

Optimal Follow Up Designs for Fractional Partial Differential Equations with Application to a Convection-Advection Model.

Edward L. Boone^a, Ryad A. Ghanam^{b,*}, Albert H. Lee^a

^a*Department of Statistical Sciences and Operations Research
Virginia Commonwealth University,
Richmond, Virginia, USA.*

^b*Virginia Commonwealth University in Qatar,
Education City, Doha, Qatar.*

Abstract

As the mathematical properties of Fractional Partial Differential Equations are rapidly being developed, there is an increasing desire by researchers to employ these models in real world data oriented contexts. The main barrier to employing these models is the choice of the fractional order α . Recently, [6] show how to both estimate and make inferences about α from a Bayesian perspective. However, for experimental settings one needs to be able to design experiments that will provide optimal information about α . This work demonstrates how to use information based criteria, namely A, D and E optimality, to choose sampling locations in follow-up designs. Specifically, the simultaneous addition of one, two and three measurement locations is considered for a simple example. The simultaneous addition of four and five measurement locations is also considered across a variety of values of α . The results show that each of the criteria provide different optimal measurement locations as the number of additional measurement locations is increased. Indicating that the choice of optimality criteria should not be decided arbitrarily. ©2023 All rights reserved.

Keywords: Bayesian Estimation, Optimal Experimental Design, Modeling error, Analytical approximate solution, Mittag-Leffler function.

2010 MSC: MSC 35R11, MSC 35C05, MSC 35C10, MSC 35E15, MSC 35G25, MSC 65M15 and MSC 65G99.

1. Introduction

Fractional Partial Differential Equations (FPDE) are growing in interest among researchers confronted with modeling flow through porous materials [3, 12, 16, 17, 18, 23, 24, 28, 14, 6]. The reason behind that is there are many phenomena in nature which are not described adequately by regular derivatives and so there is a need for fractional calculus. Among them are crowded systems, such as protein diffusion within cells, [40], and diffusion through porous media. Here fractional differential equations sometimes provide a better description, see [17, 11]. Though fractional calculus has a long history, its application in many fields of science and engineering is relatively new. Recently, there has been many works related to fractional calculus and its application. For example, Sadek et al. [31, 32, 33, 34, 35, 36] considered

*Corresponding author

Email addresses: elboone@vcu.edu (Edward L. Boone), raghanam@vcu.edu (Ryad A. Ghanam), leeah2@vcu.edu (Albert H. Lee)

doi:[10.31559/glm2023.13.4.2](https://doi.org/10.31559/glm2023.13.4.2)

Received: 24 Jul 2023 Revised: 2 Sept 2023 Accepted: 12 Sept 2023

their applications to Riccati differential equations of fractional order, controllability and observability for fractional linear dynamical systems. In recent articles Ghanam and Boone [14] took a new approach for which they were interested in estimation and hypothesis testing applied to fractional fluid transport models. In particular, estimation of various parameters associated with the model such as the fractional derivative parameter α as well as measurement error. Once empirical data is collected the parameters can be estimated using techniques discussed in [6]. However, generating this empirical data may require experimentation on samples of the porous materials under study.

In this article we consider the mathematical model presented in [6] 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 an 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, \dots, 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 .

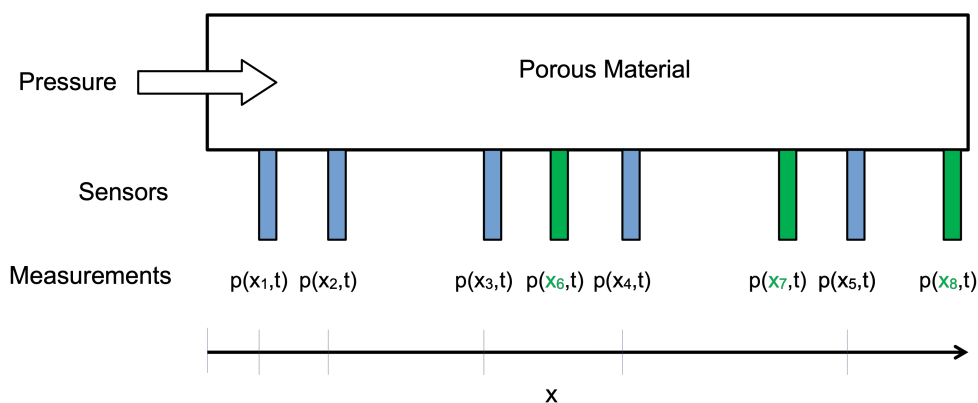


Figure 1: Example experimental device with pressure input, sensors at various locations and resulting $p(x, t)$ measurements. Blue sensors correspond to the *base* design and the red sensors are the additional sensor as suggested via the *follow-up* design.

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 [27] and [19]. 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 blue sensors reflect the *base* design and the green sensors reflect the addition of three sensors once data has 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 [27] and [19].

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 of Fractional Calculus

In this section, we state some definitions and results from fractional calculus, see [21, 37, 7, 8].

Definition 2.1. Riemann-Liouville Fractional Integral of order α for an absolutely integrable function $f(t)$ is defined by

$$({}_0I_t^\alpha f)(t) := \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} f(\tau) d\tau, \quad t > 0, \alpha > 0, \quad (2.1)$$

when the right hand side exists, and $\Gamma(\alpha)$ is defined by

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt. \quad (2.2)$$

Definition 2.2. Riemann-Liouville Fractional Derivative of order $\alpha > 0$ for an absolutely integrable function $f(t)$ is defined by

$$({}_0D_t^\alpha f)(t) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^t \frac{f(\tau)}{(t-\tau)^\alpha} d\tau, \quad t > 0, \quad 0 < \alpha < 1. \quad (2.3)$$

Definition 2.3. Caputo Fractional Derivative of order $\alpha > 0$ for a function $f(t)$, whose first derivative is absolutely integrable, is defined by

$$({}_0^*D_t^\alpha f)(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{f'(\tau)}{(t-\tau)^\alpha} d\tau, \quad t > 0, \quad 0 < \alpha < 1. \quad (2.4)$$

