ORIGINAL PAPER

Check for
updates

Inverse Population Dynamics Problem Employing a Low Cost Integral Transform Solution and Bayesian Inference with Approximation Error Model

Maiquison S. Friguis[1] · Diego C. Knupp[2] ⓘ **· Luiz A. S. Abreu[2]** ⓘ **· Leonardo T. Stutz[2] · Antônio J. Silva Neto[2]**

Abstract

Population dynamics modeling is a subject of major relevance, especially when involving human or livestock disease vectors. Such importance is due to the fact that there are several diseases that are spread by some particular species, and the knowledge on the behavior of such populations is relevant when it is intended to create public policies to control their proliferation. This work describes a problem of population dynamics with diffusive behavior, with impulsive culling and delayed reproduction. The direct problem describes the space and time evolution of the population density when the model parameters are known, and the solution of the partial differential equation is obtained with the Generalized Integral Transform Technique (GITT), a hybrid numerical-analytical approach. For practical purposes it is crucial to fit the model to a population of interest, by estimating the equation coefficients through an inverse problem approach. In this work the inverse problem is illustrated within the Bayesian framework employing a low-cost direct problem solution, and the Approximation Error Model is used to take in account the error introduced by the low-cost solution.

Keywords Diffusive populations · Impulsive culling · Integral transforms · Bayesian inference · Approximation error model

List of Symbols

b Birth rate function
B_j Impulsive culling sites intensities
C Logistic parameter growth
d Death rate function
D Dispersion constant coefficient

✉ Diego C. Knupp
   diegoknupp@iprj.uerj.br

1   Instituto Federal Fluminense Campus Cabo Frio, Estrada Cabo Frio - Búzios, Cabo Frio, RJ 28909-971, Brazil

2   LEMA - Laboratory Patrícia Oliva Soares of Experimentation and Numerical Simulation in Heat and Mass Transfer, Department of Mechanical Engineering and Energy, Polytechnic Institute, Rio de Janeiro State University, IPRJ/UERJ, Rua Bonfim 25, Vila Amelia, Nova Friburgo, RJ 28625-570, Brazil

g	Source term
K	Constant capacity
L	Space domain width
M	Subregions of integration
N	Truncation order
N_{cs}	Number of culling sites
N_D	Number of experimental data
N_R	Reduced truncation order
N_i	Norms
N_S	Number of samples
P	Logistic parameter growth
u	Adult population density
u_0	Initial adult population density
U_C	High order solution
U_R	Reduced order solution
x_j	Location of culling sites
$\mathbf{Y}$	Vector of measurements

Greek Symbols

δ	Dirac delta function
$\epsilon_i(.)$	Modeling error
e_i	Experimental noise for ith measurement
λ	Eigenvalue corresponding to the eigenfunction ψ
μ	Premature death rate
μ_e	Mean
π	Probability density
τ	Reproduction time delay
σ	Standard deviation
ψ	Eigenfunction

Subscripts and Superscripts

$\sim$	Normalized eigenfunction

Introduction

Since the pioneering works of Malthus [1] and Verhulst [2] many other studies have been proposed aiming at the analysis of population dynamics, many of them based on the movement of the population considering the dispersion and taking into account logistic growth rates, seeking the factors that can alter in any way birth rates and mortality. In [3] the random-walk is adopted as a starting point for the study of dispersal in living organisms, while in [4,5] the main models involving this mechanism are reviewed. Considering population control mechanisms [6], extinction conditions are obtained through impulsive culling. Most of those works present different models and their simplified analytical or numerical solutions to specific cases are carried out. However, most works end up ignoring the practical applications derived from such models. In this sense, a way to create a bridge between theory and practice is the inverse analysis, once the parameters estimation approach can be used for fitting such models to a given population of interest, allowing for the practical use of these models in forecasting and optimal mechanisms of population control [7–10]. The inverse problem analysis involves, in

general, intensive iterative procedures. Therefore, it is of major importance to have a precise solution methodology with good computational efficiency in solving the mathematical model, generally given by partial differential equations [11].

Aiming at a robust implementation for the population dynamics problem modeled as partial differential equations, Knupp et al. [9] employed a combination of hybrid numerical-analytical solution methodology, known as the Generalized Integral Transform Technique (GITT), with the Differential Evolution optimization method for the inverse problem related to population growth models with time delay and impulsive culling, using the maximum likelihood procedure. It was observed the need of prior information regarding one or more model parameters, which should be obtained by means of other prior independent procedure. In such cases, it is critically important to consider how the uncertainties present in the supposedly known values of these parameters affect the estimation of the others. In this scenario, however, the most attractive framework is the Bayesian approach, which combines the likelihood function with prior information in order to yield a representation for the parameters posterior distribution [12]. The Bayesian framework has been successfully employed in different areas, as can be seen in these works [13–19], just to cite a few examples.

In certain conditions, where the experimental error follows a normal distribution and the prior information can be modeled as a normal distribution, it is possible to derive the *Maximum* a posteriori objective function that may be used to obtain single point estimates for the unknowns, besides the approximation of the confidence intervals. Nevertheless, this represents only part of the information on the unknowns, and the estimated confidence intervals must be used cautiously, since this commonly adopted approach is exact only for linear problems, being just an approximation for nonlinear inverse problems, as the one to be considered in this work. The posterior probability distribution may be further explored using random sampling methods, such as the Markov Chain Monte Carlo techniques. These approaches are more computationally intensive, but on the other hand they allow to approach the true posterior distribution upon the appropriate modeling of the prior information and the experimental errors, regardless of their statistical distribution form.

In the present work, the inverse analysis of population dynamics with time delay and impulsive culling is tackled, with two main contributions: (i) in order to adequately take into account the prior information considered available for one or more model parameters, as well as the uncertainties involved, the inverse analysis is carried out within the Bayesian framework and implemented via Markov Chain Monte Carlo methods with the Metropolis-Hastings (MH) algorithm; (ii) in order to overcome the difficulty posed by the high number of simulations required by the Markov Chain Monte Carlo methods, it is necessary a fast and accurate direct problem solution implementation. In this sense, it is proposed a computationally low cost solution, constructed upon the integral transformation of the original problem, and employing semi-analytical integrations and reduced-order expansions of the nonlinear terms. The use of this approximated solution, however, introduces errors which should be taken in account in the inverse problem solution. Thus, in this work it is also investigated the Approximation Error Model (AEM) approach in the Bayesian formulation [20].

Direct Problem Formulation and Solution Methodology

Consider the time dependent density of an adult population u defined in a one-dimensional space with diffusive behaviour according to Fick's law, and constant dispersion coefficient D, reproduction time delay τ, premature death rate μ, and birth and death rates given by $b(u)$ and $d(u)$, respectively. The governing equation is given by [6],

$$\frac{\partial u(x,t)}{\partial t} = D\frac{\partial^2 u(x,t)}{\partial x^2} + e^{-\mu\tau}b(u) - d(u) - \sum_{j=1}^{N_{cs}} B_j u(x_j, t)\delta(x - x_j) \qquad (1)$$

defined in $0 < x < L$ for $t > 0$ and the positions x_j represent the location of the culling sites, with culling intensities B_j. The total number of culling sites is given by N_{cs}, and δ denotes the Dirac-delta function.

