u(t, x; P) = \int_G v(t, x; g) dP(g). \tag{3}$$

- G is collection of admissible growth rates, P is probability measure on G, and $v(t, x; g)$ is solution of (SS) with g– model assumes pop. made up of *collections of subpopulations* –individuals in same subpopulation have same growth rate

- Based on work in [BF], solutions to GRD model exhibit both dispersion and bifurcation of the population density in time. Here assume that the *admissible growth rates* g have the form

$$g(x; b, \gamma) = b(\gamma - x)$$

for $x_0 \leq x \leq \gamma$ and zero otherwise, where b is the *intrinsic growth rate* of the mosquitofish and $\gamma = x_1$ is the maximum size. This choice based on work in [BBKW], where idea of other properties related to the growth rates varying among the mosquitofish is discussed.

- Under assumption of varying intrinsic growth rates and maximum sizes, assume that b and γ are *random variables* taking values in the compact sets B and Γ, respectively. A reasonable assumption is that both are *bounded closed intervals*.

- Thus we take

$$G = \{g(\cdot; b, \gamma) | b \in B, \gamma \in \Gamma\}$$

so that G is also compact in, for example, $C[x_0, X]$ where $X = \max(\Gamma)$. Then $\mathcal{P}(G)$ is compact in the Prohorov metric and we are in framework outlined above. In illustrative examples, choose a growth rate parameterized by the intrinsic growth rate b with $\gamma = 1$, leading to a one parameter family of varying growth rates g among the individuals in the population. We also assume here that $\mu = 0$ and $K = 0$ in order to focus on only the distribution of growth rates; however, distributions could just as well be placed on μ and K.

- Next, introduce two different approaches that can be used in inverse problem for estimation of distribution of growth rates of the mosquitofish

- *First approach*, which has been discussed and used in [BF] and [BFPZ], involves the use of *delta distibutions*. We assume that probability distributions $\mathcal{P}^M$ placed on growth rates are *discrete* corresponding to a collection G^M with the form $G^M = \{g_k\}_{k=1}^M$ where $g_k(x) = b_k(1-x)$, for $k = 1, \ldots, M$. Here the $\{b_k\}$ are a *discretization* of B. For each subpopulation with growth rate g_k, there is a corresponding probability p_k that an individual is in subpopulation k. The population density $u(t, x; P)$ in (3) is then approximated by

$$u(t, x; \{p_k\}) = \sum_k v(t, x; g_k) p_k,$$

where $v(t, x; g_k)$ is the subpopulation density from (SS) with growth rate g_k. We denote this delta function approximation method as DEL(M), where M is number of elements used in this approximation.

- While it has been shown that DEL(M) provides a reasonable approximation to (3), a better approach might involve techniques that will provide a smoother approximation of (3) in the case of continuous probability distributions on the growth rates. Thus, as a *second approach*, we chose to use an approximation scheme based on piecewise linear splines. Here we have assumed that P is a continuous probability distribution on the intrinsic growth rates. We approximate the density $P' = \frac{dP}{db} = p(b)$ using piecewise linear splines, which leads to the following approximation for $u(t, x; P)$ in (3):

$$u(t, x; \{a_k\}) = \sum_k a_k \int_B v(t, x; g) l_k(b) db,$$

where $g(x; b) = b(1 - x)$, $p_k(b) = a_k l_k(b)$ is the probability density for an individual in subpopulation k and l_k represents the piecewise linear spline functions.

- This spline based approximation method is denoted by SPL(M,N), where M is the number of basis elements used to approximate the growth rate probability distribution and N is the number of quadrature nodes used to approximate the integral in the formula above. One can use the composite trapezoidal rule for the approximation of these integrals [QSS].

- One can use the approximation methods DEL(M) and SPL(M,N) in the inverse problem for the estimation of the growth rate distributions. The least squares inverse problem to be solved is

$$\begin{aligned} \min_{P \in \mathcal{P}^M(G)} J(P) &= \sum_{i,j} |u(t_i, x_j; P) - \hat{u}_{ij}|^2 \qquad (4) \\ &= \sum_{i,j} (u(t_i, x_j; P))^2 - 2u(t_i, x_j; P)\hat{u}_{ij} + (\hat{u}_{ij})^2, \end{aligned}$$

where $\{\hat{u}_{ij}\}$ is the data and $\mathcal{P}^M(G)$ is the finite dimensional approximation to $\mathcal{P}(G)$. When using DEL(M), the finite dimensional approximation $\mathcal{P}^M(G)$ to the probability measure space $\mathcal{P}(G)$ is given by

$$\mathcal{P}^M(G) = \left\{ P \in \mathcal{P}(G) |\ P' = \sum_k p_k \delta_{g_k},\ \sum_k p_k = 1 \right\},$$

where δ_{g_k} is the delta function with an atom at g_k. However,

when using SPL(M,N), the finite dimensional approximation $\mathcal{P}^M(G)$ is given by

$$\mathcal{P}^M(G) = \left\{ P \in \mathcal{P}(G) |\ P' = \sum_k a_k l_k(b),\ \sum_k a_k \int_B l_k(b) db = 1 \right\}.$$

- Furthermore, we note that this least squares inverse problem (4) becomes a quadratic programming problem [BF, BFPZ]. Letting $\mathbf{p}$ be the vector that contains $p_k, 1 \leq k \leq M$, when using DEL(M) or $a_k, 1 \leq k \leq M$, when using SPL(M,N), we let $\mathbf{A}$ be the matrix with entries given by

$$A_{km} = \sum_{i,j} v(t_i, x_j; g_k) v(t_i, x_j; g_m),$$

$\mathbf{b}$ the vector with entries given by

$$b_k = -\sum_{i,j} \hat{u}_{ij} v(t_i, x_j; g_k),$$

and

$$c = \sum_{i,j} (\hat{u}_{ij})^2,$$

where $1 \leq k, m \leq M$. In the place of (4), we now minimize

$$F(\mathbf{p}) \equiv \mathbf{p}^T \mathbf{A}\mathbf{p} + 2\mathbf{p}^T \mathbf{b} + c \tag{5}$$

over $\mathcal{P}^M(G)$. We note when using DEL(M) we also had to include the constraint

$$\sum_k p_k = 1,$$

while when using SPL(M,N) we had to include the constraint

$$\sum_k a_k \int_B l_k(b)db = 1.$$

However, in both cases, we were able to include these constraints along with non-negativity constraints on the $\{p_k\}$ and $\{a_k\}$ in the programming of these two inverse problems.

Other Size-Structured Population Models

The Sinko-Streifer (SS) model [SS] and its variations have been widely used to describe numerous age and size-structured populations (see [BBDS1, BBDS2, BBKW, BF, BFPZ, BeAnd, Kot, Metz] for exMPLE).

- Shrimp growth models Dispersion in size observed in experimental data (Figures 2 and 3) for early growth of shrimp–data: different raceways at Shrimp Mariculture Research Facility, Texas Agricultural Experiment Station in Corpus Christi, TX. Initial sizes were very similar–variability observed in aggregate type longitudinal data. Reasonable model for population must account for variability in size distribution data–perhaps a result of variability in individual growth rates across population [CSLFJ]. Models developed in [Shrimp, Shrimp-exp].

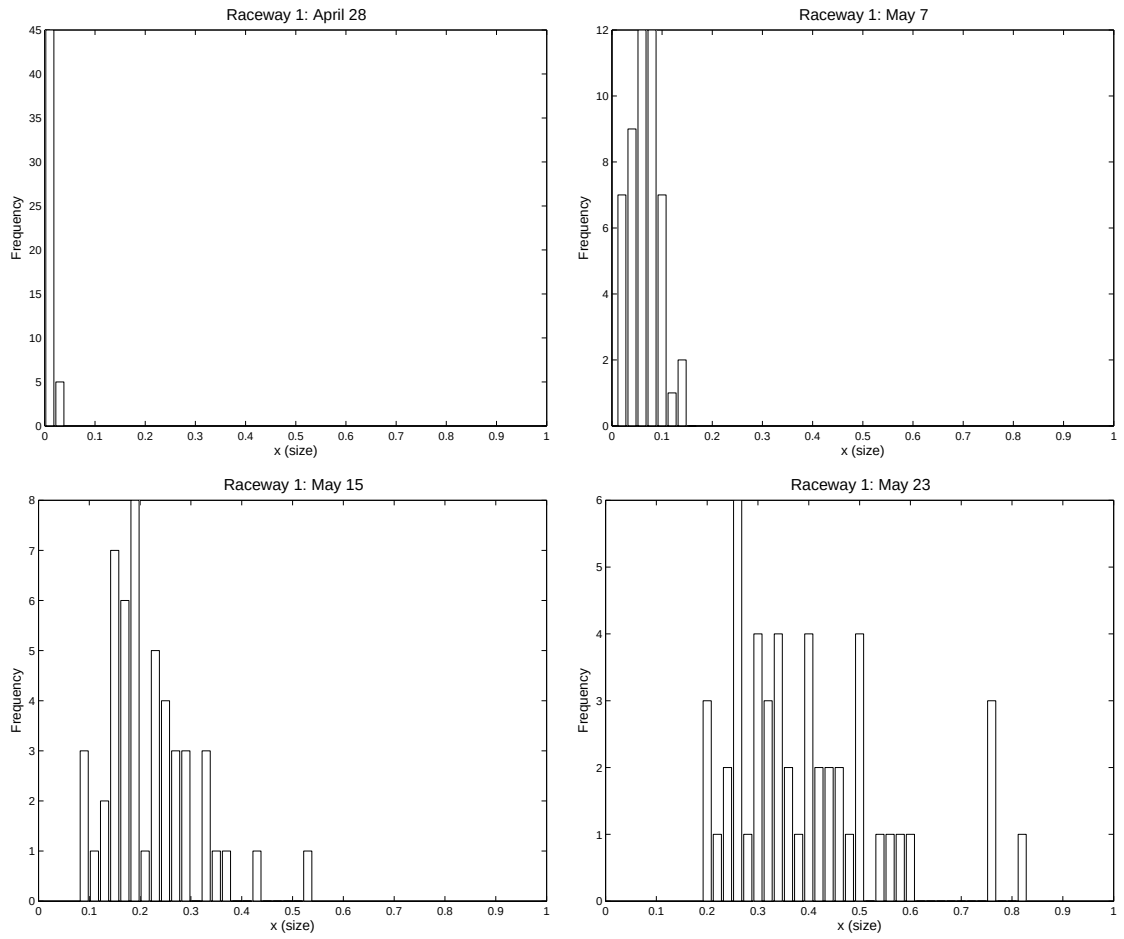


Figure 2: Histograms for longitudinal data for size (in grams) for Raceway 1.

