International Journal of Applied Mathematics

Volume 38 No. 1 2025, 23–45
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)
doi: http://dx.doi.org/10.12732/ijam.v38i1.3

DESIGNING OPTIMAL SEQUENTIAL EXPERIMENTS FOR CONVECTION-ADVECTION MODELS USING FRACTIONAL PARTIAL DIFFERENTIAL EQUATIONS

Edward L. Boone [1] **, Ryad A. Ghanam** [2]

[1] Department of Statistical Sciences ‘

and Operations Research

Virginia Commonwealth University

Richmond, VA 23284, USA

e-mail: elboone@vcu.edu

[2] Department of Liberal Arts and Sciences

Virginia Commonwealth University in Qatar

Doha, QATAR

e-mail: raghanam@vcu.edu

Received: December 15, 2024

.

Abstract

Researchers are becoming increasingly interested in using Fractional Partial Differential Equation (FPDE) models for physical systems such as modeling the flow of a gas through porous materials. These models rely on the fraction of the differentiation α, which needs to be estimated from empirical data. Experimentation is required in order to generate empirical data, specifically, distance x from the pressure source and t the time since the pressure was initially applied to the system which will generate an output pressure measurement $p(x,t)$. While sampling times are easy to choose when a sensor is in place, the location of sensors from the pressure source are typically arbitrarily chosen. This work shows how to design experiments using a sequential design with a base design and sequentially adding sampling design points by finding the optimal sensor locations along x. In this paper, we considered three methods of optimizations: A-optimality, D-optimality and E-optimality. For the A-optimality, we minimize the of the sum of the marginal variances of all the parameters. For the D-optimality we maximize the determinant of the information matrix. For the E-optimality, we o minimize the largest eigenvalue of the inverse of the information matrix. To estimate the parameters, a Bayesian framework is utilized combined with a sequential design approach to search through the possible locations for the next sensor in the follow up design. Two simple FPDE parameterizations are used to illustrate the method with an initial sensor location design of six sensors and with five additional sensors locations determined sequentially. The simple examples suggest that the parameter values influence the location of the next best sensor location.

Math. Subject Classification: 26A33, 34A08, 35R11, 65M15, 65G99

Key Words and Phrases: fractional derivative, Bayesian estimation, optimal experimental design, modeling error

1. Introduction

Fractional Partial Differential Equations (FPDEs) have garnered increasing interest among researchers modeling complex phenomena, such as flow through porous materials [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. These equations are particularly valuable because many natural processes are

not adequately described by classical derivatives, necessitating the use of fractional calculus. Examples include crowded systems, such as protein diffusion within cells [11], and diffusion in porous media, where fractional differential equations often provide more accurate descriptions [12].

Although fractional calculus has a long history, its application across various fields of science and engineering is relatively recent. For instance, Ghanam and Boone[9] proposed a new approaches to estimation and hypothesis testing for fractional fluid transport models. Their work focused on estimating key parameters, such as the fractional derivative parameter α, and accounting for measurement error. Once empirical data is collected, these parameters can be estimated using techniques outlined in [10]. However, generating such empirical data often requires experimental investigations on the porous materials being studied.

In this article we consider the mathematical model presented in [10] for which a parameter estimation for the fractional derivative order α is obtained and apply that to solve an optimization problem related to design of experiment. In this work, we consider how to place sensors for observation along the distance from the pressure source axis x using a follow-up study approach. Figure 1 shows and example of how an experimental set up may look. Pressure is applied to one end of the porous material and at specified distances $x_1, x_2, ..., x_k$ sensors are placed to measure the pressure output at those locations at various times t from the pressure being applied, denoted $p(x_i, t)$. Notice that the choice of locations of the sensors may or may not be optimal for learning about the parameters that govern the process. The question posed here is: How can a follow-up study framework allow for better estimation of the parameters, specifically the fraction parameter α.

The follow-up approach to design of experiments consist of two parts, first, a *base* design is conducted where sensors are placed and initial data is obtained and the information is used to help determine where the additional sensors are placed for the second stage of experimentation, the *follow-up* design. The advantage of follow-up studies is that researchers can place the additional sensors in locations that will allow for the most information about the parameters to be learned from the data. Experimental Design is a broad field in statistics which is covered well by [13] and [14]. For this study, a base design with six locations chosen arbitrarily, and follow-up designs with 1, 2, 3, 4 and 5 possible new locations are considered. In Figure 1, the purple sensors reflect the *base* design and the green sensors reflect the addition of three sensors once data has

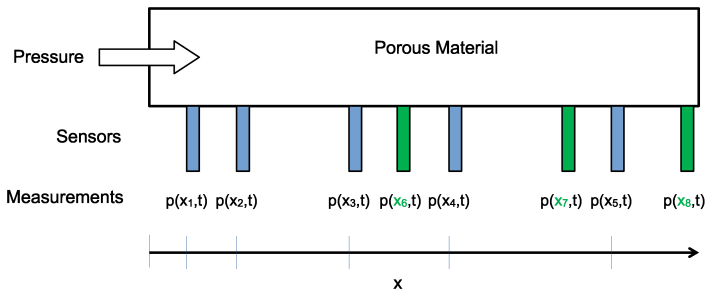

Figure 1. Example experimental device with pressure input, sensors at various locations and resulting $p(x,t)$ measurements. Blue sensors correspond to the *base* design and the green sensors are the additional sensor as suggested via the *follow-up* design.

been collected from the base sensor. These three green sensors are the *follow-up* design, in which the locations are chosen to minimize the volume of the Variance-Covariance matrix of the model parameters. In this optimization problem, we use three different techniques: A, D and E optimality and make a comparison across these methods. Each of these criteria use an information approach, specifically, Fisher's information. For more information about these methods we refer the reader to [13] and [14].

The outline of the paper is as follows: In Section 2, we give the various definitions of fractional derivatives and some properties related to their Laplace transforms. In Section 3, we state the mathematical model under consideration and define the problem under investigation. In Sections 4 and 5, we provide simulated examples the design of experiment optimality criteria. In Sections 6 and 7, we provide the new results and discussion.

2. **Preliminaries**

In this section, we give some definitions and preliminaries about fractional calculus. For more reference about fractional calculus, we recommend [15, 16, 17, 18] for the readers.

Riemann-Liouville Fractional Integral of order α for an absolutely integrable function $f(t)$ is defined by

$$\left({}_0I_t^{\alpha} f\right)(t) := \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} f(\tau) d\tau, \quad t > 0, \quad \alpha > 0, \tag{1}$$

when the right hand side exists.