The initial and boundary conditions are given respectively by:

$$u(x,t) = u_0(x,t), \quad -\tau \le t \le 0 \qquad (2)$$

$$\left.\frac{\partial u}{\partial x}\right|_{x=0} = 0, \quad \left.\frac{\partial u}{\partial x}\right|_{x=L} = 0 \qquad (3)$$

Assume that the rates $b(u)$ and $d(u)$ are defined according to a logistic model, being respectively given by [7]:

$$b(u) = \frac{u(x, t-\tau)}{K + Cu(x, t-\tau)} \qquad (4)$$

$$d(u) = \frac{Pu^2(x,t)}{K + Cu(x, t-\tau)} \qquad (5)$$

Table 1 Parameters for the direct problem solution

D	P	K	C	B	μ	τ	L	$u_0(x,t)$
0.0001	0.005	1.0	0.5	0.001	0.0	10.0	1.0	1.0

Table 2 Direct problem solution convergence behavior, $u(x, t = 50)$ (a) With $M = 30$ subregions for the semi analytical integration. (b) With $M = 150$ subregions for the semi analytical integration

	Complete Model				Reduced Model ($N_R = 1$)			
N	$x = 0.2$	$x = 0.4$	$x = 0.5$	sec.*	$x = 0.2$	$x = 0.4$	$x = 0.5$	sec.*
20	0.741	0.736	0.886	2.25	0.738	0.731	0.885	0.36
30	0.728	0.723	0.873	5.27	0.725	0.719	0.871	0.44
50	0.719	0.714	0.877	17.85	0.717	0.710	0.876	1.18
80	0.715	0.710	0.881	64.19	0.712	0.706	0.879	2.48
100	0.714	0.709	0.881	120.63	0.711	0.705	0.880	4.34
20	0.741	0.736	0.886	11.98	0.738	0.731	0.885	0.78
30	0.729	0.723	0.873	30.01	0.725	0.719	0.871	1.26
50	0.720	0.715	0.877	114.69	0.717	0.710	0.876	3.42
80	0.716	0.710	0.880	361.89	0.713	0.706	0.879	5.85
100	0.714	0.709	0.880	673.35	0.711	0.705	0.879	11.06

*CPU time

Table 3 Direct problem solution convergence behavior, $u(x, t = 125)$ (a) With $M = 30$ subregions for the semi analytical integration. (b) With $M = 150$ subregions for the semi analytical integration, 21

	Complete Model				Reduced Model ($N_R = 1$)			
N	$x = 0.2$	$x = 0.4$	$x = 0.5$	sec.*	$x = 0.2$	$x = 0.4$	$x = 0.5$	sec.*
20	0.608	0.573	0.690	5.77	0.608	0.563	0.680	0.30
30	0.600	0.566	0.683	16.42	0.600	0.557	0.673	0.56
50	0.594	0.560	0.689	46.33	0.595	0.552	0.680	1.17
80	0.592	0.559	0.693	153.64	0.592	0.550	0.684	2.37
100	0.591	0.558	0.694	308.13	0.591	0.549	0.685	3.58
20	0.608	0.573	0.690	28.67	0.608	0.563	0.680	0.81
30	0.600	0.566	0.683	66.41	0.600	0.557	0.673	1.42
50	0.595	0.561	0.689	254.52	0.595	0.552	0.679	3.09
80	0.592	0.559	0.692	393.88	0.592	0.550	0.683	5.87
100	0.591	0.558	0.693	690.53	0.591	0.550	0.683	8.78

*CPU time

Table 4 Test cases

	Approximation Error Model	With inverse crime	Without inverse crime	With Prior	Without Prior
Case 1			✓		✓
Case 2		✓			✓
Case 3			✓	✓	
Case 4		✓		✓	
Case 5	✓		✓		✓
Case 6	✓		✓	✓	

Table 5 Case 1—Inverse problem solution (a) Statistics obtained after discarding the burn-in states. (b) Estimates with 95% confidence intervals obtained with MCMC method. CI width (%): ratio between the 95% confidence interval range and the estimated mean value

Parameters	Exact	Mean(μ)	Standard Deviation (σ)	Coefficient of Variation(σ/μ)%
$D \times 10^4$	1	0.617138	6.35198×10^{-3}	1.02926
$P \times 10^3$	5	3.43084	1.77159×10^{-2}	0.516373
K	1	0.758786	6.38265×10^{-3}	0.841166
$C \times 10$	5	0.0181148	1.54953×10^{-2}	85.5394
$B \times 10^3$	1	0.739847	7.73914×10^{-3}	1.04605

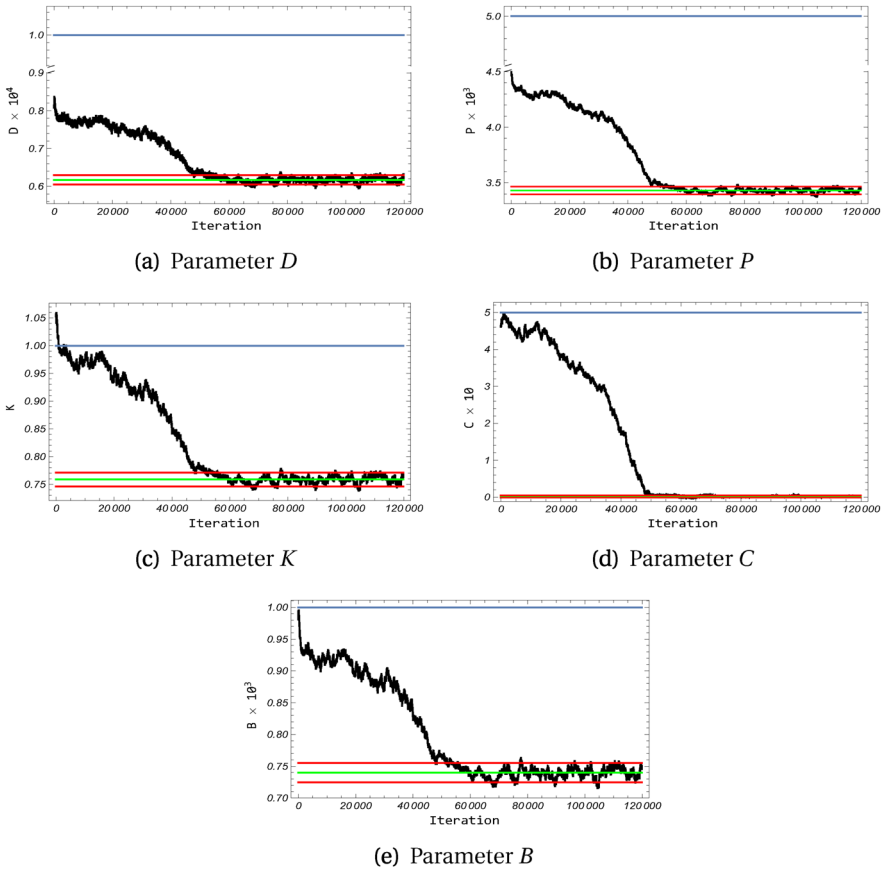| Parameters	Initial	Confidence Interval	CI width(%)
$D \times 10^4$	0.82142	[0.60469, 0.62959]	4.03464
$P \times 10^3$	4.50741	[3.39612, 3.46556]	2.02414
K	1.05749	[0.74628, 0.77130]	3.29731
$C \times 10$	4.60489	[−0.01226, 0.04849]	335.308
$B \times 10^3$	0.97761	[0.72468, 0.75501]	4.10043

Fig. 1 Case 1—Markov Chain for each parameter (black), Mean (green), 95% confidence interval (red) and exact value (blue)

In this population dynamics model, the parameters are biologically interpreted as: D is the population dispersion coefficient, P is the intrinsic growth rate, K is the carrying capacity, C is the rate of replacement of the population in carrying capacity and B is the intensity of impulsive culling sites.

For the solution of the partial differential equation, given by Eqs. (1–5), the hybrid numerical–analytical approach named Generalized Integral Transform Technique (GITT) [21] has been employed. Following the formalism for the solution of the nonlinear model given by Eq. (1), the following integral transformation pair is defined:

$$\text{Transform: } \bar{u}_i(t) = \int_0^L u(x,t)\widetilde{\psi}_i(x)dx \tag{6}$$

$$\text{Inverse: } u(x,t) = \sum_{i=0}^{\infty} \bar{u}_i(t)\widetilde{\psi}_i(x) \tag{7}$$