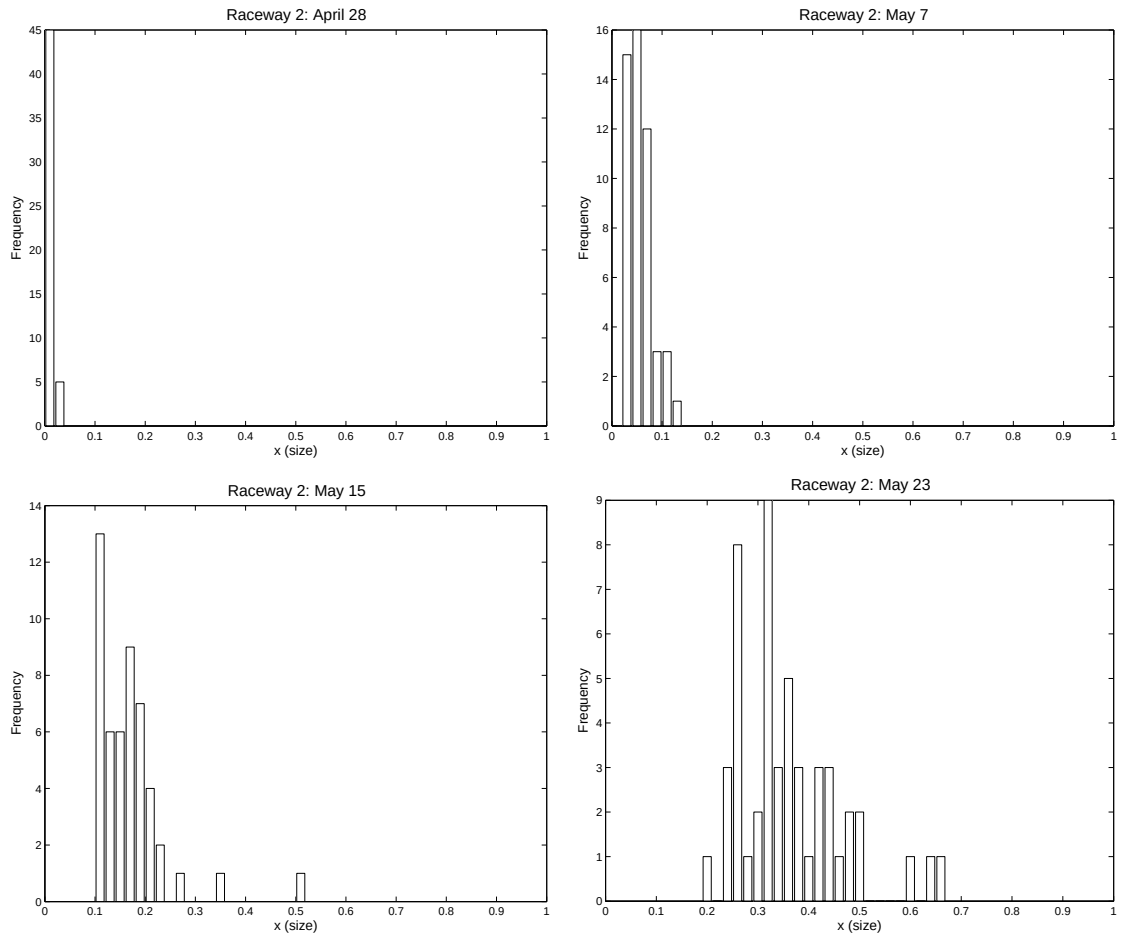
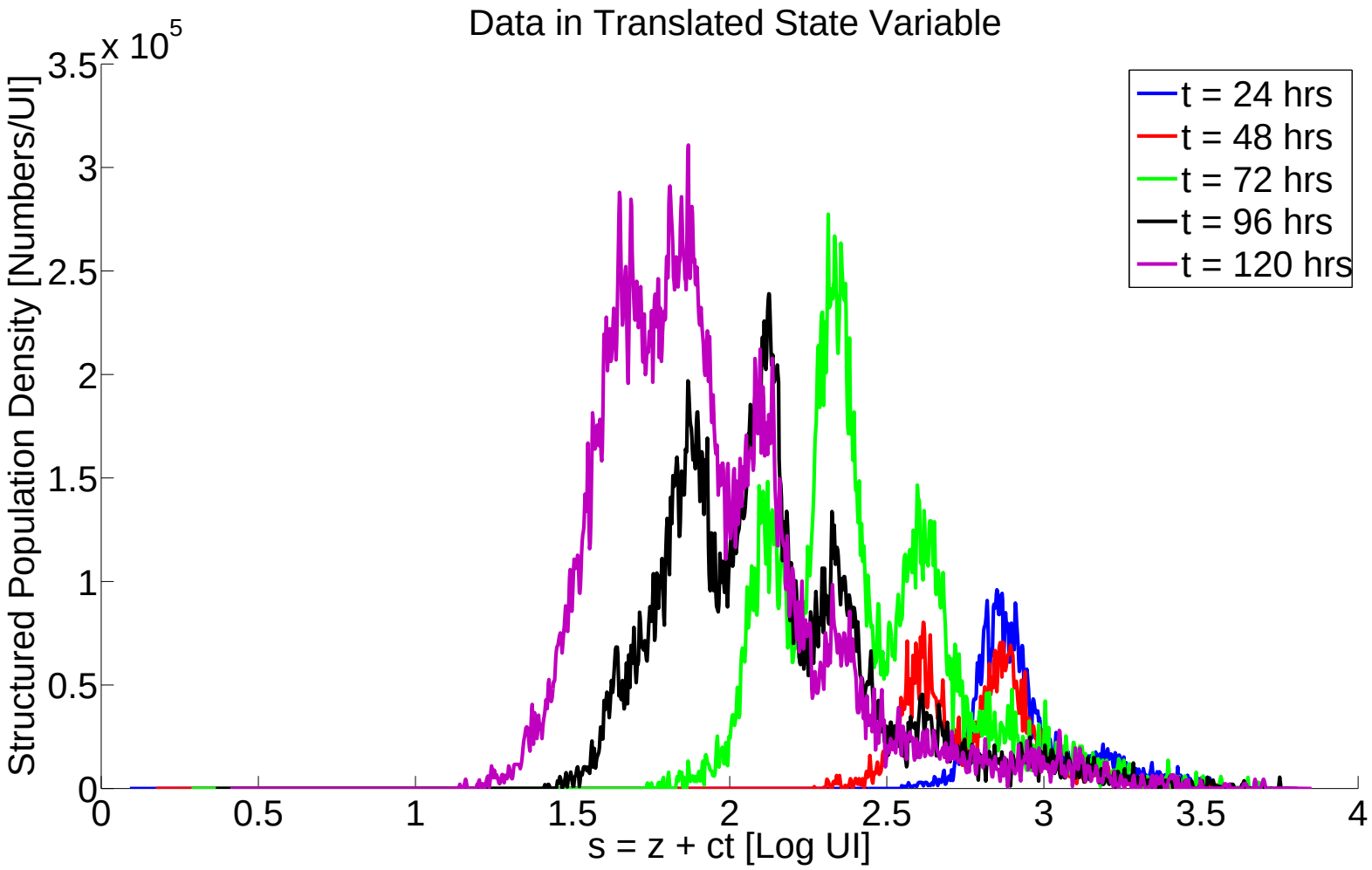


Figure 3: Histograms for longitudinal data for size (in grams) for Raceway 2.

- Florescent labeling in cell population growth More recently, extensions of these models have been employed in cell population models where size is replaced by intensity of a label or marker [Tat1, BTSBRSM].
- Original data sets shown below in ***translated log intensity*** $s = z + ct$. Note subsequent division peaks now strongly correlated with specific regions in the state variable, unlike when plotted vs. original log intensity variable z


Data in Translated State Variable
x 10^5
3.5
3
2.5
2
1.5
1
0.5
0
Structured Population Density [Numbers/UI]
0
0.5
1
1.5
2
2.5
3
3.5
4
s = z + ct [Log UI]
t = 24 hrs
t = 48 hrs
t = 72 hrs
t = 96 hrs
t = 120 hrs

Mathematical Model

- Model: dynamics of life/death process–population of CFSE-labelled cells–proposed by Bochorov et al. (2007)–variation of Bell-Anderson / Sinko-Streifer (1967) population models: x = CFSE FI (in units of intensity, UI) of a cell; $n(t,x)$ = label-structured population density (cells/UI) with FI x at time t–Population density given by

$$\begin{aligned} \frac{\partial n}{\partial t}(t,x) \quad &+ \quad \frac{\partial [v(x)n(t,x)]}{\partial x} = -(\alpha(x)+\beta(x))n(t,x) \\ &+ \quad \chi_{[x_{\min}, x_{\max}/\gamma]} 2\gamma\alpha(\gamma x)n(t,\gamma x), \end{aligned} \tag{6}$$

$v(x)$ = label loss rate, $\alpha(x)$ = cell proliferation rate, $\beta(x)$ = cell death rate, $x \in [x_{\min}, x_{\max}]$ and $t > 0$.

- cells naturally lose FI over time even in absence of division (due to catabolic activity), $v(x)$ represents natural label loss rate (UI/hr)– parameter γ is label dilution factor i.e., ratio of FI of a mother cell to FI of a daughter cell
- FACS returns data on logarithmic scale, make change of variables $z = \log_{10} x$

Resulting model when $v(x) = -cx$ (which we assume here) and $\tilde{n}(t,z) \equiv n(t,10^z)$ is

$$\begin{aligned} \frac{\partial \tilde{n}}{\partial t}(t,z) \quad + \quad & \frac{\partial[\tilde{v}(z)\tilde{n}(t,z)]}{\partial z} = -(\tilde{\alpha}(z)+\tilde{\beta}(z))\tilde{n}(t,z) \\ + \quad & \chi_{[z_{\min}, z_{\max}-\log_{10}\gamma]} 2\gamma\tilde{\alpha}(z+\log_{10}\gamma)\tilde{n}(t,z+\log_{10}\gamma), \end{aligned} \tag{7}$$

where $\tilde{v}(z) = -\tilde{c} = -c/ln10$, and $\tilde{\alpha}, \tilde{\beta}$ are appropriately defined cell proliferation and death rates, respectively.

