On the Modeling and Deconvolution of Blood or Breath Alcohol Concentration (BrAC/BAC) from Biosensor-Measured Transdermal Alcohol Concentration (TAC)

USC Dornsife
Dana and David Dornsife College of Letters, Arts and Sciences

J. Li, S.E. Luczak, I.G. Rosen
University of Southern California

USC Viterbi School of Engineering

Introduction

- All currently available methods (e.g. drink diaries, blood or urine samples, breath analyzers, etc.) to measure and record BrAC or BAC are either inaccurate or not practical in the field. Moreover, using these methods to collect these measurements continuously is essentially impossible.
- Technology developed over the last 20 years has produced alcohol biosensor devices that measure TAC, the amount of alcohol diffusing through the skin. Unfortunately, however, at present, there is no known direct quantitative correlation, in the form of an empirical conversion formula, between TAC and BAC/BrAC.
- In earlier work, we parsed the TAC to BAC/BrAC conversion into two sub-problems. The first is the fitting of a forward model in the form of a partial differential (diffusion) equation that captures the dynamics of the transport of ethanol molecules from the blood through the skin and its measurement by the sensor, and the second is the use of the estimated model to deconvolve estimated BAC/BrAC from the biosensor measured TAC.
- In this study, we compare our earlier approach to two more conventional approaches to the problem, a Fourier/frequency domain based approach and a scheme based on an ARMA input/output response surface and discuss their pros and cons.

Model

Mathematically, we modeled the transdermal transport of ethanol as

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

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

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

$$\varphi(0,\eta) = 0, \qquad 0 < \eta < 1$$

$$y(t) = \varphi(t,0), \qquad t > 0$$

where $\varphi(t,\eta)$ denotes the concentration of ethanol in the epidermal layer of the skin at depth η cm and time t seconds, $u(t)$ is BrAC/BAC and $y(t)$ is TAC.

Method

We compared the following three methods for estimating the kernel function

- **Method 1: Distributed Parameter Modeling**

The input/output system given in the Model Section below can be rewrite as a linear convolution process

$$y(t;q) = \int_0^t h(t-\tau;q)u(\tau)d\tau$$

where $h(t;q) = C(q)e^{A(q)t}B(q)$ is the unknown convolution filter or kernel, $A(q)$ is the partial differential operator appearing in the model, $e^{A(q)t}$ is the analytic semigroup generated by $A(q)$, and $B(q)$ and $C(q)$ are the unbounded input and output operators, respectively.

- **Method 2: Fourier Domain-Based Approach**

The kernel $h(t)$ can be estimated in the frequency domain via

$$\hat{h}(t) = \mathcal{F}^{-1}\{W_{LPF}(f)\,\mathcal{F}\{y(t)\}/\mathcal{F}\{u(t)\}\}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ are the forward and inverse Fourier transform respectively and $W_{LPF}(f)$ is a low-pass filter with cut-off frequency set to the half of the Nyquist sampling rate of the signal to reduce the noise.

- **Method 3: ARMA-Based Approach**

The ARMA system can be expressed as

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

where $\phi(t)$ and $\theta(t)$ are the auto-regressive and moving average coefficients respectively. $e(t)$ is the estimation error. To better incorporate delays between BrAC/BAC and TAC due to the nature of the data, we adopt the method from [PC]. Once $\theta(t)$ and $\phi(t)$ are obtained, $h(t)$ can be estimated directly from $y(t)$ by setting $u(t) = \delta(t)$. The resulting $h(t)$ is also low-pass filtered with $W_{LPF}(f)$ as in Method 2.

- **The Deconvolution Technique**

We performed the deconvolution from TAC to BrAC/BAC via solving the following optimization problem, which is independent of all three kernel estimation methods.

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

where $A(t)$ is the convolution operator formed from $\hat{h}(t)$, ∇ is the gradient operation and λ_1 and λ_2 are two regularization parameters.

