International Journal of Applied Mathematics

Volume 36 No. 1 2023, 127-144
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)
doi: http://dx.doi.org/10.12732/ijam.v36i1.11

STATISTICAL INFERENCE ON FRACTIONAL ADVECTION-DIFFUSION TRANSPORT SYSTEMS

Edward Boone[1], Ryad Ghanam[2] §

[1]Department of Statistical Sciences
and Operations Research
Virginia Commonwealth University
Richmond, VA 23284, USA

[2]Department of Liberal Arts and Sciences
Virginia Commonwealth University in Qatar
Doha, QATAR

Abstract: The field fractional partial differential equations has been greatly expanding with in the last 20 years as evidenced by the growing body of literature. The work has primarily been theoretical in nature as there has been limited availability of experimental data. This work demonstrates using simulated data, how one can utilize these models and apply them to real world data to make inferences about the parameter values, predictions at future values and even what forcing mechanism was used to generate the data. This is done using a Bayesian approach that employs Markov chain Monte Carlo techniques. The proposed approach is evaluated using simulation studies concerning credible interval coverage probabilities and highest posterior model probabilities. The simulation study shows that the proposed method is effective for many parameters and that more work is needed to better estimate some physical coefficients.

AMS Subject Classification: 35R11, 35C05, 35C10, 35E15, 35G25, 65M15, 65G99
Key Words: Bayesian estimation; validation; modeling error; analytical approximate solution

Received: December 16, 2022

§Correspondence author

1. Introduction

Many physical systems are modeled by partial differential equations [1, 2]; applications of such equations arise in many fields, such as, atmospheric chemistry [3], modeling of fluid flow in homogeneous media [4], combustion [5], and porous media [6]. It has been observed and proved that the use of fractional calculus leads to a good fit of the experimental data and therefore describes better the behavior of many materials (see [7, 8, 9, 10, 11, 12, 13, 14]). Hilfer [9, 10] showed that time fractional derivatives are equivalent to infinitesimal generators of generalized time fractional evolutions arising in the transition from microscopic to macroscopic time scales. In [9], Hilfer showed that this transition from ordinary time derivative to fractional time derivative indeed arises in physical problems. The Hilfer's idea on time fractional evolution is presented in detail in Chapter 9 of the book [12]. Bagley and Torvik [7] used the fractional calculus to generalize the Kelvin-Voigt theory and showed that it has several attractive features.

In a paper by Ghanam et al. [15], they have considered transparent boundary conditions for a diffusion problem modified by Hilfer derivative. That that paper, they considered a homogeneous fractional diffusion problem in an infinite reservoir sometimes called a "modified" diffusion equation. The equation involves a memory term in the form of a time fractional derivative. For the sake of reducing the computational domain to a bounded one they have established appropriate "artificial" boundary conditions. This is to avoid the effect of reflected waves in case of a "solid" standard boundary. Then, an equivalent problem is studied in this bounded domain. They used the Laplace-Fourier transform, the two-parameter Mittag-Leffler function and some properties of fractional derivatives go give analytical solution.

An open area of research is how to apply these methods to empirical data. One of the main barriers to application is determining the fractional parameter α. In [16], Ghanam et al. show how to estimate this parameter using a Bayesian framework. In addition this work considers whether or not a fractional calculus based model is supported by data. Their work considered a very simple FPDE. This work extends that work to more complicated FPDE across three forcing functions.

This work employs a Bayesian framework for inferences which utilizes the posterior probability distribution of the parameters and the posterior predictive distributions for predictions. Often the posterior probability density is not analytically tractable and hence numerical methods such as sampling may be used to explore the properties of the posterior probability density. Markov chain Monte Carlo (McMC) techniques are often employed to obtain a large

number of samples from the posterior distribution. All inferences are made from the samples from the posterior probability distribution. For more on using the Bayesian framework for statistical inferences, see [17],[18]. Other techniques such as Method of Moments [19] or Maximum Likelihood [20] are not considered here and the authors have not found any literature where they are applied to FPDEs as likelihood is not differentiable with respect to α.

This project considers the utilizing the Bayesian inferential framework to provide inferences for a system of FPDE across three forcing functions, specifically:

1. Parameter estimation techniques that utilizes empirical data will be developed.
2. Predictive models that utilize empirical data will be developed.
3. Comparing empirical data to each model across forcing function to determine which forcing mechanism generated the data. This will involve determining the posterior probability of the model/forcing function combination given the data.

