



Published in final edited form as:

Biostat Biom Open Access J. 2019 March ; 9(3): 57–68.

Estimating Concentration Response Function and Change-Point using Time-Course and Calibration Data

B Qiang^{1,*}, A Abdalla², S Morgan², P Hashemi², E Peña³

¹Department of Math and Stats, Southern Illinois University Edwardsville, USA

²Department of Chemistry and Biochemistry, USC, Columbia, SC 29208

³Department of Statistics, USC, USA

Abstract

In this paper the problem of determining the functional relationship between time and the concentration of a chemical substance is studied. An intervention drug is administered on the experimental unit from which the chemical substance (specimen) is measured. This drug is hypothesized to cause a change in the concentration level of the chemical substance a certain lag-time after the intervention. However, the concentration value could not be directly measured, but rather a surrogate response can be measured. In the time-course study, this surrogate response is measured using different electrodes which possess varied behaviors. To utilize these surrogate measurements arising from the different electrodes (sensors), a calibration study is undertaken which measures the surrogate response for the different electrodes at known concentration levels. Based on the time-course and calibration data sets, a statistical procedure to estimate the signal function and the lag-time is proposed. Simulation studies indicate that the proposed procedure is able to reasonably recover the signal function and the lag-time. The procedure is then applied to the real data sets obtained during an analytical chemistry experiment.

Keywords

Calibration study; Change-point; General linear model; Chemical concentration; Time-course experiment

Introduction

The general problem tackled in this paper is determining the time-behavior of the concentration of a chemical obtained from an experimental unit (e.u.) subjected to an intervention. Specifically, of interest is to study if a specific drug intervention causes a change in concentration. Concentration levels, however, are difficult to measure directly, especially when the e.u. could not be sacrificed, such as during a time-course study. Thus, the concentration levels are to be inferred indirectly through electric charge measurements which are stochastically related to the concentration levels. The measurement of the charge

This work is licensed under Creative Commons Attribution 4.0 License

*Corresponding author: B Qiang, Department of Math and Stats, Southern Illinois University Edwardsville, Edwardsville, USA.

is reliant on the use of an electrode. But different electrodes possess different behaviors when measuring the charge. Thus, to be able to use different electrodes to infer the concentration levels, a calibration study using these electrodes is performed to determine the relationship between concentration and charge for each of the electrodes.

The specific study that motivated the problem considered in this paper was performed in the laboratory of one of the authors (Dr. P. Hashemi) at the Department of Chemistry and Biochemistry, University of South Carolina. The chemical concentration of interest is that of serotonin, a substance that has been implicated in affective disorders such as depression, and the e.u.'s are mice (*in vivo*). The intervention performed on the mice is the administration of a pharmacological agent or drug. The type of data sets obtained from the study are shown in Figures 1 & 2. Figure 1 shows charge measurements over time for five different electrodes, with panels 1 and 2 showing results for the two intervention agents or drugs: pargyline and GBR 12909, respectively. Pargyline inhibits serotonin metabolism while GBR 12909 is a dopamine reuptake inhibitor. These drugs or agents inhibit serotonin metabolism and dopamine reuptake. Figure 2 presents the calibration data for ten electrodes, together with fitted mean response curves. The five electrodes used in the time-course study for each of the two drugs came from these 10 electrodes utilized in the calibration study. For instance, the electrodes used in the time-course experiment with pargyline were the electrodes labeled 2, 3, 4, 5, and 10; whereas, for GBR 12909 the electrodes were those labeled 1, 6, 7, 8, and 9.

The goal in this study is to estimate the mean concentration response function over time based on the time course study and the calibration study. Another aspect in the time-course study is that the drug intervention was administered after an initial no-drug period. Thus, another important goal is to determine the lag-time post administering the drug after which the drug takes effect, if indeed the drug has an effect. Discipline-specific details of this motivating and focused application has been previously published [1]. We point out the statistical methodology developed in this paper and the ideas contained herein have the potential of being useful in other situations where calibrated measurements are obtained [2].

Mathematical setting

In this section we describe the postulated models for both the time-course study and the calibration study. For this purpose, let us suppose that we have a specimen from the e.u. We denote by X_t the concentration level of the chemical at time t for this specimen. We denote by T the time at which the intervention is performed, that is, the time of administering the drug intervention. We shall postulate the stochastic model

$$X_t = \gamma_0 + \gamma_1 t + g(t - (T + \Delta)) + \varepsilon_t, \quad t \geq 0, \quad (1)$$

where Δ is a lag-time and $g(\cdot)$ is a continuous function with $g(t) = 0$ for $t \leq 0$. If smoothness is desired, we may impose the condition that $g(\cdot)$ is differentiable for $t > 0$. The time of intervention T is known, while the lag-time Δ , together with the regression coefficients γ_0 and γ_1 , and the function $g(\cdot)$ are unknown. The error is assumed to be white noise, with ε_t having a normal distribution with mean zero and variance τ^2 , with τ also unknown.

As discussed in section 1, the X_t could not be directly measured. If it could be directly measured and we possess the values of X_t 's, then we could estimate the model parameters, and also infer the change point T^+ , or the lag-time.

To enable the determination of the chemical concentration levels X_t 's, there are K possible electrodes that could be used for measuring the charge. At time t , n_t charge measurements using different electrodes will be taken. Thus, denoting by Y the charge and by E the electrode type, at time t , n_t pairs $(Y_{tj}, E_{tj}), j = 1, 2, \dots, n_t$, are obtained. The measured charge is affected by the chemical concentration level and the electrode type. The relationship between the charge and the concentration and electrode type is given by

$$Y = \left[\beta_0 + \sum_{k=2}^K \xi_k I\{E = k\} \right] + \left[\beta_1 + \sum_{k=2}^K \eta_k I\{E = k\} \right] X + \epsilon \quad (2)$$

where $\beta_0, \beta_1, \xi_2, \dots, \xi_k, \eta_2, \dots, \eta_k$ are unknown, and ϵ has a normal distribution with mean zero and variance σ^2 , with σ unknown. The function $I(\cdot)$ is the indicator function, taking a 1(0)-value depending on whether the argument is true(false). This is a linear model that incorporates an interaction between the concentration and the electrode type [3,4].