Proposition 2.4. The Relationship between Caputo fractional derivative and Riemann-Liouville fractional derivative is given by

$$\begin{aligned} ({}_0^*D_t^\alpha f)(t) &= {}_0D_t^\alpha f(t) - f(0^+) \frac{t^{-\alpha}}{\Gamma(1-\alpha)} \\ &= {}_0D_t^\alpha [f(t) - f(0^+)]. \end{aligned} \quad (2.5)$$

Definition 2.5. Hilfer Fractional Derivative of order $\alpha > 0$ and type β for an absolutely integrable function $f(t)$ with respect to t is defined by,

$$\left(D_t^{\alpha, \beta} f \right) (t) = \left({}_0I_t^{\beta(1-\alpha)} \frac{d}{dt} I_t^{(1-\beta)(1-\alpha)} f \right) (t), \quad t > 0, \quad 0 < \alpha < 1, 0 \leq \beta \leq 1. \quad (2.6)$$

Proposition 2.6. The Laplace transforms of fractional derivatives are given by:

$$\mathcal{L} [{}_0^*D_t^\alpha f(t); s] := s\tilde{f}(s) - s^{\alpha-1}f(0^+), \quad 0 < \alpha \leq 1, \quad (2.7)$$

where

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

$$\mathcal{L} [{}_0D_t^\alpha f(t); s] := s\tilde{f}(s) - ({}_0I_t^{1-\alpha} f)(0^+), \quad 0 < \alpha \leq 1, \quad (2.8)$$

where

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

$$\mathcal{L} \left[D_t^{\alpha, \beta} f(t); s \right] := s\tilde{f}(s) - s^{\beta(1-\alpha)} \left[{}_0I_t^{(1-\beta)(1-\alpha)} f(0^+) \right], \quad 0 < \alpha < 1, \quad (2.9)$$

where

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

It is clear from Proposition 2.6 that the differences in these Laplace transforms are in the initial data $f(0^+)$, $({}_0I_t^{1-\alpha}f)(0^+)$, and $({}_0I_t^{(1-\beta)(1-\alpha)}f)(0^+)$.

Lemma 2.7. Assume that $f(t)$ is continuous on $[0, T]$, for some $T > 0$, then

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

for $\alpha > 0$.

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

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

□

Definition 2.8. A one-parameter Mittag-Leffler Function is defined by

$$E_\alpha(z) = \sum_{k=0}^\infty \frac{z^k}{\Gamma(\alpha k + 1)}, \quad \alpha > 0. \tag{2.10}$$

Definition 2.9. A two-parameter Mittag-Leffler Function is defined by

$$E_{\alpha,\beta}(z) = \sum_{k=0}^\infty \frac{z^k}{\Gamma(\alpha k + \beta)}, \quad \alpha > 0, \beta > 0. \tag{2.11}$$

3. The Problem statement

This work focuses on a particular time-fractional advection-diffusion system given by (3.1) [25, 2], which is relevant to applications in transport through porous media. Our task is to provide experimental design so that inferences on parameters for the fraction, α , and the experimental error variance σ^2 carry the optimal amount of information.

$$\frac{\partial^\alpha p}{\partial t^\alpha} = \frac{\partial}{\partial x} \left(D \frac{\partial p}{\partial x} \right) - u \left(\frac{\partial p}{\partial x} \right), \quad t > 0, x > 0, \tag{3.1}$$

where $\frac{\partial^\alpha p}{\partial t^\alpha}$ is the Caputo fractional derivative of order α defined in Eq. (2.4)

3.1. The Model

The model considered here comes from equation (1.3) where we take $D = 1$ and $u = 1$ and the initial condition is $p(x, 0) = e^{-cx}$, ($c > 0$), and the boundary condition is $p(x, t) \rightarrow 0$ as $x \rightarrow \infty$. This yields the linear system,

$$\frac{\partial^\alpha p}{\partial t^\alpha} = \frac{\partial^2 p}{\partial x^2} - \frac{\partial p}{\partial x}, \quad t > 0, x > 0. \tag{3.2}$$

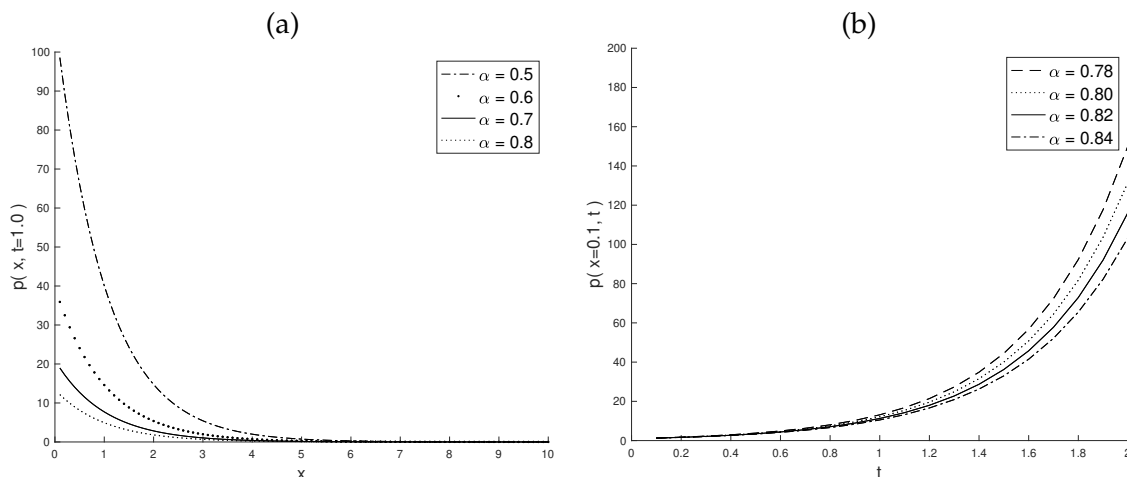


Figure 2: Plots of $p(x, t)$ across various values of α . Panel (a) shows plots of $p(x, t)$ across $\alpha = 0.5, 0.6, 0.7, 0.8$ when $t = 1.0$. Panel (b) shows plots of $p(x, t)$ across $\alpha = 0.78, 0.80, 0.82, 0.84$ when $x = 0.1$.

