Comparing a Distributed Parameter Model-Based System Identification Technique with More Conventional Methods for Inverse Problems

Jian Li, Susan E. Luczak and I. G. Rosen

Abstract. Three methods for the estimation of blood or breath alcohol concentration (BAC/BrAC) from biosensor measured transdermal alcohol concentration (TAC) are evaluated and compared. Specifically, we consider a system identification/quasi-blind deconvolution scheme based on a distributed parameter model with unbounded input and output for ethanol transport in the skin and compare it to two more conventional system identification and filtering/deconvolution techniques for ill-posed inverse problems, one based on frequency domain methods, and the other on a time series approach using an ARMA input/output model. Our basis for comparison are five statistical measures of interest to alcohol researchers and clinicians: peak BAC/BrAC, time of peak BAC/BrAC, the ascending and descending slopes of the BAC/BrAC curve, and the area underneath the BAC/BrAC curve.

Keywords. Distributed parameter systems, System identification, Filtering, Blind deconvolution, Transdermal alcohol biosensor.

2010 Mathematics Subject Classification. 35K90, 47D06, 65M32, 92C55, 93B30, 93C20, See `www.ams.org/msc`.

1 Introduction

Distributed parameter systems (i.e. partial and functional differential equations) have been used for modeling, simulation, control, estimation, identification, and optimization in a variety of applications arising in engineering, biology, chemistry, physics, space science, and economics. The appeal of using these infinite dimensional systems to describe the evolution of the state is based on a number of factors. These include: (1) their ability to yield high fidelity models with relatively low dimensional parameterization, (2) the fact that they typically result when modeling is based on first principles and physics-based empirical observations, and (3) that there is a rich, functional analytic theory upon which the analyses

of well-posedness (i.e. existence, uniqueness, continuous dependence), stability, long-term behavior, and finite dimensional and numerical approximation and convergence, can be based. In addition, although distributed parameter models result in an infinite dimensional state equation, by exercising some level of mathematical precision, they can readily be interfaced with finite dimensional control inputs and output observations, which are typically encountered in practice.

There is a large body of research on the theory and application of distributed parameter systems in control science, and in particular, their identification and estimation, reported in the literature. Rather comprehensive surveys can be found in a number of monographs by Banks, *et al.* ([1] and [3]) and in [15]. However, we know of no studies where the performance of these models once their unknown parameters have been identified, is compared to that of models obtained and identified by more conventional means. More specifically, how well do these fit models perform when they are actually used to design a controller, simulate a process, or as in the particular problem of interest to us here, solve an in general, ill-posed inverse problem, when compared to more traditional finite dimensional techniques found in the engineering literature.

In this brief paper, we present and discuss our experience using a distributed parameter system as the basis for a data analysis system for a biosensor that measures transdermal alcohol concentration (TAC), or more specifically, the number of ethanol molecules evaporating from the surface of the skin ([17], [26], [27], [28], [29]). The role of the data analysis system is to use the TAC signal from the biosensor to estimate the concentration of ethanol in the blood. Blood and breath alcohol concentration (BAC/BrAC) are currently the basis for all clinical research on the effects of alcohol on human physiology and behavior. In addition, both BAC and BrAC are used as quantitative determinants of whether an individual is legally intoxicated or driving under the influence (DUI) of alcohol. A passive, non-invasive biosensor that provides researchers and clinicians with an accurate and continuous estimate of BAC in real time from the field would be considered a major breakthrough and is currently the focus of several academic and commercial research efforts ([6], [7], [8], [9], [10], [22]). BrAC measurement taken in the field using a breath analyzer and then followed up with a blood sample and a laboratory determination of BAC are also the basis for arrest, prosecution, and determination of guilt in cases of DUI.

Our approach here is to formulate the problem as a single input single output (SISO) system in which the alcohol in the blood is the source of the input and the biosensor measured transdermal alcohol is the output signal. The objective of estimating the BrAC/BAC then takes the form of a system identification/input determination, or blind deconvolution problem. We model the forward transport of ethanol through the skin using a one-dimensional diffusion equation with in-

put and output on the boundary. Linear semigroup theory ([18], [25], [30]) and the infinite dimensional variation of constants formula are used to obtain a parametric representation for the unknown forward convolution filter. We then show how the approximation theory for linear semigroups can be used to develop finite dimensional convergent numerical approximation schemes for estimating the convolution kernel and deconvolving BAC from TAC.

The ideas outlined in the previous two paragraphs have been presented in detail elsewhere (see, for example, [6] or [22]), and we only provide a brief summary here. Of primary interest to us in this paper, however, is the comparison of this distributed parameter modeling based approach to two more conventional system identification methods, a Fourier/frequency domain based approach [20] and a scheme based on an autoregressive moving average (ARMA) input/output response surface [11]. This comparison is of particular interest to us as we are unaware of any previous studies that show how these methods perform with extremely limited data. Indeed, it is quite common in practice that only one episode of lab/clinic calibration data is available for a given subject. For our purposes here, a drinking episode is defined to be the period of time, and the BAC, BrAC or TAC data during that period of time, in between two periods of zero (or below some pre-determined level) BAC, BrAC or TAC. A more precise definition of drinking episode will be given later in Section 4 when we discuss the results of our numerical studies.

Our comparison here is based on actual clinical BrAC and TAC data collected in the field by one of the co-authors (S.E.L.) using a transdermal biosensor and a hand-held portable breath analyzer. We note in passing that alcohol researchers almost always use BrAC as a proxy for BAC which for obvious reasons is far more difficult, expensive, and invasive to measure. In carrying out our research we did not have access to, or, the resources and IRB approval to collect our own, simultaneous BAC and TAC measurements. In fact, direct BAC measurements for the purpose of research are rather rare (most researchers rely on BrAC) and we are not aware of the availability of any simultaneous direct BAC and TAC measurements within the research community.

Since the primary focus of our study here is the comparison of first principles physics-based distributed parameter modeling to more conventional response surface methods for identifying input/output models, central to our effort is the use of diffusion to model the movement of ethanol molecules from the blood through the skin to the biosensor on the surface of the skin. The three methods described here for identifying input/output models for the transdermal transport of ethanol with TAC as the output, will work equally well with either BAC or BrAC as the input. If one assumes linear actuator dynamics (to borrow a term from the control literature) in the forward model to describe the relationship between BrAC and BAC,

since the model itself in its entirety is linear, the form of the model is exactly the same whether BAC or BrAC is used as the input, with only the values of the independently identifiable parameters (conceivably) changing when the models are fit to experimental or clinical data.

The assumption of linear actuator dynamics in the relationship between BrAC and BAC is justified since this is how the breath analyzer works; quantitatively it is based on the empirically determined and accepted blood/breath partition ratio of 2300:1 (there is some debate as to whether this should be reduced to 2000:1, see, for example, [12]). However, this is of no consequence to us in our study here since we fit the complete model based on input (i.e. BrAC) and output (i.e. TAC) data. Since the model is linear, the actuator dynamics in the form of a gain in front of the BrAC input converting it to BAC, and the sensor dynamics in the form of a gain in front of the concentration of ethanol on the surface of the skin converting it to measured TAC output, are simply multiplied. This product (the two gains are not individually identifiable) is one of the parameters that are estimated when the distributed model is fit.