In order to estimate the model parameters in model (2), a calibration study is performed. See [5] for some review of calibration methods. In this study, we followed the classic inverse regression approach. First, we regress the response variable, Y , on predictor X , and estimate linear regression coefficients using least-squares method; thereafter, the value of an unknown X , is to be estimated given an observation of Y , by subtracting the estimated intercept and dividing by the estimated slope [6–9]. Inference for the calibration parameters, is not trivial because of the presence of a normally distributed estimated slope in the denominator, which causes the inverse estimator to have infinite variance [10]. In this paper, we use Delta method to construct approximate confidence intervals for the calibration parameters. A more conservative confidence interval approach based on inverting simultaneous tolerance intervals was proposed by Scheffe [9] in literature. An alternative approach to the problem is referred to as reverse regression, when X 's are treated as the response and formally regressed on Y 's (even though the X 's are measured with negligible error). Krutchkoff [4] compared inverse and reverse regression using Monte Carlo simulations. Properties and limitations of the reverse estimators were studied by Williams [10] and Halperin [3].

In this calibration study, known levels of concentration are used, and the different electrodes are used to measure the charge. We denote by L the number of concentration levels, and these levels will be denoted by $x_{01} < x_{02} < \dots < x_{0L}$. At concentration level x_{0j} , all K electrode types are utilized, and for each electrode type, there are m charge measurements. Thus, for the j^{th} concentration level, there are Km observations, and for the L concentration levels, there are $N = LKm$ charge measurements. The data could be summarized as in Table 1 and pictorially depicted as in Figure 2.

With $\xi_1 = \eta_1 = 0$, the linear models governing these charge measurements are given by

$$Y_{ijl} = (\beta_0 + \xi_j) + (\beta_1 + n_j)x_{0i} + \epsilon_{ijl} \quad (3)$$

for $i = 1, \dots, L; j = 1, \dots, K; l = 1, \dots, m$, and with the ϵ_{ijl} 's being independent and identically distributed (IID) and having a normal distribution with mean zero and variance σ^2 . These linear models could be fitted using object functions in a variety of statistical packages, such as the function `lm` or `glm` in the R statistical package [8].

Estimating parameters

Instead of using the calibration data representation presented in Table 1, for purposes of describing more concisely the estimators of parameters, we denote by $Y = (Y_1, Y_2, \dots, Y_N)^T$ the $N \times 1$ vector of charge values. The design matrix is W , which is an $N \times 2K$ matrix whose i th row is

$$W_i = [1, x_i, v_{2i}, \dots, v_{Ki}, v_{2i}x_i, \dots, v_{Ki}x_i]$$

with x_i the concentration level, and $v_{ji} = I\{E_i = j\}$, $j = 2, \dots, K$, indicates whether the electrode type is j . With $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_N)^T$ denoting the error vector, the linear model could be written via

$$Y = W\theta + \epsilon \quad (4)$$

where the $2K \times 1$ regression coefficient vector θ is

$$\theta = [\beta_0, \beta_1, \eta_2, \dots, \eta_K, \xi_2, \dots, \xi_K]^T. \quad (5)$$

In this model, $\epsilon \sim N_N(0, \sigma^2 I_N)$ where I_N is the $N \times N$ identity matrix.

The least-squares (LS) estimator of θ , which is also the maximum likelihood (ML) estimator, is given by (see, for instance, any linear theory book such as [6])

$$\hat{\theta} = [W^T W]^{-1} [W^T Y]. \quad (6)$$

This is an unbiased estimator of θ . The error variance σ^2 is unbiasedly estimated by

$$\hat{\sigma}^2 = \frac{1}{N - 2K} \|Y - W\hat{\theta}\|^2 = \frac{1}{N - 2K} Y^T [I_N - H] Y \quad (7)$$

where $H = W[W^T W]^{-1} W^T$. An unbiased estimator of the covariance matrix of $\hat{\theta}$ is

$$\widehat{Cov}(\hat{\theta}) \equiv \hat{\sigma}^2 [W^T W]^{-1}. \quad (8)$$

The validation of the assumptions underlying this linear model could be performed graphically or via the global method in [7].

Calibrated concentration estimators

The calibration problem will now be described. Suppose that, at a given time, we are given a charge measurement for a specific electrode: say, (y_0, e_0) , where e_0 is the electrode type and y_0 is the charge on that specific electrode. More generally, suppose that M pairs of charge and electrode type are available at a given time: $(y_{01}, e_{01}), (y_{02}, e_{02}), \dots, (y_{0M}, e_{0M})$. Based on these observations, what is an estimate of the concentration level, and what is an approximate $100(1-\alpha)\%$ confidence interval for the concentration level?

Calibration based on one pair (y_0, e_0)

Consider first just having one pair (y_0, e_0) of charge and electrode type observations. Denote by $x_0 \equiv (y_0, e_0)$ the concentration level that led to this charge value of y_0 for electrode type e_0 . Based on the linear model relationship, y_0 is a realization of the random variable

$$y_0 = (\beta_0 + \xi_{e0}) + (\beta_1 + \eta_{e0})x_0 + \epsilon_0$$

with the convention that $\xi_1 = \eta_1 = 0$ and $\epsilon_0 \sim N(0, \sigma^2)$. Solving for x_0 , we obtain

$$X_0(y_0, e_0, \epsilon_0; \theta) = \frac{y_0 - (\beta_0 + \xi_{e0}) - \epsilon_0}{\beta_1 + \eta_{e0}}. \quad (9)$$

Of course, the error term ϵ_0 is not observable, but it has mean 0 and variance σ^2 . The expected value of $X_0(y_0, e_0, \epsilon_0; \theta)$, given (y_0, e_0) , is

$$X_0(y_0, e_0, \epsilon_0; \theta) = E\{X_0(y_0, e_0, \epsilon_0; \theta) | (y_0, e_0)\} = \frac{y_0 - (\beta_0 + \xi_{e0})}{\beta_1 + \eta_{e0}}$$

and its conditional variance is

$$Var\{X_0\{X_0(y_0, e_0, \epsilon_0; \theta) | (y_0, e_0)\}\} = \frac{\sigma^2}{(\beta_1 + \eta_{e0})^2}.$$

Consequently, a plausible estimator of $x_0(y_0, e_0)$ is

$$\hat{x}_0(y_0, e_0; \hat{\theta}) = g_{e0}(y_0; \hat{\theta}) \equiv \frac{y_0 - (\hat{\beta}_0 + \hat{\xi}_{e0})}{\hat{\beta}_1 + \hat{\eta}_{e0}}. \quad (10)$$