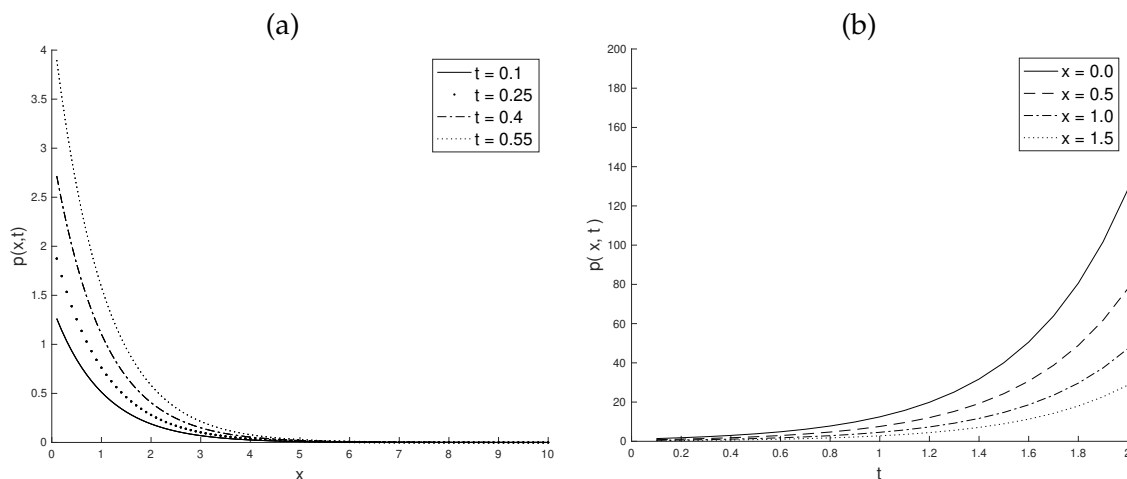


Figure 3: Plots of $p(x, t)$ across x and t when $\alpha = 0.82$. Panel (a) shows plots of $p(x, t)$ across x when $t = 0.1, 0.25, 0.4, 0.55$. Panel (b) shows plots of $p(x, t)$ across t when $x = 0.0, 0.5, 1.0, 1.5$.

For this work we will assume $c = 1$ for simplicity, and in this case $p(x, t)$ has a closed form solution given by:

$$p(x, t) = e^{-x} \sum_{k=0}^{\infty} \frac{2t^{\alpha k}}{\Gamma(\alpha k + 1)} \tag{3.3}$$

The fractional differential equation of interest is taken from [25] to model the pressure as a function of time t and location x . Figure 2 shows the behavior of the solution across various values of α . Panel (a) shows the influence of α 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. In Panel (b) the pressures is shown across t for values of $\alpha = 0.78, 0.80, 0.82, 0.84$. Again, as α increases there is a dampening effect on the pressure. Figure 3 shows the solution across values of x and t when $\alpha = 0.82$. In Panel (a) the solution 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, however, as t increases the pressure also increases. Panel (b) plots the solution across t for $x = 0.0, 0.5, 1.0, 1.5$. Again, as t increases the pressure does as well, however, as x increases the pressure is reduced.

3.2. Statistical Model

A Bayesian approach is used to estimate the parameters in the model which requires that both a likelihood be specified as well as prior distributions on the model parameters [39].

$$Y(x_i, t_j) = p(x_i, t_j)\epsilon(x_i, t_j) = \left[e^{-x_i} \sum_{k=0}^{\infty} \frac{2t_j^{\alpha k}}{\Gamma(\alpha k + 1)} \right] \epsilon(x_i, t_j), \quad (3.4)$$

where $\epsilon(x_i, t_j) \stackrel{\text{iid}}{\sim} \text{LogNormal}(1, \sigma^2)$. Here the LogNormal likelihood is chosen to ensure that pressure is always a positive value. Let $y(x_i, t_j)$ be the observed value of $Y(x_i, t_j)$ where $i = 1, \dots, n_x$ and $j = 1, \dots, n_t$. In this specification, the likelihood has two parameters, α and σ^2 . The prior distribution for α and σ^2 , $\pi(\alpha, \sigma^2)$ are specified as $\alpha \sim \text{Beta}(\alpha^*, \beta^*)$ to reflect the prior knowledge that α is bound between 0 and 1. For σ^2 the prior distribution is specified as $\sigma^2 \sim \chi^2(\text{df})$ to reflect the prior knowledge that σ^2 must be a positive value. For more on prior distribution and selection see [5].

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 $\pi(\alpha, \sigma^2 | p(\mathbf{x}, \mathbf{t}), y(\mathbf{x}, \mathbf{t}))$ can be found using Bayes' Theorem [4]:

$$\pi(\alpha, \sigma^2 | p(\mathbf{x}, \mathbf{t}), y(\mathbf{x}, \mathbf{t})) = \frac{\pi(\alpha, \sigma^2) L(y(\mathbf{x}, \mathbf{t}) | p(\mathbf{x}, \mathbf{t}), \alpha, \sigma^2)}{\int \pi(\alpha, \sigma^2) L(y(\mathbf{x}, \mathbf{t}) | p(\mathbf{x}, \mathbf{t}), \alpha, \sigma^2) d\alpha d\sigma^2}. \quad (3.5)$$

In the case considered here there is no analytic solution to $\pi(\alpha, \sigma^2 | p(\mathbf{x}, \mathbf{t}), y(\mathbf{x}, \mathbf{t}))$ and hence sampling method must be employed to draw samples from the posterior distribution, from which all inferences will be made. There are many choices for the algorithm to sample from the posterior distribution such as Acceptance Sampling, Metropolis-Hastings Sampling, Sampling Importance Resampling, etc. For more on sampling algorithms see [15], [13], [1].