In summary, the GRD model (3) represents one approach to accounting for variability in growth rates by imposing a probability distribution on the growth rates in the SS model (1). Individuals in the population grow according to a *deterministic growth model* (2), but different individuals in the population may have different parameter dependent growth rates in the GRD model. The population is assumed to consist of *subpopulations* with individuals in the same subpopulation having the same growth rate. The growth uncertainty of individuals in the population is the result of *variability in growth rates among the subpopulations*. This modeling approach, which entails a stationary probabilistic structure on a family of deterministic dynamic systems, may be most applicable when the growth of individuals is assumed to be the result of *genetic variability.*

However, a second approach that has been studied as well is based on the assumption that individual growth is a Markov diffusion stochastic process which leads to the Fokker-Planck model for shrimp population density [Allen, BTW, Gard, Okubo]. The growth process for each individual is stochastic, and each individual grows according to a stochastic growth model. In the Fokker-Planck model, the uncertainty in the growth of individuals is the result of the growth stochasticity of each individual. This modeling approach may be most applicable when the variability in the growth rate of individuals is believed to be the result of variability in environmental factors such as discussed in [GAZ, LLS, PMR]. Theoretical arguments in [GRD-FP] demonstrate that the population density from the GRD model is same as population density obtained from the Fokker-Planck model when equivalent levels of variability are used in both models. Numerical results are also presented in [GRD-FP2] to further

validate the theoretical analysis of [GRD-FP]. Therefore, one can use the computationally "easier" approach to model the population of interest when appropriately chosen forms of variability can be determined. Based on these studies, suggest use of GRD model (3) to incorporate uncertainty in the growth rates in the size-structured population model for the early growth of shrimp. A natural question arises immediately: how to collect date to carry out the minimization to determine a reasonable value for $\mathcal{P}$. In particular, what sampling size and sampling frequency should be used in experiments to adequately estimate $\mathcal{P}$?

Example 2: PBPK Models for TCE

Uncertainty is an inherent factor in mathematical models for biological systems. Model equations themselves are an approximation of the phenomena they are designed to model, introducing a degree of uncertainty that is difficult to measure. Further simplifications and approximations of a model for theoretical and computational purposes result in additional layers of uncertainty. Moreover, many biological processes are subject to variability that may not be incorporated into a mathematical model. Experimental observations also introduce uncertainty when data are used with a model to estimate parameters.

Two types of variability that are common in biological models and are well-known in the statistical literature [DG1] are intra-individual and inter-individual variability. *Intra-individual* variability is defined as variability that occurs within a given individual organism or biological process. This type of variability may result in time-dependent and/or spatially-dependent variation within an individual. Biological examples of such variability include parameters such as body weight, blood pressure, fat content and cell membrane permeabilities.

A second type of variability that is commonly found in biological modeling is *inter-individual* variability. This type of variability results from variations in individuals across a population. Biological models that are based on behavior or phenomena over a population are almost always subject to inter-individual variability. This is especially the case when a model is designed to predict or explain

experimental observations that are collected from multiple individuals.

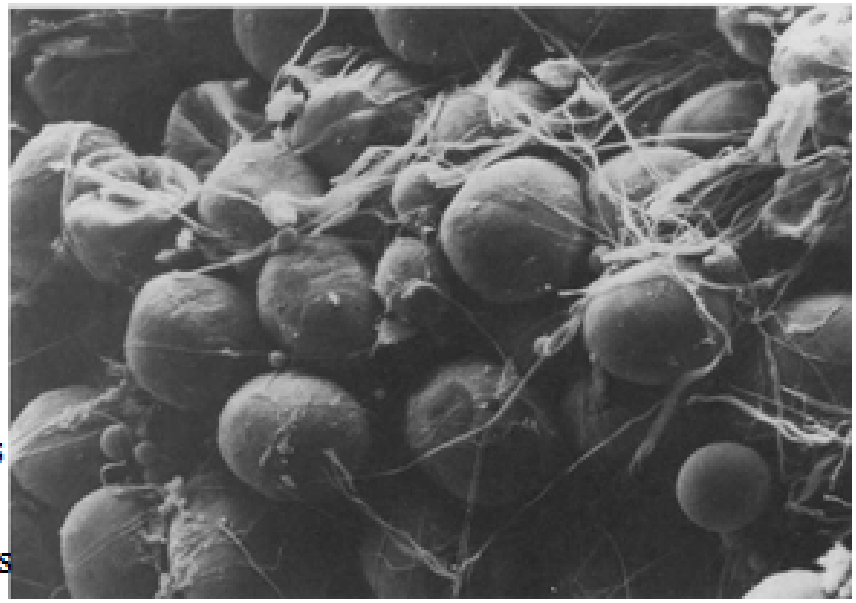
It is reasonable to expect that different individuals of a population would possess different values for biological, physical and chemical parameters. These parameters would then take on varying values across the population, so that each parameter would be associated with a probability distribution that would mathematically describe this variation. Using data from multiple individuals, one can estimate the resulting probability distributions with inverse problem techniques, thereby obtaining both the mean and variance of the uncertainty.

Examples of biological parameters that are often subject to inter-individual variability include growth and death rates, susceptibility to infection, efficacy of vaccines and other prophylactics, and age. Note that each of the examples given above

for intra-individual variability may also involve inter-individual variability depending on the type of model and experimental observations. Similarly, each of the examples for inter-individual variability also may be subject to intra-individual variability.

The next motivating example we consider here is a toxicokinetic model for the systemic transport of the *environmental contaminant trichloroethylene (TCE)*. TCE is a solvent that has been used widely in industry as a metal degreasing agent, and is now a common soil and groundwater contaminant. This highly fat-soluble compound is rapidly absorbed into the bloodstream, and has been shown to accumulate in the adipose (fat) tissue of humans and animals. Known and suspected toxic effects of TCE and its metabolites in laboratory animals and/or humans include acute effects such as dizziness, drowsiness, headaches and fatigue, as well as chronic effects such as developmental defects and lung, kidney and liver tumors.

PBPK Models for TCE in Fat Cells



Millions of cells with varying size, residence time, vasculature, geometry:
"Axial-dispersion" type adipose tissue compartments to embody uncertain physiological heterogenieties in single organism (rat) = intra-individual variability

Inter-individual variability treated with parameters (including dispersion parameters) as random variables –estimate distributions from aggregate data (multiple rat data) which also contains uncertainty (noise)

Toxicokinetic models are used in the overall risk assessment process for toxic compounds to help quantify the expected risk of toxicity to humans as a function of the level of exposure to the given chemical. In particular, physiologically based pharmacokinetic (PBPK) models predict the effective dose level of a toxic compound that is delivered to the "target" tissues (i.e., tissues that experience toxic effects) for a given external exposure level. PBPK models are compartmental models that describe the systemic transport of a compound through the tissues and organs, including the dynamics of uptake, tissue distribution, metabolism and elimination. The resulting model is a system of differential equations, with each equation representing the dynamics of tissue concentrations in a particular tissue or organ.

In [ABEP] and [BPo], three PBPK models for TCE are developed and compared, each with a different submodel for the adipose tissue compartment.

As discussed in [ABEP], preliminary simulations indicated that a perfusion-limited adipose tissue compartment does not appear to sufficiently capture the dynamics of TCE accumulation in fat as seen in experimental data. Moreover, adipose tissue is known to have highly heterogeneous physiological properties, including significant variations in fat cell size, lipid distribution, blood flow rates and cell membrane permeabilities. These characteristics further suggest that the "well-mixed," rapid equilibrium assumptions of the perfusion-limited model may be inappropriate for describing the disposition of fat-accumulating compounds such as TCE in adipose tissue.

To better capture the dynamics of TCE in fat tissue, a spatially varying axial dispersion model was developed [ABEP] to address the *intra-individual* variability that results from the heterogeneous lipid distribution and physiological characteristics of adipose tissue. This variability is built into the adipose compartmental model with a special axial dispersion term, where the "dispersion" coefficient is a measure of the degree of intra-individual variability that occurs in the fat.

In addition to the intra-individual variability that appears to affect TCE concentrations in fat tissue, *inter-individual* variability also plays a major role in toxicokinetic models in general. Experimental techniques that require measurements of chemical concentrations in tissues over time be taken from multiple individuals immediately introduces inter-individual variability into the measured observations, and must be considered in the development of mathematical models.

To further amplify on our discussions of intra- and inter-individual variability as used in this paper, we note that neither of these refer specifically to *cellular* variability here. The PBPK models we use are standard compartmental models which have already incorporated variability across cells (even though the resulting models are deterministic). While this is a usual practice in most PBPK model formulations, one could reasonably argue against the use of deterministic models for expected blood or tissue concentrations in a given individual (patient or test subject in clinical trials). There are in fact a number of ways to introduce cellular variability in one individual. In addition to usual stochastic dynamics, one might also consider probability measure dependent dynamical systems

$$\dot{x}(t) = f(t, x(t), P)$$

for the expected value x of the blood/tissue concentrations. Here the expected value is taken over an ensemble of cells. Examples of such

models can be found in the HIV cellular models of [BBH, BBPP], the molecular level modeling of biotissue in [BP] and of polymers in [BMP] along with the discussions of individual vs. aggregate dynamics estimation problems in [BBPP]. The Prohorov framework outlined below as developed in [BB] can also be used (see [BBPP]) to develop a theory for such aggregate dynamics measure-dependent models. We do not do that here. Rather, as described above and detailed below, we employ deterministic models (i.e., systems with deterministic parameters) involving spatially varying axial dispersion formulations to capture intra-individual variability in adipose TCE concentrations (which in this case is essentially *inter-cellular* variability as well variability across space). In these models the inter-individual aspects are introduced to treat variability of the deterministic parameters across populations of individuals (patients or test subjects).

As biological models have become more widely utilized and influential in a variety of fields, the need to account for variability and uncertainty in modeling has been recognized. Markov Chain-Monte Carlo methods have been developed to address issues of variability and uncertainty, and these methods have been applied to PBPK models as a part of the parameter estimation process. Monte Carlo methods are based on a Bayesian statistical approach that involves the use of experimental data to update estimates of a hypothesized "prior" probability distribution for one or more model parameters.

An alternative, probability-based method has been developed to incorporate uncertainty and variability in mathematical models. This method, which is discussed in [B, BB, HTB-HK] is centered around a probabilistic parameter estimation approach that involves the estimation of probability distributions for model parameters. Well-known theoretical results from probability theory establish the theoretical soundness of this technique, which can be implemented computationally in a straightforward manner.