Riemann-Liouville Fractional Derivative of order $\alpha > 0$ for an absolutely integrable function $f(t)$ is defined by

$$\left({}_0D_t^{\alpha} f\right)(t) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^t \frac{f(\tau)}{(t-\tau)^{\alpha}} d\tau, \qquad t > 0, \quad 0 < \alpha < 1 \tag{2}$$

Caputo Fractional Derivative of order $\alpha > 0$ for a function $f(t)$, whose first derivative is absolutely integrable, is defined by

$$\left({}_0^*D_t^{\alpha} f\right)(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{f^{'}(\tau)}{(t-\tau)^{\alpha}} d\tau, \qquad t > 0, \quad 0 < \alpha < 1 \tag{3}$$

Relationship between Riemann-Liouville and Caputo Fractional Derivative

$$\begin{aligned} \left({}_0^*D_t^{\alpha} f\right)(t) &= {}_0D_t^{\alpha} f(t) - f(0^+) \frac{t^{-\alpha}}{\Gamma(1-\alpha)} \\ &= {}_0D_t^{\alpha} \left[f(t) - f(0^+)\right]. \end{aligned} \tag{4}$$

Hilfer Fractional Derivative of order $\alpha > 0$ and type β for an absolutely integrable function $f(t)$ with respect to t is defined by

$$\left(D_t^{\alpha,\beta} f\right)(t) = \left({}_0I_t^{\beta(1-\alpha)} \frac{d}{dt} I^{(1-\beta)(1-\alpha)} f\right)(t), \tag{5}$$

where $t > 0,\ 0 < \alpha < 1,\ 0 \le \beta \le 1$.

Laplace transform of fractional derivatives

$$\mathcal{L}[{}_0^*D_t^{\alpha} f(t); s] := s\widetilde{f}(s) - s^{\alpha-1} f(0^+), \quad 0 < \alpha \le 1, \tag{6}$$

where

$$f(0^+) := \lim_{t \to 0^+} f(t).$$

$$\mathcal{L}[{}_0D_t^{\alpha} f(t); s] := s\widetilde{f}(s) - \left({}_0I_t^{1-\alpha} f\right)(0^+), \quad 0 < \alpha \le 1, \tag{7}$$

where

$$\left({}_0I_t^{1-\alpha} f\right)(0^+) := \lim_{t \to 0^+} \left({}_0I_t^{1-\alpha} f\right)(t)$$

$$\mathcal{L}\left[D_t^{\alpha,\beta} f(t); s\right] := s\widetilde{f}(s) - s^{\beta(1-\alpha)} \left[{}_0I^{(1-\beta)(1-\alpha)} f(0^+)\right], \tag{8}$$

where $0 < \alpha < 1$ and

$$\left({}_0I_t^{(1-\beta)(1-\alpha)} f\right)(0^+) := \lim_{t \to 0^+} \left({}_0I_t^{(1-\beta)(1-\alpha)} f\right)(t). \tag{9}$$

Note:
One can see that the differences in these Laplace transforms are in the initial data $f(0^+)$, $\left({}_0I_t^{1-\alpha}f\right)(0^+)$, and $\left({}_0I_t^{(1-\beta)(1-\alpha)}f\right)(0^+)$.

Lemma. *Assume that $f(t)$ is continuous on $[0,T]$, for some $T>0$, then*

$$\lim_{t\to 0^+}\left({}_0I_t^{\alpha}f\right)(t)=0,$$

for $\alpha>0$.

Proof. Let $0\le t\le T$, then

$$\begin{aligned}\left|\int_0^t (t-\tau)^{\alpha-1}f(\tau)d\tau\right| &\le \int_0^t |t-\tau|^{\alpha-1}|f(\tau)|d\tau\\ &\le M\int_0^t |t-\tau|^{\alpha-1}d\tau\\ &\le M.\frac{-(t-\tau)^{\alpha}}{\alpha}|_0^t\\ &\le \frac{M}{\alpha}t^{\alpha}\to 0 \text{ as } t\to 0.\end{aligned}$$

Remarks:

(1) Caputo derivative represents a sort of regularization in the time domain (origin) for Riemann-Liouville derivative.
(2) Hilfer fractional derivative interpolates between Riemann-Liouville fractional derivative and Caputo fractional derivative, because if $\beta=0$ then Hilfer fractional derivative corresponds to Riemann-Liouville fractional derivative and if $\beta=1$ then Hilfer fractional derivative corresponds to Caputo fractional derivative.
(3) $f(0^+)$ is required to be finite.
(4) The three derivatives are equal if f is continuous on $[0,T]$ and $f(0^+)=0$.

Mittag-Leffler Function

The Mittag-Leffler function is a generalization of the exponential function

$$e^z=\sum_{k=0}^{\infty}\frac{z^k}{k!}.$$

1-parameter Mittag-Leffler Function

$$E_\alpha(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + 1)}, \quad \alpha > 0. \tag{10}$$

2-parameter Mittag-Leffler Function

$$E_{\alpha,\beta}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + \beta)}, \quad \alpha > 0, \beta > 0. \tag{11}$$

3. **Mathematical Model**

Time-fractional advection-diffusion systems have become popular in the literature [19, 20], as they are relevant to applications of transport through porous media. One issue with using these models is the estimation of α which has been addressed by [10] using a Bayesian approach. The next step is to determine how experimental designs can improve the parameter estimation so that all inferences for the fraction of differentiation, α, and the experimental error variance σ^2 are based on the optimal amount of experimental information. The time-fractional advection-diffusion equation is given by [19]:

$$\frac{\partial^\alpha p}{\partial t^\alpha} = \frac{\partial}{\partial x}\left(D\frac{\partial p}{\partial x}\right) - U\left(\frac{\partial p}{\partial x}\right), \qquad t > 0, x > 0. \tag{12}$$

where $\frac{\partial^\alpha p}{\partial x^\alpha}$ is the Caputo fractional derivative of order α, and D and U are functions of t and x.

3.1. **Models Considered. Model 1**

For simplicity we first consider equation (12) with $D = 1$ and $U = 1$ and initial condition $p(x, 0) = e^{-cx}$, $(c > 0)$, and boundary condition is $p(x, t) \to 0$ as $x \to \infty$. In this case, let $c = 1$. This example yields the following linear system,

$$\frac{\partial^\alpha p}{\partial t^\alpha} = \frac{\partial^2 p}{\partial x^2} - \frac{\partial p}{\partial x}, \qquad t > 0, x > 0. \tag{13}$$

Under this specification and some mile conditions $p(x, t)$ has a closed form solution and is given by:

$$p(x, t) = e^{-x} \sum_{k=0}^{\infty} \frac{2t^{\alpha k}}{\Gamma(\alpha k + 1)}. \tag{14}$$

Model 2

A more complex model is considered which comes from equation (12)

where $D(p) = \frac{1}{p}$ and $U(p, p_x) = \frac{1}{p}\frac{\partial p}{\partial x}$, with initial condition $p(x, 0) = e^{-x}$, and boundary condition $p(x, t) \to 0$ as $x \to \infty$. In contrast to Model 1, this specification also yields a non-linear system:

$$\frac{\partial^\alpha p}{\partial t^\alpha} = \frac{\partial}{\partial x}\left(\frac{1}{p}\frac{\partial p}{\partial x}\right) - \frac{1}{p}\left(\frac{\partial p}{\partial x}\right)^2. \tag{15}$$