The posterior predictive distribution will be needed to determine the impact of adding a new measurement location by obtaining predictions of the outcome, Y_{new} at a new proposed location X_{new} . The posterior predictive distribution is given by:

$$\pi(Y_{\text{new}}(X_{\text{new}}, \mathbf{t}) | p(\mathbf{x}, \mathbf{t}), y(\mathbf{x}, \mathbf{t}), X_{\text{new}}, \mathbf{t}, \alpha, \sigma^2) = \int L(Y_{\text{new}}(X_{\text{new}}, \mathbf{t}) | p(X_{\text{new}}, \mathbf{t}), X_{\text{new}}, \mathbf{t}, \alpha, \sigma^2) \times \pi(\alpha, \sigma^2 | p(\mathbf{x}, \mathbf{t}), y(\mathbf{x}, \mathbf{t})) d\alpha d\sigma^2. \quad (3.6)$$

Notice that the predictive distribution is a probability distribution and hence any predictions generated from it are essentially random samples from the distribution.

4. Simulated Example

Suppose the system of interest is given by (3.3) where $\alpha = 0.82$. Further suppose that data for $y(\mathbf{x}, \mathbf{t})$ has been observed, with noise, at all combinations of $n_x = 31$ equally spaced levels of x from 0.01 to 10 and $n_t = 11$ 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 unperturbed data surface in panel (a) and the perturbed data surface in panel (b). Notice that this set parameter specification produces a quite noisy surface. In real world situations we would expect to observe perturbed data similar to that in panel (b). The goal of this work is to determine if statistical techniques can be used to adequately estimate the parameters of the underlying surface when presented with noisy data.

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

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

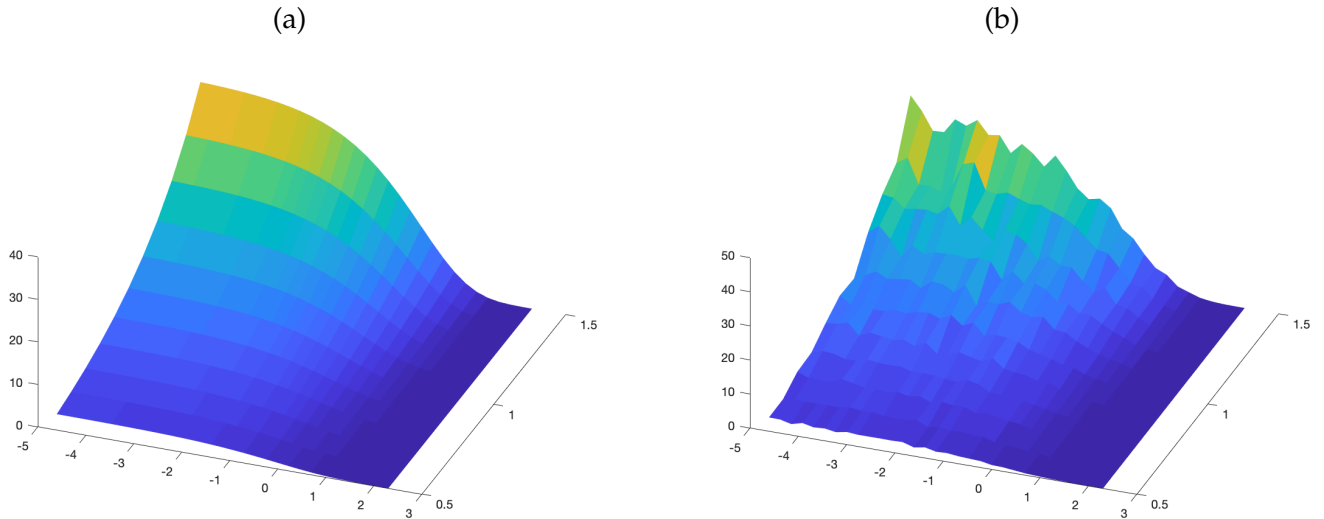


Figure 4: Simulated example of (3.3) with no noise (a), $\alpha = 0.82$ and with noise $\sigma = 0.1$ (b).

Using the LogNormal likelihood, the prior above and Bayes Formula the posterior distribution is given as:

$$\begin{aligned}
 \pi(\alpha, \sigma | y(x_i, t_i), p(x_i, t_i)) &\propto \alpha^{\alpha^*} (1 - \alpha)^{\beta^*} (\sigma^2)^{k/2-1} e^{-\sigma^2/2} \\
 &\times e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{n_x} \sum_{j=1}^{n_t} (\ln y(x_i, t_j) - p(x_i, t_j))^2} \\
 &\times \sigma^{-n_x - n_t} \prod_{i=1}^{n_x} \prod_{j=1}^{n_t} y(x_i, t_j)^{-1} \tag{4.1} \\
 &= \alpha^3 (1 - \alpha)^3 (\sigma^2)^{-1/2} e^{-\sigma^2/2} \\
 &\times e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{n_x} \sum_{j=1}^{n_t} (\ln y(x_i, t_j) - p(x_i, t_j))^2} \\
 &\times \sigma^{-n_x - n_t} \prod_{i=1}^{n_x} \prod_{j=1}^{n_t} y(x_i, t_j)^{-1}.
 \end{aligned}$$

The Metropolis-Hastings Markov chain Monte Carlo (MCMC) algorithm was employed to draw 1,000 posterior samples to be used for all inferences. Note that the sampler was tuned in order to find appropriate step-sizes and any samples from this tuning were discarded. Traceplots were examined to ensure convergence of the sampler to the region of high posterior density.

5. Design Optimality Criteria

As the goal of the study is determine new measurement locations along x which will provide better estimates for α and σ , one must define “better”. One popular approach is to work with the *information* generated by the addition of the new measurements. Fisher’s Information is often used to provide a theoretical framework for determining the information from a sample which is given by (under certain regularity conditions) [9], [39]:

$$[J(\theta)]_{i,j} = -E \left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log f(X|\theta) | \theta \right], \tag{5.1}$$

where θ is the parameter vector 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 $J(\beta) = W'\Sigma^{-1}W$. The variance-covariance of the least squares estimator for β in this general linear model is given

by $\text{Cov}(\beta) = (W'\Sigma^{-1}W)^{-1}$ [27]. In particular, notice that $\text{Cov}(\beta) = \mathcal{J}(\beta)^{-1}$ in this case. In general, if the estimator $\hat{\theta}$ is efficient then it achieves the Cramér-Rao lower bound which is given by $\text{Cov}(\hat{\theta}) = \mathcal{J}(\theta)^{-1}$. If $\hat{\theta}$ is not efficient then $\text{Cov}(\hat{\theta}) \geq \mathcal{J}(\theta)^{-1}$. Hence $\text{Cov}(\hat{\theta})^{-1}$ provides a reasonable approximation to the Fisher information matrix $\mathcal{J}(\theta)$ [9].