A distinct advantage of this approach over the Monte Carlo-based methods is an added level of flexibility in choosing the prior probability distributions. These probability-based methods can be used with preselected prior distributions as with Monte Carlo methods, or they may be used with weighted sums of Dirac delta measures that do not assume a fixed form for the probability distribution functions. A version of this method has been applied to

a population model for mosquitofish in rice paddies, and was used to successfully describe fish population dynamics by estimating distributed growth rate functions using aggregate experimental data [BFPZ].

Overview of the TCE model

Here we provide an overview of the PBPK-hybrid model for TCE as developed in [ABEP, Po-thesis]. This model utilizes standard physiologically based pharmacokinetic compartmental equations for various non-fat tissues. The fat tissue compartment is described with a spatially varying dispersion model, and is designed specifically to capture the *intra-individual* variability that results from the heterogeneous physiology of fat.

The most commonly used compartmental model in PBPK modeling is the perfusion-limited, or flow-limited compartment. This model is based on simple mass balance principles and assumptions of rapid equilibrium and spatial uniformity. Moreover, it is assumed that the blood flow rate to the tissue is much slower than the rate of transport of the compound across cell membranes. The resulting equation for

the tissue concentration C of the compound is given by

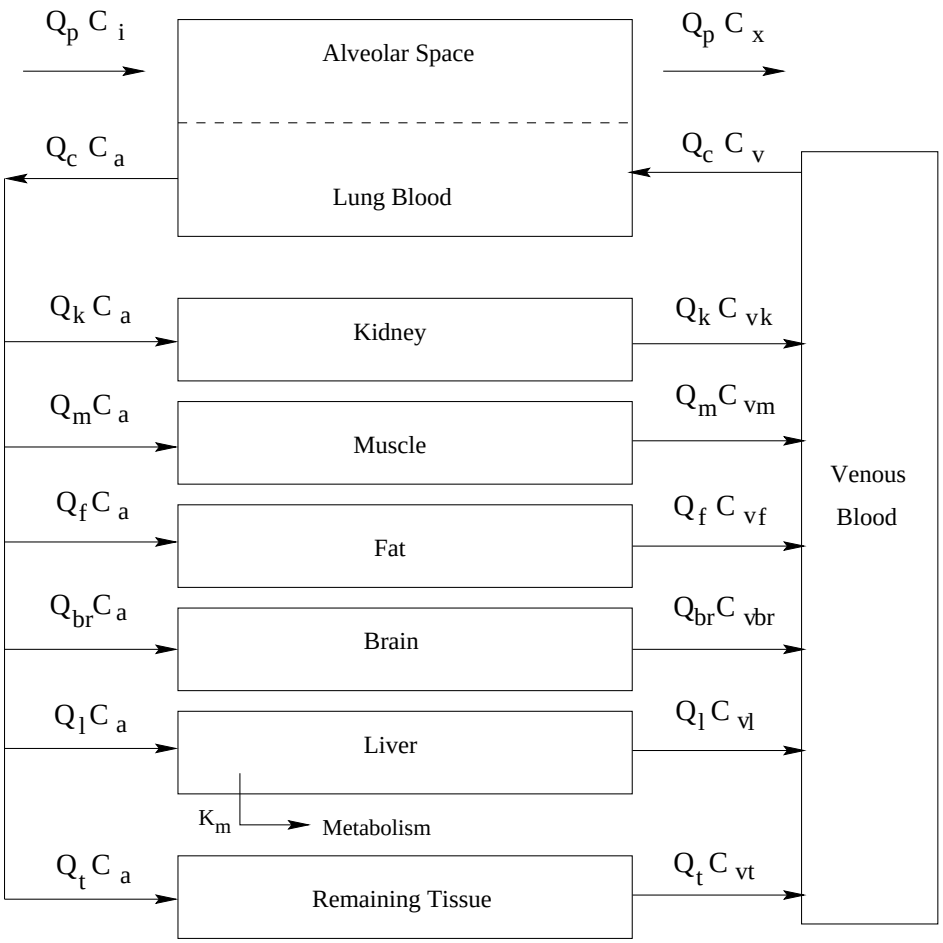
$$V\frac{dC(t)}{dt} = Q_{bl}(C_{in}(t) - C_{out}(t)),$$

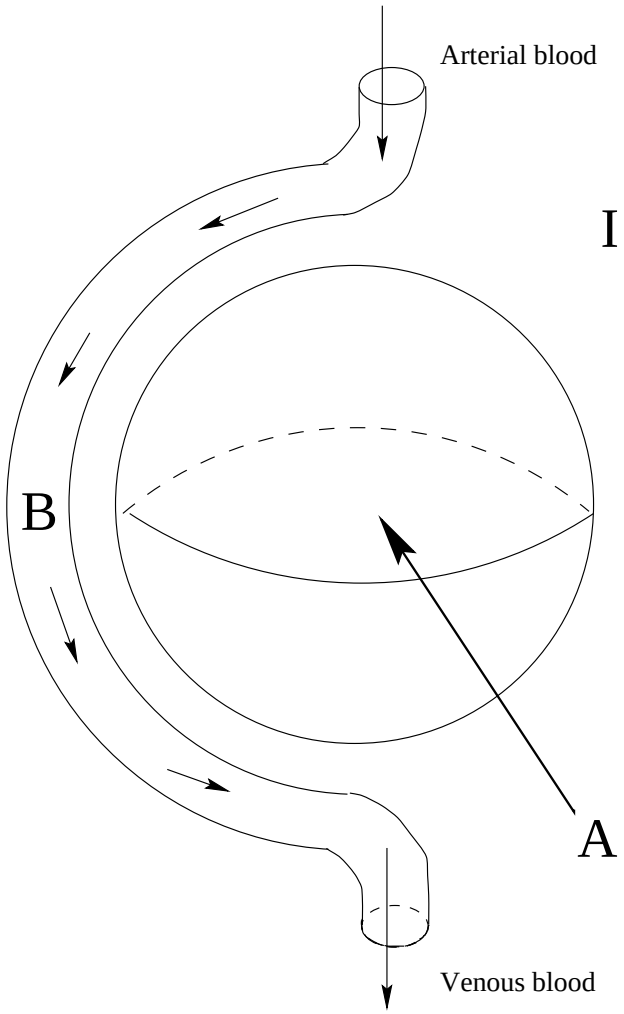
where V is the volume of the tissue, Q_{bl} is the volumetric blood flow rate to the tissue, and C_{in} and C_{out} are the concentrations of compound entering and leaving the tissue respectively. Under standard assumptions, the concentration C_{out} is equal to the concentration C of compound in the tissue divided by the blood:tissue partition coefficient.

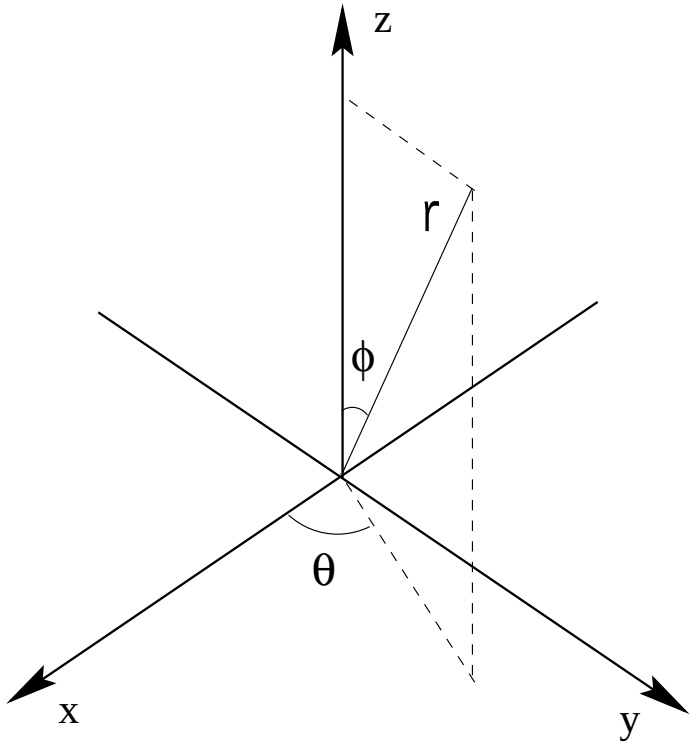
For many tissues and compounds of interest the perfusion-limited compartmental model is adequate to describe the dynamics of such compounds inside the tissues. In the case of highly lipophilic substances such as TCE, however, the standard models may not accurately capture the transport of these chemicals in the adipose tissue. The highly heterogeneous physiology of fat tissue appears to

have a major influence on the behavior of TCE in fat. Using a PBPK model for TCE in Long-Evans rats with a perfusion-limited fat compartment [Marina] (see Figure 52 for a model schematic), model simulations suggested that the standard model does not capture the concentration profile of TCE in adipose tissue as seen in experimental data [ABEP].

PBPK Model for Inhaled TCE




Arterial blood
I
B
A
Venous blood


z
r
ϕ
θ
x
y

The schematic in Figure 53 depicts a geometric representation of an adipocyte-capillary unit in adipose tissue. The adipocyte region (A) is represented by a sphere and is immersed in the interstitial fluid (I). The capillary or blood region (B) is a cylindrical tube that wraps around the adipocyte. Coordinates are in spherical coordinates (r, θ, ϕ).

To account for the spatial variation in TCE fat concentrations as suggested by the physiology of adipose tissue, an axial dispersion model was developed to replace the perfusion-limited fat tissue compartment. This model is based directly on the structure of fat tissue, which consists primarily of spherical, lipid-containing cells called adipocytes. Each adipocyte is in contact with one or more capillaries and is immersed in interstitial fluid.

The model equations are based on an axial dispersion model developed by Roberts and Rowland [RobertsRowland] for the liver.