Fortunately, this case has a closed form solution for $p(x, t)$ assuming some mild conditions.

$$p(x, t) = e^{-x}\sum_{k=0}^{\infty}\frac{(-t^\alpha)^k}{\Gamma(\alpha k + 1)}, \qquad t > 0, x > 0. \tag{16}$$

To better understand behavior of each model some solutions are provided across both x, t and α for both models. Figure 2 shows the behavior of the fPDE's solution across various values of α. Panel (a) shows the influence of α on the solution of Model 1 across x when $\alpha = 0.5, 0.6, 0.7, 0.8$ at time $t = 1.0$. Notice that as α increases the pressure is dampened or compressed towards zero. Furthermore, the value of α has a strong impact on pressure when $x = 0$. In Panel (b) the pressures is shown across t for values of $\alpha = 0.78, 0.80, 0.82, 0.84$. Notice as time increases the pressure increases showing the delay in the process due to the porosity. Similarly, for Model 2, Panel (c) the means of the solutions at α across the same values of x at time $t = 1.0$ and Panel (d) the pressures is shown across t for the same values of α as in Panel (b). Notice that as x increases the pressure decreases to zero and the value of α does not seem to have a large effect on the pressures. Also as t increases the pressure decreases as well and α does not have a strong impact on the resulting pressures.

Figure 3 shows the solution across values of x and t when $\alpha = 0.82$. In Panel (a) the solution of Model 1 is plotted across x for $t = 0.1, 0.25, 0.4, 0.55$. Here as x increases there is an exponential reduction in the pressure. Panel (b) plots the solution of Model 1 across t for $x = 0.0, 0.5, 1.0, 1.5$. Notice that as t increases the pressure increases as well. Panel (c) shows the solution of Model 2 plotted across x for $t = 0.1, 0.25, 0.4, 0.55$. Again, as x increases the pressure does as well. Panel (d) shows the solution of Model 2 across t for $x = 0.0, 0.5, 1.0, 1.5$. As t increases pressure decreases and the value of x has a dramatic effect on the pressure.

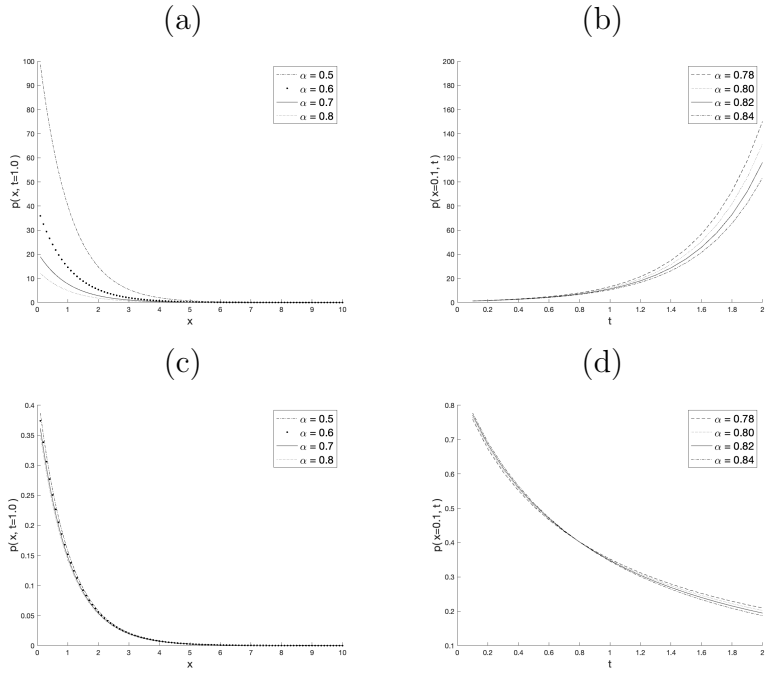

Figure 2. Plots of Model 1 and Model 2 $p(x,t)$ across various values of α. Panel (a) shows a plot of Model 1 $p(x,t)$ across $\alpha = 0.5, 0.6, 0.7, 0.8$ when $t = 1.0$. Panel (b) shows a plot of Model 1 $p(x,t)$ across $\alpha = 0.78, 0.80, 0.82, 0.84$ when $x = 0.1$. Panels (c) and (d) show the analog of Panels (a) and (b) for Model 2.

3.2. **Statistical Model.** Here a Bayesian framework is used to estimate the model parameters, which requires that both a likelihood be specified as well as prior distributions on the model parameters [21]. Likelihood is associated with the errors generated by the observation process. Let $\epsilon(x_i, t_j)$ be the error at location x_i at time t_j. Since model is interested in positive pressure, a log-normal error structure will be used. This error structure multiplicative in nature, unlike many error structures that are additive. Let $Y(x_i, t_j)$ where $i = 1, ..., n_x$ and $j = 1, ..., n_t$ be the random variable for $p(x_i, t_j)$ and $y(x_i, t_j)$ are the actual observed pressure

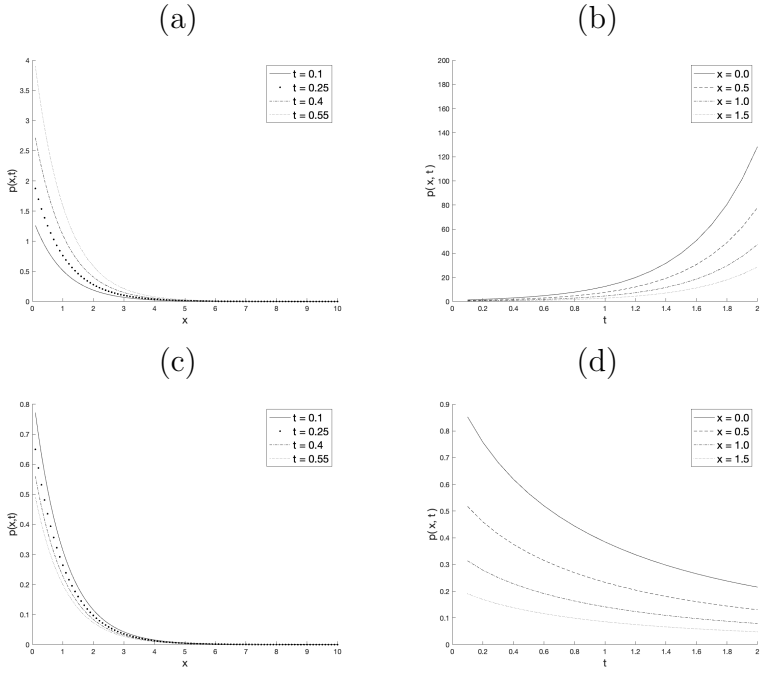

Figure 3. Plots of Model 1 $p(x,t)$ across x and t when $\alpha = 0.82$. Panel (a) shows plots of $p(x,t)$ across x when $t = 0.1, 0.25, 0.4, 0.55$. Panel (b) shows plots of $p(x,t)$ across t when $x = 0.0, 0.5, 1.0, 1.5$. Panels (c) and (d) show the analog of Panels (a) and (b) for Model 2.