We note that it is conceivable that there could be some advantages to basing the model fits on a direct measurement of BAC rather than BrAC since, for example, it would not be susceptible to confounding effects such as mouth alcohol and any nonlinear disturbances introduced by the breath analyzer. On the other hand, there is almost certainly some level of uncertainty introduced when BAC is determined in the laboratory. In addition, along with the relative unavailability of BAC data, a model that yields estimates of BrAC from TAC rather than BAC from TAC is likely to be more desirable to alcohol researchers and clinicians since it is BrAC data that they are used to looking at and working with. BrAC also has the added benefit of being a measure of arterial BAC, which is closer to the alcohol concentration that passes the blood-brain barrier than is found in veinous blood.

An outline of the remainder of the paper is as follows. In Section 2 we briefly summarize the problem of estimating BAC from TAC, its formulation as a blind deconvolution problem, and modeling it using a distributed parameter system. In Section 3 we describe the scheme for estimating the convolution kernel and the deconvolution of BAC from TAC and we also discuss frequency domain and ARMA based approaches. In Section 4 we present our numerical results and a final fifth section contains a discussion of the results and concluding remarks.

2 Estimating Blood Alcohol Concentration from Transdermal Alcohol Biosensor Measurements

There is currently no known procedure for the accurate and unobtrusive field collection of quantitatively interpretable alcohol consumption data. Collection of blood and urine samples in the field is not practical and breath analyzers and drink diaries have either heavy subject burden or large time gaps between assessment points, hinder naturalistic drinking behavior, and may not be accurate even when individuals are trying to be compliant. Breath analyzer readings are frequently too high due to mouth alcohol or too low due to not taking deep lung breaths, while drink diaries are inaccurate due to not knowing the alcohol content of a drink or the amount consumed.

Biosensor devices that measure TAC, the amount of alcohol diffusing through the skin, are proving to be a promising technology. Unfortunately, however, to date, they are primarily being used as abstinence monitors because TAC data does not consistently correlate with BrAC and BAC across individuals, devices, and environmental conditions. To wit, there is currently no well-established method for producing reliable estimates of BrAC or BAC (eBrAC and eBAC) from TAC data. As we indicated previously, the breath analyzer relies on a relatively simple model from basic chemistry (i.e., Henry's Law) for the exchange of gases between circulating pulmonary blood and alveolar air [13]. This model has been found to be reasonably robust across subjects and the environment. The transport and filtering of alcohol by the skin on the other hand, is physiologically more complex and based on a number of factors that differ across individuals (e.g., skin layer thickness, tortuosity) and drinking episodes within individuals (e.g., temperature, skin hydration, vasodilation). This means that, regardless of how reliable and accurate transdermal alcohol devices become at measuring TAC, the raw TAC data will never trivially map directly onto BrAC/BAC across individuals and drinking episodes.

2.1 A Quasi-Blind Deconvolution Problem

In order to address this problem, we have been developing a data analysis system to accompany the sensor hardware that converts TAC to BAC/BrAC. In a relatively recent series of papers [7],[8],[9], Dougherty *et al.* have looked at fitting standard linear regression models to convert TAC to BrAC. Our approach, on the other hand, is based on a first principles physics-based model (in the form of a distributed parameter system or partial differential equation (PDE)) for the transport of ethanol molecules from the blood through the epidermal layer of the skin and its measurement by the sensor. The use of the physics-based model as opposed to

generic input-output response surface techniques, such as standard linear regression, neural networks, etc. allows us to keep the dimension of the parameter space relatively low, therefore reduce the likelihood of overfitting. The result is a forward input/output model that can be formulated as either a continuous or discrete time convolution in which the convolution kernel or filter, while unknown, is defined in terms of the parameters that appear in the physics based model. Obtaining an estimate for the BAC or BrAC can then be formulated as a blind or quasi blind (since the filter is defined parametrically) deconvolution problem. The current protocol for using the system involves simultaneously collecting calibration BrAC and TAC (by using the sensor) in either the lab or clinic for a single drinking episode before sending the patient or research subject out in the field with just the TAC sensor. Estimated BrAC from the TAC for all subsequent drinking episodes recorded in the field is then obtained by using the calibration data to fit the parameters that determine the filter, and then using the fit filter to deconvolve estimated BrAC from the TAC.

2.2 A Distributed Parameter Model for the Transdermal Transport of Ethanol

Let $\varphi(t, \eta)$ denote the concentration in moles/cm^2 of ethanol in the interstitial fluid in the epidermal layer of the skin at depth η cm and time t in seconds. Let L denote the skin thickness in cm. We model the transport of ethanol through this medium as a diffusion process

$$\frac{\partial \varphi}{\partial t}(t, \eta) = D\frac{\partial^2 \varphi}{\partial \eta^2}(t, \eta), \quad 0 < \eta < L, \quad t > 0 \qquad (2.1)$$

where $D > 0$ denotes the diffusivity in units of cm^2/sec. For boundary conditions, at the skin surface ($\eta = 0$), we model evaporation of the alcohol vapor using a Robin boundary condition; the flux (from right to left) is proportional to ethanol at the boundary of the epidermal layer,

$$D\frac{\partial \varphi}{\partial \eta}(t, 0) = \alpha\varphi(t, 0), \quad t > 0 \qquad (2.2)$$

where $\alpha > 0$ denotes the constant of proportionality in units of cm/sec. At the interface of the dermal layer (which has a blood supply) and the epidermal layer (which does not contain blood vessels) ($\eta = L$) we impose a Neumann boundary condition. That is that the flux (from right to left) is proportional to the BAC,

$$D\frac{\partial \varphi}{\partial \eta}(t, L) = \beta u(t), \quad t > 0 \qquad (2.3)$$

where the parameter $\beta > 0$ characterizes the exchange of ethanol molecules between the blood and the interstitial fluid in units of moles / (cm $\times$ sec $\times$ BAC (or BrAC) units), and u denotes the concentration of ethanol in the blood as given in BAC (or BrAC) units. The boundary condition (2.3) serves as our input condition. We assume that there is no alcohol in the epidermal layer at time $t = 0$ which yields the initial conditions

$$\varphi(0, \eta) = 0, \quad 0 < \eta < L \tag{2.4}$$

We model the processing by the TAC sensor of the ethanol evaporating from the surface of the skin via a linear relation which serves as an output condition

$$y(t) = \gamma \varphi(t, 0), \quad t > 0 \tag{2.5}$$

where γ denotes the constant of proportionality in units of TAC units $\times$ cm^2/mole.