2. The Problem statement

In this section we will consider the following fractional partial differential equation:

$$\begin{array}{c} \frac{\partial U}{\partial t} - K\, D^{\alpha,\beta} \frac{\partial^2 U}{\partial x^2} = f(x,t) \\ U(x,0) = g(x), \quad -L \le x \le 0 \\ U(-L,t) = 0, \quad 0 \le t \le T \\ \frac{\partial}{\partial x} D^{\alpha,\beta} U(0,t) = 0, \quad 0 \le t \le T, \end{array} \tag{1}$$

where $(D^{\alpha,\beta} f)(t)$ is the Hilfer fractional derivative of $f(t)$ of order α and type β defined by

$$(D^{\alpha,\beta} f)(t) = \left(I^{\beta(1-\alpha)} \frac{d}{dt} I^{(1-\alpha)(1-\beta)} f \right)(t), \ 0 < \alpha < 1, \ 0 \le \beta \le 1, \tag{2}$$

with

$$\left(I^{\alpha} f\right)(t) = \frac{1}{\Gamma(\alpha)} \int_0^{\infty} (t-s)^{\alpha-1} f(s) ds, \ t > 0, \ \alpha > 0. \tag{3}$$

Equation (1) has been studied and solved by Ghanam et al. [15] and explicit solutions have been provided. In fact, the solution for the special case when

$f(x,t) = f(t)\, l(x)$ is given by

$$U(x,t) = \Sigma_{n=0}^{\infty} \left(T_n(0) E_{1-\alpha,1}(-K\lambda_n^2\, t^{1-\alpha}) + C_n B(n,t) \right) \cos \frac{(2n+1)\pi x}{2L}, \quad (4)$$

where

$$\begin{gathered} T_n(0) = \frac{\int_{-L}^{0} g(x) \cos \frac{(2n+1)\pi x}{2L} dx}{\int_{-L}^{0} \cos^2 \frac{(2n+1)\pi x}{2L} dx} \\ E_{\alpha,\beta}(z) = \Sigma_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n+\beta)} \\ B(n,t) = \int_0^t f(t-s)\, E_{1-\alpha,1}(-K\, \lambda_n^2 s^{1-\alpha}) ds \\ C_n = \frac{2}{L} \int_{-L}^{0} l(x) \cos \frac{(2n+1)\pi x}{2L} d\, x. \end{gathered} \quad (5)$$

In this paper, we will consider the special case of Equation (1) with its solution Equation (4) for the case the case $f(t) = 1$. We will now consider three models based on the nature of the forcing function $l(x)$.

2.1. Model 1: A constant forcing function

In this model we consider the following:

$$\begin{gathered} \frac{\partial U}{\partial t} - K\, D^{\alpha,\beta} \frac{\partial^2 U}{\partial x^2} = C \\ U(x,0) = g(x), \quad -L \le x \le 0 \\ U(-L,t) = 0, \quad 0 \le t \le T \\ \frac{\partial}{\partial x} D^{\alpha,\beta} P(0,t) = 0, \quad 0 \le t \le T. \end{gathered} \quad (6)$$

We note that, in the above model we take $l(x) = C$ where C is constant. This yields the following:

$$C_n = \frac{4C(-1)^{n+1}}{(2n+1)\pi}. \quad (7)$$

2.2. Model 2: A linear forcing function

In this model we consider the following:

$$\begin{gathered} \frac{\partial U}{\partial t} - K\, D^{\alpha,\beta} \frac{\partial^2 U}{\partial x^2} = x \\ U(x,0) = g(x), \quad -L \le x \le 0 \\ U(-L,t) = 0, \quad 0 \le t \le T \\ \frac{\partial}{\partial x} D^{\alpha,\beta} U(0,t) = 0, \quad 0 \le t \le T. \end{gathered} \quad (8)$$

We note that in the above model we take $l(x)$ to be a linear function in the form $l(x) = x$,

$$C_n = -\frac{8\left(\frac{\pi(2n+1)(-1)^n}{2} - 1\right) L}{\pi^2 (2n+1)^2}. \quad (9)$$

Model 3: A wave forcing function

In this model we consider the following:

$$\begin{array}{c} \frac{\partial U}{\partial t} - K\, D^{\alpha,\beta} \frac{\partial^2 U}{\partial x^2} = \sin(x) \\ U(x,0) = g(x), \quad -L \le x \le 0 \\ U(-L,t) = 0, \quad 0 \le t \le T \\ \frac{\partial}{\partial x} D^{\alpha,\beta} P(0,t) = 0, \quad 0 \le t \le T. \end{array} \tag{10}$$