Results & Conclusions

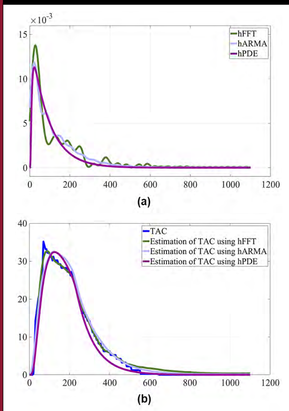

Figure 1

The three estimated impulse responses $\hat{h}_{PDE}(t)$, $\hat{h}_{FFT}(t)$ and $\hat{h}_{ARMA}(t)$ are shown in Fig. 1a. The convolution results on the single training episode of data is shown in Fig. 1b. Qualitatively, we can see that $\hat{h}_{FFT}(t)$ exhibits oscillations and consequently is less smooth than either $\hat{h}_{ARMA}(t)$ and $\hat{h}_{PDE}(t)$ successively. In addition, the corresponding recovered TAC is closer to the true signal using $\hat{h}_{FFT}(t)$ than it is using $\hat{h}_{ARMA}(t)$ or $\hat{h}_{PDE}(t)$. Quantitatively we evaluate the TAC recovery performance by measuring the error, in the l_2 norm sense, between the three estimated TAC curves and the true TAC. We found 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, following the same trend as that from visualizations.

Fig. 2 shows results of the prediction of BrAC in the other ten drinking episodes of data which were not used in training, although we do have contemporaneous BrAC for evaluation. To quantitatively test the performance of the predictions and compare the three kernels, we consider the following five metrics: (1) peak BrAC (peak), (2) the time at which the the peak BrAC occurs (delay), (3) the incline slope of the BrAC curve (slope 1), (4) the decline slope of the BrAC curve (slope 2) and (5) the area under the BrAC curve (AUC).

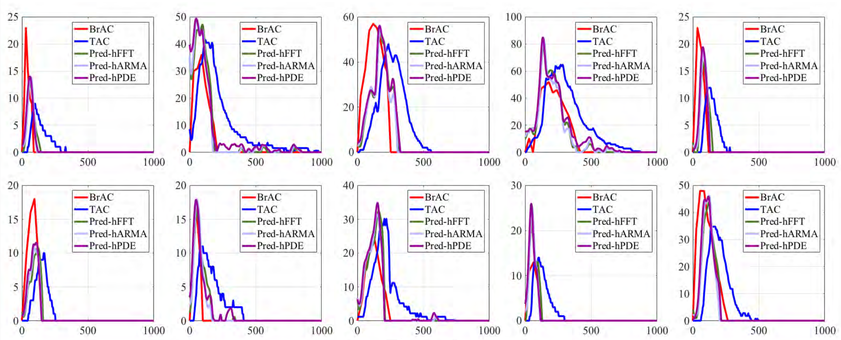

Figure 2

We define a relative error for any of the metrics as

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

where $M_{\widehat{BrAC}}$ and M_{BrAC} are the metrics computed from the estimated BrAC and the true BrAC, respectively. Fig. 3 shows the statistics of the relative errors corresponding to the five metrics in the form of box-plots across ten episodes for (a) $\hat{h}_{PDE}(t)$, (b) $\hat{h}_{FFT}(t)$ and (c) $\hat{h}_{ARMA}(t)$, respectively. The associated table of statistics is shown in Table I.

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 estimation of the two slopes, although it has slightly larger bias.

		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

Table I

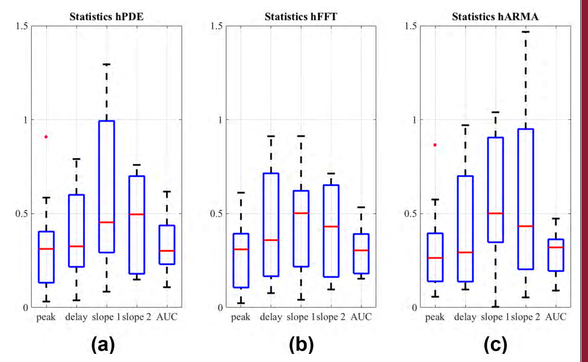

Figure 3

It also worth mentioning that the Fourier-based method is highly computationally efficient but theoretically has infinitely many parameters to estimate. On the other hand, the PDE-based method requires only the estimation of two parameters but requires more computational effort and time to fit the data. The ARMA-based method lies in between those two in terms of computational efficiency and number of parameters.

Reference & Acknowledgements

[PC] Perrott, M. H. and Cohen, R. J., An Efficient Approach to ARMA Modeling of Biological Systems with Multiple Inputs and Delays, IEEE Trans. Biomed. Eng., vol. 43, no. 1, pp. 1–14, 1996.

This project was supported by grants ABMRF and NIAAA (R21AA17711)