As it stands, the model given by the equations (2.1) - (2.5) is determined by five parameters: D, L, α, β and γ. However, not all five of the parameters are independent nor are they uniquely identifiable from the input/output data. Without loss of generality and by converting to what are essentially dimensionless quantities, the number of unknown parameters to be fit can be reduced to two, which we denote by the vector $q = [q_1 \ q_2]^T$ [6]. For simplicity, leaving the names of the variables unchanged (although their units certainly are), in this way, our input/output model becomes

$$\frac{\partial \varphi}{\partial t}(t, \eta) = q_1 \frac{\partial^2 \varphi}{\partial \eta^2}(t, \eta), \quad 0 < \eta < 1, \quad t > 0 \tag{2.6}$$

$$q_1 \frac{\partial \varphi}{\partial \eta}(t, 0) - \varphi(t, 0) = 0, \quad t > 0 \tag{2.7}$$

$$q_1 \frac{\partial \varphi}{\partial \eta}(t, 1) = q_2 u(t), \quad t > 0 \tag{2.8}$$

$$\varphi(0, \eta) = 0, \quad 0 < \eta < 1 \tag{2.9}$$

$$y(t) = \varphi(t, 0), \quad t > 0 \tag{2.10}$$

The input/output model (2.6) - (2.10) is an example of a distributed parameter system. We note that there is a feature of the model that makes it somewhat non-standard: the input and the output are on the boundary. In the next section, when the model is reformulated abstractly as an abstract evolution equation in an infinite dimensional Hilbert space, this results in the input and output operators being unbounded in the standard space where an initial boundary value problem such as (2.6) - (2.10) is typically set. This requires some care (see, for example, [5] and [21]) when we re-cast the problem using techniques from functional analysis, specifically, linear semigroup theory. Also, reformulating the problem in discrete time is some help in this regard as well.

3 Determining the Convolution Kernel or Impulse Response Function

In this section, we describe three techniques for obtaining and estimating the impulse response function for the input/output model for the dynamical system described in the previous section. We first consider an approach based on the first principles physics-based distributed parameter model discussed in Section 2.2. We then look at two more conventional approaches: a method based on frequency domain techniques and one based on an ARMA model.

3.1 Distributed Parameter Modeling and the Impulse Response Function

In this section, we provide a brief summary of how tools from functional analysis and in particular, the theory of linear semigroups of operators can be used to transform the distributed parameter model presented in the previous section into a linear input/output model in the form of a convolution. We also describe how the resulting impulse response function, or convolution kernel or filter, which is defined in terms of operators on infinite dimensional spaces, can be approximated via matrix representations for linear operators defined on a sequence of approximating finite dimensional subspaces. A more detailed presentation of the results discussed here including relevant theorems and proofs can be found in [6] and [22].

Abstract Parabolic Input/Output Systems, their Associated Impulse Response Functions, and their Identification

Let V and H be Hilbert spaces that satisfy the dense and continuous embeddings $V \hookrightarrow H \hookrightarrow V^*$, where V^* denotes the space of continuous linear functionals that is dual to V. Let $\langle \cdot, \cdot \rangle$ denote the H inner product. For $q \in \{Q, d\}$, a compact metric space, let $a(q; \cdot, \cdot) : V \times V \to R$ be a bilinear form satisfying the following three conditions:

(i) (Boundedness)

$$|a(q; \psi_1, \psi_2)| \leq \rho ||\psi_1|| \, ||\psi_2||, \quad \psi_1, \psi_2 \in V$$

(ii) (Coercivity)

$$a(q; \psi, \psi) + \lambda |\psi|^2 \geq \mu ||\psi||^2, \quad \psi \in V$$

(iii) (Continuous Dependence)

$$|a(q_1; \psi_1, \psi_2) - a(q_2; \psi_1, \psi_2)| \leq d(q_1, q_2) ||\psi_1|| \, ||\psi_2||, \psi_1, \psi_2 \in V, q_1, q_2 \in Q$$

For $q \in \{Q, d\}$ let $b(q; \cdot) : V \to R$ and $c(q; \cdot) : V \to R$ be linear and continuous and consequently, therefore, it follows that $b(q; \cdot) = b(q)$ for some $b(q) \in V^*$ and $c(q; \cdot) = c(q)$ for some $c(q) \in V^*$. We then consider the input/output system in weak form given by

$$\langle \dot{\varphi}, \psi \rangle + a(q; \varphi, \psi) = b(q; \psi)u, \quad \psi \in V, \quad y = c(q; \varphi) \quad (3.1)$$

or equivalently

$$\langle \dot{\varphi}, \psi \rangle + a(q; \varphi, \psi) = \langle b(q), \psi \rangle_{V^*,V} u, \quad \psi \in V, \quad y = \langle c(q), \varphi \rangle_{V^*,V} \quad (3.2)$$

where $\langle \cdot, \cdot \rangle_{V^*,V}$ denotes the natural extension of the H inner product to the duality pairing between V and V^*. If we set $W(0, T) = \{\psi : \psi \in L_2(0, T, V),\ \dot{\psi} \in L_2(0, T, V^*)\}$ and $u \in L_2(0, T)$ it can be shown [16] that the system (3.1) or (3.2) admits a unique solution $\varphi \in W(0, T)$ that depends continuously on $u \in L_2(0, T)$. It follows that $W(0, T) \subseteq C(0, T, H)$ and that $y \in L_2(0, T)$.

For $q \in Q$, the q-dependent bilinear form on $V \times V$, $a(q; \cdot, \cdot) : V \times V \to R$, defines a bounded linear operator $A(q) \in \mathcal{L}(V, V^*)$ by $\langle A(q)\psi_1, \psi_2 \rangle = -a(q; \psi_1, \psi_2)$, for $\psi_1, \psi_2 \in V$. Then, if we let $\mathcal{H}$ denote either of the spaces H or V^*, we can consider the linear operator $A(q)$ to be the unbounded linear operator, $A(q) : D_q \subseteq \mathcal{H} \to \mathcal{H}$ where $D_q = V$ in the case $\mathcal{H} = V^*$, and $D_q = \{\psi \in V : A(q)\psi \in \mathcal{H}\}$ in the case $\mathcal{H} = H$. It can then be shown [2], [3], [30] that $A(q)$ is a closed, densely defined unbounded linear operator on $\mathcal{H}$ and it is the infinitesimal generator of an analytic semigroup of bounded linear operators, $\{e^{A(q)t} : t \geq 0\}$ on $\mathcal{H}$.

For $q \in Q$, define the bounded linear operators $B(q) : R \to V^*$ and $C(q) : V \to R$ by $\langle B(q)v, \psi \rangle_{V^*,V} = b(q; \psi)v = \langle b(q), \psi \rangle_{V^*,V}v$, and $C(q)\psi = c(q; \psi) = \langle c(q), \psi \rangle_{V^*,V}$, respectively, for $\psi \in V$ and $v \in R$. The input/output system can now be written formally in the standard form in $\mathcal{H}$ as

$$\begin{aligned} \dot{x}(t) &= A(q)x(t) + B(q)u(t) \\ y(t) &= C(q)x(t), \quad t > 0 \end{aligned} \quad (3.3)$$

where the state $x(t) = \varphi(t, \cdot) \in \mathcal{H}$. Assuming that the system is initially at rest (i.e. that $\varphi(0, \cdot) = 0$), and using the fact that $\{e^{A(q)t} : t \geq 0\}$ is an analytic semigroup on $\mathcal{H}$ and therefore that $e^{A(q)t}\psi \in D_q \subseteq V$, for $\psi \in V^*$, we obtain from the abstract variation of constants formula that

$$y(t) = \int_0^t C(q)e^{A(q)(t-s)}B(q)u(s)ds = \int_0^t h(q; t - s)u(s)ds \quad (3.4)$$

where $h(q; t) = C(q)e^{A(q)t}B(q),\ t > 0$.