We note that in the above model we take $l(x)$ to be an oscillating function of the form $l(x) = \sin(x)$,

$$\begin{array}{rl} C_n & = \frac{4(\pi(n+\frac{1}{2})+L)\cos(L-\pi n-\frac{1}{2}\pi)}{-4(n+\frac{1}{2})^2\pi^2+4L^2} \\ & + \frac{(\pi(-n-\frac{1}{2})+L)\cos L+\pi n+\frac{1}{2}\pi)-2L)}{-4(n+\frac{1}{2})^2\pi^2+4L^2}\,. \end{array} \tag{11}$$

3. Prior Distributions, Likelihood and Computation

For notation, $U(x,t)$ is the underlying process at location x and time t, $\tilde{U}(x,t)$ is an observation of the process at location x at time t and $\hat{U}_\theta(x,t)$ is the solution to the system U at location x and time t using parameters θ. Here $\hat{U}_\theta(x,t)$ will be used as the mean process with with an additive observation error ϵ, namely:

$$\tilde{U}(x,t) = U(x,t) + \epsilon, \tag{12}$$

where $\epsilon \sim N(0,\sigma^2)$ and ϵ is independent across different observations.

A Bayesian estimation approach will be used and hence Bayes Theorem [21] is essential to obtain the posterior probability distribution for the parameters θ given the observed data $\tilde{U}(x,t)$, namely $p(\theta|\tilde{U}(x,t))$ is given by [20, 17, 19]:

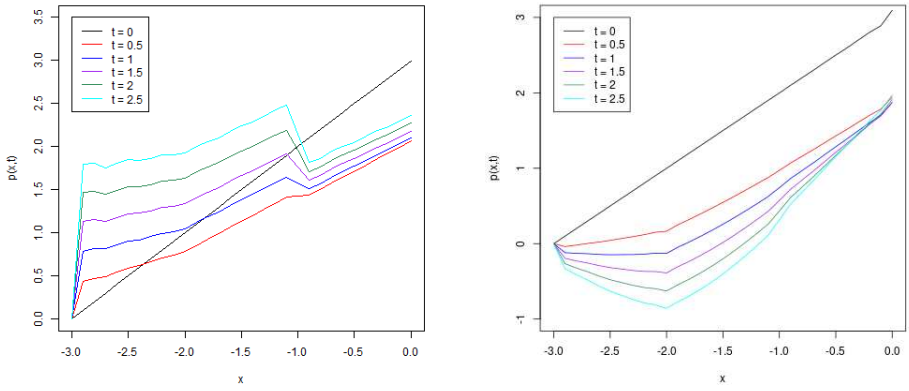
$$p(\theta|\tilde{U}(x,t)) = \frac{p(\theta)L(\tilde{U}(x,t)|\theta)}{\int_\Theta p(\theta)L(\tilde{U}(x,t)|\theta)d\theta}\,, \tag{13}$$

where $p(\theta)$ is the prior probability distribution for the parameters specified *a priori* to experimentation and $L(\tilde{U},t)|\theta)$ is the likelihood of the data given θ. Hence, one needs to specify both the prior distribution for θ and the likelihood. The normal likelihood will be utilized for all three models, specifically:

$$L(U(x,t)|\theta) = Normal(\hat{U}_\theta(x,t)|\theta,\sigma^2), \tag{14}$$

where the mean is $\hat{U}_\theta(x,t)$ is the solution to the differential equation at location x and time t when using parameters θ and σ^2 is the variance parameter which will also need to be estimated. For more on Bayesian inference in general, see [18].

(a) (b)



(c)

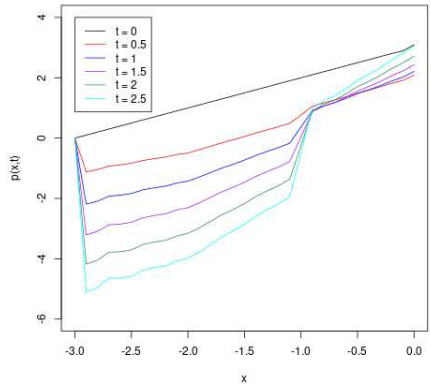


Figure 1: Plots of solutions of Model 1 (a), Model 2 (b) and Model 3 (c) across x at various time points. Here $\alpha = 0.7$, $K = 1$ and $C = 1$ (C is only for Model 1).

3.1. Prior distribution specification

Model 1 has parameters $\theta = \{\alpha, K, C, \sigma\}$ and the prior distributions for these parameters are specified as:

$$\alpha \sim Beta(3, 3) \tag{15}$$

$$\begin{aligned} K &\sim LogNormal(0,1) \\ C &\sim LogNormal(0,1) \\ \sigma &\sim Exp(1). \end{aligned}$$

Note that α must be in the interval $[0,1]$ and the Beta distribution ensures that requirement is met. Similarly, K and C must be on the interval $(0,\infty)$ hence the LogNormal distribution will guarantee that negative values will not be considered. Since σ^2 is a variance parameter it has support on $(0,\infty)$ and hence an Exponential distribution is reasonable.