In our case, smaller variance of the joint posterior distribution of α and σ is preferable. Hence we wish to minimize the variance-covariance matrix of α and σ . This issue is how does one determine if one matrix is better than another? Typically one would look at the quadratic form associated with the matrix. If C_1 and C_2 are matrices where $x'C_1x < x'C_2x$ for all $x \in \mathbb{R}^n$, then we would say that C_2 dominates C_1 and hence C_1 would be considered smaller. However, this approach is difficult to implement in practice. Instead other criterion have been developed such as considering the trace of the matrices. If $\text{trace}(C_1) < \text{trace}(C_2)$ then C_1 is smaller than C_2 . In the variance-covariance context the trace is simply the sum of the marginal variances of each of the random variables. Another popular technique is to compare the determinants of the matrices. If $|C_1| < |C_2|$ then C_1 is smaller than C_2 . Similarly, one could consider the largest eigenvalues of each of the matrices and rank on those.

Let Ξ be the set of all experimental designs (measurement locations) under consideration with ξ_j be an individual design. Three popular criteria to evaluate experimental designs based on Fisher Information are A-optimality, D-optimality and E-optimality. [27, 19, 20]:

- A-optimal: An experimental design, ξ^* is considered *A-optimal* if:

$$\text{trace}(\mathcal{J}(\alpha, \sigma|\xi^*))^{-1} < \text{trace}(\mathcal{J}(\alpha, \sigma|\xi_j))^{-1}$$

for all $\xi_j \in \Xi$. D – optimal : An experimental design, ξ^* is considered *D-optimal* if:

$$|\mathcal{J}(\alpha, \sigma|\xi^*)^{-1}| < |\mathcal{J}(\alpha, \sigma|\xi_j)^{-1}|$$

for all $\xi_j \in \Xi$.

- E-optimal: An experimental design, ξ^* is considered *E-optimal* if:

$$\max \text{Eigenvalue}(\mathcal{J}(\alpha, \sigma|\xi^*)^{-1}) < \max \text{Eigenvalue}(\mathcal{J}(\alpha, \sigma|\xi_j)^{-1})$$

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

$$\widehat{\text{Cov}}_j(\alpha, \sigma|y(x_i, t_i), p(x_i, t_i), Y_{new, \xi_j}, \xi_j, \mathbf{t}) = \frac{1}{m-1} \sum_{k=1}^m \left[\begin{pmatrix} \alpha_k \\ \sigma_k \end{pmatrix} - \begin{pmatrix} \bar{\alpha} \\ \bar{\sigma} \end{pmatrix} \right] \left[\begin{pmatrix} \alpha_k \\ \sigma_k \end{pmatrix} - \begin{pmatrix} \bar{\alpha} \\ \bar{\sigma} \end{pmatrix} \right]', \quad (5.2)$$

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

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 observations and the 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} .
 - 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** $\xi_j \in \Xi$ **do**
 - 6: Sample Y_{new, ξ_j} from $\pi(Y_{new} | X_{new}, \mathbf{t}) | p(\mathbf{x}, \mathbf{t}), y(\mathbf{x}, \mathbf{t}), X_{new}, \xi_j, \mathbf{t})$ using MCMC samples from step 4.
 - 7: Obtain MCMC samples for α and σ from $\pi(\alpha, \sigma^2 | p(\mathbf{x}, \mathbf{t}), y(\mathbf{x}, \mathbf{t}), Y_{new, \xi_j}, X_{new}, \xi_j, \mathbf{t})$.
 - 8: Calculate $C_j = \widehat{\text{Cov}}_j(\alpha, \sigma | y(x_i, t_i), p(x_i, t_i), Y_{new, \xi_j}, \xi_j, \mathbf{t})$.
 - 9: Calculate and store $A_j = \text{tr}(C_j)$, $D_j = |C_j|$ and $E_j = \max \text{Eigen}(C_j)$.
 - 10: **end for**
 - 11: Examine A_j , D_j and E_j to determine the optimal design ξ_j for each criteria.
-

Note that in Step 5 in Algorithm 1 there may be a huge 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 $\binom{\Lambda - \lambda_B}{\lambda_F}$ possible follow-up designs exist. For example, if there are 48 possible measurement locations and six are chosen in the base design and five additional measurement locations will be chosen, then there are $\binom{48 - 6}{5} = 850,558$ possible follow-up designs. This is a large number of possible designs to calculate, hence a search algorithm such as *Simulated Annealing* can be employed. While this does not enumerate all the designs in Ξ it searches to find the one that meets the optimality criteria. This will necessitate that each criterion evaluated separately instead of simultaneously in step 5(d). For more on simulated annealing see [22].

Both the MCMC sampler and the Simulated Annealing algorithms are programmed in MATLAB 2022b. To explore 1,000 candidate design combinations from the design space, computation took approximately 1,200 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.

6. New Results for Design of Experiments

To consider the effect of σ and the number of additional measurement locations on the follow-up design location(s) for the A, D and E criteria when $\alpha = 0.82$ is performed using a base design of 0.2, 1.0, 2.0, 5.0, 7.0, and 10.0. On all studies 1,000 MCMC samples are used for parameter estimation. For one additional location all 42 candidate locations are considered, for two additional locations all 946 candidate combinations are considered and for three additional locations simulated annealing is used with 1,000 steps. Table 6 shows the location results for the follow-up designs for A, D and E criteria when $\alpha = 0.82$ and $\sigma \in \{0.01, 0.1\}$. Notice that for one additional sensor and $\sigma = 0.1$ all three criteria produce a follow up measurement location at 4.2. Similarly for $\sigma = 0.01$ all three criteria produce the same follow up measurement location at 6.2. When two additional measurement locations are allowed and $\sigma = 0.1$ both the D and E criteria produce follow-up locations at 0.4 and 4.2. However, the A criterion produces a different pair of locations at 1.2 and 2.6. This shows that the three criteria do not necessarily produce the same values which would suggest that the choice of criteria can be arbitrary. For the three additional location scenarios the criteria do not produce the same set of locations. However, the D and E criteria do produce one of the locations to be 2.6 in the $\sigma = 0.01$ case. Of further note is that for the A optimality criterion the choice of locations are the same across $\sigma = 0.01$ and $\sigma = 0.1$.