Let a sampling time $\tau > 0$ be given and consider zero order hold inputs of the form $u(t) = u_i, \ t \in [i\tau, (i+1)\tau), \ i = 0, 1, 2, \ldots$ (typically $u_i = u(i\tau), \ i = 0, 1, 2, \ldots$, where u is a given continuous time input). Set $x_i = x(i\tau)$ and let $y_i = y(i\tau), \ i = 0, 1, 2, \ldots$. It then follows in the usual way that

$$x_{i+1} = \hat{A}(q)x_i + \hat{B}(q)u_i, \quad y_i = Cx_i, \quad i = 0, 1, 2, \ldots \quad (3.5)$$

with $\varphi_0 = 0 \in V$, where $\hat{A}(q) = e^{A(q)\tau} \in \mathcal{L}(V, V)$, and $\hat{B}(q) = \int_0^\tau e^{A(q)s}B(q)ds \in \mathcal{L}(R, V)$. Boundedness of the operators $\hat{A}(q)$ and $\hat{B}(q)$ follows from the fact that $\{e^{A(q)t} : t \geq 0\}$ is an analytic semigroup on V, H and V^* [2], [3], [16], [30]. If $A(q) : D_q \subset V^* \to V^*$ is invertible with bounded inverse, it follows that $\hat{B}(q) = \int_0^\tau e^{A(q)s}B(q)ds = A(q)^{-1}e^{A(q)s}B(q)\big|_0^\tau = (\hat{A}(q) - I)A(q)^{-1}B(q)$ and therefore that

$$
\begin{aligned}
y_i &= \sum_{j=0}^{i-1} C\hat{A}(q)^{i-j-1}\hat{B}(q)u_j \\
&= \sum_{j=0}^{i-1} C\hat{A}(q)^{i-j-1}(\hat{A}(q) - I)A(q)^{-1}B(q)u_j, \quad i = 0, 1, 2, \ldots
\end{aligned}
\quad (3.6)
$$

or that

$$y_i(q) = \sum_{j=0}^{i-1} \hat{h}_{i-j}(q)u_j, \quad i = 0, 1, 2, \ldots \quad (3.7)$$

where $\hat{h}_i(q) = C\hat{A}(q)^{i-1}(\hat{A}(q) - I)A(q)^{-1}B(q), i = 0, 1, 2, \ldots$

Given training data, $\{\tilde{u}_j, \tilde{y}_j\}, \ j = 0, 1, 2, \ldots, N$, we formulate the identification problem as a nonlinear least squares fit to data. That is we seek $q \in \{Q, d\}$ which minimizes the cost functional

$$J(q) = \sum_{i=0}^{N} |y_i(q) - \tilde{y}_i|^2 \quad (3.8)$$

where $u_i = \tilde{u}_i, \ i = 0, 1, 2, \ldots, N$.

Finite Dimensional Approximation, the Adjoint Method, and Differentiating the Matrix Exponential

Computing the value of the cost functional $J(q)$ and its gradient for a given value of $q \in \{Q, d\}$ as would be required to solve the optimization problem posed above requires the finite dimensional approximation of the infinite dimensional operators

that appear in the definition of $\hat{h}_i(q)$. For $n = 1, 2, \ldots$, let $H^n = \mathrm{span}\{\psi_j^n\}_{j=0}^n \subset V$, and let $P^n : H \rightarrow H^n$ denote the orthogonal projection of H onto H^n with respect to the H inner product. We assume that the $\{\psi_j^n\}_{j=0}^n$ are such that $\lim_{n\to\infty} P^n\psi = \psi$ in H for $\psi \in H$ and in V for $\psi \in V$. For $n = 1, 2, \ldots$, and $q \in Q$, define $A^n(q) \in \mathcal{L}(H^n, H^n)$ to be the finite dimensional linear operator whose matrix representation is given by $\lfloor A^n(q) \rfloor_{i,j} = -[\langle \psi_i^n, \psi_j^n \rangle]^{-1}[a(q; \psi_i^n, \psi_j^n)]$, for $i, j = 0, 1, 2, \ldots, n$. In what follows, we do not distinguish between the finite dimensional operators and their matrix representations with respect to the basis, $\{\psi_j^n\}_{j=0}^n$, for the approximating finite dimensional subspaces, H^n, defined above. In fact all the equations below are valid for both the finite dimensional operators and their matrix representations. Under certain conditions, in particular when $A(q)^{-1}$ exists, and λ in (ii) above is non-positive, it is not difficult to show that $A^n(q) = (P_a^n A(q)^{-1})^{-1}$, where P_a^n is the orthogonal projection of V onto H^n with respect to the inner product $\langle \cdot, \cdot \rangle_a = a(q; \cdot, \cdot)$ on V. We set $\hat{A}^n(q) = e^{A^n(q)\tau}$ and obtain the finite dimensional approximating discrete time input/output system given by

$$x_{i+1}^n = \hat{A}^n(q)x_i^n + (\hat{A}^n(q) - I)A^n(q)^{-1}\hat{P}^n B(q)u_i$$
$$y^n(q) = Cx_i^n, \quad i = 1, 2, \ldots \tag{3.9}$$

where $x_0^n = 0 \in H^n$, and $\hat{P}^n$ denotes the standard bounded extension of P^n to V^*. It follows that

$$y_i^n(q) = \sum_{j=0}^{i-1} C(\hat{A}^n(q))^{i-j-1}(\hat{A}^n(q) - I)\hat{P}^n B(q)u_j$$
$$= \sum_{j=0}^{i-1} \hat{h}_{i-j}^n(q)u_j, \quad i = 0, 1, 2, \ldots \tag{3.10}$$

where $\hat{h}_i^n(q) = C(\hat{A}^n(q))^{i-1}(\hat{A}^n(q) - I)\hat{P}^n B(q), \ i = 1, 2, \ldots$. With the help of the Trotter-Kato semigroup approximation theorem, it can be shown that $\hat{h}_i^n(q) \rightarrow \hat{h}_i(q)$ as $n \to \infty$, uniformly in q for $q \in \{Q, d\}$ and uniformly in i for i in bounded subsets of $\mathbb{Z}$. We then seek $q \in \{Q, d\}$ which minimizes the cost functional

$$J^n(q) = \sum_0^N |y_i^n(q) - \tilde{y}_i|^2 \tag{3.11}$$

where $u_i = \tilde{u}_i, \ i = 0, 1, 2, \ldots, N$.