Model 2 has parameters $\theta = \{\alpha, K, \sigma\}$ and the prior distributions for these parameters are specified as:

$$\begin{aligned} \alpha &\sim Beta(3,3) \\ K &\sim LogNormal(0,1) \\ \sigma &\sim Exp(1). \end{aligned} \tag{16}$$

Similar logic is used in specifying the prior distributions for the parameters as in Model 1.

Model 3 has parameters $\theta = \{\alpha, K, \sigma\}$ and the prior distributions for these parameters are specified as:

$$\begin{aligned} \alpha &\sim Beta(3,3) \\ K &\sim LogNormal(0,1) \\ \sigma &\sim Exp(1). \end{aligned} \tag{17}$$

Similar logic is used in specifying the prior distributions for the parameters as in Model 1. As both Model 2 and Model 3 have the same parameter sets, for consistency the same specification is used.

3.2. Computational Details

As the solution to the differential equation, $\hat{U}_\theta(x,t)$, must be approximated numerically there is no closed form for the posterior distribution and hence inferences on the parameters cannot be obtained in a straightforward manner. To overcome this, McMC techniques will be employed to obtain samples from the posterior distribution, $p(\theta|\tilde{U}(x,t))$, from which all inferences will be made.

The first computational issue is estimating the solution to the differential equation $\hat{U}_\theta(x,t)$, given in equations (6,8,10) using the appropriate C_n given in equations (7,9,11), respectively. These were programmed and calculated in

Julia 1.8.2 one solution across the observed x and t values take approximately 8.81 seconds as it involves several integrals and an infinite sum. All solutions were monitored to ensure the values converged to with in 0.0001 of the true value.

The next computational issue is to obtain samples of the parameters from the posterior distribution. The Metropolis-Hastings McMC scheme will be used to obtain the samples. Note that to obtain one sample from the posterior distribution the solution to the PFDE must be calculated at least four times hence the estimation procedure is very computationally intensive. The posterior distribution is proportional to the quantity given below:

$$\begin{aligned} p(\theta|\tilde{U}(x,t)) &\propto \alpha^2(1-\alpha)^2(CK)^{-1}e^{-1/2(\ln(C)^2+\ln(K)^2)-\sigma} \\ &\times e^{-\frac{1}{2\sigma^2}(\tilde{U}(x,t)-\hat{U}_\theta(x,t))'(\tilde{U}(x,t)-\hat{U}_\theta(x,t))}. \end{aligned} \tag{18}$$

To explore the posterior distribution 1,000 samples were generated using the Metropolis-Hasting framework. Traceplots were examined to ensure convergence of the sampler to a stationary process. The sampler was programmed in Julia 1.8.2 and used no external McMC packages. For each of the models the sampling process took between 7 to 9 hours to complete on a computer with a AMD Ryzen 3 1200 Quad-Core processor at 3.10 GHz and 16GB of RAM. All inferences are made from the 1,000 samples. Traceplots of the posterior samples were examined to ensure convergence of the sampler to a region of high posterior density. Furthermore, several short chains were obtained before sampling to tune the step values of the Metropolis-Hastings sampler. These tuning samples were discarded as burn-in samples. For more on McMC methods in general see [22, 23].

3.3. Estimation and Prediction

To verify the method and algorithm a simulated dataset was generated for each model and the estimation algorithm was employed to generate posterior samples for each of these models. For Model 1 the true parameterisation is $\alpha = 0.7$, $K = 1$, $C = 1$ and $\sigma^2 = 0.1$, for Model 2 $\alpha = 0.7$, $K = 1$ and $\sigma^2 = 0.1$ and for Model 3 $\alpha = 0.7$, $K = 1$ and $\sigma^2 = 0.1$. Table **??** gives the true value, posterior mean and 95% credible interval for each of the parameters in each of the models. Notice that the posterior means are close to the true values and the true value falls within it's respective 95% credible interval. This shows that the proposed estimation method is capable of correctly estimating the parameters.

Figure 2 provides histograms of the posterior samples from each Model and their corresponding parameters. The vertical red line corresponds to the true underlying value of the parameter. Notice that only a few of the parameters show a symmetric distribution, namely Model 1 K (c), Model 2 σ (g). Many parameters exhibit a bimodal shape, namely Model 1 α (a) and C (b), Model 2 α (e) and K (f), and Model 3 α (h) and K (i). The fact that these do not demonstrate a unimodal pattern one can see how Maximum Likelihood techniques and their variants would be difficult to fit as there are likely many local maxima for this problem.