measurements. For each model the structure is given by:

$$\text{Model 1: } Y(x_i,t_j) = p(x_i,t_j)\epsilon(x_i,t_j) = \left[e^{-x_i}\sum_{k=0}^{\infty}\frac{2t_j^{\alpha k}}{\Gamma(\alpha k+1)}\right]\epsilon(x_i,t_j). \tag{17}$$

$$\text{Model 2: } Y(x_i,t_j) = p(x_i,t_j)\epsilon(x_i,t_j) = \left[e^{-x_i}\sum_{k=0}^{\infty}\frac{\left(-t_j^{\alpha}\right)^k}{\Gamma(\alpha k+1)}\right]\epsilon(x_i,t_j), \tag{18}$$

where $\epsilon(x_i,t_j) \overset{iid}{\sim} LogNormal(1,\sigma^2)$. In this specification,$\alpha$ and σ^2 are the two parameters to be estimated. In order to obtain a posterior distribution of the parameters, prior distributions must be specified. Here the prior distribution for α and σ^2, $\pi(\alpha,\sigma^2)$ are specified as $\alpha \sim Beta\,(\alpha^*,\beta^*)$

to reflect the prior knowledge that α is bound between 0 and 1 (Recall that $\alpha = 0$ implies no differentiation and $\alpha = 1$ implies no fractional derivative is needed). Noting that $\sigma^2 > 0$ a reasonable prior distribution specification is $\sigma^2 \sim \chi^2(df)$. For more on prior distribution and selection, see [22].

For notation let $\mathbf{x}$ be all x_i values and $\mathbf{t}$ be all the values of t_j and $y(\mathbf{x}, \mathbf{t})$ be all the corresponding values of $y(x_i, t_j)$ observed and $p(\mathbf{x}, \mathbf{t})$ be all the solutions at the corresponding location and time values. The posterior distribution $\pi\left(\alpha, \sigma^2 | p(\mathbf{x}, \mathbf{t}), y(\mathbf{x}, \mathbf{t})\right)$ can be found using Bayes' Theorem [23]:

$$\pi\left(\alpha, \sigma^2 | p(\mathbf{x}, \mathbf{t}), y(\mathbf{x}, \mathbf{t})\right) = \frac{\pi(\alpha, \sigma^2) L(y(\mathbf{x}, \mathbf{t}) | p(\mathbf{x}, \mathbf{t}), \alpha, \sigma^2)}{\int \pi(\alpha, \sigma^2) L(y(\mathbf{x}, \mathbf{t}) | p(\mathbf{x}, \mathbf{t}), \alpha, \sigma^2) d\alpha d\sigma^2}. \tag{19}$$

In this case no analytic solution to $\pi\left(\alpha, \sigma^2 | p(\mathbf{x}, \mathbf{t}), y(\mathbf{x}, \mathbf{t})\right)$ exists and hence numerical methods need to be utilized to make inferences from the posterior distribution. Sampling from the posterior distribution is a popular approach and using the posterior samples to make inferences. There are many choices for the algorithm to sample from the posterior distribution such as Acceptance Sampling, Metropolis-Hastings Sampling, Sampling Importance Resampling, Slice Sampling, etc. For more on sampling algorithms see [24], [25], [26].

The posterior predictive distribution will be needed to generate new pressure values, Y_new associated with a additional sensor being placed at unobserved location X_new. These values will be used to augment the observed data set to determine which sensor location should be added next. The posterior predictive distribution is given by:

$$\begin{aligned} &\pi\left(Y_{new}(X_{new}, \mathbf{t}) | p(\mathbf{x}, \mathbf{t}), y(\mathbf{x}, \mathbf{t}), X_{new}, \mathbf{t}\right) \\ &= \int L\left(Y_{new}(X_{new}, \mathbf{t}) | p(X_{new}, \mathbf{t}), X_{new}, \mathbf{t}, \alpha, \sigma^2\right) \\ &\times \pi\left(\alpha, \sigma^2 | p(\mathbf{x}, \mathbf{t}), y(\mathbf{x}, \mathbf{t})\right) d\alpha d\sigma^2. \end{aligned} \tag{20}$$

Notice that the predictive distribution is a probability distribution and hence any predictions generated from it are essentially random samples from this distribution, not simply an estimate of the mean at that location.

4. **Simulated Example**

As (3.1) is unfamiliar to many readers suppose $\alpha = 0.82$ and from that system pressure data $y(\mathbf{x}, \mathbf{t})$ has been observed, with noise, at all combinations of $n_x = 31$ equally spaced levels of x from 0.01 to 10 and