A key feature of their model is its *aggregate structure*, using a single cellular unit with the dispersion term to represent the intra-individual variability that occurs across the millions of cells in the tissue. As detailed in [ABEP], we have adapted their model to describe the geometry of adipose tissue and the transport of TCE within the fat. The resulting system of partial differential equations is given by

$$\begin{aligned} V_B \frac{\partial C_B}{\partial t} &= \frac{V_B}{r_2 \sin\phi} \frac{\partial}{\partial \phi} \left[\sin\phi \left(\frac{\mathcal{D}_B}{r_2} \frac{\partial C_B}{\partial \phi} - v C_B \right) \right] \\ &+ \lambda_I \mu_{BI} (f_I C_I(\theta_0) - f_B C_B) \\ &+ \lambda_A \mu_{BA} (f_A C_A(\theta_0) - f_B C_B) \end{aligned} \tag{8}$$

$$-\frac{\mathcal{D}_B}{r_2}\frac{\partial C_B}{\partial \phi}(t,\phi)+vC_B(t,\phi)\Big|_{\phi=\varepsilon_1} = \frac{Q_c}{1000A_B}C_a(t) \quad (9)$$

$$-\frac{\mathcal{D}_B}{r_2}\frac{\partial C_B}{\partial \phi}(t,\phi)+vC_B(t,\phi)\Big|_{\phi=\pi-\varepsilon_2} = \frac{Q_c}{1000A_B}C_v(t) \quad (10)$$

$$\begin{aligned} V_I\frac{\partial C_I}{\partial t} &= \frac{V_I D_I}{r_1^2}\left[\frac{1}{\sin^2\phi}\frac{\partial^2 C_I}{\partial\theta^2}+\frac{1}{\sin\phi}\frac{\partial}{\partial\phi}\left(\sin\phi\frac{\partial C_I}{\partial\phi}\right)\right] \\ &+ \delta_{\theta_0}(\theta)\chi_B(\phi)\lambda_I\mu_{BI}(f_B C_B - f_I C_I) \\ &+ \mu_{IA}(f_A C_A - f_I C_I) \quad (11) \\ C_I(t,\theta,\phi) &= C_I(t,\theta+2\pi,\phi) \quad (12) \\ \frac{\partial C_I}{\partial\theta}(t,\theta,\phi) &= \frac{\partial C_I}{\partial\theta}(t,\theta+2\pi,\phi) \quad (13) \\ C_I(t,\theta,0) &< \infty \quad (14) \\ C_I(t,\theta,\pi) &< \infty \quad (15) \end{aligned}$$

$$
\begin{aligned}
V_A \frac{\partial C_A}{\partial t} &= \frac{V_A D_A}{r_0^2}\left[\frac{1}{\sin^2\phi}\frac{\partial^2 C_A}{\partial \theta^2} + \frac{1}{\sin\phi}\frac{\partial}{\partial \phi}\left(\sin\phi \frac{\partial C_A}{\partial \phi}\right)\right] \\
&+ \delta_{\theta_0}(\theta)\chi_B(\phi)\lambda_A \mu_{BA}(f_B C_B - f_A C_A) \\
&+ \mu_{IA}(f_I C_I - f_A C_A) \qquad (16) \\
C_A(t,\theta,\phi) &= C_A(t,\theta+2\pi,\phi) \qquad (17) \\
\frac{\partial C_A}{\partial \theta}(t,\theta,\phi) &= \frac{\partial C_A}{\partial \theta}(t,\theta+2\pi,\phi) \qquad (18) \\
C_A(t,\theta,0) &< \infty \qquad (19) \\
C_A(t,\theta,\pi) &< \infty. \qquad (20)
\end{aligned}
$$

The capillary equation (8) describes the transport of TCE in the

capillary region of the adipose tissue and utilizes the dispersion term

$$\frac{V_B}{r_2 \sin\phi} \frac{\partial}{\partial\phi} \left[\sin\phi \frac{\mathcal{D}_B}{r_2} \frac{\partial C_B}{\partial\phi} \right]$$

with dispersion coefficient $\mathcal{D}_B$. This term accounts for the variability in physiological properties that occurs across the population of fat cells, with a large dispersion coefficient indicating a high degree of variability. Mathematically, the dispersion term is equivalent to a standard diffusion term, although the dispersion term is used specifically to approximate the observed physiological phenomena of varying path lengths, flow velocities and compound transit times that occur within a tissue.

The boundary conditions (9) and (10) connect the adipose capillary region to the systemic arterial and venous blood compartments using flux balance. Transport of TCE between the capillary region and the other two adipose subcompartments (interstitial and adipocyte) is

modeled in the PDE (8). The variables $C_B(t)$, $C_I(t)$ and $C_A(t)$ denote concentrations of TCE in the capillary, interstitial and adipocyte regions respectively, while $C_a(t)$ and $C_v(t)$ represent the systemic arterial and venous blood concentrations of TCE.

The interstitial region is modeled with the two-dimensional PDE (11) and boundary conditions (12) – (15). The adipocyte region equations (16) – (20) are similar in structure to the interstitial equations, and describe the diffusion of TCE around the surface of the adipocyte as well as the transport of TCE between the three adipose subcompartments. The boundary conditions are standard periodic and finiteness boundary conditions that are commonly used for the diffusion equation on a spherical domain. A detailed derivation and description of the dispersion model is given in [ABEP, Po-thesis].

Adipose model parameters include the dispersion coefficient $\mathcal{D}_B$

(m^2/hour); diffusion coefficients D_I and D_A (m^2/hour); the fractions f_B, f_I, f_A of unbound TCE in each adipose region; cell membrane permeability coefficients μ_{BA}, μ_{IA}, μ_{BI} (liters/hour); blood flow parameters v (m/hour) and $\mathcal{F}$; and inter-region transport parameters λ_I and λ_A.

The adipose model equations (8) – (20) are coupled with standard compartmental equations for the lung, arterial blood, venous blood, liver, brain, kidney, muscle and remaining non-fat tissue to obtain a whole-body PBPK-hybrid model. Uptake of TCE is via inhalation into the lungs, and metabolism is modeled with a Michaelis-Menten

term in the liver. The resulting equations are given by

$$
\begin{aligned}
V_v \frac{dC_v}{dt} &= Q_m C_m/P_m + Q_t C_t/P_t + Q_f C_B(\cdot, \pi - \varepsilon_2) + Q_{br} C_{br}/P_{br} \\
&+ Q_l C_l/P_l + Q_k C_k/P_k - Q_c C_v \qquad (21) \\
C_a &= \frac{Q_c C_v + Q_p C_c}{Q_c + \frac{Q_p}{P_b}} \qquad (22) \\
V_m \frac{dC_m}{dt} &= Q_m (C_a - C_m/P_m) \qquad (23) \\
V_t \frac{dC_t}{dt} &= Q_t (C_a - C_t/P_t) \qquad (24) \\
V_{br} \frac{dC_{br}}{dt} &= Q_{br} (C_a - C_{br}/P_{br}) \qquad (25) \\
V_l \frac{dC_l}{dt} &= Q_l (C_a - C_l/P_l) - \frac{v_{max} C_l/P_l}{k_M + C_l/P_l} \qquad (26) \\
V_k \frac{dC_k}{dt} &= Q_k (C_a - C_k/P_k), \qquad (27)
\end{aligned}
$$

where $C_v(t)$, $C_{br}(t)$, $C_k(t)$, $C_m(t)$, $C_l(t)$ and $C_t(t)$ denote TCE concentrations in the venous blood, brain, kidney, muscle, liver and remaining tissue compartments, respectively. The chamber air concentration $C_c(t)$ is specified as part of the experiment and is used as a forcing function in the arterial blood equation (22). For the results we present in this paper, we set the chamber air concentration to 2000 parts per million TCE for one hour, followed by zero ppm TCE until the final time t_f (in hours).

Model parameters include tissue volumes V (in liters), volumetric blood flow rates to the tissues Q (liters/hour), and blood:tissue partition coefficients P, each labeled with a subscript corresponding to the appropriate tissue. The cardiac output and ventilation rates (in liters/hour) are denoted by Q_c and Q_b respectively, and the blood:air partition coefficient is denoted as P_b. The standard

Michaelis-Menten metabolic parameters are denoted by v_{max} (mg/hour) and k_M (mg/liter). See [ABEP, Po-thesis] for complete discussion of the model equations and parameters.

Theoretical results relating to well-posedness of the whole-body PBPK-hybrid model are presented in [BPo]. In particular, we have shown the existence of a unique weak solution for a general class of nonlinear parabolic equations that includes the TCE model as a special case. Moreover, we established the well-posedness of the deterministic estimation problem for the TCE model, and in [Po-thesis] we have addressed the well-posedness of probability-based parameter estimation methods applied to the TCE model. Numerical methods and simulations for this model with deterministic parameters are given in [BPnum], and results for the standard PBPK models are compared to those for the PBPK-hybrid model.

Examples of Metrics on Probabilities

We summarize some functional analysis fundamentals. **Reference:** H.T. Banks, *A Functional Analysis Framework for Modeling, Estimation and Control in Science and Engineering*, CRC/Taylor&Francis, 2011.

- The space of continuous functions on $[a,b]$ is denoted by $C[a,b]$. It has a norm defined by $|\varphi| = \sup_{\xi\in[a,b]} |\varphi(\xi)|$ for $\varphi \in C[a,b]$. Then we have the following theorem from [DS].

  Theorem 1. *The following equivalences hold:*

  $$C^*[a,b] \cong rba[a,b] \cong NBV[a,b]$$

  where $NBV[a,b]$ stands for normalized bounded variation functions and $rba[a,b]$ stands for regular bounded additive set

functions.

Therefore we can talk about the weak* topology for NBV$[a,b]$; however, $C[a,b]$ is not the dual of another space, NBV$[a,b]^* =?$ (i.e., we do not have satisfactory characterizations or representations). Thus we cannot talk about the weak* topology in $C[a,b]$.

Discussion of rba$[a,b]$: Suppose μ is a regular set function. Then by definition of *regular*, given a set F and $\epsilon > 0$, there exists a closed set E and an open set O such that $E \subset F \subset \bar{O}$ and $\mu(\bar{O} - E) < \epsilon$. Additive means that for any two Borel subsets A and B contained in $[a,b]$ such that $A \cap B = \emptyset$, and $\mu(A \cup B) = \mu(A) + \mu(B)$.

Discussion of NBV$[a,b]$: BV$[a,b]$ has a norm $|f| = |f(a^+)| +$

$\mathrm{var}(f)$ where $\mathrm{var}(f)$ is the total variation

$$\mathrm{var}(f) = \sup_{\pi} \sum_{i=1}^{N} |f(x_i) - f(x_{i-1})|,$$

and π denotes the partitions of $[a, b]$. Note that $\mathrm{var}(f)$ provides a semi-norm on BV and BV can be written as $\mathrm{BV}_0 \oplus N$ where $\mathrm{BV}_0 = \{f \in \mathrm{BV} | f(a) = 0\}$ and N is the 1-D space of constant functions.

$\mathrm{NBV}[a, b]$ is normalized by making f right-continuous at the interior points and $f(a^+) = 0$ with $|f| = \mathrm{var}(f)$.