Int. J. Appl. Comput. Math (2021) 7:189

Page 7 of 25   **189**

(a) Parameter D

(b) Parameter P

(c) Parameter K

(d) Parameter C

(e) Parameter B
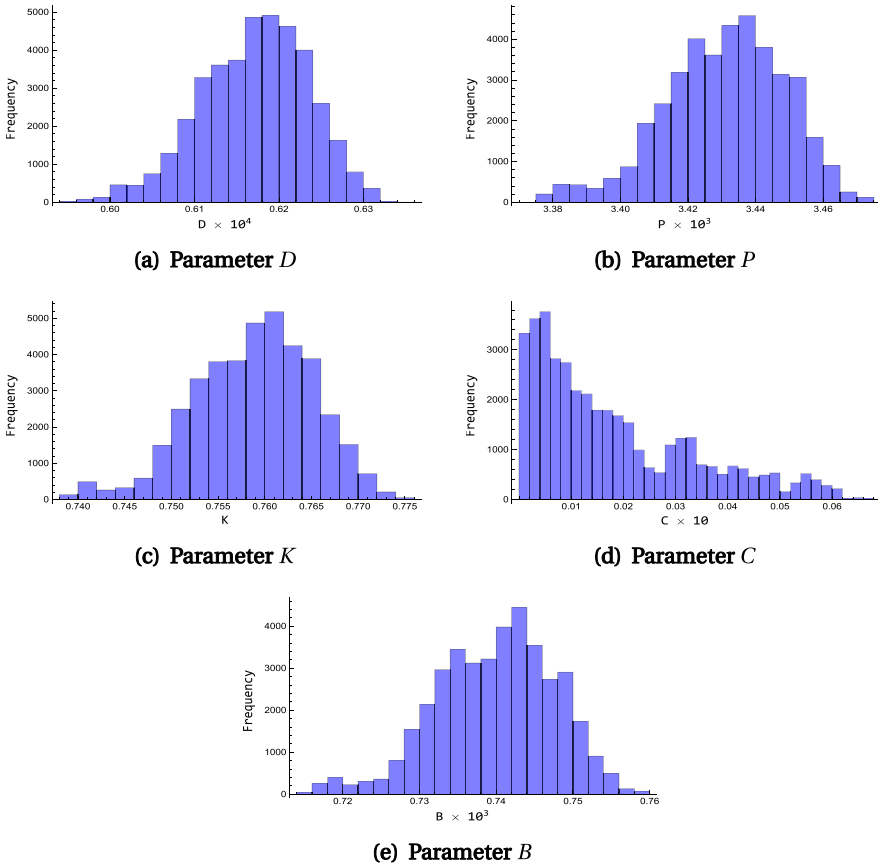
Fig. 2 Case 1—posterior distribution

where $\widetilde{\psi}_i(x)$ are the normalized eigenfunctions:

$$\widetilde{\psi}_i(x) = \frac{\psi_i(x)}{\sqrt{N_i}} \tag{8}$$

with the normalization integral, N_i, given by:

$$N_i = \int_0^L \psi_i^2(x)dx \tag{9}$$

The eigenfunctions are obtained considering an auxiliary eigenvalue problem, derived from the direct application of separation of variables to the linear homogeneous purely diffusive version of Eq. (1). Hence, the following Sturm–Liouville eigenvalue problem is employed:

$$D\frac{d^2\psi_i(x)}{dx^2} + \lambda_i^2\psi_i(x) = 0 \,,\; 0 < x < L \tag{10}$$

Table 6 Case 2—Inverse problem solution (a) Statistics obtained after discarding burn-in states. (b) Estimates with 95% confidence intervals obtained with MCMC method. CI width (%): ratio between the 95% confidence interval range and the estimated mean value

Parameters	Exact	Mean (μ)	Standard Deviation (σ)	Coefficient of Variation (σ/μ)%
$D \times 10^4$	1	1.04804	7.63162×10^{-3}	0.73
$P \times 10^3$	5	5.62001	3.25257×10^{-2}	0.56
K	1	1.21984	1.43334×10^{-2}	1.17
$C \times 10$	5	13.6429	4.29008×10^{-1}	3.14
$B \times 10^3$	1	1.04939	6.42819×10^{-3}	0.61

Parameters	Initial	Confidence Interval	CI width(%)
$D \times 10^4$	1.315	[1.03309, 1.063]	2.85
$P \times 10^3$	5.866	[5.55626, 5.68376]	2.27
K	1.568	[1.19175, 1.24793]	4.60
$C \times 10$	13.099	[12.802, 14.4837]	12.32
$B \times 10^3$	1.2581	[1.03679, 1.06199]	2.40

with boundary conditions:

$$\left.\frac{d\psi_i(x)}{dx}\right|_{x=0} = 0, \quad \left.\frac{d\psi_i(x)}{dx}\right|_{x=0} = 0 \tag{11}$$

Equation (10), with the boundary conditions given in Eq. (11), leads to an infinite set of solutions for the eigenfunctions $\psi_i(x)$, for discrete values of λ_i:

$$\psi_i(x) = \cos(\lambda_i x), \quad i = 0, 1, 2, \ldots \tag{12}$$

where the eigenvalues λ_i are given by:

$$\lambda_i = \frac{i\pi\sqrt{D}}{L}, \quad i = 0, 1, 2, \ldots \tag{13}$$

The integral transformation of the original problem is carried out by operating on Eq. (1) with $\int_0^L \widetilde{\psi}_i(x)\,(.)\,dx$, and employing the inverse formula, Eq. (7), into the source-terms, yielding the following set of ordinary differential equations for the transformed potentials $\overline{u}_i$,

$$\frac{d\overline{u}_i(t)}{dt} + \lambda_i^2 \overline{u}_i(t) = \overline{g}_i(t, \mathbf{u}), \quad i = 0, 1, 2, \ldots \tag{14}$$

with

$$\mathbf{u} = (\overline{u}_0, \overline{u}_1, \overline{u}_2 \ldots) \tag{15}$$

and

$$\overline{g}_i(t, \mathbf{u}) = \overline{g}_{i,1}(t, \mathbf{u}) + \overline{g}_{i,2}(t, \mathbf{u}) + \overline{g}_{i,3}(t, \mathbf{u}) \tag{16}$$

where

$$\overline{g}_{i,1} = \int_0^L \frac{\widetilde{\psi}_i(x)e^{-\mu\tau}\sum_{j=0}^{\infty}\overline{u}_j(t-\tau)\widetilde{\psi}_j(x)}{K + C\sum_{j=0}^{\infty}\overline{u}_j(t-\tau)\widetilde{\psi}_j(x)}dx \tag{17}$$