Table 1: True value, posterior mean and 95% credible intervals for Models 1, 2 and 3 for parameters: α, K, C (Model 1 only) and σ^2.

Model	Parameter	True	Mean	95% CI
Model 1	α	0.7	0.693	(0.671, 0.715)
	K	1.0	1.015	(0.995, 1.041)
	C	1.0	0.999	(0.968, 1.035)
	σ^2	0.1	0.108	(0.095, 0.122)
Model 2	α	0.7	0.706	(0.644, 0.774)
	K	1.0	1.012	(0.970, 1.044)
	σ^2	0.1	0.095	(0.086, 0.106)
Model 3	α	0.7	0.688	(0.670, 0.701)
	K	1.0	0.997	(0.981, 1.007)
	σ^2	0.1	0.097	(0.085, 0.112)

As prediction is often very important for researchers and practitioners using the same simulated data sets from the estimation study a predictive performance study is considered. Here the posterior predictive distribution of each model is created using the posterior samples. The posterior predictive distribution is given by:

$$p(\check{U}(x,t)|\tilde{U}(x,t)) = \int_{\Theta} p(\theta|\tilde{U}(x,t))L(\check{U}(x,t)|\tilde{U}(x,t))d\theta \tag{19}$$

where $\check{U}(x,t)$ is the predicted value of $U(x,t)$. Using the posterior samples for each model predicted values were generated for the data with corresponding 95% predictive intervals. The predicted values are the median of the posterior predictive distribution and the predictive intervals are found by obtaining the 2.5% and 97.5% quantiles from the posterior predictive distribution. Figure 3

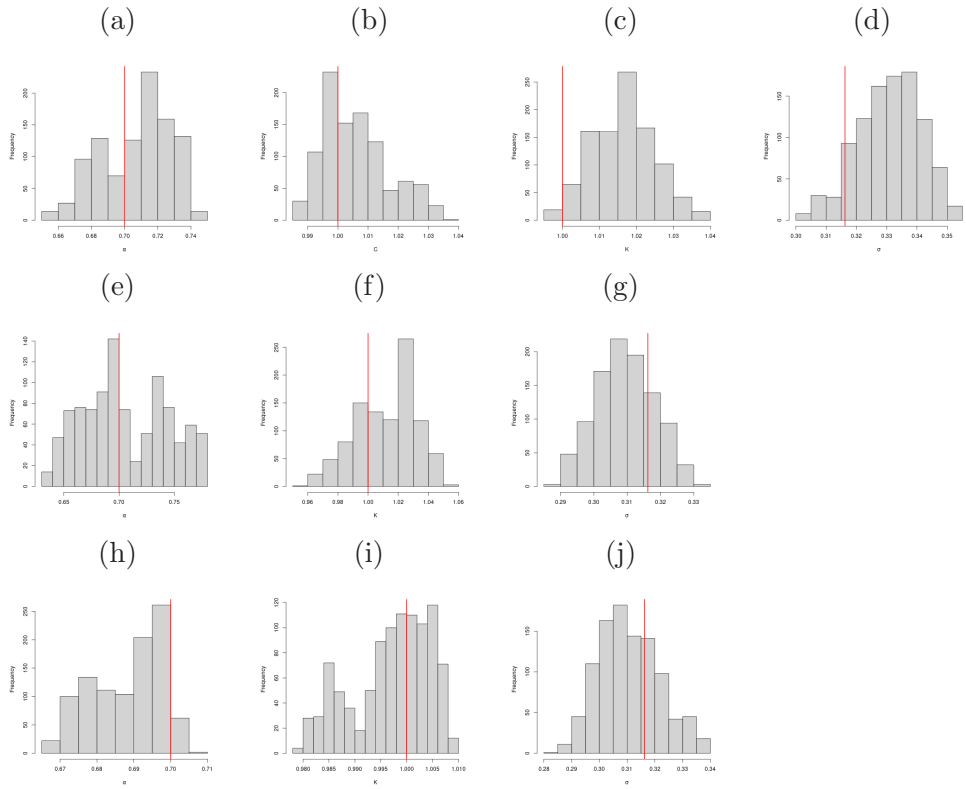


Figure 2: Histograms of posterior samples from Model 1, α (a), C (b), K (c) and σ (d). For Model 2 parameters α (e), K (f) and σ (g). For Model 3 parameters α (h), K (i) and σ (j). Red line corresponds to the true value of the parameter.

shows the data, the predicted value and the 95% predictive intervals for each model across x and across t. When looking across x, $t = 1$ and when considering across t, $x = 1$. Notice that for all of the Models the predictive interval captures the data. This demonstrates that the estimation paradigm has potential use in prediction as well.