Suppose $f \in \mathrm{NBV}[a, b]$; then we can write

$$f(x) = f(a) + p_f(x) - n_f(x),$$

where p_f and n_f are nondecreasing monotone functions. Recall

$$p_f(x) = \sup_{\pi} \sum_{i=1}^{N} |f(x_i) - f(x_{i-1})|^+.$$

Thus we can get a signed measure (recall Lebesgue-Stieltjes measures)

$$\mu_f = \mu_{p_f} - \mu_{n_f},$$

where μ_f is regular. From above we have $C^*[a,b] \cong \mathrm{rba}(a,b)$, which means there is a one-to-one correspondence

$$x^* \leftrightarrow \mu,$$

given by

$$x^*(\varphi) = \int \varphi d\mu.$$

We also have $C^*[a,b] \cong \mathrm{NBV}[a,b]$, which means we have

$$x^* \leftrightarrow \mu_f,$$

given by

$$x^*(\varphi) = \int \varphi d\mu_f = \left(\int \varphi df \right)$$

where integrals are Lebesque-Stieltjes integrals-see [A] and [DS].

Weak* Convergence and the Prohorov Metric

Recall two types of measure dependent problems: individual dynamics and aggregate dynamics.

Type II: Individual Dynamics and Aggregate Data

Such examples include the mosquitofish and shrimp examples outlined above. The dynamics for $u(t, x; \gamma)$, $\gamma \in \Gamma$ have a distribution P over Γ so that

$$u(t, x) = \int_{\Gamma} v(t, x; \gamma) dP(\gamma).$$

If P is discrete,

$$\int_T v(t,x;\gamma)dP(\gamma) = \sum^N v(t,x;\gamma_j)p_j.$$

If P is (absolutely) continuous, we can use whatever quadrature rule we wish, and

$$\int v(t,x;\gamma)p(\gamma)d(\gamma) \approx \sum_{j=1}^N v(t,x;\gamma_j)p(\gamma_j)\triangle\gamma_j.$$

More generally, for a general probability distribution P, we have

$$u(t,x) = \int_T v(t,x;\gamma)dP(\gamma),$$

as the expected value of the density

$$u(t,x) = u(t,x;P) = \mathcal{E}[v(t,x;\cdot)|P]$$

.

Since in typical inverse problems, we must use the data d_{ij} to estimate P itself, we are therefore required to understand the sensitivity of $u(t, x; P)$ as a function of P. Thus, we see that the data for parametric estimation or inverse problems is $d_{ij} \sim u(t_i, x_j)$ where $u(t, x)$ is the total population density given by $u(t, x; P)$.

Type III: Aggregate Dynamics

In this case, individual dynamics are not available. Instead, the dynamics themselves depend on probability measure, e.g.,

$$\frac{\partial u}{\partial t} + \frac{\partial^2 u}{\partial x^2} = f(u, P).$$

Here, as an example we can refer to the polarization of inhomogeneous dielectric materials in the theory of electromagnetics [BBL, BG1, BG2]. To obtain the macroscopic polarization, we sum over all the parameters. We cannot separate dynamics to obtain individual dynamics, and therefore we have an example, where the dynamics for the E and H fields depend explicitly on the probability measure P.

For inverse problems, data $d_{ij} \approx u(t_i, x_j) = u(t_i, x_j; P)$ in Type I,

where $d_{ij} \approx E(t_i, \bar{x}_j; P)$ in Type II. Notice that $P \in \mathcal{P}(Q)$, which is the space of probability measures, where Q consists of "parameters", and $\mathcal{P}(Q)$ is a set of probability distributions over Q.

For example, in ordinary least square problems, we have

$$J(P) = \sum_{ij} |d_{ij} - u(t_i, x_j; P)|^2 \quad \text{or} \quad \sum_{ij} |d_{ij} - E(t_i, \bar{x}_j; P)|^2,$$

to be minimized over $P \in \mathcal{P}(Q)$ or over some subset of $\mathcal{P}(Q)$. We need (for computational and theoretical considerations) a sense of closeness of two probability distributions P_1 and P_2. The ideas of local minima, continuity, "gradient", etc., depend on a metric (or topology) for P_1 and P_2.

More fundamentally for Type II systems, the "well-posedness of systems", including existence and continuous dependence of solutions on "parameters", depends on the concept of P_1 and P_2 being close, i.e., $P \to E(t, \bar{x}; P)$ is continuous.

We consider $\mathcal{P}(Q)$ as the space of probability measures or distributions on Q. Let P_1 and P_2 be probability distributions. If on the real line $(-\infty, \infty)$ or $(0, \infty)$ we will sometimes not distinguish between the probability measure P and its cumulative distribution function F.

We need to understand the concept of $\rho(P_1, P_2)$. There are several that metrize the weak* topology on $\mathcal{P}$ including:

$$\begin{aligned} \text{Levy} &- \rho_L(P_1, P_2) \\ \text{Prohorov} &- \rho_{PR}(P_1, P_2) \\ \text{Bounded Lipschitz} &- \rho_{BL}(P_1, P_2) \end{aligned}$$

and also some that do not metrize the weak* topology on $\mathcal{P}$:

$$\begin{aligned} \text{Total variation} &- \rho_{TV}(P_1, P_2) \\ \text{Kolmogorov} &- \rho_K(P_1, P_2). \end{aligned}$$

For relevant material, see [P], [B], and [H].

Prohorov Metric

Previous developments in probability theory provide helpful results in the pursuit of a possible complete computational methodology. One of the most important tools in probability theory is the Prohorov metric, which we will define. Let Q be a metric space with metric d. Given a closed subset A of Q, define the ϵ-neighborhood of A as

$$\begin{aligned} A^\epsilon &= \{q \in Q : d(\hat{q}, q) \leq \epsilon \text{ for some } \hat{q} \in A\} \\ &= \{q \in Q : \inf_{\hat{q} \in A} d(\hat{q}, q) \leq \epsilon\}. \end{aligned}$$

We define the Prohorov metric $\rho : \mathcal{P}(Q) \times \mathcal{P}(Q) \to R^+$ by

$$\rho(P_1, P_2) \equiv \inf\{\epsilon > 0 : P_1(A) \leq P_2(A^\epsilon) + \epsilon, \ A \text{ closed } \subset Q\}.$$

This can be shown to be a metric on $\mathcal{P}(Q)$ and has many properties including

$(a.)$ $(\mathcal{P}(Q), \rho)$ is a complete metric space;

$(b.)$ If Q is compact, then $(\mathcal{P}(Q), \rho)$ is a compact metric space.

Note that the definition of ρ is not intuitive. For example, we do not necessarily know what $P_k \to P$ in ρ means. We have the following important characterizations [B]. Given $P_k, P \in \mathcal{P}(Q)$, the following convergence statements are equivalent:

1. $\rho(P_k, P) \to 0$;
2. $\int_Q f dP_k(q) \to \int_Q f dP(q)$ for all bounded, continuous $f : Q \to R^1$;
3. $P_k[A] \to P[A]$ for all Borel sets $A \subset Q$ with $P[\partial A] = 0$.

Thus, we immediately obtain the following results:

- Convergence in the ρ metric is equivalent to convergence in distribution or so-called "weak" convergence of measures.
- Let $C_B^*(Q)$ denote the topological dual of $C_B(Q)$, where $C_B(Q)$

is the usual space of bounded continuous functions on Q with the supremum norm. If we view $\mathcal{P}(Q) \subset C_B^*(Q)$, convergence in the ρ topology is equivalent to weak* convergence in $\mathcal{P}(Q)$. Note the misnomer "weak convergence of measures" used by probabilists.

More importantly,

$$\rho(P_k, P) \to 0 \text{ is equivalent to } \int_Q x(t_i; q) dP_k(q) \to \int_Q x(t_i; q) dP(q),$$

and $P_k \to P$ in ρ metric is equivalent to

$$\mathcal{E}[x(t_i; q)|P_k] \to \mathcal{E}[x(t_i; q)|P]$$

or "convergence in expectation". This yields that

$$P \to J(P) = \sum_{i=1}^{n} |\mathcal{E}[x(t_i; q)|P] - \hat{d}_i|^2$$

is continuous in the ρ topology. Continuity of $P \to J(P)$ allows us to

assert the existence of a solution to $\min J(P)$ over $P \in \mathcal{P}(Q)$.

These results give us the following theorem.

Theorem 2. *The Prohorov metric metrizes the weak* topology of* $\mathcal{P}(Q)$.

Now we may consider whether there are other (equivalent) metrics that metrize the weak* topology on $\mathcal{P}$. But first we point out an application in statistics where the need for metrics on distributions arises.

Robust Statistics

Inference procedures - "robust" to derivations in underlying assumptions - "distributional robustness" insensitivity to "small" derivations in distributions. There is a need for distributional metrics.

- "Robust" < "Nonparametric" < "Distribution free" procedures and statistical tests.
- Statisticians refer to the "weak" topology on $\mathcal{P}$: (actually, mathematically it is weak* topology on measures).
- "Weak" topology on $\mathcal{P}$: weakest topology on $\mathcal{P}$ such that for every bounded continuous ψ ($\psi \in \mathcal{C}_B(\Omega)$) the map

$$P \to \int_Q \psi dP$$

  is continuous, i.e. consider $\mathcal{P} \subset C_B^*(\Omega)$.

It is desirable to know when $\mathcal{P}(Q)$ is a Polish space. A Polish space is a complete, separate, metrizable space.

Levy Metric

On for $Q = R^1$, where $P_1 \sim$ cdf F_1,

$$\rho_L(P_1, P_2) = \inf\{\epsilon | \forall x, P_1(x - \epsilon) - \epsilon \leq P_2(x) \leq P_1(x + \epsilon) + \epsilon\}.$$

We can argue that it is symmetric and indeed defines a metric.

Remarks

- $\sqrt{2}\rho_L(P_1, P_2)$ is maximum distance between graphs of P_1, P_2 measured along 45^o direction. It is a theorem that the Levy metric metrizes the weak topology of $\mathcal{P}$.
- The Prohorov metric is more difficult to visualize but is applicable when Q is any complete separable metric space (Polish space), not just the real line (Levy case). For example, when Q is a function space such as growth rates or mortality rates in Sinko-Streiffer (mosquito fish problem), the Prohorov metric is

applicable whereas the Levy metric is not.

We have already noted that both Levy and Prohorov metrics metrize the weak topology on $\mathcal{P}$, the former only when $Q = R^1$. Moreover, we can argue that for Q complete separable metric space, then $(\mathcal{P}(Q), \rho_{PR})$ is a complete separable metric space.