Table 1: Selected follow-up designs for the A, D, and E criteria. Base design is 0.2, 1.0, 2.0, 5.0, 7.0, and 10.0. For this simulation $\alpha = 0.82$, $\sigma = 0.1$ and $\sigma = 0.01$.

Variance	Additional	Criteria	Design
$\sigma = 0.1$	1	A	4.2
		D	4.2
		E	4.2
	2	A	1.2, 2.6
		D	0.4, 4.2
		E	0.4, 4.2
	3	A	6.0, 8.2, 9.6
		D	2.6, 4.4, 7.6
		E	2.2, 6.2, 8.0
$\sigma = 0.01$	1	A	6.2
		D	6.2
		E	6.2
	2	A	0.4, 6.4
		D	6.0, 8.4
		E	4.4, 9.8
	3	A	6.0, 8.2, 9.6
		D	0.4, 0.8, 2.6
		E	1.6, 2.6, 5.6

To examine the effect of α on the optimal design location for four and five simultaneously added locations a simulation study was performed by varying α at values 0.3, 0.4, 0.5, 0.6 and 0.7 and keeping $\sigma = 0.1$. Using the approach for the one, two and three additional sensor study above, a simulated dataset was generated for each value of α with base design locations 0.2, 1.0, 2.0, 5.0, 7.0 and 10.0. Using each of these datasets the A, D and E optimal locations were found by employing the MCMC combined with simulated annealing approach with 1,000 search steps for the MCMC estimates and 1,000 simulated annealing steps.

Figure 5 shows the results of the study with the A optimal locations denoted by a blue circle (\circ), D optimal denoted by a black triangle (Δ) and E-optimal denoted by a red plus sign (+). Notice that the A optimal points tend to be placed at higher and higher values as alpha is increased. This is less prominent in the E optimal locations which do seem to increase as well. The D optimal solutions exhibit a similar pattern of increasing the follow up design points, however, to a much smaller magnitude than both A or E optimality.

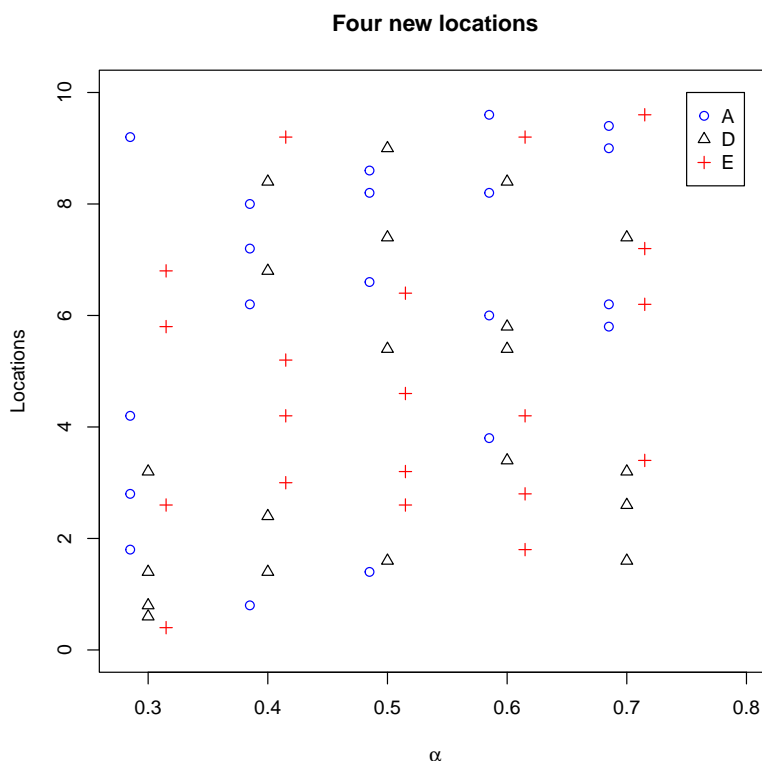


Figure 5: Follow-up design locations for $\alpha = 0.3, 0.4, 0.5, 0.6, 0.7$ with four simultaneously added locations for A(\circ), D(Δ) and E($+$) optimality criteria. Base design is 0.2, 1.0, 2.0, 5.0, 7.0, and 10.0 with $\sigma = 0.1$.

Figure 6 shows the results of the study with the A optimal locations denoted by a blue circle (\circ), D optimal denoted by a black triangle (Δ) and E-optimal denoted by a red plus sign ($+$). Notice that the A optimal design points are consistently placed above 1.8 and spread across the remaining design points at random with high values near 9.0 and 10.0 common. This is in contrast to the E optimal locations which are much higher for $\alpha = 0.3$ but spread out more evenly across the design space for the other values of α . Of particular interest is the D optimal solutions that are above 3.5 for $\alpha = 0.5$ and 0.6 and rather randomly placed across the other values of α . The differences here between the four and five location follow up designs may be due to the base design choice which contains considerable information so additional points can be spread across a wider range of values.