The cost functional J^n is minimized iteratively. It is clear that for a given value of q, the value of $J^n(q)$ can now easily be computed. The gradient of $J^n(q)$

is computed using the adjoint method [14]. For $i = 0, 1, 2, \ldots, N$, set $v_i^n = 2C^T(Cx_i^n - \tilde{y}_i), \in R^{n+1}$ and define the adjoint system

$$z_{i-1}^n = [\hat{A}^n(q)]^T z_i^n + v_{i-1}^n, \quad i = N, N-1, \ldots, 1, \quad z_N^n = v_N^n \qquad (3.12)$$

The gradient of $J^n(q)$ can then be computed as

$$\overrightarrow{\triangledown} J^n(q) = \sum_{i=1}^N [z_i^n]^T \left\{ \frac{\partial \hat{A}^n(q)}{\partial q} x_{i-1}^n \right.$$
$$- \frac{\partial \hat{A}^n(q)}{\partial q}(\hat{A}^n(q) - I)\hat{P}^n B(q)\tilde{u}_{i-1}\hat{P}^n B(q)\tilde{u}_{i-1} \qquad (3.13)$$
$$\left. -(\hat{A}^n(q) - I)\frac{\partial}{\partial q}(\hat{A}^n(q) - I) \right\}$$

The tensor $\dfrac{\partial \hat{A}^n(q)}{\partial q}$ can be computed at the same time that the matrix (representation for the operator) $\hat{A}^n(q)$ is computed by making use of the sensitivity equations. For $t \geq 0$ and $q \in Q$, set $\Phi^n(q;t) = e^{A^n(q)t}$. Then $\Phi^n(q;\cdot)$ is the unique principal fundamental matrix solution to the initial value problem

$$\dot{\Phi}^n(q;\cdot) = A^n(q)\Phi^n(q;\cdot), \quad \Phi^n(q;0) = I \qquad (3.14)$$

Setting $\Psi^n(q;t) = \partial \phi^n(q;t)/\partial q$, differentiating with respect to q, interchanging the order of differentiation, and using the product rule, we obtain

$$\dot{\Psi}^n(q;\cdot) = A^n(q)\Psi^n(q;\cdot) + \frac{\partial A^n(q)}{\partial q}\Phi^n(q;\cdot), \quad \Psi^n(q;0) = 0 \qquad (3.15)$$

Combining these two initial value problems and solving we obtain

$$\begin{bmatrix} \dfrac{\partial \hat{A}^n(q)}{\partial q} \\ \hat{A}^n(q) \end{bmatrix} = \begin{bmatrix} \Psi^n(q;\tau) \\ \Phi^n(q;\tau) \end{bmatrix} = e^{\begin{bmatrix} A^n(q) & (\partial A^n(q)/\partial q) \\ 0 & A^n(q) \end{bmatrix}\tau} \begin{bmatrix} 0 \\ I \end{bmatrix} \qquad (3.16)$$

Application to the System Discussed in Section 2.2

Let Q be a closed and bounded subset of R^2 endowed with the Euclidean metric, let $H = L_2(0,1)$ together with the standard inner product $\langle \psi_1, \psi_2 \rangle = \int_0^1 \psi_1(x)\psi_2(x)dx$, and norm denoted by $|\cdot|$, and let V be the Sobolev space $V = H_1(0,1)$ together with its standard inner product $\langle\langle \psi_1, \psi_2 \rangle\rangle = \int_0^1 \psi_1(x)\psi_2(x)dx + \int_0^1 \psi_1'(x)\psi_2'(x)dx$

and norm denoted by $|| \cdot ||$. Then we have the usual dense and continuous embeddings $V \hookrightarrow H \hookrightarrow V^*$, where V^* denotes the space of distributions dual to V. The forms and functions $a(q; \cdot, \cdot) : V \times V \to R, b(q; \cdot) : V \to R$ and $c(\cdot) : V \to R$ are given by

$$a(q; \psi_1, \psi_2) = \psi_1(0)\psi_2(0) + q_1 \int_0^1 \psi_1'(x)\psi_2'(x)dx, \quad \psi_1, \psi_2 \in V \quad (3.17)$$

$b(q; \psi) = q_2\psi(1)$, and $c(\psi) = \psi(0)$, for $\psi \in V$. It follows that $b(q) = q_2\delta(\cdot - 1) \in V^*$ and $c(q) = \delta \in V^*$, where δ denotes the Dirac delta distribution, or unit impulse at zero.

With regard to finite dimensional approximation, for $n = 1, 2, \ldots$, let $\{\psi_j^n\}_{j=0}^n$ denote the set of standard linear B-spline on the interval $[0, 1]$ defined with respect to the usual uniform mesh, $\{j/n\}_{j=0}^n$, and set $H^n = \text{span}\{\psi_j^n\}_{j=0}^n \subset V$ (note the ψ_j^n are the usual "pup tent" or "chapeau" functions of height one and support of width $2/n$, $\lfloor (j - 1)/n, (j + 1)/n \rfloor \bigcap \lfloor 0, 1 \rfloor$). If $P^n : H \to H^n$ denotes the orthogonal projection of $H = L_2(0, 1)$ onto H^n, it is well known (see for example [23]) that $\lim_{n\to\infty} P^n\psi = \psi$ in H for $\psi \in H$ and in V for $\psi \in V$.

3.2 Frequency Domain Techniques and the Impulse Response Function

In this section, we solve the kernel estimation problem using a signal processing approach. Non-parametrically we can rewrite equation (3.7) in a traditional convolution equation in the continuous domain as follows:

$$y(t) = h(t) * u(t) = \int_0^t h(t - \tau)u(\tau)d\tau \quad (3.18)$$

where $*$ denotes the convolution operation.

The convolution Theorem [4] states that if $h(t)$ has a Fourier transform $H(f)$ and $u(t)$ has a Fourier transform $U(f)$, then $h(t) * u(t)$ has a Fourier transform $H(f) \cdot U(f)$. Therefore, the convolution equation in Section 3.1 can be written in a frequency domain representation as

$$Y(f) = \mathcal{F}\{y(t)\} = \mathcal{F}\{h(t) * u(t)\} = \mathcal{F}\{h(t)\} \cdot \mathcal{F}\{u(t)\} = H(f) \cdot U(f) \quad (3.19)$$

where $\mathcal{F}\{\cdot\}$ is the forward Fourier transform. Then we can estimate the spectrum of $h(t)$, i.e. $H(f)$, directly as follows

$$\hat{H}(f) = \frac{Y(f)}{U(f)} = \frac{\mathcal{F}\{y(t)\}}{\mathcal{F}\{u(t)\}} \quad (3.20)$$

The estimated spectrum $\hat{H}(f)$ must be low-pass filtered with a certain cut-off frequency before taking the inverse Fourier transform back to the time domain for the following two reasons. First, the estimated spectrum $\hat{H}(f)$ is very noisy due to the limited number of samples in the training dataset. Second, according to the Nyquist-Shannon sampling theorem [24], if the sampling frequency F_s is twice higher than or equal to the maximum frequency of the actual signal, then that band limited signal can be exactly recovered by low-pass filtering with cut-off frequency at $F_s/2$. In other words, the signal of interest in the frequency domain contains no more information beyond $F_s/2$ provided that the Nyquist-Shannon condition is satisfied. That means the non-zero entries beyond $F_s/2$ in $\hat{H}(f)$ comes purely from the noise component of the signal and therefore we can safely reduce the noise via a low-pass filtering with cut-off frequency at $F_s/2$ without losing any information about the signal. Mathematically, this procedure can be expressed as

$$\hat{H}'(f) = \hat{H}(f) \cdot W_{LPF}(f) \tag{3.21}$$

where $W_{LPF}(f)$ is the low-pass filtering window

$$W_{LPF}(f) = \begin{cases} 1, & f < F_s/2 \\ 0, & f \geq F_s/2 \end{cases} \tag{3.22}$$

Finally, the estimated convolution kernel, denoted as $\hat{h}_{FFT}(t)$, can be obtained from

$$\hat{h}_{FFT}(t) = \mathcal{F}^{-1}\{\hat{H}'(t)(f)\} \tag{3.23}$$

where $\mathcal{F}^{-1}\{\cdot\}$ is the standard inverse Fourier transform.