(a) Parameter D

(b) Parameter P

(c) Parameter K
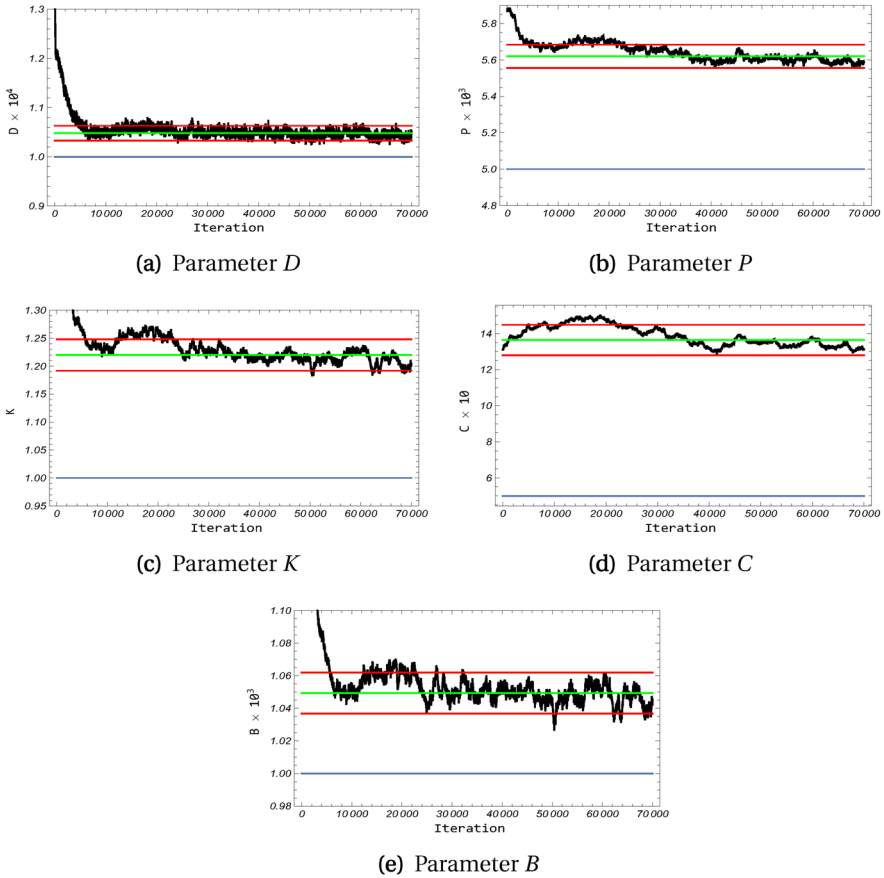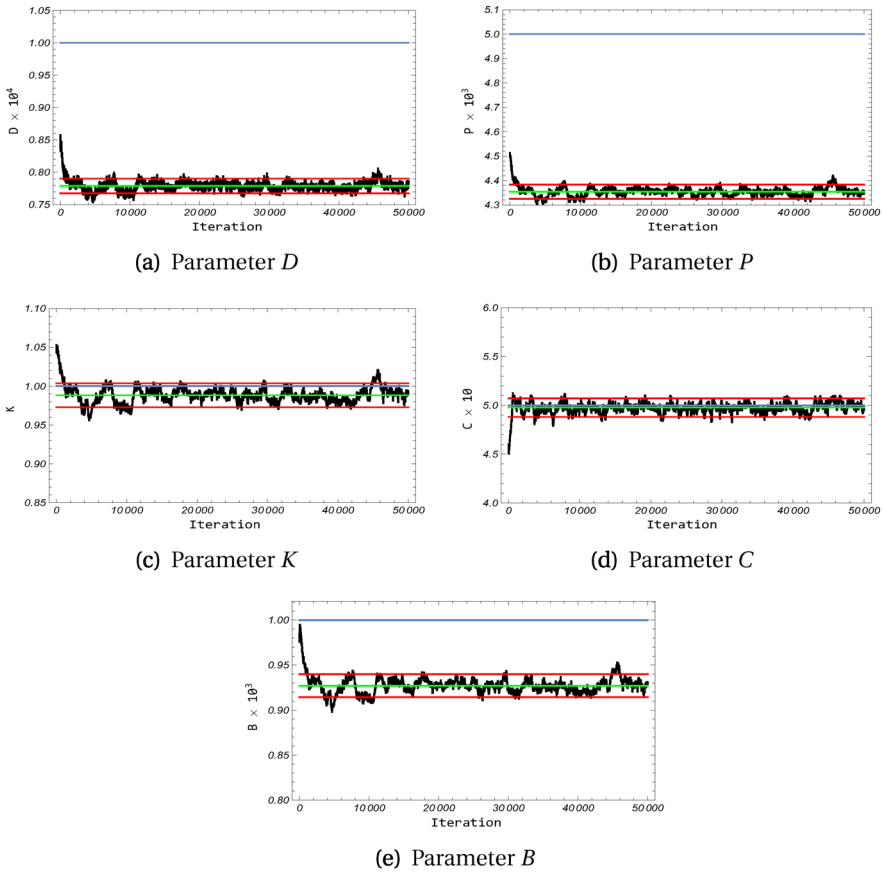

(d) Parameter C

(e) Parameter B

Fig. 3 Case 2—Markov Chain for each parameter (black), Mean (green), 95% confidence interval (red) and exact value (blue)

$$\overline{g}_{i,2} = \int_0^L \frac{\widetilde{\psi}_i(x)\, P\left(\sum_{j=0}^\infty \overline{u}_j(t)\widetilde{\psi}_j(x)\right)^2}{K + C \sum_{j=0}^\infty \overline{u}_j(t-\tau)\widetilde{\psi}_j(x)}\, dx \tag{18}$$

$$\begin{aligned}
\overline{g}_{i,3} &= \int_0^L \widetilde{\psi}_i(x) \sum_{k=1}^{N_{cs}} B_k \left(\sum_{j=0}^\infty \overline{u}_j(t)\widetilde{\psi}_j(x_k)\right)\delta(x - x_k)\,dx \\
&= \sum_{k=1}^{N_{cs}} B_k \widetilde{\psi}_i(x_k)\left(\sum_{j=0}^\infty \overline{u}_j(t)\widetilde{\psi}_j(x_k)\right)
\end{aligned} \tag{19}$$

The transformed initial conditions are obtained by the integral transformation of Eq. (2), yielding:

$$\bar{u}_i(t) = \int_0^L \tilde{\psi}_i(x) u_0(x, t)\, dx, \quad -\tau \le t \le 0, \ i = 0, 1, 2, \dots \tag{20}$$

Equations (14–20) form an infinite set of coupled ordinary differential equations, which is unlikely to allow for explicit analytical solutions. Nonetheless, if the system and the infinite expansions appearing in Eqs. (17–19) are truncated to a finite order N, the system can be

(a) Parameter D

(b) Parameter P

(c) Parameter K

(d) Parameter C

(e) Parameter B
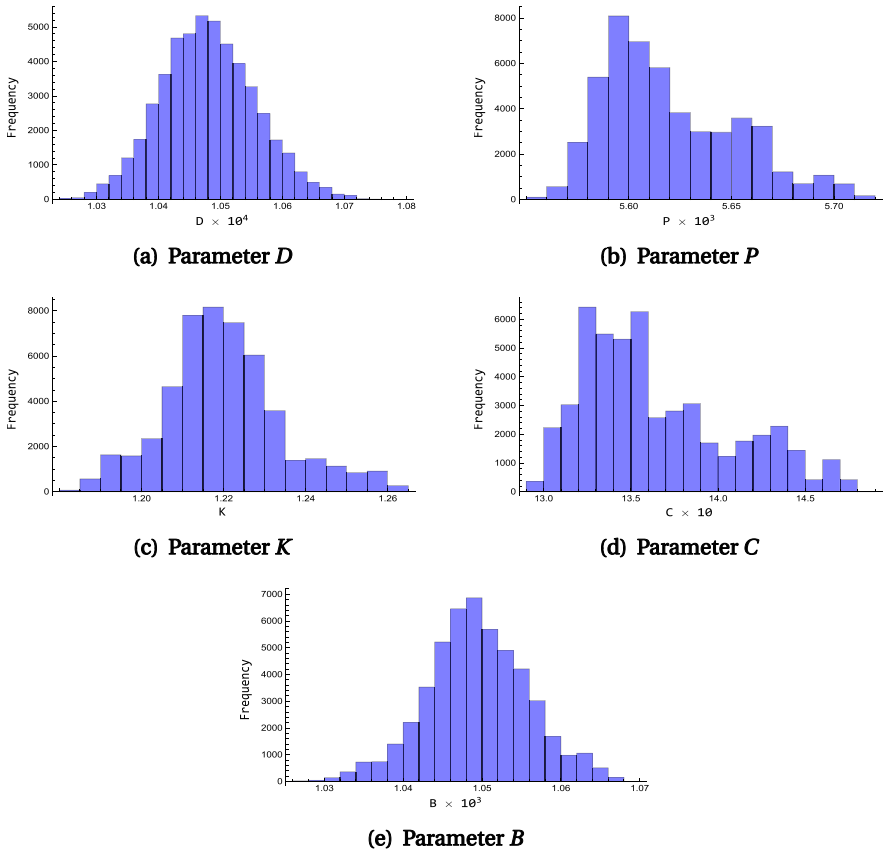
Fig. 4 Case 2—posterior distribution

numerically solved by reliable built-in routines provided by well-established computational platforms, such as the *NDSolve* routine within the Wolfram Mathematica environment.

The most computationally intensive task within this solution procedure is the calculation of Eqs. (17) and (18), since the integrals involved most probably cannot be solved analytically, and numerical techniques can become costly due to the oscillatory behavior of the eigenfunctions. Thus, the alternative semi-analytical integration procedure, such as proposed in Ref. [22], is employed. Assuming that t he function to be integrated is of the form $f(x, t, u)$, one can write the following approximation:

$$\int_{x_0}^{x_1} f(x, t, u)\psi_i(x)dx \approx \sum_{j=1}^{M} \int_{x_{j-1}}^{x_j} \widehat{f}_j(x, t, u)\psi_i(x)dx \tag{21}$$

where $\hat{f}_j(x, t, u)$ are simpler representations of the original function $f(x, t, u)$, defined in M subregions. In this work $\hat{f}$ are first order representations of f within each subregion, i.e. $\hat{f}_j$ is a linear approximation of f within $[x_{j-1}, x_j]$, for $j = 1, 2, \ldots, M$. Hence, the integrals on the RHS of Eq. (21) allows for analytical integration.