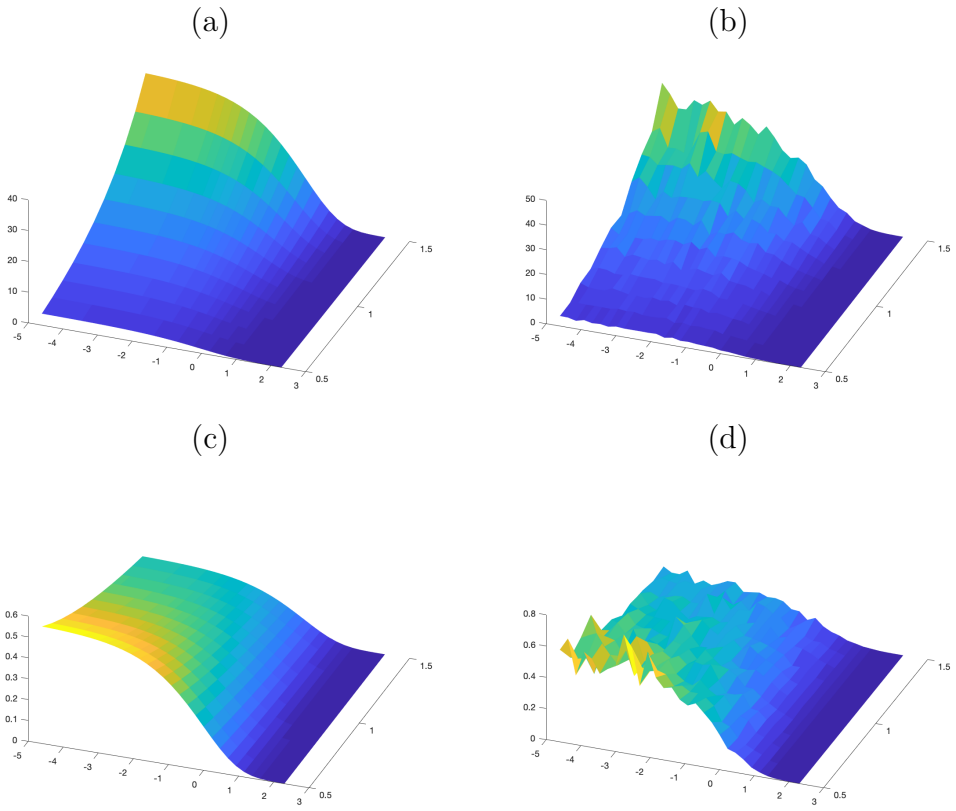

Figure 4. Simulated example of Model 1 ((3.1)) with no noise (mean only) (a), $\alpha = 0.82$ and with noise (mean with noise) $\sigma = 0.1$ (b) and Model 2 ((3.1) with no noise (mean only) (c) and with noise (mean with noise) $\sigma = 0.1$ (d)

$n_t = 11$ equally spaced times t from 0.5 to 1.5 and the noise is multiplicative following a LogNormal distribution with mean 1 and $\sigma = 0.1$. Figure 4 shows the true mean surface in panel (a) and the observed data in panel (b). Similarly, (3.1) is shown in panels (c) and (d). Notice that this value for σ induces a considerable amount of noise to the system. Furthermore, the systems produce quite different mean surfaces.

The prior distributions for α and σ were specified as follows for both Model 1 and Model 2:

$$\begin{aligned} \alpha &\sim Beta(3,3) \\ \sigma^2 &\sim \chi^2(1). \end{aligned}$$

Using the LogNormal likelihood, this prior distribution specification and Bayes Formula the posterior distribution is:

$$\begin{aligned}\pi\left(\alpha,\sigma|y(x_i,t_i),p(x_i,t_i)\right) &\propto \alpha^{\alpha^*}(1-\alpha)^{\beta^*}(\sigma^2)^{k/2-1}e^{-\sigma^2/2}\\ &\times e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n_x}\sum_{j=1}^{n_t}(\ln y(x_i,t_j)-p(x_i,t_j))^2}\\ &\times \sigma^{-n_x-n_t}\prod_{i=1}^{n_x}\prod_{j=1}^{n_t}y(x_i,t_j)^{-1} \qquad (21)\\ &= \alpha^3(1-\alpha)^3(\sigma^2)^{-1/2}e^{-\sigma^2/2}\\ &\times e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n_x}\sum_{j=1}^{n_t}(\ln y(x_i,t_j)-p(x_i,t_j))^2}\\ &\times \sigma^{-n_x-n_t}\prod_{i=1}^{n_x}\prod_{j=1}^{n_t}y(x_i,t_j)^{-1}.\end{aligned}$$

To obtain samples from the posterior distribution the Metropolis-Hastings Markov chain Monte Carlo (MCMC) algorithm was utilized. Here 1,000 posterior samples were drawn to be used for all inferences. Note that the sampler was tuned by finding candidate density variances at produce high acceptance probabilities and all samples from this tuning process were discarded. To verify convergence to a region of high posterior probability, Trace-plots of the posterior samples were examined.

5. **Design Optimality Criteria**

To evaluate each additional sensor location an *information* based approach is utilized. Many of these criteria are based on Fisher's Information as a theoretical framework for determining the information in a sample (under certain regularity conditions) is given by [27, 21]:

$$\left[\mathcal{I}(\boldsymbol{\theta})\right]_{i,j} = -E\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j}logf(X|\boldsymbol{\theta})|\boldsymbol{\theta}\right], \qquad (22)$$

where $\boldsymbol{\theta}$ is the vector of parameters and f is the probability density associated with the likelihood. In many cases, such as standard linear models, one can find an explicit form for this analytically using Fisher's Information. For example, the model $Z \sim N(W\beta, \Sigma)$, the Fisher information matrix is given by $\mathcal{I}(\beta) = W'\Sigma^{-1}W$. The variance-covariance of the least squares estimator for β in this general linear model is given by $Cov(\beta) = (W'\Sigma^{-1}W)^{-1}$ [13]. In particular, notice that $Cov(\beta) = \mathcal{I}(\beta)^{-1}$ in this case. In general, if the estimator $\hat{\theta}$ is efficient then it achieves the Cramér-Rao lower bound which is given by $Cov(\hat{\theta}) = \mathcal{I}(\theta)^{-1}$. If $\hat{\theta}$ is not efficient then $Cov(\hat{\theta}) \geq \mathcal{I}(\theta)^{-1}$. Hence $Cov(\hat{\theta})^{-1}$ provides a reasonable approximation to the Fisher information matrix $\mathcal{I}(\theta)$ [27].

Denote the set of all experimental designs (measurement locations) as Ξ and ξ_j be an individual design. As there are many different criterion to evaluate each experimental design this work focuses on three popular criteria based on Fisher Information are [13, 14, 28]:

- A-optimal: An experimental design, ξ^* is considered *A-optimal* if the $trace\left(\mathcal{I}(\alpha,\sigma|\xi^*)\right)^{-1} < trace\left(\mathcal{I}(\alpha,\sigma|\xi_j)\right)^{-1}$ for all $\xi_j \in \Xi$.
- D-optimal: An experimental design, ξ^* is considered *D-optimal* if the $|\mathcal{I}(\alpha,\sigma|\xi^*)^{-1}| < |\mathcal{I}(\alpha,\sigma|\xi_j)^{-1}|$ for all $\xi_j \in \Xi$.
- E-optimal: An experimental design, ξ^* is considered *E-optimal* if the $maxEigenvalue\left(\mathcal{I}(\alpha,\sigma|\xi^*)^{-1}\right) < maxEigenvalue\left(\mathcal{I}(\alpha,\sigma|\xi_j)^{-1}\right)$ for all $\xi_j \in \Xi$.