3.3 ARMA Modeling and the Impulse Response Function

We can also model the underlying system as an ARMA process as follows

$$\phi(t)y(t) = \theta(t)u(t) + e(t) \tag{3.24}$$

where $\phi(t)$ and $\theta(t)$ are the auto-regressive (AR) and moving average (MA) co-efficients respectively, $e(t)$ is the estimation error or as it is sometimes called, the residue or noise, $u(t)$ is the driving process or the input to the system, and $y(t)$ is the observed process or the output of the system. In our case, BrAC is the $u(t)$ and TAC is the $y(t)$.

The system can be identified and the impulse response can be obtained provided $\theta(t)$ and $\phi(t)$ are accurately estimated. There are various ways to estimate the parameters $\theta(t)$ and $\phi(t)$. For example, the AR coefficients $\phi(t)$ can be estimated

by solving the well-known Yule-Walker equation using Levinson's recursion. The MA coefficients $\theta(t)$ can be estimated via Durbin's algorithm. If the data are jointly Gaussian, then a maximum likelihood estimation technique can also be applied to estimate the coefficients. In this effort, to better incorporate delays between BrAC and TAC due to the nature of the data, we adopt the method from [19] which is briefly summarized here.

- We set the maximal possible AR and MA order to be 60 and 120, respectively, heuristically based on the data. In fact, it turns out that our estimation results are not very sensitive to the choices of these two maximal orders.

- We decrease the AR order one at a time forming a set of lower order models and use Minimum Description Length (MDL) criterion to evaluate and select the best one, the one with the lowest MDL score. This implicitly estimates the AR order while the MA order is fixed at the maximal order of choice at this step.

- We set the AR order to be the best one obtained in the previous step and remove MA coefficients one at a time from the one with lowest S/N ratio defined in equation (18) in [19], forming a set of lower order models and then use MDL to select the best one. This step estimates the MA order, thereby establishing the overall ARMA order.

- After finding the optimal orders, we fit the ARMA model based on a least-squares approach to obtain $\theta(t)$ and $\phi(t)$.

Once $\theta(t)$ and $\phi(t)$ are obtained from the steps above, the impulse response $h(t)$ can be estimated directly from $y(t)$ by setting $u(t) = \delta(t)$, where $\delta(t)$ is the standard Dirac delta function. The estimated $\hat{h}(t)$, termed $\hat{h}_{ARMA}(t)$ for this method, is also low-pass filtered with the same cut-off frequency $F_s/2$ as described in Section 3.2.

4   Results

In this section, we applied the three methods to the actual data described below in Section 4.1 and then evaluated their performance. We first performed a test on the training session to check how well the three models could be fit to the data. This was done by convolving the estimated $\hat{h}(t)$ with the training BrAC and comparing the result against the training TAC. Second, to further test the performance of the estimation, we used $\hat{h}(t)$ to deconvolve the testing TACs in another ten drinking episodes of data that were not used in training and then compared the deconvolution results with the contemporaneously collected BrACs. The deconvolution was done via solving an optimization problem described in Section 4.2. Then statistics
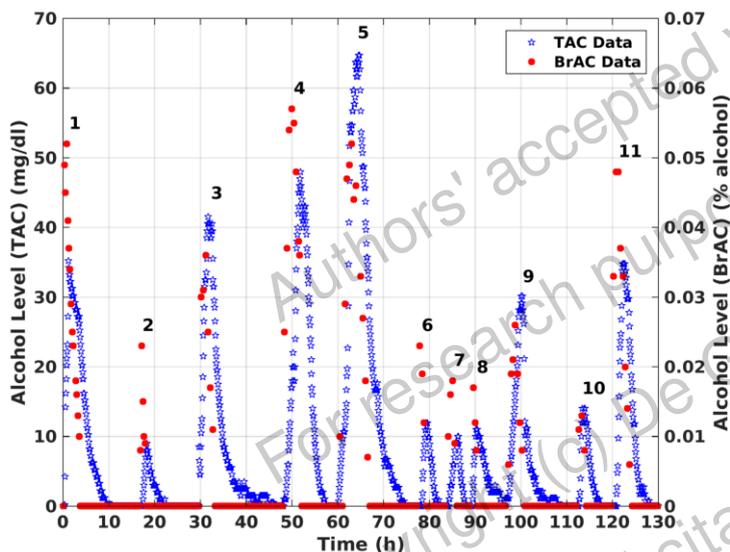
Figure 1. The 11 drinking episodes including the sensor measured transdermal alcohol concentration (TAC) and the contemporaneously measured breath alcohol concentration (BrAC)

were drawn based on the comparison results from the ten testing episodes of data and the numerical results are shown in Section 4.3. In the numerical studies to be discussed below, a drinking episode was defined to be the period of time and the biosensor generated TAC measurements during that period of time, occurring in between two periods of time each demarcated by two consecutive TAC measurements of zero (or at or below a baseline noise level, typically taken to be 2.0-5.0 mg/dl).

4.1 Experimental Data

One of the co-authors (S.E.L.) wore a WrisTAS™ 7 alcohol biosensor for 18 days while also collecting breath measurements. The WrisTAS™ 7 measures the local ethanol vapor concentration over the skin surface at 5-minute intervals. It looks like a digital watch. The participant consumed her first drink in the laboratory with BrAC being measured and recorded every 15 minutes from the start of the drinking session until BrAC returned to 0.000. She then wore the TAC device in the field and consumed alcohol ad libitum for the following 17 days. For each drinking episode, the subject would take BrAC readings every 30 minutes until the BrAC returned to 0.000. Fig. 1 shows the entire 18 day TAC signal along

with the contemporaneous BrAC measurements. Note that the TAC measurements provided by the sensor are in units of milligrams per deciliter (mg/dl) (the scale on the left), while the BrAC measurements are in units of percent alcohol (the scale on the right). Both TAC and BrAC signals were linearly interpolated into a same temporal resolution with interval of 1 minute for kernel estimation as well as deconvolution process which will be discussed below.
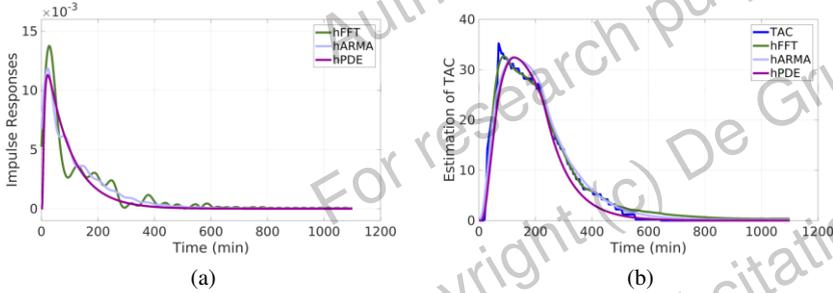

Figure 2. (a) The estimated impulse response $\hat{h}_{PDE}(t)$ (purple), $\hat{h}_{FFT}(t)$ (green) and $\hat{h}_{ARMA}(t)$ (light blue). (b) The estimated TAC from convolution with training BrAC using $\hat{h}_{PDE}(t)$ (purple), $\hat{h}_{FFT}(t)$ (green) and $\hat{h}_{ARMA}(t)$ (light blue). The true TAC is shown in blue.