It should be remembered that this work is aimed at the corresponding inverse problem analysis within the Bayesian framework, exploring the posterior densities of the sought parameters

(a) Parameter D

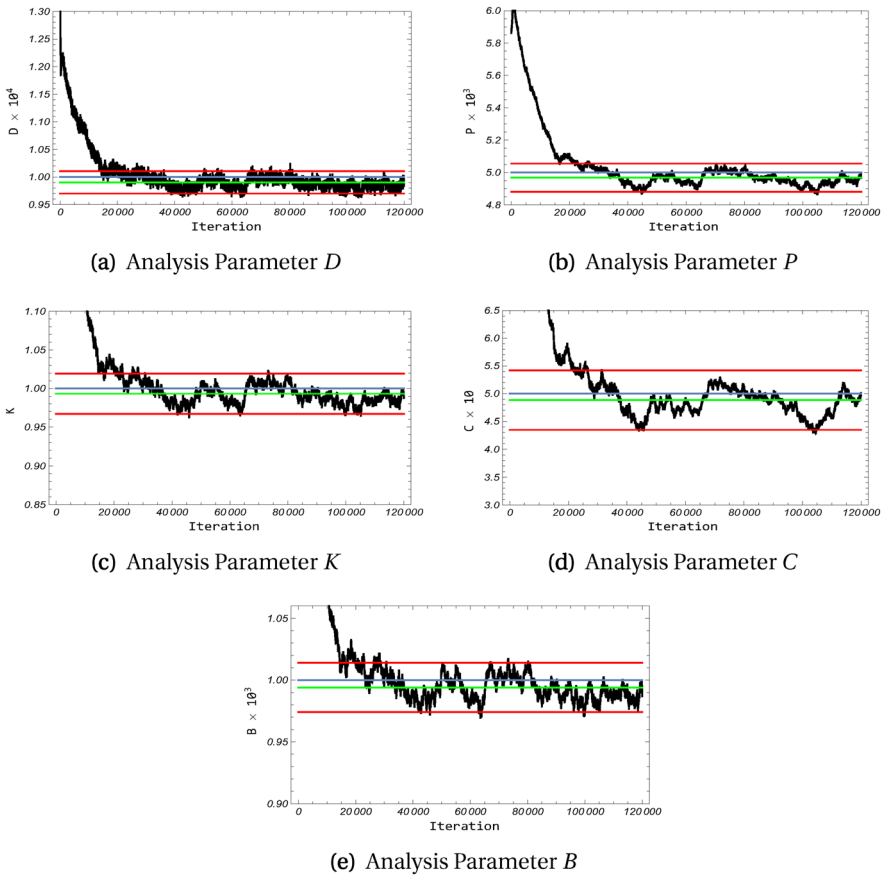(b) Parameter P

(c) Parameter K

(d) Parameter C

(e) Parameter B

Fig. 5 Case 3—Markov Chain for each parameter (black), Mean (green), 95% confidence interval (red) and exact value (blue)

with the Markov Chain Monte Carlo method, which may require thousands numerical simulations of the direct problem solution. So, even considering the semi-analytical integration proposed, the computational cost can still be very high. Based on the better convergence characteristics of the integrals of eigenfunction expansions [23], one of the contributions of the present work is the proposition of a low-cost solution of the direct problem, based on a reduced truncation order, $N_R \leq N$, of the expansions appearing in the integrals of the nonlinear terms, in Eqs. (17, 18), as follows:

$$\overline{g}_{i,1} = \int_0^L \frac{\widetilde{\psi}_i(x)e^{-\mu\tau}\sum_{j=0}^{N_R}\overline{u}_j(t-\tau)\widetilde{\psi}_j(x)}{K + C\sum_{j=0}^{N_R}\overline{u}_j(t-\tau)\widetilde{\psi}_j(x)}dx \tag{22}$$

$$\overline{g}_{i,2} = \int_0^L \frac{\widetilde{\psi}_i(x)P\left(\sum_{j=0}^{N_R}\overline{u}_j(t)\widetilde{\psi}_j(x)\right)^2}{K + C\sum_{j=0}^{N_R}\overline{u}_j(t-\tau)\widetilde{\psi}_j(x)}dx \tag{23}$$

This approach can yield a significant reduction of the computational effort associated with the numerical integrations, without considerable precision loss, as we shall demonstrate later.

(a) Parameter *D*

(b) Parameter *P*

(c) Parameter *K*

(d) Parameter *C*

(e) Parameter *B*
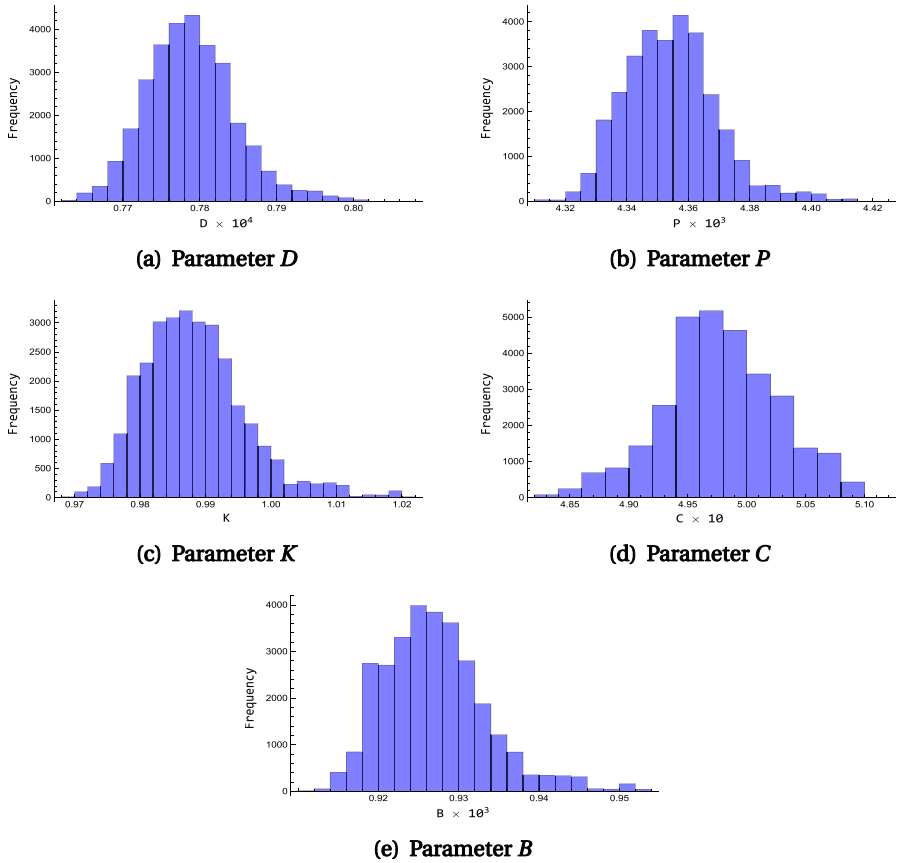
Fig. 6 Case 3—posterior distribution

Table 7 Case 4—Inverse problem solution (a) Statistics obtained after discarding the burn-in states. (b) Estimates with 95% confidence intervals obtained with MCMC method. CI width (%): ratio between the 95% confidence interval range and the estimated mean value

Parameters	Exact	Mean (μ)	Standard Deviation (σ)	Coefficient of Variation(σ/μ)%
$D \times 10^4$	1	0.990093	1.03×10^{-2}	4.09
$P \times 10^3$	5	4.96818	4.45×10^{-2}	3.51
K	1	0.993104	1.32×10^{-2}	5.24
$C \times 10$	5	4.88469	2.73×10^{-1}	21.98
$B \times 10^3$	1	0.994054	1.01×10^{-2}	4.01

Parameters	Initial	Confidence Interval	CI width(%)
$D \times 10^4$	1.315	[0.972988, 1.01346]	4.37
$P \times 10^3$	5.866	[4.88089, 5.05547]	3.86
K	1.568	[0.967087, 1.01912]	5.75
$C \times 10$	13.099	[4.34784, 5.42155]	24.23
$B \times 10^3$	1.2581	[0.974098, 1.01401]	4.32

(a) Analysis Parameter D

(b) Analysis Parameter P

(c) Analysis Parameter K

(d) Analysis Parameter C

(e) Analysis Parameter B

Fig. 7 Case 4—Markov Chain for each parameter (black), Mean (green), 95% confidence interval (red) and exact value (blue)

Inverse Problem Formulation and Solution Methodology

Consider known the rates of premature death, μ, and reproduction time delay, τ, for the population under investigation, such as discussed in Ref. [9]. Hence, in this work the inverse problem will be formulated as a problem of estimating the following parameters:

$$\mathbf{P} = \{D, P, K, C, B\}^T \tag{24}$$

where we have assumed that all the possible culling sites have the same culling intensity B.