3.4. Model Section

The McMC approach presented can also be used in the model selection setting. For example, if one does not know the mechanism that generated the observed

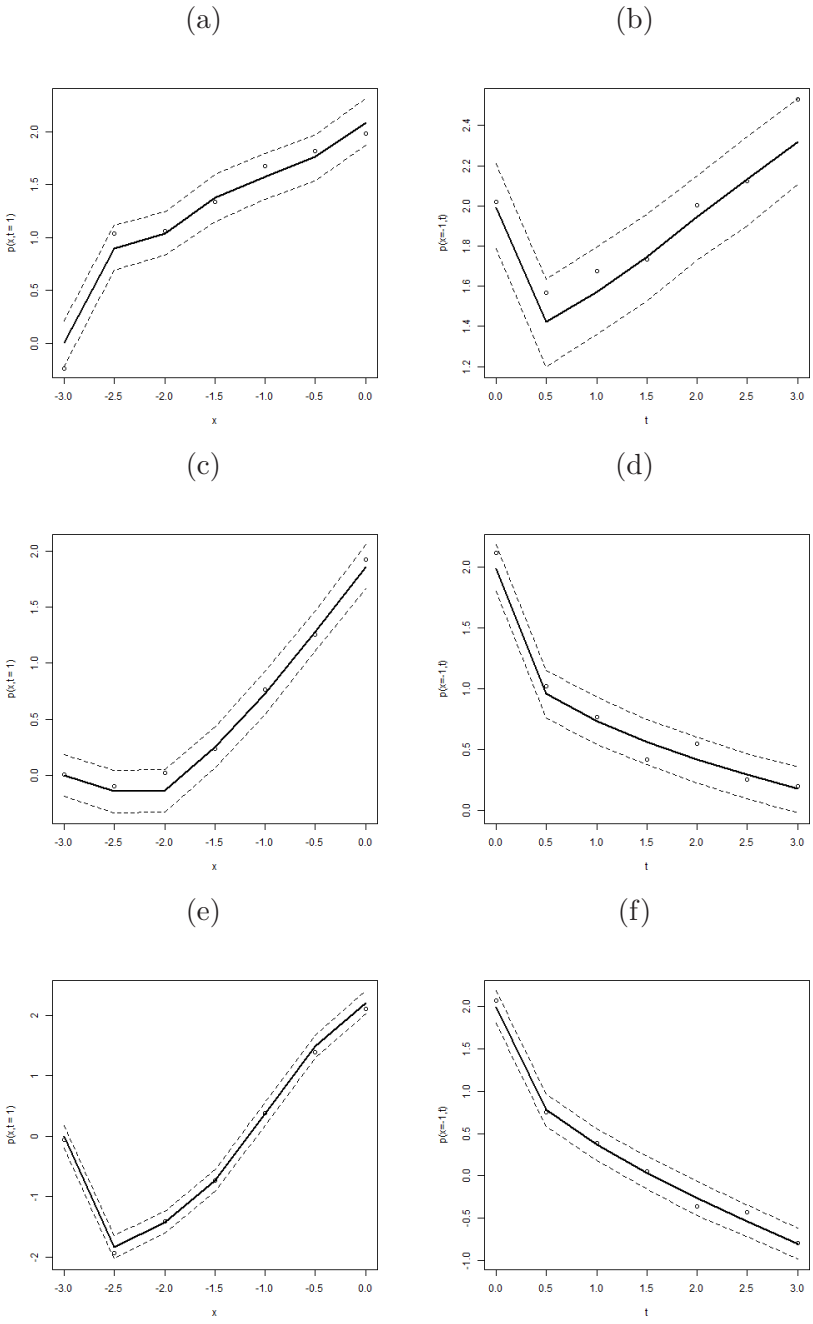


Figure 3: Posterior predictive plots for across x and t for Model 1, panels (a) and (b), Model 2, panels (c) and (d) and Model 3 panels (e) and (f).

data a model selection approach can to determine the probability that the data were generated by one of the candidate models. This notion of highest posterior model probability is presented in [24, 25]. Suppose there are m models, $M_1, M_2, ..., M_m$ under consideration. Given a dataset D, one can determine the posterior model probability of each model by using Bayes' Theorem:

$$P(M_i|D) = \frac{P(M_i)P(D|M_i)}{\sum_{j=1}^{m} P(M_j)P(D|M_j)} \tag{20}$$

where $P(M_i)$ is the prior probability of model M_i and $P(D|M_i)$ is the marginal probability of the data D given model M_i and is found by:

$$P(D|M_i) = \int_{\Theta} p(\theta|M_i)L(D|\theta, M_i)d\theta. \tag{21}$$

In Section 4 a simulation study is performed to show the ability of the method to select the correct model. Note that when none of the models have an overwhelming probability one can utilize all of the models by averaging the predictions across all models using $P(M_i|D)$ as the weighting for each prediction.