Let Y_{new}, ξ_j denote the new outcome vector $p(X_{new}, \mathbf{t})$ where X_{new} is defined by experimental design ξ_j. In this case an analytic for Fisher's Information is not available for

$$\pi\left(\alpha,\sigma|y(x_i,t_i),p(x_i,t_i),Y_{new,\xi_j},\xi_j,\mathbf{t}\right).$$

Here the empirical posterior variance-covariance matrix is a substituted as a reasonable approximation to Fisher's information matrix. Let $(\alpha_1,\sigma_1),(\alpha_2,\sigma_2),...,(\alpha_m,\sigma_m)$ be the m MCMC samples drawn from $\pi\left(\alpha,\sigma|y(x_i,t_i),p(x_i,t_i),Y_{new,\xi_j},\xi_j,\mathbf{t}\right)$. The variance-covariance matrix of the posterior samples from design ξ_j is estimated by:

$$\begin{aligned}&\widehat{Cov}_j\left(\alpha,\sigma|y(x_i,t_i),p(x_i,t_i),Y_{new,\xi_j},\xi_j\mathbf{t}\right)\\&=\tfrac{1}{m-1}\sum_{k=1}^{m}\left[\begin{pmatrix}\alpha_k\\\sigma_k\end{pmatrix}-\begin{pmatrix}\bar{\alpha}\\\bar{\sigma}\end{pmatrix}\right]\left[\begin{pmatrix}\alpha_k\\\sigma_k\end{pmatrix}-\begin{pmatrix}\bar{\alpha}\\\bar{\sigma}\end{pmatrix}\right],\end{aligned}\tag{23}$$

where $\bar{\alpha}=\frac{1}{m}\sum_{k=1}^{m}\alpha_k$ and $\bar{\sigma}=\frac{1}{m}\sum_{k=1}^{m}\sigma_k$. Hence we can reformulate the optimality criteria by simply substituting

$$\widehat{Cov}_j\left(\alpha,\sigma|y(x_i,t_i),p(x_i,t_i),Y_{new,\xi_j},\xi_j\mathbf{t}\right)\ \text{for}\ \left(\mathcal{I}(\alpha,\sigma|\xi_j)\right)^{-1}.$$

Algorithm 1 gives a step-by-step presentation of method proposed here. Note that in Step 7 the MCMC samples are based on both the base design and predicted observations for the follow-up design under consideration.

Algorithm 1 Algorithm to determine A, D and E optimal designs.

1: Set prior distribution parameters for α and σ.
2: Set base design sampling locations $\mathbf{x}$ and sampling times $\mathbf{t}$ and define the candidate locations Ξ where $\Xi \cap \mathbf{x} = \emptyset$.
3: Conduct the experiment using the base design and taking measurements $y(\mathbf{x}, \mathbf{t})$.
4: Obtain MCMC samples for α and σ
5: **for** $j \in 1$ to $|\Xi|$ **do**
6: **for** $\xi_j \in \Xi$ **do**
7: Sample Y_{new,ξ_j} from
8: $\pi\left(Y_{new}(X_{new}, \mathbf{t})|p(\mathbf{x}, \mathbf{t}), y(\mathbf{x}, \mathbf{t}), X_{new}, \xi_j, \mathbf{t}\right)$ using MCMC samples from step 4.
9: Obtain MCMC samples for α and σ from $\pi\left(\alpha, \sigma^2|p(\mathbf{x}, \mathbf{t}), y(\mathbf{x}, \mathbf{t}), Y_{new,\xi_j}, X_{new}, \xi_j, \mathbf{t}\right)$.
10: Calculate $C_j = \widehat{Cov}_j\left(\alpha, \sigma|y(x_i, t_i), p(x_i, t_i), Y_{new,\xi_j}, \xi_j \mathbf{t}\right)$.
11: Calculate and store $A_j = tr(C_j)$, $D_j = |C_j|$ and $E_j = maxEigen\left(C_j\right)$.
12: **end for**
13: Examine A_j, D_j and E_j to determine the optimal design ξ_j for each criteria.
14: Find $x^*_{new} = max_{x_{new}} q\left(C_j\right)$.
15: Update $\mathbf{x} = x^*_{new} \cup \mathbf{x}$.
16: Update $\Xi = \Xi \setminus \mathbf{x}$.
17: **end for**

Note that in Step 5 in Algorithm 1 there may be a large number of design points. Suppose there are Λ possible measurement locations and λ_B are chosen for the base design and λ_F will be chosen for follow-up design points, then $\lambda_F = \Lambda - \lambda_B$ possible follow-up designs exist. As this number is typically low in one dimension it could be quite large in higher dimensions. At each sequential stage the number of follow-up design points decreases by one.

6. **Results for Design of Experiments**

To illustrate the outcome associated with this experimental design technique a simulated dataset using $\alpha = 0.82$ and $\sigma = 0.1$ were created for both Model 1 and Model 2. For both models the base design points were $\mathbf{x} = \{0.2,\ 1.0,\ 2.0,\ 5.0,\ 7.0, 10.0\}$ at which the initial pressures were

collected. The sequential algorithm was then applied to the dataset for five sequentially added design points from all candidate set $\Xi_0 = \{0.2i : i = 1, 2, ...50\}$ with the base design points removed $\Xi = \Xi_0 \setminus \mathbf{x}$. The algorithm was applied for A, D and E optimality criterion.

The MCMC sampler is programmed in MATLAB 2024a. In this example there are 44 follow-up candidates at the start and since we are adding five follow-up points to the design this results in $44 + 43 + 42 + 41 + 40 = 210$ follow-up combinations to be evaluated. To explore all 210 sequential candidate design combinations, computation took approximately 190 seconds using an Intel Core i5 Quad-Core processor at 3.8 GHz and 32GB of RAM. MCMC diagnostics such as trace-plots and ACF plots were examined throughout the process, which showed MCMC sampler converged to the region of high posterior density with high acceptance probability.

The results for Model 1 are presented in Figure 5 which shows the underlying process as the blue curve with the base design are blue vertical dashed lines and the follow up design points denoted with vertical red lines. Note that the vertical red lines have step and location denote at the top. In Panel (a) shows the results for A-optimality with $x = 7.1$ being chosen first in the area of low pressure. The second location $x = 3$ is in the region of higher pressure, followed by the third location at $x = 2.2$ at an even higher pressure area. The fourth location is at $x = 4.2$ near the lower pressure region and finally the fifth location at $x = 1.4$ which is in the area of highest pressure. Panel (b) shows the results for D-optimality. Notice that the first three locations are near to each other in a region of low pressure at $x = 3.2, 2.6$ and 2.2 in that order. The remaining two locations are at $x = 8.8$ and $x = 8.0$ which are both in areas of low pressure. This agrees with general observations that D optimality tends to push points to the extremes of the domain. Panel (c) shows the results for E-optimality which has the first additional location at $x = 7.2$ in the area of low pressure followed by $x = 3.2$ and $x = 2.2$ all in the region of low probability. The fourth location at $x = 4.2$ in area where the pressure is approaching zero. The final location is at $x = 8$ in the region of low pressure.

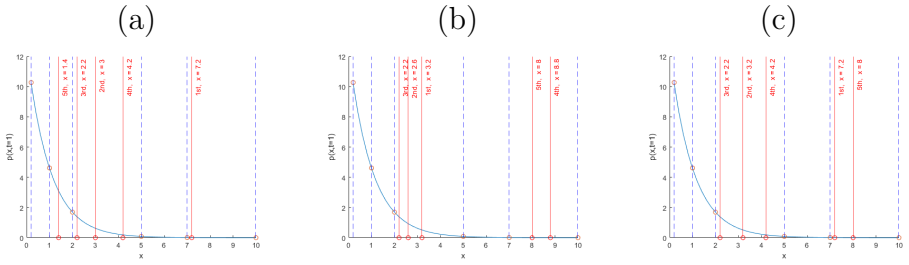

Figure 5. Follow-up design locations (red vertical lines) for **Model 1** where $\alpha = 0.82$ and $\sigma = 0.1$ with five sequentially added locations for A optimality (a), D optimality (b) and E optimality (c) criteria. Base design is 0.2, 1.0, 2.0, 5.0, 7.0, and 10.0. Here $t = 1$ for visualization purposes.