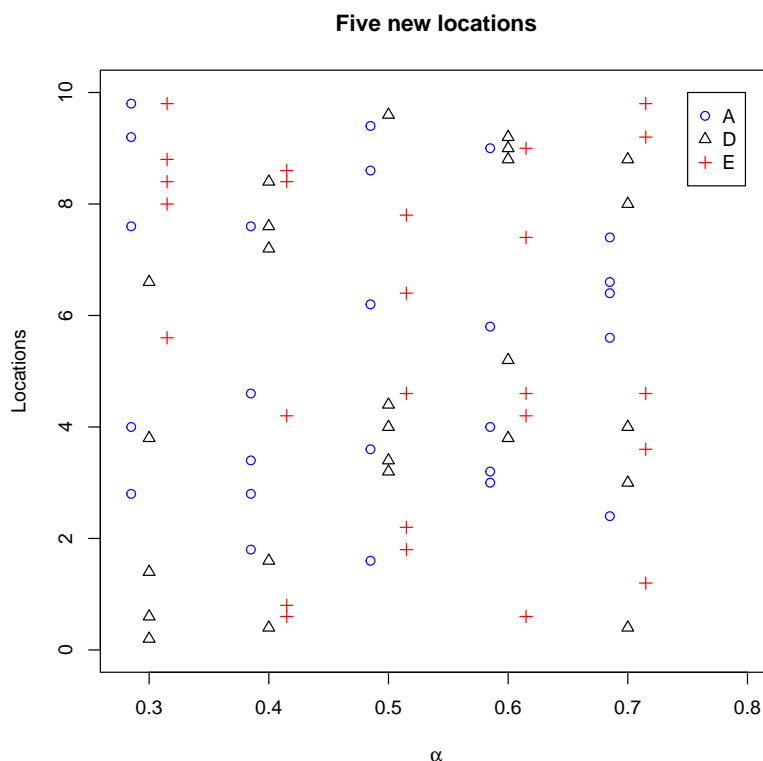


Figure 6: Follow-up design locations for $\alpha = 0.3, 0.4, 0.5, 0.6, 0.7$ with five simultaneously added locations for A(\circ), D(Δ) and E($+$) optimality criteria. Base design is 0.2, 1.0, 2.0, 5.0, 7.0, and 10.0 with $\sigma = 0.1$.

7. Discussion

This study shows how to use MCMC and Simulated Annealing techniques to simultaneously place new experimental measurement (sensor) locations for FPDEs. By first considering a specific example where $\alpha = 0.82$ with two different variances and one, two and three additional measurement locations. For the one additional measurement location, all three optimality criteria give the same new location, only depending on the model variance. As additional measurement locations are added, the three optimality criteria provide different optimal locations. 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. As the number of measurement locations increases each of the optimality criteria provide different locations.

Of particular interest is when the number of additional measurement locations is increased the more spread out the locations become across the design space. This could be an artifact of the base design chosen as it may not contain enough information to guide placement. An additional study should be conducted to determine if replication of base design measurements provides a stronger pattern in the follow-up measurement location across the values of α .

Another area that should be explored is sequentially adding follow-up measurement locations instead of the simultaneous approach taken here. In a sequential approach, a new measurement location will be chosen based on each of the criteria and that location will be kept and additional locations will be chosen one by one. For example when $\alpha = 0.82$ and $\sigma = 0.1$, Table 6 shows that all three methods chose 4.2 as the next optimal location if only one sensor is added (6.2 was chosen when $\sigma = 0.01$). How do the additional measurement locations change across various values for α for each of the different criteria? While the approach seems to be straightforward there are several scenarios that need to be addressed. If the sensors can be reused then one could gain additional measurements at already established sensor locations and hence there is a built in replication occurring, albeit unequal replication. If the sensor cannot be reused

then at each step no additional information from other sensors can be obtained. While this may seem trivial, in many applications sensors may be single use and should be studied.

Note that the criteria used here focus on *estimating* the parameter α . However, in many applications estimation may not be the goal. Instead, a researcher may be interested in whether or not using a FPDE is warranted. In the case when $\alpha = 1$, a standard PDE could be used. If the goal is to conduct a hypothesis test $H_0 : \alpha = 1$ versus $H_A : \alpha \in (0, 1)$, then the criteria should be changed accordingly. A Bayes factor approach to testing this has been presented in [6]. A similar study for the simultaneous addition of sensors could be performed.

More complex models and models with larger parameter spaces should be included in future work where numeric solvers more than likely need to be employed to solve such complex fractional differential equations and systems. Many real world applications of fractional calculus and differential equations exist where Bayesian methods and estimation is important. These range from viscoelastic diffusion in complex fluids [30], anomalous diffusion [10], fractional order control problems [26], biological systems [29], and a lot more [38]. 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.

Acknowledgment

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

References

- [1] J. Albert, Bayesian Computation with R, Second Edition, Springer, New York (2009). 3.2
- [2] I. Ali I, N.M Malik, A Realistic Transport Model with Pressure-Dependent Parameters for Gas Flow in Tight Porous Media with Application to Determining Shale Rock Properties, Transport In Porous Media Vol. 124 No. 3, pp 723-742, (2018). 3
- [3] R. L. Bagley, P. J. Torvik, A theoretical basis for the application of fractional calculus to viscoelasticity, Journal of Rheology 27 No. 3 (1983), 201–210. 1
- [4] T. Bayes, R. Price, An Essay towards solving a Problem in the Doctrine of Chance. By the late Rev. Mr. Bayes, communicated by Mr. Price in a letter to John Canton, A.M.F.R.S. Philosophical Transactions of the Royal Society of London. 53: 370–418, (1763). 3.2
- [5] J.O. Berger, Statistical Decision Theory and Bayesian Analysis, Second Edition, Springer, New York (1985). 3.2
- [6] E. Boone, R Ghanam, N. Malik & J Whitlinger, Using the Bayesian Framework for Inference in Fractional Advection-Diffusion Transport System. International Journal of Applied Mathematics. Volume 33 No. 5 (2020), 783-803. (document), 1, 7
- [7] Caputo M. *Diffusion of fluids in porous media with memory*, Geothermics, 28:113-130 (1999). 2
- [8] M. Caputo, Diffusion with space memory modelled with distributed order space fractional differential equations, Annals of Geophysics, 2:46, 223-234, (2003). 2
- [9] G. Casella G, R.L. Berger, Statistical Inference, 2nd Edition. Duxbury, Belmont, CA. (2002). 5, 5
- [10] W. Deng, X. Wu, W. Wang. Mean exit time and escape probability for the anomalous processes with the tempered power-law waiting times. EPL 117:10009 (2017). 7
- [11] K. Diethelm , N.J. Ford, Analysis of fractional differential equations, J. Math. Anal. Appl. 265:229-248 (2002). 1
- [12] S.A. Fomin, V.A. Chugunov, T. Hashida, Non-Fickian mass transport in fractured porous media, Adv. Water Resources 34 No. 2 (2011), 205–214. 1
- [13] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson , A. Vehtari, D.B. Rubin, Bayesian Data Analysis, Third Edition, CRC Press, Boca Raton, FL (2013). 3.2
- [14] 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. 1
- [15] W.R. Gilks, S. Richardson, D.J. Spiegelhalter, Markov chain Monte Carlo in Practice, Chapman & Hall, CRC Press, Boca Raton (1996). 3.2
- [16] R. Hilfer, Applications of Fractional Calculus in Physics, edited by R. Hilfer, World Scientific Publishing Company, Singapore, 2000, p. 87 and p. 429. 1
- [17] R. Hilfer, Experimental evidence for fractional time evolution in glass materials, Chem. Physics 284 (2002), 399–408. 1
- [18] R. Hilfer, On fractional relaxation, Fractals 11 (2003), 251–257. 1
- [19] K. Hinkelmann, O. Kempthorne, Design and Analysis of Experiments, Volume 1: Introduction to Experimental Design, 2nd Edition. New York, NY: Wiley-Interscience, (2007). 1, 5