This is the calibrated estimate of the concentration when given a charge value of y_0 obtained using the electrode type e_0 . By the Delta-Method [2], this will be approximately unbiased

for the expectation of $x_0(y_0, e_0)$. We seek an approximation of its variance by using the delta method. For $j = 1, 2, \dots, K$, define the gradients

$$b_j(y_0; \theta) = \frac{\partial}{\partial \theta} \left[\frac{y_0 - (\beta_0 + \xi_j)}{\beta_1 + \eta_j} \right].$$

By the Delta-Method, an estimate of the variance of $\hat{x}_0(y_0, e_0)$ is

$$\begin{aligned} \text{Var} \left[\widehat{x}(y_0, e_0; \hat{\theta}) \right] &\equiv V_0^2(y_0, e_0) = \\ &\hat{\sigma}^2 \left[b_{e_0}(y_0, \hat{\theta})^T [W^T W]^{-1} b_{e_0}(y_0, \hat{\theta}) + \frac{1}{(\hat{\beta}_1 + \hat{\eta}_{e_0})^2} \right]. \end{aligned} \tag{11}$$

An approximate $100(1-\alpha)\%$ confidence interval for the concentration level, having observed a charge value of y_0 using electrode type e_0 , is given by

$$\left[\hat{x}_0(y_0, e_0; \hat{\theta}) \pm (t_{N-2K; \alpha/2}) V_0(y_0, e_0) \right], \tag{12}$$

Where $t_{N-2K; \alpha/2}$ is the $(1-\alpha/2)$ th quantile of a t-distribution with degrees-of-freedom $N-2K$.

Calibration based on many pairs (y_0, e_0)

Next we consider the situation where several charge measurements are taken using possibly different electrode types. Let

$$(y_0, e_0) = \{(y_{0m}, e_{0m}) : m = 1, 2, \dots, M\}.$$

Here y_{0m} is the m th charge measurement which is obtained using electrode type e_{0m} . From the preceding subsection, based on this particular measurement we obtain an estimate of the concentration level, given by

$$\hat{x}_0(y_0, e_0; \hat{\theta})$$

which has an approximate estimated variance of $V_{0m}^2(y_{0m}, e_{0m})$ whose expression is obtained via (11). The M estimates of the concentration level obtained for each of the elements in (y_0, e_0) will not be independent of each other, owing to the fact that they will all depend on $\hat{\theta}$. In fact, through the Delta-Method, we could obtain their approximate covariance matrix. However, for the sake of simplicity and practicality, we ignore the dependence among these M estimates. Under this simplified assumption, whose appropriateness will be examined later through simulation studies, we could then combine the M estimates by simply taking into account the possibly varying estimates of their variances. We recall the following

well-known distribution theory result, which is easily proved using a Lagrange multiplier minimization approach.

Theorem 1

Let W_1, W_2, \dots, W_m be independent random variables with common mean μ and respective variances $\tau_1^2, \tau_2^2, \dots, \tau_m^2$. Among all linear combinations $\sum_{l=1}^m c_l W_l$ with $\sum_{l=1}^m c_l = 1$, so that the mean of the linear combination is still μ , the one with the smallest variance coincides with the choice of coefficients given by

$$c_l^* = \frac{(1/\tau_l^2)}{\sum_{l=1}^m (1/\tau_l^2)}, \quad l = 1, 2, \dots, m.$$

The variance of this optimal linear combination is

$$V = \left[\sum_{l=1}^m c_l^* W_l \right] = \left[\frac{1}{m} \sum_{l=1}^m \frac{1}{\tau_l^2} \right]^{-1}.$$

Using this result, we obtain our combined estimate of the concentration level from the M estimates via

$$\hat{x}_0(y_0, e_0; \hat{\theta}) = \sum_{m=1}^M \hat{c}_m \hat{x}_0(y_{0m}, e_{0m}; \hat{\theta}), \tag{13}$$

where the weights are given by

$$\hat{c}_m = \frac{[1/V_{0m}^2(y_{0m}, e_{0m})]}{\sum_{q=1}^M [1/V_{0q}^2(y_{0q}, e_{0q})]}, \quad m = 1, 2, \dots, M. \tag{14}$$

The approximate variance of $\hat{x}_0(y_0, e_0; \hat{\theta})$ is given by

$$V_{\hat{x}_0}^2(y_0, e_0) = \left[\frac{1}{M} \sum_{m=1}^M \frac{1}{V_{0m}^2(y_{0m}, e_{0m})} \right]^{-1}, \tag{15}$$

As a consequence, an approximate $100(1-\alpha)\%$ confidence interval for the concentration level, having observed (y_0, e_0) , is given by

$$\left[\hat{x}_0(y_0, e_0; \hat{\theta}) \pm (t_{N-2K; \alpha/2}) V_{\hat{x}_0}(y_0, e_0) \right]. \tag{16}$$

Time-course study

As mentioned in section, 1 the main goal of the study is to determine the behavior of the concentration of the chemical over a period of time, where at some point T during the monitoring period, a drug intervention is performed. We suppose that over the period $[0, T^*]$ charge measurements using possibly different electrode types are performed at specified times $0 < t_1 < t_2 < t_3 < \dots < t_L < T^*$. The charge and electrode observations at time t_l are given by (y_{0l}, e_{0l}) . Based on these observation vectors, we find an estimate of the concentration level at time t_l to be

$$\hat{x}_{0l}(y_{0l}, e_{0l}),$$

together with its estimate of its variance $V_{0l}^2(y_{0l}, e_{0l}; \hat{\theta})$. For each time t_l , we could also construct the approximate $100(1 - \alpha)\%$ confidence interval for the concentration value given by

$$[\hat{x}_{0l}(y_{0l}, e_{0l}; \hat{\theta}) \pm (t_{N - 2K; \alpha/2})V_{0l}(y_{0l}, e_{0l})].$$