The results for Model 2 are presented in Figure 6 in the same manner as for Model 1. One interesting item that can be seen immediately is that the results for A (panel a) and E (panel c) optimal are identical with follow up locations at $x = 1.6, 3.2, 2.4, 1.8, 8.0$ in that order. This is in contrast to Model 1 where the A and E results were different. Further note that the added points for A and E tend to be where the function is above zero with one point in the region where the function is near zero. For the D optimality criteria (panel b) three points are in the region where the function is near zero and two points are in the region above zero. Of interest is the order. First $x = 3.6$ is selected which is in a region above zero, followed by $x = 7.4$ in a region where the function is near 0, then $x = 2.8$ is selected which is in a region where the function is above zero, next $x = 6.6$ is selected in region near zero and finally $x = 5.4$ which is in a region near zero. For this example, D optimality seems to like to spread the points between the regions above zero and regions near zero.

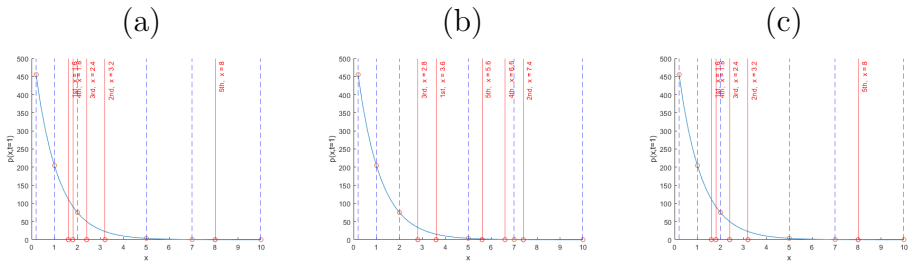

Figure 6. Follow-up design locations (red vertical lines) for **Model 2** where $\alpha = 0.82$ and $\sigma = 0.1$ with five sequentially added locations for A optimality (a), D optimality (b) and E optimality (c) criteria. Base design is 0.2, 1.0, 2.0, 5.0, 7.0, and 10.0. Here $t = 1$ for visualization purposes.

The above example considered only when $\alpha = 0.82$, however it showed some interesting patterns in that the A and E optimality conditions produced identical results for Model 2. By expanding the values of α one can see if this pattern persists or if other interesting patterns emerge. Table 6 shows the result of a simulation study where α is varied across the values $0.3, 0.4, 0.5, 0.6, 0.7$ when $\sigma = 0.01$ for each of the optimality criteria. For each model and α combination, one dataset is generated on the base design at $x = 0.2, 1.0, 2.0, 5.0, 7.0,$ and 10.0 and then all three optimality criteria are applied to the dataset to ensure there is no additional variability due to the initial dataset. One item to notice directly is that for Model 2 both A and E optimality criteria produce the exact same additional design points in the same order. Even for Model 1 when values of α are low both A and E produce quite similar results. When $\alpha = 0.7$ the results between A and D differ slightly at Steps 4 and 5. However, for D optimality there doesn't seem to be as clear of a pattern. This may be due to the fact that the information matrix is relatively orthogonal and α is the dominant term in the diagonal and hence dominant in terms of eigenvalues. This would explain this phenomenon. A large scale sensitivity study could be conducted to see when this pattern is present. However, this is beyond the scope of this study.

Table 1. Selected follow-up designs for the A, D, and E criteria on Models 1 and 2. Base design is 0.2, 1.0, 2.0, 5.0, 7.0, and 10.0. Here is varied across $\alpha = 0.3, 0.4, 0.5, 0.6, 0.7$ and $\sigma = 0.01$.

Model	Criteria	α	Step 1	Step 2	Step 3	Step 4	Step 5
Model 1	A	0.3	7.2	3.0	2.2	4.2	1.6
		0.4	3.6	3.0	2.2	4.4	1.6
		0.5	7.2	3.0	2.2	2.8	3.8
		0.6	7.8	3.0	2.2	4.2	6.4
		0.7	1.8	3.4	9.0	4.2	6.4
	D	0.3	3.2	2.6	2.2	8.8	8.0
		0.4	3.6	7.2	9.0	6.4	5.4
		0.5	3.2	9.0	7.4	4.0	4.8
		0.6	6.8	7.6	2.8	8.8	7.4
		0.7	8.0	5.4	9.2	3.8	4.6
	E	0.3	7.2	3.0	2.2	4.2	1.6
		0.4	3.6	3.0	2.2	4.4	1.6
		0.5	7.2	3.0	2.2	2.8	3.8
		0.6	7.8	3.0	2.2	4.2	6.4
		0.7	1.8	3.4	2.4	4.4	1.2
Model 2	A	0.3	7.8	2.8	6.8	5.2	3.4
		0.4	4.8	2.8	2.6	3.0	7.6
		0.5	7.4	1.8	6.8	8.6	3.4
		0.6	7.8	2.8	6.8	2.2	7.4
		0.7	8.6	2.8	6.8	6.4	3.4
	D	0.3	4.8	1.8	0.6	6.0	3.4
		0.4	6.4	2.8	8.6	8.8	0.6
		0.5	7.4	6.0	8.2	5.8	3.8
		0.6	4.8	1.2	8.0	3.2	5.4
		0.7	2.2	4.4	7.2	6.6	9.8
	E	0.3	7.8	2.8	6.8	5.2	3.4
		0.4	4.8	2.8	2.6	3.0	7.6
		0.5	7.4	1.8	6.8	8.6	3.4
		0.6	7.8	2.8	6.8	2.2	7.4
		0.7	8.6	2.8	6.8	6.4	3.4

7. **Discussion**

This study shows how to use MCMC techniques to sequentially place new experimental measurement (sensor) locations for FPDEs. First considering a specific example where the fractional derivative parameter $\alpha = 0.82$ with two different variances. For a more general exploration $\alpha = 0.3, 0.4, 0.5, 0.6$ and 0.7 is considered with four and five additional measurement locations. In the limited cases considered here for Model 2 both the A and E criteria produced identical results indicating further study should be performed.

The work here was performed on simpler models where the analytic solutions could be obtained. More complex models and models with larger parameter spaces should be included in future work where numeric solvers need to be employed to solve such complex fractional differential equations and systems. Determining the accuracy of the numeric solver and its impact on inferences should be examined to ensure that the choice of numeric solver does not unduly influence any parameter inferences or predictive distributions.