4. Simulation Studies

To examine the ability of the approach for statistical inference on parameter estimates a simulation study was conducted. This study varies α and the variance of the observation error σ. Here α is varied across 0.3, 0.5 and 0.7 and σ is varied across 0.1 and 0.2 for each of the models. The other parameters are set to $K = 1$ across all models and $C = 1$ for Model 1, in order to keep the size of the simulation study reasonable. For each model and parameter combination 100 data sets were generated and the proposed estimation method is employed to create 95% posterior credible intervals for the each parameters. The coverage probability is calculated for each of the parameters and models using:

$$\hat{\theta} = \frac{1}{100} \sum_{i=1}^{100} \mathbf{I}\left[\theta \in (\theta_{L,i}, \theta_{U,i})\right], \tag{22}$$

where $\mathbf{I}$ is an indicator function θ is the parameter of interest, $\theta_{L,i}$ and $\theta_{U,i}$ are the lower and upper bounds for the 95% posterior credible interval for the ith dataset.

Table 2 shows the results of this simulation study where all three models are considered and their associated parameter estimated coverage probabilities are

presented. One very clear item to observe is that the parameter K was never in any of the calculated posterior credible intervals for any model or specified parameter combination. As this seemed very odd, further inspection into the individual credible intervals revealed that all of the credible intervals were slightly lower than the true value of $K = 1$. This phenomenon can be observed in Table 1 where most of the region of the 95% posterior credible intervals K were below its value true value. Another odd phenomenon to notice in Table 2 is that α, C and σ had 100% coverage. This gives an indication that the posterior credible intervals are wider than would be expected in general. Hence there seems to be a lack of correct uncertainty calibration in the approach.

Table 2: Coverage probability study for 95% credible intervals for Models 1, 2 and 3 for parameters: α, K, C (Model 1 only) and σ^2 for $\alpha = 0.3, 0.5, 0.7$ and $\sigma = 0.1, 0.2$. Both $K = 1$ and $C = 1$ were used as true values where appropriate. Results based on 100 simulated datasets for each parameter combination.

	Parameters		95% CI Coverage			
Model	σ	α	$\hat{\alpha}$	$\hat{K}$	$\hat{C}$	$\hat{\sigma}$
Model 1	0.1	0.3	1.00	0.00	1.00	1.00
		0.5	1.00	0.00	1.00	1.00
		0.7	1.00	0.00	1.00	1.00
	0.2	0.3	1.00	0.00	1.00	1.00
		0.5	1.00	0.00	1.00	1.00
		0.7	1.00	0.00	1.00	1.00
Model 2	0.1	0.3	1.00	0.00		1.00
		0.5	1.00	0.00		1.00
		0.7	1.00	0.00		1.00
	0.2	0.3	1.00	0.00		1.00
		0.5	1.00	0.00		1.00
		0.7	1.00	0.00		1.00
Model 3	0.1	0.3	1.00	0.00		1.00
		0.5	1.00	0.00		1.00
		0.7	1.00	0.00		1.00
	0.2	0.3	1.00	0.00		1.00
		0.5	1.00	0.00		1.00
		0.7	1.00	0.00		1.00

To determine the ability of utilizing a highest posterior model probability approach to correctly identify the process that the generate by, a simulation study was conducted. A total of 90 detests were generated, 30 from Model 1 with $\alpha = 0.7,\ K = 1,\ C = 1$ and $\sigma = 0.1$, 30 from Model 2 with $\alpha = 0.7,\ K = 1$ and $\sigma = 0.1$ and 30 from Model 3 with $\alpha = 0.7,\ K = 1$ and $\sigma = 0.1$. It should be noted that even though the parameter values are the same the structure of the solution of the FPDE is very different. For each of the 90 simulated data setss all three models were fit to the data and the posterior model probability was calculated using equation (20). For each model and data set, the McMC sampler was utilized to obtain 1,000 samples from the posterior distribution, which in turn were used to approximate (21). The same prior distributions were specified for each model as those in previous sections. The model with the highest posterior model probability was said to be the "correct model."

Table 3 shows the confusion matrix for this study. The cell in the first row and second column has the value 0 meaning that of all the data sets generated by Model 2, 0 were classified as Model 1. Similarly, for the first row first column there is a value of 30, meaning that for data sets generated by Model 1, 30 were classified as model 1. It is clear to see that across all 90 data sets, the model with the highest posterior model probability was the model that actually generated the data. This demonstrates the ability of the posterior model probability approach to select the correct model.