Note, however, that these confidence intervals are not adjusted for multiplicity. On the other hand, based on the time course data and the resulting concentration estimates at each of the times of observations, given by

$$\{(t_l, \hat{x}_{0l}(y_{0l}, e_{0l}; \hat{\theta}), V_{0l}(y_{0l}, e_{0l})) : l = 1, 2, \dots, L\},$$

we could fit an appropriate regression curve that takes into account the possibly differing variability of the $\hat{x}_{0l}(y_{0l}, e_{0l}; \hat{\theta})$'s. One of the simplest models that could be fitted to this data, under the hypothesis that upon the drug intervention at time T the concentration curve should change after a lag-time of Δ , specifies that

$$\hat{x}_{0l} = \omega_0 + \omega_1 t_l + k_1 U_l + k_2 U_l^2 + V_{0l\epsilon l}, \tag{17}$$

and

$$U_l = \max(0, t_l - (T + \Delta)), l = 1, 2, \dots, L,$$

and ϵ_l has mean zero and variance τ^2 . Note that the τ^2 in this model is not the same as the τ^2 in the time-course model in (1). Here, the time of intervention T is known, whereas the lag-time Δ is not known. For a specified Δ , this model is easily fitted using the `lm` command in R with the `weights` option enabled, which performs a weighted linear regression fit [8]. The coefficient of determination could be obtained and denoted by $R^2(\Delta)$. The coefficient of determination could be plotted with respect to possible values of Δ . The value of Δ that maximizes this coefficient of determination $R^2(\Delta)$ is a plausible estimate for the lag-time Δ . Thus, a possible estimator of the lag-time is

$$\hat{\Delta} = \operatorname{argmax}_{\Delta} R^2(\Delta). \quad (18)$$

This could be computed by fitting the model over a sequence of values of Δ and searching for the value of Δ that maximizes $R^2(\Delta)$. This computational method is the approach we implemented in this paper. The confidence interval described above at a given time point is appropriate if we only have data at one time point. However, in the time-course study, we actually obtain a series of concentration estimates for each of the time points considered. These pairs of time and concentration estimates are then used to estimate the time course model parameters. In the process of fitting the linear-parabolic curve, we could then also obtain pointwise confidence intervals for the concentration value at each of the time points, where we utilize the estimate of the error standard deviation. We surmise that the resulting point-wise confidence intervals is a better indicator of where the concentration values are. In the real-data application of the procedure in section 6, this point-wise confidence interval approach will be presented.

A real-data illustration

This section provides a more detailed description of the statistical methods performed in the data analysis in some portions of paper [1]. The data sets were obtained in the Hashemi laboratory at University of South Carolina. We limit our consideration to the intervention drug pargyline, which was hypothesized to have an impact on the concentration level of serotonin, in contrast to the intervention drug GBR 12909. Figures 1–2 present the time-course data and the fitted linear models. These figures are adapted with permission from [1] and are copyrighted from American Chemical Society. In the next section, we then present simulation studies to provide us some ideas on the properties of this procedure.

A time-course study was performed leading to the data set with charge measurements over time (from 0 minutes to 121 minutes), with the intervention drug pargyline (10 mg/kg) administered at 60 minutes, via interperitoneal injection. The time-course data with pargyline as the intervention drug is depicted in the first panel of Figure 1. The different symbols correspond to the 5 different electrodes (out of 10 possible electrodes) used in measuring the charges. A calibration study was also performed where, for known concentration levels, charge measurements were obtained for each of 10 possible electrodes. The charges of each electrode are measured under three concentration values (25, 50, and 100 nM) each with 4 replicates. The data set obtained from this calibration study are the solid points in Figure 2.

A linear model, as described in (2), was fitted using the calibration data. The fitted linear models, using the calibration data, for each electrode type are depicted as the lines in Figure 2 for each of the 10 electrodes. A summary of the estimated model parameters are in Table 2. These linear models with interaction terms provide excellent fits to the observed calibration data with coefficient of determinations all above 99%. The interaction effects are all significant.

Using these fitted linear models, given the charge measurements at each time point from the time course study, estimates of the concentration levels at each time point were obtained. This procedure yielded pairs of values of time and concentration estimates which are the solid points in Figure 3. A functional continuous model, as described in Section 5, was fitted to these pairs of time and concentration values using weighted regression via the `lm` command in the R environment. The estimate of the time-lag after which the drug takes effect is $\hat{\Delta} = 3.32$ minutes. As mentioned earlier, this estimate was obtained by maximizing the coefficient of determination with respect to the possible values of Δ [see discussion prior to (18)]. The resulting fitted linear-parabolic model whose equation is given in (18) is shown in Figure 3.

$$\hat{C}(t) = 64.584 + 0.0063t + 0.5163U(t) - 0.0044U(t)^2 \quad (19)$$

where $U(t) = \max(0, 63.32 - t)$. This plot is depicted together with the concentration estimates at each time point, which are the open circles in the plot. Point-wise confidence interval at each time point is also included in the plot. These 95% point-wise confidence intervals were constructed when the functional model was fitted to the pairs of time and concentration values using the `predict.lm` command in R. Details pertaining to the fitted functional relationship between time and concentration are summarized in Table 3. Observe that the coefficient for the time effect is not significantly different from 0 ($p = .429$) which is to be expected since without drug intervention the serotonin concentration is not expected to change. We also mention that the final fitted model has an R^2 equal to 95.59%, indicating an excellent fit of the linear-parabolic model for relating concentration to time for this study.

Simulation studies

In order to study the properties of the procedure described above for estimating the lag-time parameter Δ , we performed a computer simulation study. All simulation programs were coded in R and function objects in R such as `lm`, `rnorm`, etc. were utilized. The main purpose of this study was to determine if we are able to effectively estimate the lag-time parameter Δ using a time-course data and a calibration data. In the simulation, we generate the time-course data using the model

$$x_t = \gamma_0 + \gamma_1 t + k_1 \max(0, t - (T + \Delta)) + k_2 [\max(0, t - (T + \Delta))]^2 + \tau Z_t^{(1)} \quad (20)$$

where $\gamma_0 = 20$, $\gamma_1 = .25$, $k_1 = 10$, $k_2 = -.12$, $T = 60$, $\tau = 10$ and with $Z_t^{(1)}$'s variables.

The time index took values in $\{0, 1, 2, \dots, 120\}$. Thus, for each $t \in \{1, 2, \dots, 120\}$ an x_t was generated according to this model.