4.2   A Deconvolution Technique

A deconvolution is a process that estimates the input $u(t)$ given the output $y(t)$ and the estimated kernel $\hat{h}(t)$. The deconvolution result $\hat{u}(t)$ is the answer to the question: what is the profile of BrAC given the measurements in the form of TAC? In order to fairly compare the three methods, we need a deconvolution technique that is independent of all three kernel estimation methods, i.e. the deconvolution process is not biased toward any of the three kernel estimation methods. Therefore, we performed the deconvolution via solving the following optimization problem:

$$\hat{u}(t) = \underset{u(t)}{\operatorname{argmin}} ||A(t)u(t) - y(t)||_{l_2}^2 + \lambda_1 ||u(t)||_{l_2}^2 + \lambda_2 ||\nabla u(t)||_{l_2}^2, \ s.t. \ u(t) \geq 0$$

(4.1)

where $A(t)$ is the convolution operator (Toeplitz convolution matrix in discrete domain) formed from $\hat{h}(t)$, ∇ is the gradient operator and λ_1 and λ_2 are two regularization parameters that both were chosen to be 0.2 empirically to impose the magnitude constraint and the smoothness of the data. We also restrict the result to be non-negative due to the nature of the data and the estimation problem. This op-
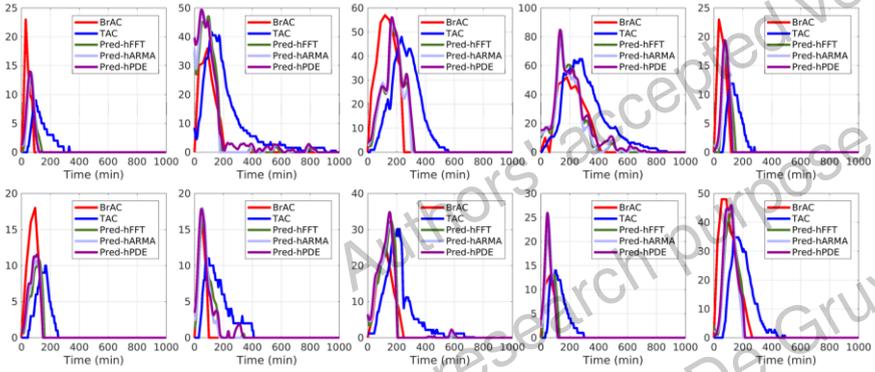
Figure 3. Prediction of BrAC (red) via deconvolution from TAC (blue) using $\hat{h}_{PDE}(t)$ (purple), $\hat{h}_{FFT}(t)$ (green) and $\hat{h}_{ARMA}(t)$ (light blue) on the ten drinking episodes not used in training (top row: episodes 2-6, bottom row: episodes 7-11).

timization problem is a constrained convex problem thus can be solved efficiently via a non-negative least squares technique (Matlab function LSQNONNEG).

4.3 Experimental Results

First, we examined the kernel function estimated from the single training data using the three methods and evaluated the performance of the TAC prediction based on the estimated kernel functions under this training setting. The three estimated impulse responses $\hat{h}_{PDE}(t)$, $\hat{h}_{FFT}(t)$ and $\hat{h}_{ARMA}(t)$ are shown in Fig. 2a. The convolution results on the single training episode of data are shown in Fig. 2b.

Qualitatively, we can see that $\hat{h}_{FFT}(t)$ has more curvature and is less smooth than $\hat{h}_{ARMA}(t)$, and than $\hat{h}_{PDE}(t)$, successively. In addition, the corresponding recovered TAC is closer to the true one using $\hat{h}_{FFT}(t)$ than it is using either $\hat{h}_{ARMA}(t)$ or $\hat{h}_{PDE}(t)$. Quantitatively we evaluate the TAC recovery performance by measuring the error, in l_2 (or L_2) norm sense, between the three estimated TAC curves and the true TAC. It turns out that $\hat{h}_{FFT}(t)$ yields an error of 32.388, $\hat{h}_{ARMA}(t)$ yields an error of 44.898 and $\hat{h}_{PDE}(t)$ yields an error of 65.373. This is the same trend observed by inspection.

Furthermore, we evaluated the prediction performance of the three methods on testing datasets. Fig. 3 shows results of the prediction of BrAC in the other ten episodes of data which are totally unavailable during the training session. To quantitatively test the performance of the predictions and compare the three kernels, we take the following five measures: (1) the height of the peak (peak), (2)

the time delay when peak occurs (delay), (3) the ascending slope (slope 1), (4) the descending slope (slope 2) and (5) the area under the curve (AUC). We define a relative error of any of measurement as

$$\frac{|M_{\widehat{BrAC}} - M_{BrAC}|}{M_{BrAC}} \times 100\% \tag{4.2}$$

where $M_{\widehat{BrAC}}$ and M_{BrAC} are the measurements obtained from the estimated BrAC and the true BrAC.

Fig. 4 shows the statistics of the relative errors of the five measures as box-plots across ten episodes for $\hat{h}_{PDE}(t)$, $\hat{h}_{FFT}(t)$ and $\hat{h}_{ARMA}(t)$, respectively. The associated table of statistics is shown in Table 1.
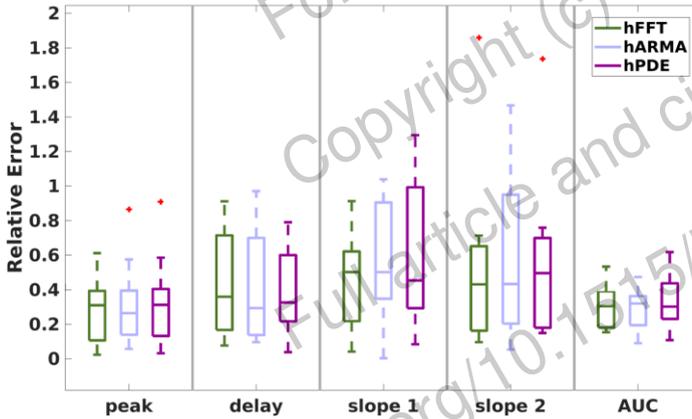

Figure 4. Box plots of the five statistics: the relative errors of peak, delay, slope 1, slope 2 and AUC of the prediction result using $\hat{h}_{PDE}(t)$ (purple), $h_{FFT}(t)$ (green) and $\hat{h}_{ARMA}(t)$ (light blue)

5 Discussion and Concluding Remarks

Overall, the three methods are very comparable. The ARMA-based method yields the most accurate estimation of the peak while the PDE-based method produces the best estimation of the delay. The Fourier-based method has the least variance out of the three, especially for the estimations of the two slopes, although it has slightly larger bias.

Table 1. Five error statistics of the prediction
using (a) $\hat{h}_{PDE}(t)$, (b) $\hat{h}_{FFT}(t)$ and (c)
$\hat{h}_{ARMA}(t)$

		Peak	Delay	Slope 1	Slope 2	AUC
$\hat{h}_{PDE}(t)$	Mean	0.325	0.387	0.579	0.567	0.333
	Median	0.312	0.326	0.454	0.496	0.302
	S.D.	0.268	0.256	0.417	0.468	0.169
$\hat{h}_{FFT}(t)$	Mean	0.291	0.418	0.456	0.519	0.302
	Median	0.310	0.359	0.502	0.431	0.304
	S.D.	0.182	0.315	0.294	0.519	0.129
$\hat{h}_{ARMA}(t)$	Mean	0.319	0.421	0.528	0.552	0.286
	Median	0.264	0.294	0.501	0.433	0.320
	S.D.	0.251	0.338	0.354	0.444	0.123

In terms of parameter estimation, the Fourier-based method is very computationally efficient but theoretically has infinitely many parameters to estimate. On the other hand, the PDE-based method requires the estimation of only two parameters but requires longer computational time to fit the data. The ARMA-based method lies in between those two in terms of computational efficiency and the number of parameters.