	True		
HPMP	Model 1	Model 2	Model 3
Model 1	30	0	0
Model 2	0	30	0
Model 3	0	0	30

Table 3: Confusion matrix for the True model that generated the data and the Highest Posterior Model Probability (HPMP) selected using the presented approach. For each model 30 datasets were generated for a total of 90 datasets.

5. Discussion

This work shows both the challenges and successes of using a Bayesian parameter estimation approach for the fractional parameter α, the coefficient K and the model error variance σ^2 and in Model 1 the constant forcing pressure C.

The proposed method works well for α, σ^2 and C, however does not perform well for the coefficient K as it is slightly underestimated. Furthermore, the proposed method performs well in terms of predictive performance in spite of the underestimation of K. Additionally, the approach works well with model selection for situations when one may be uncertain about which process was used to generate the data. Overall, the proposed method shows promise for researchers in the fields of Engineering, Chemistry, and Physics to employ fractional partial differential equation modeling to real world data sets.

As mentioned before, the proposed method does not quantify the uncertainty associated with the constant K as evidenced by the poor coverage probabilities. This may be attributable to the approximation used when calculating the solution. The solution given in equation (5) involves nested infinite sums which must be approximated using finite sums in which the error in the inner sum will accumulate across the outer sum. This notion is further evidenced by the fact that the estimated values of K were slightly under the true value. A larger scale study needs to be performed to understand the impact of the accuracy of the approximations on the coverage probabilities for K.

Future work includes studying the impact of approximation accuracy of the infinite sums on credible interval coverage probabilities for all parameters. In addition, [26] consider utilizing internal model error in conjunction with external model error to account for model misspecification such as approximation inaccuracy. Due to the computational efforts needed by this approach one could speed up the calculations while still accurately quantifying the uncertainty. One other area of exploration is the design of experiments for generating real world data. At what distances x and what times t should the researcher focus on to be able to best estimate the parameters and make inferences while still keeping the number of experimental units low. This shows there is a large amount of work that can still be done in this area to improve domain researchers ability to employ these models to solve their problems from a data science perspective.

Acknowledgements

The authors would like to thank Virginia Commonwealth University in Qatar and Qatar Foundation for their support of this research through the Faculty Research Grant and the Mathematical Data Sciences Lab at VCU Qatar.

References

[1] W. Hundsdorfer, J. Verwer, *Numerical solution of Time-Dependent Advection-Diffusion-Reaction Equations*, Springer-Verlag, Berlin-Heidelberg (2003).

[2] R.B. Bird, W.E. Stewart, E.N. Lightfoot, *Transport Phenomena*, John Wiley and Sons, Inc., New York (2002).

[3] T.E. Graedel, P.J. Crutzen, *Atmosphere, Climate and Change*, Henry Holt and Company(1997).

[4] K. Aziz, A. Settari, *Petroleum Reservoir Simulation*, Applied Science Publishers Ltd., London (1979).

[5] I. Glassman, R.A. Yetter, *Combustion*, Academic Press, London-San Diego-Burlington (2008).

[6] 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, *Transport In Porous Media*, **124**, No. 3 (2018), 723-742.

[7] 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.

[8] 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.

[9] R. Hilfer, *Applications of Fractional Calculus in Physics* (Ed. by R. Hilfer), World Scientific Publishing Co., Singapore, 2000, p. 87 and p. 429.

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

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

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

[13] F. Mainardi, *Fractional Calculus and Waves in Linear Viscoelasticity*, Imperial College Press, London (2010, 2022).

[14] R.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.

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

[16] 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*, **33**, No 5 (2020), 783-803. http://dx.doi.org/10.12732/ijam.v33i5.4.

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

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

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

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

[21] 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** (1763), 370–418.

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

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

[24] J.A. Hoeting, D. Madigan, A.E. Raftery, C.T. Volinsky, Bayesian model averaging: A tutorial. *Statistical Science*, **14**, No 4 (1999), 382-417.

[25] R.E. Kass, L. Wasserman, A reference Bayesian test for nested hypothesis and its relationship to the Schwarz criterion, *Journal of the American Statistical Association*, **90**, No 431 (1995), 928-934.

[26] E. Boone, J, Hannig, R. Ghanam, S. Ghosh, F. Ruggeri, S. Prudhomme, Model validation of a single degree-of-freedom oscillator: A case study. *Stats.*, **2022**, No 5 (2022), 1195–1211. https://doi.org/10.3390/stats5040071.