At time t and for the generated x_t -value, a charge measurement was undertaken using three electrode types: $m = 1, 2, 3$. The charge value was generated according to the model

$$Y_m = (\beta_0 + \xi_m) + (\beta_1 + \eta_m)x_t + Z_{tm}^{(2)} \quad (21)$$

Where $Z_{tm}^{(2)}$'s are IID random variables. The parameter values were set to $\beta_0 = 1$, $\beta_1 = 1.5$, $\xi = (0,1,2)$, $\eta = (0,.5,1.5)$. Thus, this leads to the 'observed time-course data' given by

$$\{(t, y_{tm}), m = 1, 2, 3; t = 0, 1, 2, \dots, 120\}.$$

A calibration experiment was also simulated. In this calibration experiment, five specific concentration values $x \in \{0, 10, 25, 50, 100\}$ were utilized. At each of these concentration values, five charge values were generated for each of the three electrode types according to the model

$$Y_{jmk} = (\beta_0 + \xi_m) + (\beta_1 + \eta_m)x_j + Z_{jmk}^{(3)}, k = 1, 2, \dots, 5,$$

where the parameter values are the same as in the preceding model, and $x_1 = 0$, $x_2 = 10$, $x_3 = 25$, $x_4 = 50$, $x_5 = 100$ and $Z_{jmk}^{(3)}$'s are IID $N(0,1)$ random variables.

On the basis of the time-course data

$$\{(t, y_{tm}), m = 1, 2, 3; t = 0, 1, 2, \dots, 120\}$$

and the calibration data

$$\{(x_j, y_{jmk}), j = 1, 2, \dots, 5; m = 1, 2, 3; k = 1, 2, \dots, 5\},$$

the goal is to see if we could reasonably recover the signal function given by

$$g(t) = E\{X_t\} = \gamma_0 + \gamma_1 t + k_1 \max(0, t - (T + \Delta)) + k_2 [\max(0, t - (T + \Delta))]^2 \quad (22)$$

and obtain good estimates of the model parameters using the procedure described in the preceding sections, in particular to obtain a reasonable estimate of $\hat{\Delta}$.

We performed the above-described simulation experiment using 5000 replications. The estimates of $\hat{\Delta}$ are provided in the frequency histogram in Figure 4. In searching for the optimizing $\hat{\Delta}$, we used increments of 0.1, so that the possible values of $\hat{\Delta}$ contains only one decimal digit. The mean of these $\hat{\Delta}$'s is 5.03828 with a standard error of 0.4990. Thus, this mean is quite close to the true value of Δ which is 5.0. Regarding the other model parameters pertaining to the time-course portion, we summarize the results in Table 4. Histograms of these 5000 estimates for the four model parameters are depicted in Figure 5. Observe that the means of these estimates for each of the regression parameters are close to their true values.

Another parameter in the time-course model is τ , which is the standard deviation of the error component. When we fit the linear-parabolic model on the time-course and calibration data, an estimate of the error standard deviation is also obtained. However, this is not an estimate of τ since this estimates a larger value than τ owing to the additional error contamination

arising from the calibration study and the charge measurements at each time and for each electrode type. For instance, in the simulation study, the histogram of the 5000 estimates of the standard deviation of the error term provided in Figure 6 all exceed the true value of $\tau = 10$. The mean of these standard deviation estimates is 23.56979 with a standard deviation of 2.723542. On the basis of this modest simulation study, the statistical procedure developed in earlier sections using the calibration and time-course data sets appears successful in recovering the functional relationship between time and concentration and is also able to infer properly about the lag-time at which the drug starts to become effective.

In order to further study the properties of the statistical procedure on the real data, we performed another simulation study using the estimated model parameters based on the data from the Hashemi laboratory. The time course data was generated using model (20), with $T = 60$, $\beta_0 = 3.3$, $\tau = 0.3$, and with $Z_t^{(t)}$'s IID $N(0,1)$ variables. The values of other parameters were set to be the same as the estimated ones in Table 3. The time index took values in $\{0,1,2,\dots,121\}$. Charge measurements were obtained using five electrodes $m = 2, 3, 4, 5, 10$. The charge values were generated according to the model (21) with parameter values

$$\beta_0 = 0.6950, \beta_1 = 0.0116;$$

$$\xi = (2.1707, 2.6107, 2.4622, 1.42622, 1.4289, 1.3953);$$

$$\eta = (0.0270, -0.0054, 0.0152, 0.0082, 0.0024),$$

and

$$Z_{tm}^{(2)} \text{ and IID } N(0, 0.095^2).$$

These correspond to the estimates of the model parameters in section 6. To simulate the calibration data, we generated charge values with concentration taking values $x \in \{25, 50, 100\}$ for 10 electrodes, i.e. $m = 1, 2, \dots, 10$, according to

$$Y_{jmk} = (\beta_0 + \xi_m) + (\beta_1 + \eta_m)x_j + Z_{jmk}^{(3)}, k = 1, 2, 3, 4,$$

where the parameter values were set identical to the estimated values in Table 2, and with $x_1 = 25$, $x_2 = 50$, $x_3 = 100$, and $Z_{jmk}^{(3)}$'s being IID from a $N(0, 0.095^2)$ distribution. Using the time-course data and the calibration data, it is of interest to see if we could reasonably recover the signal function in (22) and also obtain good estimates of the model parameters.

The resulting 5000 estimates of $\hat{\Delta}$ are provided in the frequency histogram in Figure 7. The mean of these 5000 $\hat{\Delta}$'s is 3:3008 with a standard error of 1:9341. The mean is indeed very close to the true value which is $\Delta = 3.3$, indicating the possible unbiasedness of the

estimator, though of course a simulation study could not establish this theoretical property. Note, however, that there were some negative estimates that arose and this could be a consequence of a wrong (we hypothesize larger values) specification of the error variances in the models. Details of the estimates for the other parameters are summarized in Table 5. Histograms of the estimates for these four model parameters are depicted in Figure 8.

Concluding Remarks

Motivated by a study in an analytical chemistry laboratory dealing with the concentration of the chemical substance serotonin, we developed in this paper a statistical procedure for estimating the functional relationship between time and concentration level of a substance, together with the change point, based on a time-course study that measures a concentration-surrogate variable (the charge in this application) and data from a calibration study. The novel aspect of this procedure is the presence of several measuring electrodes which have their unique behaviors when utilized to measure the surrogate variable (the charge in application). Simulation studies were performed to examine the properties of the proposed procedures, and the simulation results indicate that the procedure is able to reasonably recover the signal function and also the change-point in the signal function.