- [20] B. Jones, K. Allen-Moyer, P. Goos, A-optimal versus D-optimal design of screening experiments, *Journal of Quality Technology*, vol 53:4, pp. 369–382, (2021). DOI: 10.1080/00224065.2020.1757391 5
- [21] A.A. Kilbas, H.M. Srivastava, J.J. Trujillo, *Theory and Applications of Fractional differential equations*, North-Holland, Elsevier Science Publisher, (2006). 2
- [22] S. Kirkpatrick, Jr. Gelatt , M.P. Vecchi, Optimization by Simulated Annealing. *Science*. 220 (4598): 671–680, 1983. 5
- [23] J. Klafter, S.C. Lim, R. Metzler, *Fractional Dynamics*, Recent Advances, World Scientific, Singapore, (2011). 1
- [24] F. Mainardi, *Fractional Calculus and Waves in Linear Viscoelasticity*, Imperial College Press, London (2010). 1
- [25] N.A. Malik, I. Ali, B. Chanane , Numerical Solutions of Non-linear Fractional Transport Models in Unconventional Hydrocarbon Reservoirs using Variational Iteration Method, *Proceedings of the 5th International Conference on Porous Media and its Applications in Science and Engineering*, Kona Hawaii, (2014). 3, 3.1
- [26] P.D. Mandic, T.B. Scegara, M.P. Lazarevic, Dominant pole placement with fractional order PID controllers: D-decomposition approach. *ISA Trans* 67:76–86, (2017). 7
- [27] D. Montgomery, *Design and Analysis of Experiments*, 10th Edition. New York, NY: Wiley, (2019). 1, 5
- [28] R. Nigmatullin, The realization of the generalized transfer equation in a medium with fractal geometry, *Physica Status Solidi (b)* 133 No. 1 (1986), 425–430. 1
- [29] C. Pinto, A. Carvalho. Fractional order model for HIV dynamics. *J Comput Appl Math* 312:240–56 (2017). 7
- [30] J.F. Reverey, J. Jeon, H. Bao , M. Leippe , R. Metzler, C. Selhuber-Unkel , Superdiffusion dominates intracellular particle motion in the super crowded space of pathogenic *acanthamoeba castellanii*. *Sci Rep* 5 (11690), (2015). 7
- [31] L. Sadek, A.S. Bataineh, T. Alaoui, I. Hashim, I. The Novel Mittag-Leffler–Galerkin Method: Application to a Riccati Differential Equation of Fractional Order. *Fractal Fract.* (2023), 7, 302. <https://doi.org/10.3390/fractalfract7040302>. 1
- [32] L. Sadek, O. Sadek, Alaoui, H. T., Abdo, M. S., Shah, K. et al. Fractional Order Modeling of Predicting COVID-19 with Isolation and Vaccination Strategies in Morocco. *CMES-Computer Modeling in Engineering Sciences*, 136(2), 1931–1950 (2023). 1
- [33] Sadek L. Controllability and observability for fractal linear dynamical systems. *Journal of Vibration and Control*. 2022;0(0). doi:10.1177/10775463221123354 1
- [34] L. Sadek , A Cotangent Fractional Derivative with the Application. *Fractal and Fractional*. (2023); 7(6):444. <https://doi.org/10.3390/fractalfract7060444> 1
- [35] L. Sadek, Stability of conformable linear infinite-dimensional systems. *Int. J. Dynam. Control* 11, 1276–1284 (2023). <https://doi.org/10.1007/s40435-022-01061-w> 1
- [36] L. Sadek, B. Abouzaid, E.M. Sadek. et al. Controllability, observability and fractional linear-quadratic problem for fractional linear systems with conformable fractional derivatives and some applications. *Int. J. Dynam. Control* 11, 214–228 (2023). <https://doi.org/10.1007/s40435-022-00977-7> 1
- [37] T Sandev, R Metzler, Z. Tomovski, Fractional diffusion equation with a generalized Riemann-Liouville time fractional derivative, *J. Phys. A: Math. Theor.* 44 (255203) (21pp), (2011). 2
- [38] H. Sun , Y. Zhang , D. Baleanu , W. Chen, Y. Chen Y. A new collection of real world applications of fractional calculus in science and engineering, *Comm. Nonlin. Science and Numerical Simulation* 64:213-231, (2018). 7
- [39] D. Wackerly, W. Mendenhall, R.L. Scheaffer, *Mathematical Statistics with Applications*, 7th Edition. Thomson Brooks/Cole, Belmont CA. (2008). 3.2, 5
- [40] M. Weiss, M Elsner, F. Kartberg, T. Nilsson, Anomalous Subdiffusion Is a Measure for Cytoplasmic Crowding in Living Cells, *Biophysical J.*, 87:3518-3524 (2004). 1