Given a set of N_D experimental data, obtained from transient measurements at one or more distinct locations:

$$\mathbf{Y} = \left\{ Y_1, Y_2, \ldots, Y_{N_D} \right\}^T \tag{25}$$

In the Bayesian approach, the inverse problem is formulated as a problem of statistical inference based on the following principles [12]:

(i) The sought parameters are modeled as random variables;
(ii) The randomness describes our degree of information;

(a) Histogram Parameter D

(b) Histogram Parameter P

(c) Histogram Parameter K

(d) Histogram Parameter C

(e) Histogram Parameter B
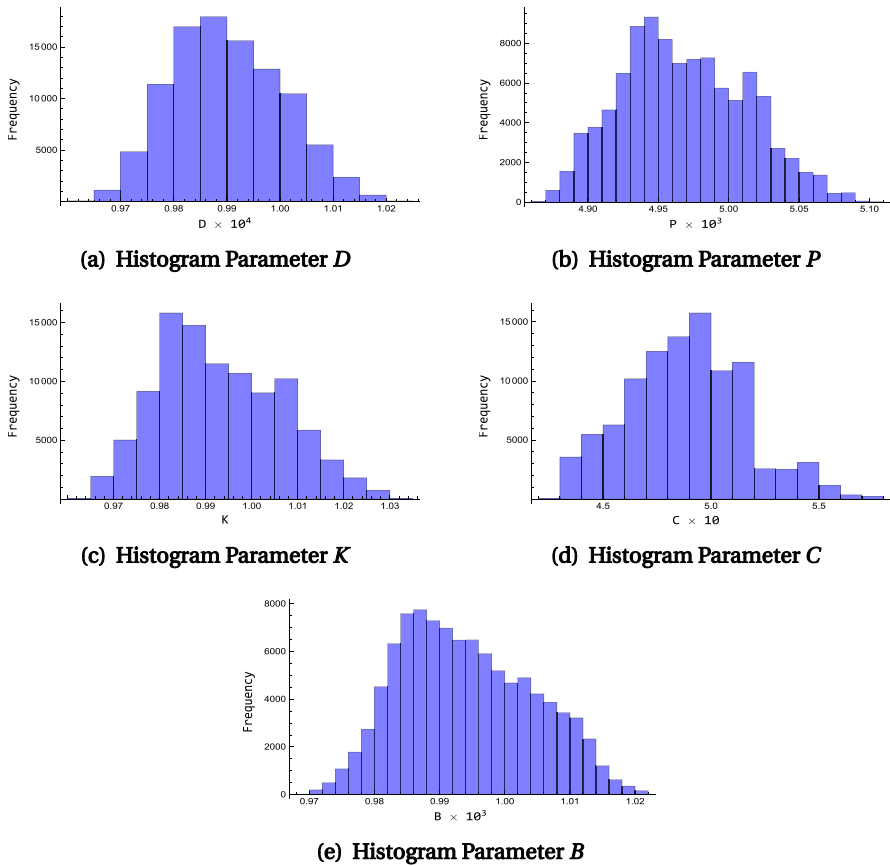
Fig. 8 Case 4—posterior distribution

(iii) The degree of information is coded into probability distributions;
(iv) The solution of the inverse problem is the posterior probability distribution.

Thus, in the Bayesian approach all possible information is incorporated into the model, as prior probability density functions, in order to reduce the amount of uncertainty present in the problem. Denoting the prior probability density function as $\pi\,(\mathbf{P})$, the Bayes' theorem formulation for inverse problems can be expressed as:

$$\pi\,(\mathbf{P}|\mathbf{Y}) = \frac{\pi\,(\mathbf{Y}|\mathbf{P})\,\pi\,(\mathbf{P})}{\pi\,(\mathbf{Y})} \qquad (26)$$

where $\pi\,(\mathbf{P}|\mathbf{Y})$ is the posterior probability distribution, $\pi\,(\mathbf{Y})$ is the marginal density and $\pi\,(\mathbf{Y}|\mathbf{P})$ is the likelihood function.

Note that this method can be implemented in various ways. In this work it is explored the Markov Chain Monte Carlo method (MCMC). While other methods result in a point estimate for the unknowns, the MCMC directly exploits the posterior probability distribution through its sampling strategy [12,24].

The implementation of the MCMC method with the Metropolis-Hastings algorithm can be performed using the following steps:

(a) Error Measurements in $x = 0.3$

(b) Error Measurements in $x = 0.4$

(c) Error Measurements in $x = 0.5$
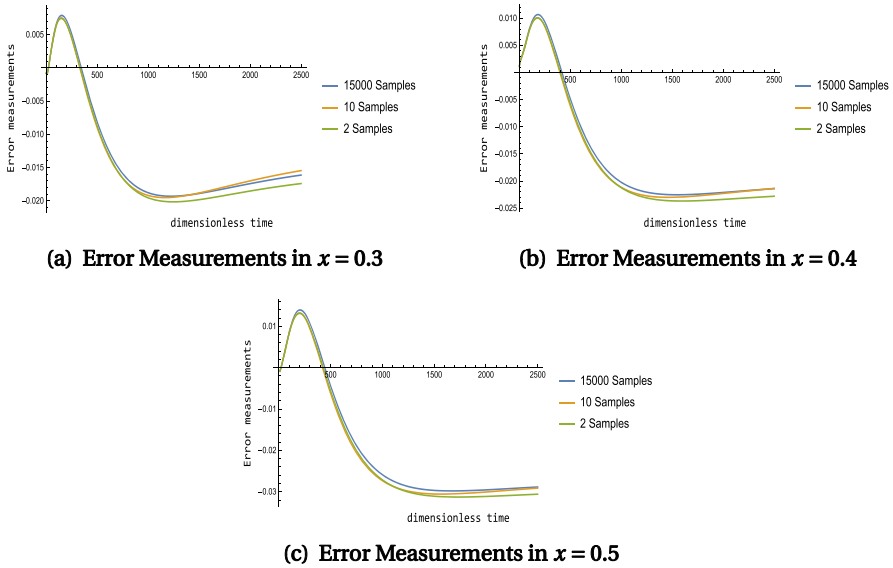
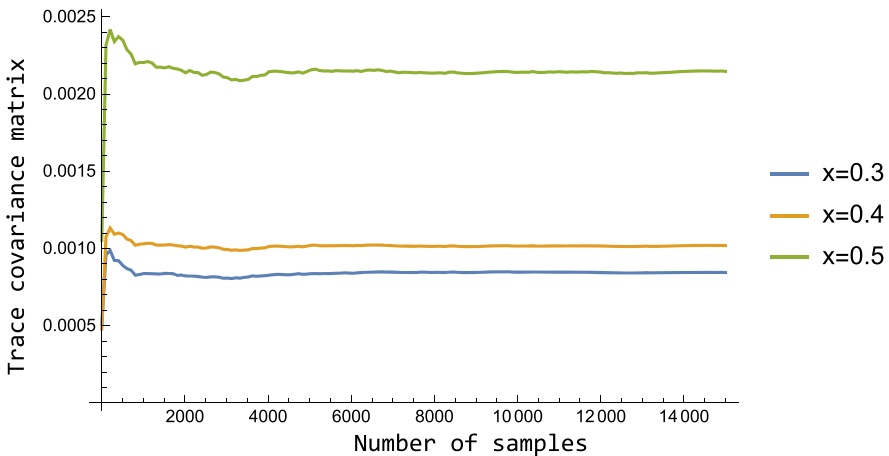**Fig. 9** Convergence of the mean of the approximation error

Fig. 10 Convergence trace of the covariance matrix

1. Initialize the chain iterations counter $i = 0$, and choose a starting value $\mathbf{P}^0$.
2. Generate a candidate value $\mathbf{P}'$ from an arbitrary auxiliary distribution $q(\mathbf{P}'|\mathbf{P}^i)$. In this work, a random walk implemented with a uniform distribution is considered. Hence:

$$q(\mathbf{P}'|\mathbf{P}^i) = U(\mathbf{P}^i - \delta, \mathbf{P}^i + \delta) \tag{27}$$

where U represents a uniform distribution with support $[\mathbf{P}^i - \delta, \mathbf{P}^i + \delta]$.