The procedure is also applied to the time-course data and the calibration data from the focus application obtained in the analytical chemistry experiment. Further theoretical studies and simulations will be desired to examine more carefully the procedure especially when applied to more complex settings such as when the $g(\cdot)$ portion in the signal function in equation (1) is not parabolic or when it is non-parametrically specified. Of interest for further research are situations where the error distributions in the regression models are not normally-distributed, which may entail the use of non-parametric regression estimation methods.

Acknowledgements

E. Peña and B. Qiang acknowledge grant NIH P30GM10333601A1 and the Biometry Core of the Center of Colon Cancer at the University of South Carolina. P. Hashemi and A. Abdalla acknowledge The University of South Carolina Start-Up Funds, The Eli Lilly Young Investigator Award in Analytical Chemistry, and NIH MH106563. We also thank Professor James Lynch and statistics graduate students Taeho Kim, Shiwen Shen, Jeff Thompson, and Lu Wang for their comments provided during regular research meetings.

Abbreviations:

EU	Experimental Unit
IID	Identically Distributed
LS	Least-Squares
ML	Maximum Likelihood

References

1. Aya Abdalla, Christopher WA, Pavithra P, Srimal S, Beidi Qiang, et al. (2017) In vivo ambient serotonin measurements at carbon-fiber microelectrodes. *Analytical Chemistry* 89(18): 9703–9711. [PubMed: 28795565]
2. George Casella, Berger Roger L (1990) *Statistical inference*. The Wadsworth & Brooks/Cole Statistics/Probability Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, USA.
3. Max Halperin (1970) On inverse estimation in linear regression. *Technometrics* 12(4): 727–736.
4. Krutchkoff RG (1967) Classical and inverse regression methods of calibration. *Technometrics* 9(3): 425–439.
5. Irma Lavagnini, Magno Franco (2007) A statistical overview on univariate calibration, inverse regression, and detection limits: Application to gas chromatography/mass spectrometry technique. *Mass Spectrometry Reviews* 26: 1–18, 2007. [PubMed: 16788893]
6. Neter J, Kutner M, Nachtsheim C, Wasserman W (1996) *Applied Linear Statistical Models*. Irwin.
7. Edsel A Peña, Slate Elizabeth H (2006) Global validation of linear model assumptions. *J Amer Statist Assoc* 101(473): 341–354.
8. R Core Team (2016) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
9. Henry Scheffe (1973) A statistical theory of calibration. *The Annals of Statistics* 1(1): 1–37.
10. Williams EJ (1969) A note on regression methods in calibration. *Technometrics* 11(1): 189–192.

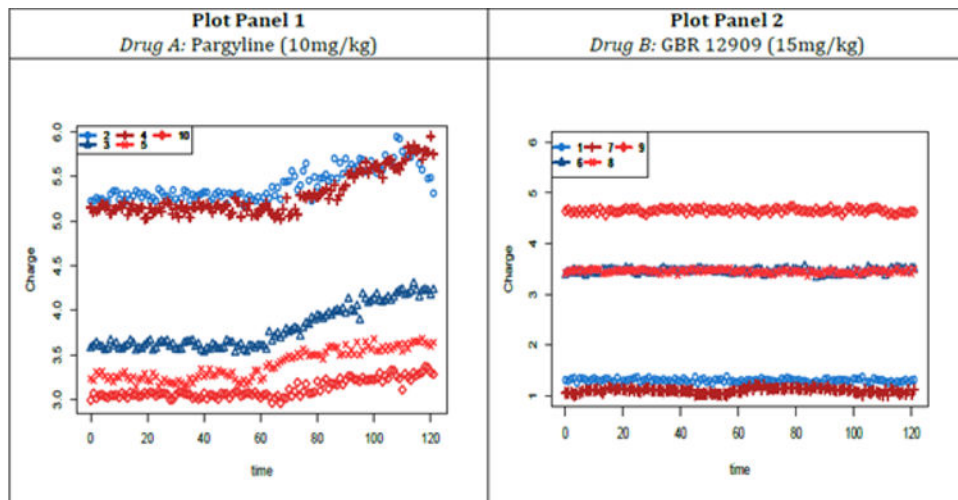


Figure 1: Charge measurements (in pC) on the specimens obtained in the time-course study associated with 5 electrodes over time (in minute) for each of the two drug interventions. Plot panel 1 is for the drug pargyline, while plot panel 2 is for the drug GBR 12909. Different symbols/numbers and colors correspond to different measuring electrodes. Adapted with permission from [1]. Copyright 2017 American Chemical Society.

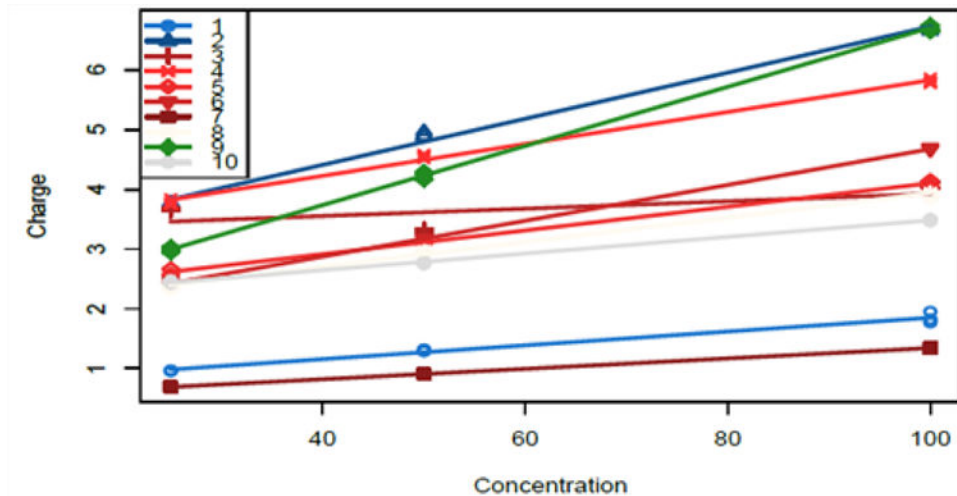


Figure 2: Concentration (in nM) and charge (in pC) measurements obtained in the calibration study for the 10 electrodes used in the time-course studies together with the fitted values based on a linear model with interaction. Different symbols and colors correspond to different electrodes. Adapted with permission from [1]. Copyright 2017 American Chemical Society.