Separability: Q_o is countably dense $\subset \Omega$.

$$\begin{aligned} \mathcal{P}_o &= \{\text{measures with finite support in } Q_o \text{ with rational masses}\} \\ &= \{P = \sum_{\text{finite}} p_i \delta_{q_i} | p_i \text{ rational}, \{q_i\} \subset Q_o\} \end{aligned}$$

Bounded Lipschitz metric:

Assume without loss of generality that distance function d on Q is bounded by 1. If necessary, replace by $\tilde{d}(q_1, q_2) = \frac{d(q_1,q_2)}{1+d(q_1,q_2)}$. Define

$$\begin{aligned} \rho_{BL}(P_1, P_2) &= \sup_{\psi \in \Psi} \left| \int_Q \psi dP_1 - \int_Q dP_2 \right| \\ \Psi &= \{\psi \in \mathcal{C}(Q) : |\psi(q_1) - \psi(q_2)| \leq d(q_1, q_2)\} \end{aligned}$$

Theorem 0.1. *For all* $P_1, P_2 \in \mathcal{P}(Q)$,

$$\rho_{PR}(P_1, P_2)^2 \leq \rho_{BL}(P_1, P_2) \leq 2\rho_{PR}(P_1, P_2).$$

Thus, ρ_{PR} and ρ_{BL} define the same topology and hence ρ_{BL} also metrizes the weak topology on $\mathcal{P}(Q)$.

Other metrics

- Total variation:

$$\rho_{TV}(P_1, P_2) = \sup_{A \in \mathcal{B}} |P_1(A) - P_2(A)|$$

- Kolmogorov: $(Q = R^1)$

$$\rho_K(P_1, P_2) = \sup_{x \in R'} |P_1(x) - P_2(x)|$$

These metrics do not metrize the weak topology, but do satisfy

$$\begin{aligned} \rho_L \leq \quad \rho_{PR} \quad \leq \rho_{TV} \\ \rho_L \leq \quad \rho_K \quad \leq \rho_{TV}\,. \end{aligned}$$

Why is the knowledge about ρ_{BL} versus ρ_{PR} useful? We have that

$$\begin{aligned}
\Delta_k &\equiv \left| \int G(g) dP_k(g) - \int G(g) dP(g) \right| \\
&= \left| \int G(g) f_k(g) dg - \int G(g) f(g) dg \right| \\
&= \left| \int G(g) [f_k(g) - f(g)] dg \right| \\
&\leq \rho_{BL}(P_k, P) \\
&\leq 2\rho_{PR}(P_k, P).
\end{aligned}$$

Therefore, if we know $P_k \to P$ in ρ_{BR}, it may be useful in direct estimates. But $\Delta_k \to 0$ is Prohorov metric convergence itself if $G \in \mathcal{C}_B(Q)$.

References

[HTB] H.T. Banks, Remarks on uncertainty assessment and management in modeling and computation, *Mathematical and Computer Modelling*, **33** (2001), 39–47.

[BB] H.T. Banks and K.L. Bihari, Modelling and estimating uncertainty in parameter estimation,*Inverse Problems*, **17** (2001), 95–111.

[BBH] H.T. Banks, D.M. Bortz and S.E. Holte, Incorporation of variability into the mathematical modeling of viral delays in HIV infection dynamics, *Mathematical Biosciences*, **183** (2003), 63–91.

[BBPP] H.T. Banks, D.M. Bortz, G.A. Pinter and L.K. Potter, Modeling and imaging techniques with potential for application

in bioterrorism, CRSC-TR03-02, January, 2003; Chapter 6 in *Bioterrorism: Mathematical Modeling Applications in Homeland Security*, (H.T. Banks and C. Castillo-Chavez, eds.), Frontiers in Applied Math, **FR28**, SIAM, Philadelphia, PA, pp. 129–154, 2003.

[BBKW] H.T. Banks, L.W. Botsford, F. Kappel, and C. Wang, Modeling and estimation in size structured population models, LCDS-CCS Report 87-13, Brown University; *Proceedings 2nd Course on Mathematical Ecology*, (Trieste, December 8-12, 1986) World Press, Singapore, pp. 521–541, 1988.

[BG1] H.T. Banks and N.L. Gibson, Well-posedness in Maxwell systems with distributions of polarization relaxation parameters, CRSC-TR04-01, NCSU, January, 2004; *Applied Math. Letters*, **18** (2005), 423–430.

[BG2] H.T. Banks and N.L. Gibson, Electromagnetic inverse problems involving distributions of dielectric mechanisms and parameters, CRSC-TR05-29, August, 2005; *Quarterly of Applied Mathematics*, **64** (2006), 749–795.

[BI] H.T. Banks and D.W. Iles, *A comparison of stability and convergence properties of techniques for inverse problems,* LCDS 86-3, Brown University, Providence, RI, January 1986.

[BI2] H.T. Banks and D.W. Iles, On compactness of admissible parameter sets: Convergence and stability in inverse problems for distributed parameter systems, in *Control Problems for Systems Described by Partial Differential Equations and Applications*, (I. Lasiecka, R. Triggiani, eds.), LN in Control and Inf. Sci., Vol. 97, pp. 130–142, 1987.

[BI] H.T. Banks and B.G. Fitzpatrick, Estimation of growth rate

distributions in size-structured population models, CAMS Tech. Rep. 90-2, University of Southern California, January, 1990; *Quart. Appl. Math.*, **49** (1991), 215–235,

[BFPZ] H.T. Banks, B.G. Fitzpatrick, L.K. Potter, and Y. Zhang, Estimation of probability distributions for individual parameters using aggregate population data, CRSC-TR98-6, (January, 1998); In *Stochastic Analysis, Control, Optimization and Applications,* (Edited by W. McEneaney, G. Yin and Q. Zhang), Birkhäuser Verlag, Basel, pp. 353–371, 1998.

[BK] H.T. Banks and K. Kunisch, *Estimation Techniques for Distributed Parameter Systems*, Birkhaüser, Boston, 1989.

[BPo] H.T. Banks and L.K. Potter, Probabilistic methods for addressing uncertainty and variability in biological models: Application to a toxicokinetic model, CRSC-TR02-27, N.C.

State University, September, 2002; *Math. Biosci.*, **192** (2004), 193–225.

[BTSBRSM] H.T. Banks, W. Clayton Thompson, Karyn L. Sutton, Gennady Bocharov, Dirk Roose, Tim Schenkel and Andreas Meyerhans, Estimation of cell proliferation dynamics using CFSE data, CRSC-TR09-17, August, 2009; *Bull. Math. Biology*, submitted.

[BS] S.L. Beal and L.B. Sheiner, Estimating population kinetics, *CRC Critical Rev. in Biomed. Engr.*, **8** (1982), 195–222.

[Bi] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.

[BF2] Richard L. Burden and J. Douglas Faires, *Numerical Analysis* third edition, PWS Publishers, Boston, 1985.

[DGa1] M. Davidian and A.R. Gallant, Smooth nonparametic

maximum likelihood estimation for population pharmacokinetrics, with application to quinidine, *J. Pharm. and Biopharm.*, **20** (1992), 529–556.

[DGa2] M. Davidian and A.R. Gallant, The nonlinear mixed effects model with a smooth random effects density, *Biometrika*, **80** (1993), 475–488.

[DG1] M. Davidian and D.M. Gilitan, *Nonlinear Models for Repeated Measurement Data*, Chapman & Hall, London, 1995.

[DG2] M. Davidian and D.M. Gilitan, Nonlinear models for repeated measurement data: An overview and update, *J. Agricultural, Biological, and Environ. Statistics*, **8** (2003), 387–419.

[E] M.V. Evans, W.K. Boyes, P.J. Bushnell, J.H. Raymer and J.E. Simmons, *A physiologically based pharmacokinetic modle for trichloroethylene (TCE) in Long-Evans rats*, preprint, 1999.

[Ga] A.R. Gallant, *Nonlinear Statistical Models*, J. Wiley & Sons, New York, 1987.

[G] N.L. Gibson, *Terahertz-Based Electromagnetic Interrogation Techniques for Damage Detection*, PhD Thesis, N. C. State University, Raleigh, August, 2004.

[HT] R.V. Hogg and E.A. Tanis *Probability and Statistical Inference* Prentice Hall, Upper Saddle River, 2001.

[J] R.I. Jennrich, Asymptotic properties of non-linear least squares estimators, *Ann. Math. Statist.*, **40** (1969), 633–643.

[L1] B.G. Lindsay, The geometry of mixture likelihoods: A general theory, *Ann. Statist.*, **11** (1983), 86–94.

[L2] B.G. Lindsay, *Mixture Models: Theory, Geometry and Applications*, Vol. **5**, NSF-CBMS Regional Conf. Series in Probability and Statistics, Inst. Math. Stat., Haywood, CA,

1995.

[LL] B.G. Lindsay and M.L. Lesperance, A review of semiparametric mixture models, *J. Statist. Plannning and Inference*, **47** (1995), 29–39.

[Ma] A. Mallet, A maximum likelihood estimation method for random coefficient regression models, *Biometrika*, **73** (1986), 645–656.

[SRM] L.B. Sheiner, B. Rosenberg and K. Melmon, Modeling of individual pharmacokinetics for computer-aided drug dosage, *Comp. Biomed. Res.*, **5** (1972), 441–459.

[Po] L.K. Potter, *Physiologically Based Pharmacokinetic Models for the Systemic Transport of Trichloroethylene*, Ph.D. Thesis, N.C. State University, August, 2001; (http://www.lib.ncsu.edu).

[QSS] A. Quarteroni, R. Sacco, and F. Saleri *Numerical Mathematics*

Springer, New York, 2000.

[S1] A. Schumitzky, Nonparametric EM algorithms for estimating prior distributions, *Applied Mathematics and Computation*, **45** (1991), 141–157.

[S2] A. Schumitzky, The nonparametric maximum likelihood approach to pharmacokinetic population analysis, *Proceedings of the Western Simulation Multiconference-Simulation in Health Care*, Soc. for Computer Simulation, 1993.

[Wahba1] G. Wahba, Bayesian "confidence intervals" for the cross-validated smoothing spline, *J. Roy. Statist. Soc. B*, **45** (1983), 133–150.

[Wahba2] G. Wahba, (Smoothing) splines in nonparametric regression, Dept of Statistics-TR 1024, Univ. Wisconsin, September, 2000; In *Encyclopedia of Environmetrics*, (A.

El-Shaarawi and W. Piegorsch, eds.), Wiley, Vol. 4, pp. 2099–2112, 2001.

[Allen] L.J.S. Allen, *An Introduction to Stochastic Processes with Applications to Biology*, Pearson Education, Inc., Upper Saddle River, New Jersey, 2003.