3. Calculate the following acceptance probability α, for the chosen candidate:

$$\alpha = \min \left[1, \frac{\pi(\mathbf{P}'|\mathbf{Y}) q\left(\mathbf{P}^i|\mathbf{P}'\right)}{\pi\left(\mathbf{P}^i|\mathbf{Y}\right) q\left(\mathbf{P}'|\mathbf{P}^i\right)} \right] \tag{28}$$

4. Following the Metropolis-Hastings algorithm, draw a random number Λ from $\Lambda \sim U(0, 1)$, where U is the uniform distribution with support [0,1].
5. *If* $\Lambda \leq \alpha$ *then* accept the new value and make $\mathbf{P}^{i+1} = \mathbf{P}'$.
6. *Else* reject the candidate, and make $\mathbf{P}^{i+1} = \mathbf{P}^i$.
7. Increase the counter i to $i + 1$, and go back to step 2.

This procedure is carried out to yield a Markov chain with N_{MCMC} states, in the form $\left\{\mathbf{P}^1, \mathbf{P}^2, \ldots, \mathbf{P}^{N_{MCMC}}\right\}$. This sequence of generated samples is a representation of the posterior distribution. Hence, inference on the sought posterior distribution is obtained from inference on the generated samples. For inference purpose, it is worth noting that the initial states of the Markov chain must be ignored while the chain has not converged to the equilibrium. These ignored states are commonly referred to as the burn-in period.

Approximation Error Model

While the solution of the inverse problem considering only Gaussian additive errors is given by:

$$\mathbf{Y} = U_C(\mathbf{P}) + \mathbf{e} \tag{29}$$

where U_C is considered the solution of the direct problem, taken as the solution that best represents the nature of the problem and $\mathbf{e}$ is a vector containing measurement errors, which in most cases are well represented by a normal distribution with zero mean and standard deviation σ_e.

On the other hand, in the Approximation Error Model approach (AEM), the modeling error is treated as an additional noise to the conventional error model. The approximation error modeling can be written as [25].

$$\mathbf{Y} = U_R(\mathbf{P}) + [U_C(\mathbf{P}) - U_R(\mathbf{P})] + \mathbf{e} \tag{30}$$

where $U_R(P)$ is the solution of the problem considering the reduced model, and one can define:

$$\epsilon(\mathbf{P}) = U_C(\mathbf{P}) - U_R(\mathbf{P}) \tag{31}$$

The approach of modeling error can be written as:

$$\mathbf{Y} = U_R(\mathbf{P}) + \eta(\mathbf{P}) \tag{32}$$

with $\eta(\mathbf{P}) = \epsilon(\mathbf{P}) + \mathbf{e}$.

The posterior distribution analysis procedure is the same as the conventional error model, with only a small change in the likelihood composition.

$$\pi(\mathbf{Y}|\mathbf{P}) = 2\pi^{-\frac{N_d}{2}} |\mathbf{W}_*|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\mathbf{R}_*^T \mathbf{W}_*^{-1} \mathbf{R}_* \right] \tag{33}$$

where $\mathbf{W}_*$ is the covariance matrix and $\mathbf{R}_*$ is the residual vector, given by:

$$\mathbf{R}_* = \mathbf{Y} - U_R(\mathbf{P}) - \epsilon_*(\mathbf{P}) \tag{34}$$

where $\epsilon_*(\mathbf{P})$ is the mean value of $\epsilon(\mathbf{P})$. To obtain $\mathbf{W}_*$ and $\epsilon_*(\mathbf{P})$ it was used:

1. N_S samples are generated for the parameter vector **P**.
2. U_R and U_C solutions are calculated for each set **P**.
3. N_S samples are calculated for the modeling error $\epsilon(\mathbf{P})$.
4. The mean and covariance matrix are calculated for $\epsilon(\mathbf{P})$

Results and Discussion

Some numerical results are now reported, first for the direct problem solution, considering the solution methodology described in "Direct Problem Formulation and Solution Methodology" section, and next for the inverse problem solution, considering the methodology described in "Inverse Problem Formulation and Solution Methodology" section. For the numerical examples, we consider the parameter values presented in Table 1, which are the same employed in Refs. [7,9], considering four culling sites, i.e. $N_{CS} = 4$ in Eq. (1), with equal culling intensities, $B_1, \ldots, B_4 = B$, located at $x_1 = 0.2$, $x_2 = 0.4$, $x_3 = 0.6$ and $x_4 = 0.8$. For the solution of the direct and inverse problems, the computational implementations were run with a computer equipped with CPU Intel Core i5-6200U, 2.3 GHz processor, with 7.85 GB RAM.

The convergence of the direct problem solution was analyzed for the order of the semi analytical integration (M), see Eq. (21) and the truncation order of the eigenfunctions expansions (N), considering both the complete solution, in which $N_R = N$ in Eqs. (22) and (23), and the approximated solution, with $N_R < N$, aiming at the computational cost reduction. Besides investigating the convergence behavior of the solutions, it was also analyzed the computational time required. Comparing Table 2a, b, one may observe a very good performance of the semi-analytical integration procedure, achieving convergence of two to three significant digits for $M = 30$ in comparison to the more refined case, with $M = 150$, while yielding considerable lower CPU times. Furthermore, the results presented in Table 2 demonstrate good convergence behaviors regarding also the truncation order of the eigenfunctions expansions (N), for both the complete and the approximated solutions, with a remarkable CPU time reduction observed in the approximated solution, without significant accuracy loss, keeping the results converged with two to three significant digits. It is worth noting that the less refined solution presented in Table 2, with $N = 20$, $N_R = 1$ and $M = 30$, was achieved with only 0.36 s and keeps the relative error in the order of 3% in comparison with the most refined case presented, with $N = 100$, $M = 150$ and $N_R = N$, which took 673.35 s to compute. Similar results are presented in Table 3 for the time instant ($t = 125$) with even better results observed for the approximated solution, demonstrating the good agreement with the more refined solutions also for increasing time instants.

For the inverse problem solution, the synthetic experimental data were simulated in the following form:

$$Y_i = U_i(\mathbf{P}_{\text{exact}}) + e_i, \quad e_i \sim N(0, 0.004) \tag{35}$$

where U_i represents the direct problem solution at a given location x and time instant t, using the exact values for the parameters **P**. In order to investigate the influence of the approximated solution into the estimations, two sets of experimental data were considered. The first set was generated considering the approximated solution with $N = 20$ and $N_R = 1$, and the second set was generated using the most refined solution considered, with $N = N_R = 150$. Although the most appropriate test scenario is the use of the complete solution for simulating the experimental data, in order to avoid the so called inverse crime [12], we also tested the

Table 8 Case 5—Inverse problem solution (a) Statistics obtained after discarding the burn-in states. (b) Estimates with 95% confidence intervals obtained with MCMC method. CI width (%): ratio between the 95% confidence interval range and the estimated mean value

Parameters	Exact	Mean (μ)	Standard Deviation (σ)	Coefficient of Variation(σ/μ)%
$D \times 10^4$	1	0.890586	1.81×10^{-2}	4.09
$P \times 10^3$	5	4.80234	6.81×10^{-2}	3.51
K	1	0.940803	1.84×10^{-2}	5.24
$C \times 10$	5	0.317024	1.86×10^{-1}	21.98
$B \times 10^3$	1	1.05132	2.11×10^{-2}	4.01

Parameters		Initial	Confidence Interval	CI width(%)
$D \times 10^4$		1.03073	[0.8552, 0.9260]	7.94
$P \times 10^3$		5.70019	[4.6687, 4.9359]	5.56
K		1.17517	[0.9046, 0.9770]	7.70
$C \times 10$		5.25067	[−0.0491, 0.6832]	231
$B \times 10^3$		1.20620	[1.0099, 1.0928]	7.89