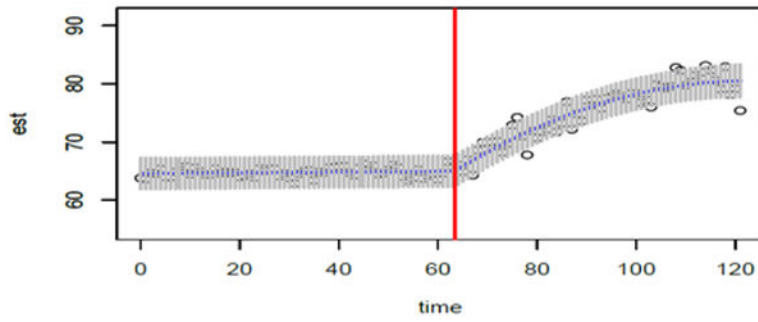


Figure 3: Plot of the concentration over time with the fitted linear-parabolic model and point-wise confidence intervals for the intervention drug pargyline.

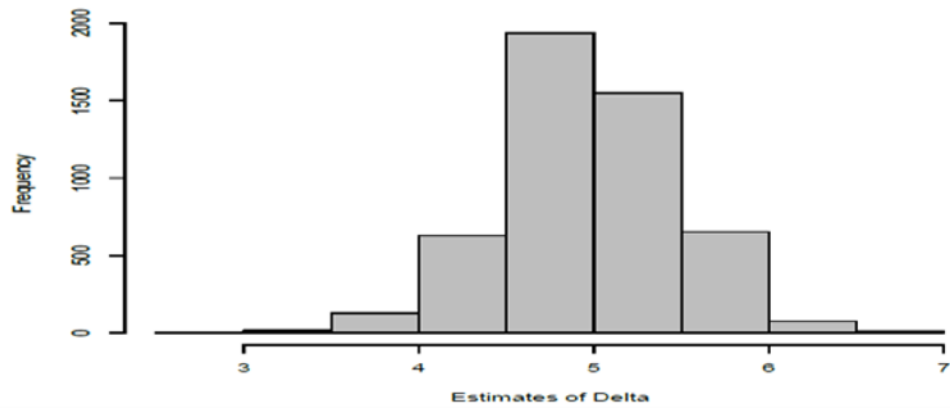


Figure 4: Frequency histogram of the estimates of from the simulation study. The true value was and there were 5000 replications.

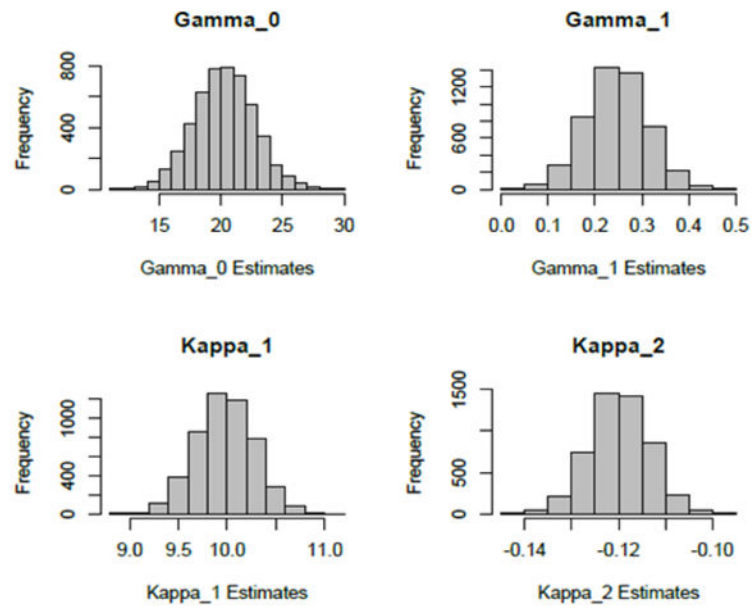


Figure 5: Frequency histograms of the 5000 estimates of the model parameters obtained from the simulation study. The true values of the parameters are $\gamma_0 = 20.0$, $\gamma_1 = 0.25$, $k_1 = 10.0$, $k_2 = -0.12$.

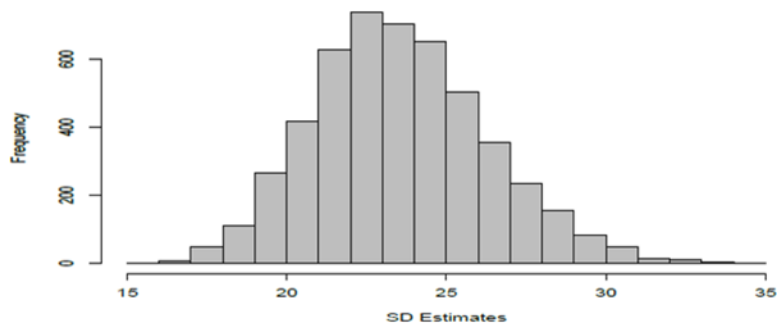


Figure 6: Histogram of the 5000 estimates of the standard deviation of the error term of the time-concentration linear-parabolic regression model.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

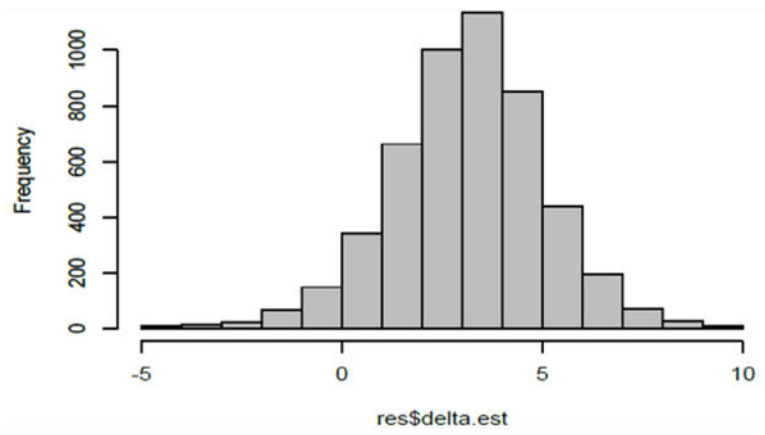


Figure 7: Frequency histogram of the estimates of from the simulation study with 5000 replications. The true value was $= 3.3$.