[BBDS1] H.T. Banks, J.E. Banks, L.K. Dick, and J.D. Stark, Estimation of dynamic rate parameters in insect populations undergoing sublethal exposure to pesticides, CRSC-TR05-22, NCSU, May, 2005; *Bulletin of Mathematical Biology*, **69** (2007), 2139–2180.

[BBDS2] H.T. Banks, J.E. Banks, L.K. Dick, and J.D. Stark, Time-varying vital rates in ecotoxicology: selective pesticides and aphid population dynamics, *Ecological Modelling*, **210** (2008), 155–160.

[BaBi] H.T. Banks and K.L. Bihari, Modeling and estimating uncertainty in parameter estimation, *Inverse Problems*, **17** (2001), 95–111.

[Shrimp] H.T.Banks, V.A. Bokil, S. Hu, F.C.T. Allnutt, R. Bullis, A.K. Dhar and C.L. Browdy, Shrimp biomass and viral infection for production of biological countermeasures, CRSC-TR05-45, NCSU, December, 2005; *Mathematical Biosciences and Engineering*, **3** (2006), 635–660.

[BBPP] H.T. Banks, D.M. Bortz, G.A. Pinter and L.K. Potter, Modeling and imaging techniques with potential for application in bioterrorism, Chapter 6 in *Bioterrorism: Mathematical Modeling Applications in Homeland Security*, (H.T. Banks and C. Castillo-Chavez, eds.), Frontiers in Applied Mathematics **FR28**, SIAM, Philadelphia, 2003, pp. 129–154.

[BBKW] H.T. Banks, L.W. Botsford, F. Kappel, and C. Wang, Modeling and estimation in size structured population models, LCDS-CCS Report 87-13, Brown University; *Proceedings 2nd Course on Mathematical Ecology*, (Trieste, December 8-12, 1986) World Press (1988), Singapore, 521–541.

[BD] H.T. Banks and J.L. Davis, A comparison of approximation methods for the estimation of probability distributions on parameters, CRSC-TR05-38, NCSU, October, 2005; *Applied Numerical Mathematics*, **57** (2007), 753–777.

[BDTR] H.T. Banks and J.L. Davis, Quantifying uncertainty in the estimation of probability distributions with confidence bands, CRSC-TR07-21, NCSU, December, 2007; *Mathematical Biosciences and Engineering*, **5** (2008), 647–667.

[GRD-FP] H.T. Banks, J.L. Davis, S.L. Ernstberger, S. Hu, E.

Artimovich, A.K. Dhar and C.L. Browdy, A comparison of probabilistic and stochastic formulations in modeling growth uncertainty and variability, CRSC-TR08-03, NCSU, February, 2008; *Journal of Biological Dynamics*, **3** (2009) 130–148.

[Shrimp-exp] H.T. Banks, J.L. Davis, S.L. Ernstberger, S. Hu, E. Artimovich and A.K. Dhar, Experimental design and estimation of growth rate distributions in size-structured shrimp populations, CRSC-TR08-20, NCSU, November, 2008; *Inverse Problems,* to appear.

[GRD-FP2] H.T. Banks, J.L. Davis, and S. Hu, A computational comparison of alternatives to including uncertainty in structured population models, CRSC-TR09-14, June, 2009; in *Three Decades of Progress in Systems and Control,* Springer, 2009, to appear.

[BF] H.T. Banks and B.G. Fitzpatrick, Estimation of growth rate distributions in size-structured population models, CAMS Tech. Rep. 90-2, University of Southern California, January, 1990; *Quarterly of Applied Mathematics,* **49** (1991), 215–235.

[BFPZ] H.T. Banks, B.G. Fitzpatrick, L.K. Potter, and Y. Zhang, Estimation of probability distributions for individual parameters using aggregate population data, CRSC-TR98-6, NCSU, January, 1998; In *Stochastic Analysis, Control, Optimization and Applications,* (Edited by W. McEneaney, G. Yin and Q. Zhang), Birkhaüser, Boston, 1989.

[BP] H.T. Banks and G.A. Pinter, A probabilistic multiscale approach to hysteresis in shear wave propagation in biotissue, CRSC-TR04-03, NCSU, January, 2004; *SIAM J. Multiscale Modeling and Simulation,* **2** (2005), 395–412.

[BTW] H.T. Banks, H.T. Tran and D.E. Woodward, Estimation of variable coefficients in Fokker Planck equations using moving finite elements, CAMS Tech. Rep. 90-9, August, 1990, University of Southern California; *SIAM Journal of Numerical Analysis*, **30** (1993), 1574–1602.

[BeAnd] G.I. Bell and E.C. Anderson, Cell growth and division I. A mathematical model with application to cell volume distributions in mammalian suspension culures, *Biophys. J.*, **7** (1967), 330–351.

[B] P. Billingsley, *Convergence of Probability Measures,* Wiley, New York, 1968.

[CB] G. Casella and R.L. Berger, *Statistical Inference,* Duxbury, California, 2002.

[CSLFJ] F.L. Castille, T.M. Samocha, A.L. Lawrence, H. He, P.

Frelier, and F. Jaenike, Variability in growth and survival of early postlarval shrimp (*Penaeus vannamei* Boone 1931), *Aquaculture,* **113** (1993), 65–81.

[G] A.R. Gallant, *Nonlinear Statistical Models,* J. Wiley & Sons, New York, 1987.

[Gard] T.C. Gard, *Introduction to Stochastic Differential Equations*, Marcel Dekker, New York, 1988.

[GAZ] A. Gross, S. Abutbul, and D. Zilberg, Acute and chronic effects of nitrite on white shrimp, Litopenaeus vannamei, cultured in low-salinity brackish water, *Journal of the World Aquaculture Society,* **35** (2004), 315–321.

[J] R.I. Jennrich, Asymptotic properties of non-linear least squares estimators, *The Annals of Mathematical Statistics,* **40** (1969), 633–643.

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

[LLS] S. Laramore, C.R. Laramore, and J. Scarpa, Effect of low salinity on growth and survival of postlarvae and juvenile *litopenaeus vannamei*, *Journal of the World Aquaculture Society,* **32** (2001), 385–392.

[Tat1] T. Luzyanina, D. Roose, T. Schinkel, M. Sester, S. Ehl, A. Meyerhans and G. Bocharov, Numerical modeling of label-structured cell population growth using CFSE distribution data, *Theor. Biol. and Med. Model.*, **4** (2007), 1–26.

[Metz] J.A.J. Metz and O. Diekmann, *The Dynamics of Physiologically Structured Populations,* Lec. Notes in Biomath, Vol. **68**, Springer-Verlag, Berlin, 1986.

[Okubo] A. Okubo, *Diffusion and Ecological Problems: Mathematical*

Models, Biomathematics, Vol. **10**, Springer-Verlag, Berlin, 1980.

[PMR] J. Ponce-Palafox, C.A. Martinez-Palacios, and L.G. Ross, The effects of salinity and temperature on the growth and survival rates of juvenile white shrimp, *penaeus vannamei boone 1931*, *Aquaculture,* **157** (1997), 107–115.

[SW] G.A.F. Seber and C.J. Wild, *Nonlinear Regression,* John Wiley & Sons, New York, 1989.

[SS] J.W. Sinko and W. Streifer, A new model for age-size structure for a population, *Ecology,* **48** (1967), 910–918.

[WHO] United Nations Environment Programme, International Labour Organisation, World Health Organization, Trichloroethylene, World Health Organization, Geneva, 1985.

[BPnum] H.T. Banks and L.K. Potter, Model predictions and comparisons for three toxicokinetic models for the systemic

transport of trichloroethylene, CRSC-TR01-23, August, 2001; *Mathematical and Computer Modeling*, **35** (2002), 1007–1032.

[ABEP] R. A. Albanese, H. T. Banks, M. V. Evans, L. K. Potter, Physiologically based pharmacokinetic models for the transport of trichloroethylene in adipose tissue, Bulletin of Mathematical Biology 64 (2002) 97–131.

[BMP] H. T. Banks, N. G. Medhin, G. A. Pinter, Nonlinear reptation in molecular based hysteresis models for polymers, Tech. Rep. CRSC-TR03-45, Center for Research in Scientific Computation, North Carolina State University, December, 2003; *Quarterly Applied Math.*, **62** (2004), 767-779.

[HTB-HK] H. T. Banks, Incorporation of uncertainty in inverse problems, in: Y.C. Hon, et al., (Eds.), *Recent Development in Theories and Numerics: Intl. Conf. Inverse Problems* (Hong

Kong, Jan. 9-12, 2002), World Scientific Press, 2003, pp. 26–36.

[Po-thesis] L. K. Potter, Physiologically based pharmacokinetic models for the systemic transport of trichloroethylene, Ph.D. thesis, North Carolina State University, Raleigh, NC, www.lib.ncsu.edu (August 2001).

[Marina] M. V. Evans, W. K. Boyes, P. J. Bushnell, J. H. Raymer, J. E. Simmons, A physiologically based pharmacokinetic model for trichloroethylene (TCE) in Long-Evans rats, Preprint (1999).

[BPo] H. T. Banks, L. K. Potter, Well-posedness results for a class of toxicokinetic models, Tech. Rep. CRSC-TR01-18, Center for Research in Scientific Computation, North Carolina State University, July 2001 (www.ncsu.edu/crsc/reports.html); Dynamic Systems and Applications (in Press) .

[RobertsRowland] M. S. Roberts, M. Rowland, A dispersion model of

hepatic elimination: 1. Formulation of the model and bolus considerations, Journal of Pharmacokinetics and Biopharmaceutics 14 (1986) 227–260.

[BBL] H.T. Banks, M.W. Buksas, T. Lin. *Electromagnetic Material Interrogation Using Conductive Interfaces and Acoustic Wavefronts*. SIAM FR 21, Philadelphia, 2002.

[BI] H.T. Banks and K. Ito, A unified framework for approximation and inverse problems for distributed parameter systems, *Control: Theory and Adv. Tech.*, **4** (1988), 73–90.

[DS] N. Dunford and J.T. Schwartz. *Linear Operators*, Vol. I, 1966.

[HP] E. Hille and R. Phillips. *Functional Analysis and Semigroups*. AMS Colloquium Publications, Vol. 31, 1957.

[H] P. Huber. *Robust Statistics*. Wiley, New York, (1981).

[K] T. Kato. *Perturbations Theory for Linear Operators*, 1966.

[Kr] R. Kress. *Linear Integral Equations*. Springer-Verlag, New York, 1989.

[P] Yu. V. Prohorov, Convergence of random processes and limit theorems in probability theory, *Theor. Prob. Appl.*, **1** (1956), 157-214.

[Pa] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer-Verlag, New York, 1983.