It is also worth noting that, from the testing results (Fig. 3) we can see that the peak of TAC is sometimes higher than the peak of BrAC but sometimes lower, reflecting the large variability of the transport and filtering of alcohol by the skin, even within individuals. An LTI model probably may not be capable of fully capturing the dynamics of the process, thus yielding a good estimation of the BrAC given this variability. However, considering that usually only one lab calibration dataset is available for the training, without making any further assumptions of the model, the LTI system based model, including the three models we discussed in this paper, might be a reasonable, simple model to obtain a relative robust estimation of the BrAC.

Out of the three methods, this paper can potentially provide a guideline for the choice of method depending on what property people are looking for in their particular applications. For example, if a court is more interested in an accurate estimation of the timing of the BAC, in case of the determination of guilt of DUI, then the ARMA-based option might be the right choice.

Bibliography

[1] H Thomas Banks, *Functional Analysis Framework for Modeling, Estimation and Control in Science and Engineering*, Taylor and Francis, 2012.

[2] H Thomas Banks and Kazufumi Ito, *A Unified Framework for Approximation in Inverse Problems for Distributed Parameter Systems.*, DTIC Document, Report, 1988.

[3] H Thomas Banks and Karl Kunisch, *Estimation Techniques for Distributed Parameter Systems*, Birkhauser, Boston, 1989.

[4] Ronald Newbold Bracewell, *The Fourier transform and its applications*, 31999, McGraw-Hill New York, 1986.

[5] RF Curtain and D Salamon, Finite-dimensional compensators for infinite-dimensional systems with unbounded input operators, *SIAM Journal on Control and Optimization* **24** (1986), 797–816.

[6] Zheng Dai, I.G. Rosen, Chuming Wang, Nancy Barnett and Susan E. Luczak, Using drinking data and pharmacokinetic modeling to calibrate transport model and blind deconvolution based data analysis software for transdermal alcohol biosensors, *Mathematical Biosciences and Engineering* **13** (2016), 911–934.

[7] D. M. Dougherty, N. E. Charles, A. Acheson, R. M. John, S.and Furr and N. Hill-Kapturczak, Comparing the detection of transdermal and breath alcohol concentrations during periods of alcohol consumption ranging from moderate drinking to binge drinking, *Experimental and Clinical Psychopharmacology* **20** (2012), 373–381.

[8] D. M. Dougherty, N. Hill-Kapturczak, Y. Liang, T. E. Karns, S. E. Cates, S. L. Lake and J. D. Roache, Use of continuous transdermal alcohol monitoring during a contingency management procedure to reduce excessive alcohol use, *Drug and Alcohol Dependence* **142** (2014), 301–306.

[9] D. M. Dougherty, T. E. Karns, J. Mullen, Y. Liang, S. L. Lake, J. D. Roache and N. Hill-Kapturczak, Transdermal alcohol concentration data collected during a contingency management program to reduce at-risk drinking, *Drug and Alcohol Dependence* **148** (2015), 77–84.

[10] Miguel A Dumett, IG Rosen, J Sabat, A Shaman, L Tempelman, C Wang and RM Swift, Deconvolving an estimate of breath measured blood alcohol concentration from biosensor collected transdermal ethanol data, *Applied Mathematics and Computation* **196** (2008), 724–743.

[11] G. C. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control*, Prentice Hall, Englewood Cliffs, 1984.

[12] Alan Wayne Jones, The relationship between blood alcohol concentration (BAC) and breath alcohol concentration (BrAC): a review of the evidence, *Road safety web publication* **15** (2010).

[13] Dominick A Labianca, The chemical basis of the Breathalyzer: A critical analysis, *J. Chem. Educ* **67** (1990), 259–261.

[14] A F J Levi and I G Rosen, A Novel Formulation Of The Adjoint Method In The Optimal Design Of Quantum Electronic Devices, *SIAM Journal on Control and Optimization* **48** (2010), 3191–3223.

[15] Han-Xiong Li and Chenkun Qi, *Spatio-Temporal Modeling of Nonlinear Distributed Parameter Systems*, Springer, Dordrecht, 2011.

[16] J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, 170, Springer, 1971.

[17] E Nyman and A Palmlöv, The elimination of ethyl alcohol in sweat, *Acta Physiologica* **74** (1936), 155–159.

[18] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Applied Mathematical Sciences, Springer, 1983.

[19] Michael H. Perrott and Richard J. Cohen, An efficient approach to ARMA modeling of biological systems with multiple inputs and delays, *IEEE Transactions on Biomedical Engineering* **43** (1996), 1–14.

[20] R. Pintelon and J. Schoukens, *System Identification: A Frequency Domain Approach, Second Edition*, Wiley IEEE Press, 2012.

[21] Anthony J Pritchard and Dietmar Salamon, The linear quadratic control problem for infinite dimensional systems with unbounded input and output operators, *SIAM Journal on Control and Optimization* **25** (1987), 121–144.

[22] I. G. Rosen, Susan E. Luczak and Jordan Weiss, Blind deconvolution for distributed parameter systems with unbounded input and output and determining blood alcohol concentration from transdermal biosensor data, *Applied Mathematics and Computation* **231** (2014), 357–376.

[23] M. H. Schultz, *Spline Analysis*, Prentice-Hall, 1973.

[24] Claude E. Shannon, Communication in the Presence of Noise, *Proceedings of the IRE* **37** (1949), 10–21.

[25] Ralph E Showalter, *Hilbert Sspace Methods in Partial Differential Equations*, Courier Corporation, 2010.

[26] Robert Swift, Transdermal Alcohol Measurement for Estimation of Blood Alcohol Concentration, *Alcoholism: Clinical and Experimental Research* **24** (2000), 422–423.

[27] Robert Swift, Direct measurement of alcohol and its metabolites, *Addiction* **98** (2003), 73–80.

[28] Robert M. Swift, Transdermal measurement of alcohol consumption, *Addiction* **88** (1993), 1037–1039.

[29] Robert M Swift, Christopher S Martin, Larry Swette, Anthony LaConti and Nancy Kackley, Studies on a wearable, electronic, transdermal alcohol sensor, *Alcoholism: Clinical and Experimental Research* **16** (1992), 721–725.

[30] Hiroki Tanabe, *Equations of evolution*, 6, Pitman Publishing, 1979.

Author information

Jian Li, Department of Electrical Engineering Systems
University of Southern California
3740 McClintock Ave. EEB 424
Los Angeles 90089, USA.
E-mail: jli981@usc.edu

Susan E. Luczak, Department of Psychology
University of Southern California, USA.
E-mail: luczak@usc.edu

I. G. Rosen, Modeling and Simulation Laboratory, Department of Mathematics
University of Southern California, USA.
E-mail: grosen@math.usc.edu