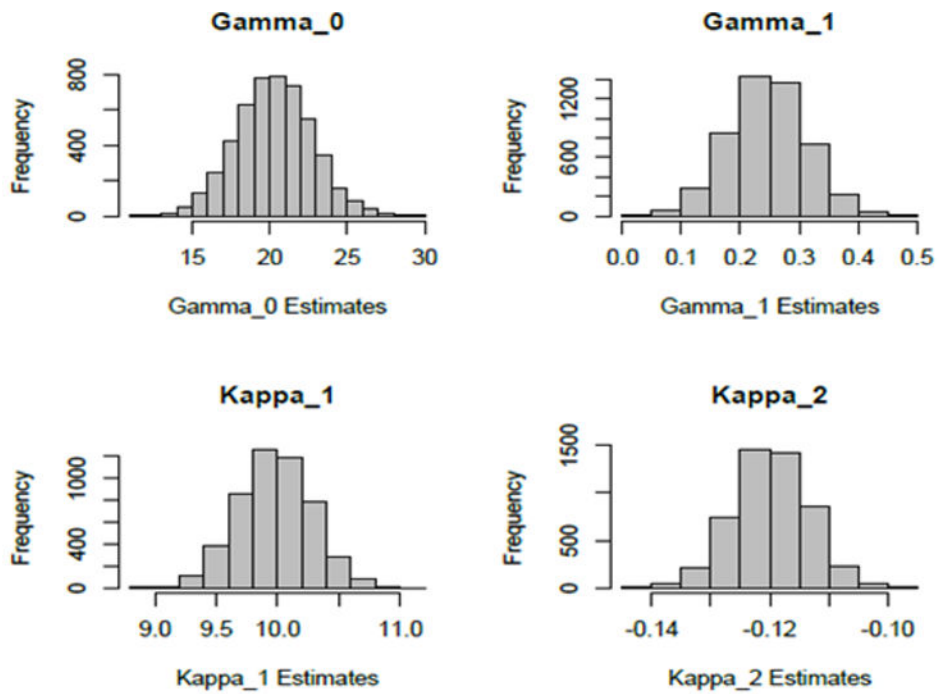


Figure 8: Frequency histograms of the 5000 estimates of the model parameters obtained from the simulation study. The true values of the parameters are $\gamma_0 = 64.5839$, $\gamma_1 = 0.006311$, $k_1 = 0.5163$, $k_2 = -0.004390$.

Table 1:

Format of the calibration data from the calibration study.

Electrode Type	Concentration Levels			
	x_{01}	x_{02}	...	x_{0L}
1	Y_{111}, \dots, Y_{11m}	Y_{211}, \dots, Y_{11m}	...	Y_{L11}, \dots, Y_{L1m}
2	Y_{121}, \dots, Y_{12m}	Y_{221}, \dots, Y_{22m}	...	Y_{L21}, \dots, Y_{L2m}
i	i	i	i	i
K	Y_{1K1}, \dots, Y_{1Km}	Y_{2K1}, \dots, Y_{2Km}	...	Y_{LK1}, \dots, Y_{LKm}

Estimates, including their estimated standard errors (SE), of the parameters of the calibration model. Corresponding p-values are provided in the last column.

Table 2:

Coefficient	Estimates	SE of Estimates	p-value
β_0	0.6949750	0.0581590	$< 2 \times 10^{-16}$
β_1	0.0115531	0.0008793	$< 2 \times 10^{-16}$
ξ_2	2.1706500	0.0822493	$< 2 \times 10^{-16}$
ξ_3	2.6106750	0.0822493	$< 2 \times 10^{-16}$
ξ_4	2.4621500	0.0822493	$< 2 \times 10^{-16}$
ξ_5	1.4289000	0.0822493	$< 2 \times 10^{-16}$
ξ_6	0.9786500	0.0822493	$< 2 \times 10^{-16}$
ξ_7	-0.2215625	0.0822493	0.00829
ξ_8	1.1935250	0.0822493	$< 2 \times 10^{-16}$
ξ_9	1.0626500	0.0822493	$< 2 \times 10^{-16}$
ξ_{10}	1.395275	0.0822493	$< 2 \times 10^{-16}$
η_2	0.027039	0.0012435	$< 2 \times 10^{-16}$
η_3	-0.0053756	0.0012435	3.64×10^{-05}
η_4	0.0151561	0.0012435	$< 2 \times 10^{-16}$
η_5	0.0082419	0.0012435	1.74×10^{-09}
η_6	0.0184476	0.0012435	$< 2 \times 10^{-16}$
η_7	-0.0028579	0.0012435	0.02363
η_8	0.0089354	0.0012435	1.22×10^{-10}
η_9	0.0379233	0.0012435	$< 2 \times 10^{-16}$
η_{10}	0.0023726	0.0012435	0.05926

EEstimates, together with estimates of the standard error (SE), of the coefficients in the time-course model. Corresponding p-values are provided in the last column.

Table 3:

Regressor	Estimate	SE of Estimate	p-Value
1	64.583863	0.315068	$< 2 \times 10^{-16}$
t	0.006311	0.007949	0.429
$\mathcal{U}(t)$	0.516323	0.033507	$< 2 \times 10^{-16}$
$\mathcal{U}(t)^2$	-0.004390	0.000540	4.86×10^{-13}

Table 4:

Means and standard deviations of the 5000 estimates obtained in the simulation study of the four model parameters γ_0 , γ_1 , k_1 , k_2 in the time-course portion of the model. The true parameter values are provided in the second column.

Parameter	True Value	Mean of Estimates	SD of Estimates
γ_0	20.00	20.28285	2.45459
γ_1	0.25	0.24523	0.06517
k_1	10.00	9.98088	0.29866
k_2	-0.12	-0.11980	0.00629

Table 5:

Means and standard deviations of the 5000 estimates obtained in the simulation study of the four model parameters γ_0 , γ_1 , k_1 , k_2 , in the time-course portion of the model. The true parameter values are provided in the second column.

Coefficient	True Value	Mean of Estimates	SD of Estimates
γ_0	64.5839	64.5718	0.6921
γ_1	0.006311	0.005909	0.01278
k_1	0.5163	0.5169	0.05266
k_2	-0.004390	-0.004436	0.001013