Additional real world applications of fractional calculus and differential equations where Bayesian methods and estimation could prove useful should be considered. These range from viscoelastic diffusion in complex fluids [30], anomalous diffusion [31], fractional order control problems [32], biological systems [33], and many more [34]. In addition, inverse problems should be considered as they are frequently encountered by researchers and practitioners. For example, one might observe a pressure field and want to know what where the inputs that generated that field. While this is challenging, it may allow for several fruitful lines of research. Such as what are the best sensor locations on the output pressure to determine the associated pressure inputs.

Acknowledgment

The authors would like to thank Qatar Foundation and Virginia Commonwealth University in Qatar for funding this project through the Faculty Research Grant and the Mathematical data Science Lab.

References

[1] R. L. Bagley, P. J. Torvik, A theoretical basis for the application of fractional calculus to viscoelasticity, *J. Rheol.*, **27** (1983), 201–210.

[2] S. A. Fomin, V. A. Chugunov, T. Hashida, Non-Fickian mass transport in fractured porous media, *Adv. Water Resour.*, **34** (2011), 205–214.

[3] R. Hilfer, In: *Applications of Fractional Calculus in Physics*, Edited by R. Hilfer, World Scientific, Singapore (2000), p. 87, p. 429.

[4] R. Hilfer, Experimental evidence for fractional time evolution in glass materials, *Chem. Phys.*, **284** (2002), 399–408.

[5] R. Hilfer, On fractional relaxation, *Fractals*, **11** (2003), 251–257.

[6] J. Klafter, S. C. Lim, R. Metzler, *Fractional Dynamics, Recent Advances*, World Scientific, Singapore (2011).

[7] F. Mainardi, *Fractional Calculus and Waves in Linear Viscoelasticity*, Imperial College Press, London (2010).

[8] R. R. Nigmatullin, The realization of the generalized transfer equation in a medium with fractal geometry, *Phys. Status Solidi B*, **133** (1986), 425–430.

[9] R. Ghanam, N. Malik, N. Tatar, Transparent boundary conditions for a diffusion problem modified by Hilfer derivative, *J. Math. Sci. Univ. Tokyo*, **21** (2014), 129-152.

[10] E. Boone, R. Ghanam, N. Malik, J. Whitlinger, Using the Bayesian framework for inference in fractional advection-diffusion transport system, *Int. J. Appl. Math.*, **33** (2020), 783-803; DOI: 10.12732/ijam.v33i5.4.

[11] M. Weiss, M. Elsner, F. Kartberg, T. Nilsson, Anomalous subdiffusion is a measure for cytoplasmic crowding in living cells, *Biophys. J.*, **87** (2004), 3518-3524.

[12] K. Diethelm, N. J. Ford, Analysis of fractional differential equations, *J. Math. Anal. Appl.*, **265** (2002), 229-248.

[13] D. Montgomery, *Design and Analysis of Experiments*, 10th Edition, Wiley, New York (2019).

[14] K. Hinkelmann, O. Kempthorne, *Design and Analysis of Experiments, Volume 1: Introduction to Experimental Design*, 2nd Edition, Wiley-Interscience, New York (2007).

[15] A. A. Kilbas, H. M. Srivastava, J. J. Trujillo, *Theory and Applications of Fractional Differential Equations*, Elsevier Science, North-Holland (2006).

[16] T. Sandev, R. Metzler, Z. Tomovski, Fractional diffusion equation with a generalized Riemann-Liouville time fractional derivative, *J. Phys. A: Math. Theor.*, **44** (2011), 255203.

[17] M. Caputo, Diffusion of fluids in porous media with memory, *Geothermics*, 28 (1999), 113-130.

[18] M. Caputo, Diffusion with space memory modelled with distributed order space fractional differential equations, *Ann. Geophys.*, **46** (2003), 223-234.

[19] N. A. Malik, I. Ali, B. Chanane, Numerical solutions of non-linear fractional transport models in unconventional hydrocarbon reservoirs using variational iteration method, In: *Proc. 5th Int. Conf. on Porous Media and its Applications in Science and Engineering, Kona Hawaii*, June 22-27 (2014).

[20] I. Ali, N. A. Malik, A realistic transport model with pressure-dependent parameters for gas flow in tight porous media with application to determining shale rock properties, *Transp. Porous Media*, **124** (2018), 723-742.

[21] D. Wackerly, W. Mendenhall, R. L. Scheaffer, *Mathematical Statistics with Applications*, 7th Edition, Thomson Brooks/Cole, Belmont, CA (2008).

[22] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, 2nd Edition, Springer, New York (1985).

[23] T. Bayes, R. Price, An essay towards solving a problem in the doctrine of chance, *Phil. Trans. R. Soc. Lond.*, **53** (1763), 370-418.

[24] W. R. Gilks, S. Richardson, D. J. Spiegelhaler, *Markov Chain Monte Carlo in Practice*, Chapman & Hall/CRC Press, Boca Raton (1996).

[25] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, D. B. Rubin, *Bayesian Data Analysis*, 3rd Edition, CRC Press, Boca Raton, FL (2013).

[26] J. Albert, *Bayesian Computation with R*, 2nd Edition, Springer, New York (2009).

[27] G. Casella, R. L. Berger, *Statistical Inference*, 2nd Edition, Duxbury, Belmont, CA (2002).

[28] B. Jones, K. Allen-Moyer, P. Goos, A-optimal versus D-optimal design of screening experiments, *J. Qual. Technol.*, **53** (2021), 369-382.

[29] S. Kirkpatrick, C. D. Gelatt Jr., M. P. Vecchi, Optimization by Simulated Annealing, *Science*, **220** (1983), 671–680.

[30] J. F. Reverey, J.-H. Jeon, H. Bao, M. Leippe, R. Metzler, C. Selhuber-Unkel, Superdiffusion dominates intracellular particle motion in the supercrowded space of pathogenic acanthamoeba castellanii, *Sci. Rep.*, **5** (2015), 11690.

[31] W. Deng, X. Wu, W. Wang, *Mean exit time and escape probability for the anomalous processes with the tempered power-law waiting times*, *EPL*, **117** (2017), 10009.

[32] P. D. Mandic, T. B. Scekara, M. P. Lazarevic, Dominant pole placement with fractional order PID controllers: D-decomposition approach, *ISA Trans*, 67 (2017), 76–86.

[33] CMA Pinto, ARM Carvalho, *Fractional order model for HIV dynamics*, *J. Comput. Appl. Math.*, **312** (2017), 240–256.

[34] H. G. Sun, Y. Zhang, D. Baleanu, W. Chen, Y. Q. Chen, A new collection of real world applications of fractional calculus in science and engineering, *Commun. Nonlinear Sci. Numer. Simul.*, **64** (2018), 213-231.