Table 9 Case 6—Inverse problem solution (a) Statistics obtained after discarding the burn-in states. (b) Estimates with 95% confidence intervals obtained with MCMC method. CI width (%): ratio between the 95% confidence interval range and the estimated mean value

Parameters	Exact	Mean(μ)	Standard Deviation (σ)	Coefficient of Variation(σ/μ)%
$D \times 10^4$	1	1.00561	1.65566×10^{-2}	1.64643
$P \times 10^3$	5	4.98208	4.47747×10^{-2}	0.89871
K	1	0.99157	2.12822×10^{-2}	2.14631
$C \times 10$	5	4.81839	2.22816×10^{-1}	4.6243
$B \times 10^3$	1	0.99108	1.58183×10^{-2}	1.59607

| Parameters		Initial	Confidence Interval	CI width(%)
$D \times 10^4$		1.08432	[0.973156, 1.03806]	6.4539
$P \times 10^3$		5.27353	[4.89433, 5.06984]	3.5229
K		1.10727	[0.94986, 1.03328]	8.41337
$C \times 10$		4.94680	[4.38167, 5.2551]	18.1269
$B \times 10^3$		1.08386	[0.960073, 1.02208]	6.2565

case involving the inverse crime, i.e., the model employed in simulating the experimental data is the same employed in the inverse problem solution, aiming at evaluating how the quality of estimates vary when considering the given approximation. In Ref. [9] a thorough sensitivity analysis was performed for this problem, which showed that all parameters but C could be simultaneously estimated without the use of any prior information. Hence, in this work, we considered the use of prior information for this parameter, whereas for the others no prior information is adopted. The prior information for C was modeled by a normal distribution with mean in the exact value for the parameter C, and standard deviation of 5% of the mean value. Nonetheless, the inverse problem solution without prior information for this parameter
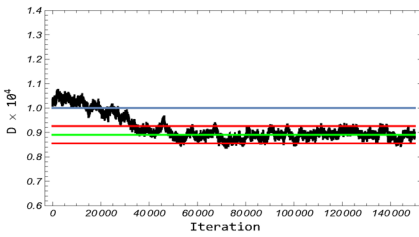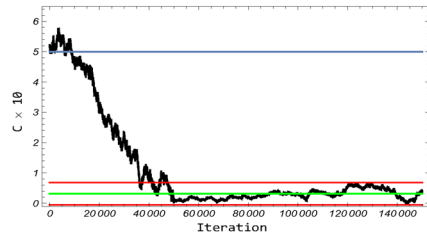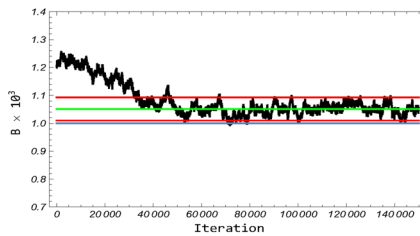
(a) Analysis Parameter *D*
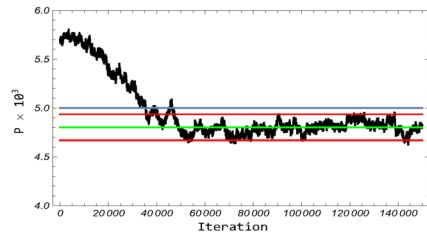
(b) Analysis Parameter *P*
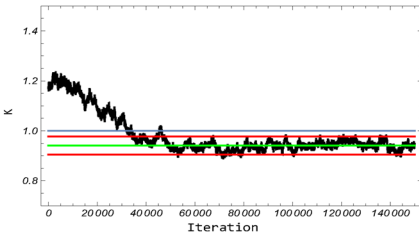
(c) Analysis Parameter *K*

(d) Analysis Parameter *C*

(e) Analysis Parameter *B*

Fig. 11 Case 5—Markov Chain for each parameter (black), Mean (green), 95% confidence interval (red) and exact value (blue)

was also tried here, in order to revisit the conclusions achieved in Ref. [9], now within the Bayesian framework. In summary, we can separate the test cases considering: (i) nature of the simulated experimental data; and (ii) use of a priori information for the parameter C. In this sense, the test cases considered are shown in Table 4. For all cases considered, the direct problem solution within the inverse problem procedure was obtained using the approximated solution with $N = 20$, $M = 30$ sub-regions in the semi-analytical integration, (21) and approximated solution $N_R = 1$, (23).

In Case 1 the Markov chain was constructed with a total of 120,000 states, being the first 80,000 states discarded as the burn-in. Table 5 summarizes the calculated statistics quantities. One may observe that the absence of prior associated with lack of accuracy of the solution in reproducing the experimental data (the experimental data were simulated with the complete solution in this case) leads to poor estimates. Figure 1 shows that the Markov chains converge to incorrect values, and it should be noticed that the parameter C was estimated to a meaningless value, in which the confidence interval encompasses the zero value, suggesting an overparametrized model. This result is in agreement with the sensitivity
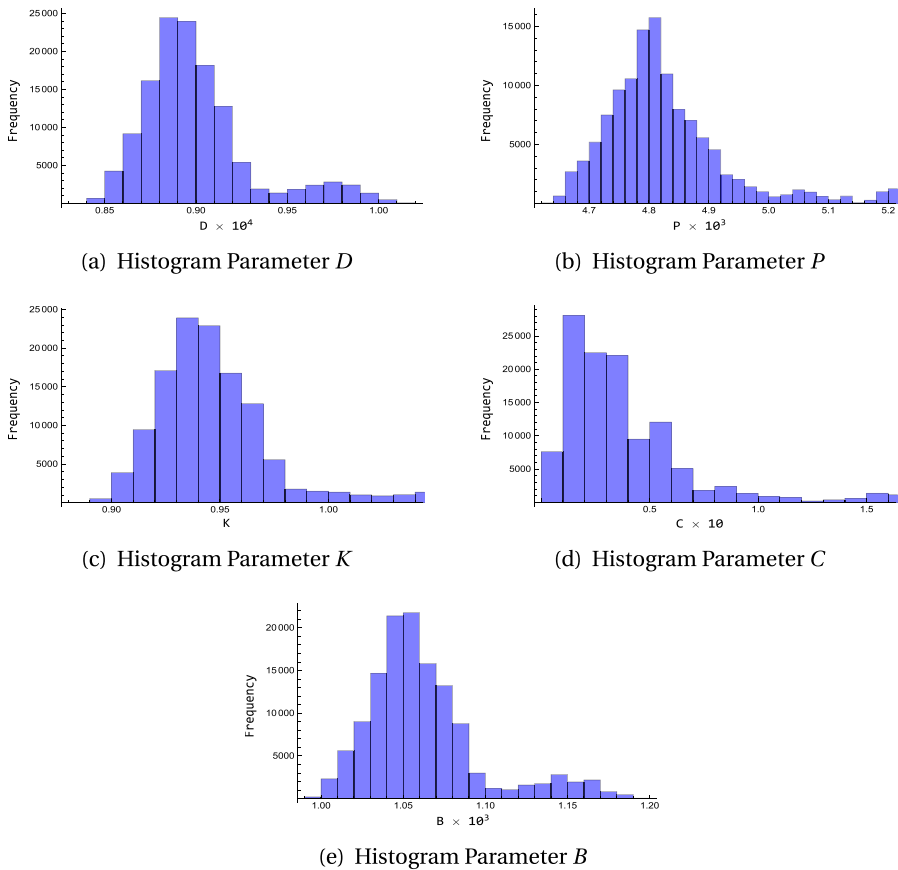
(a) Histogram Parameter *D*

(b) Histogram Parameter *P*

(c) Histogram Parameter *K*

(d) Histogram Parameter *C*

(e) Histogram Parameter *B*

Fig. 12 Case 5—posterior distribution

analysis performed in Ref. [9]. Figure 2 consolidates this analysis by presenting the histogram of the posterior distribution of the sought parameters.

In Case 2 it was considered 70,000 states for the Markov chains, and the first 20,000 states were discarded as the burn-in. The results presented in Table 6 show that the estimated confidence intervals are close to the expected values with low levels of uncertainty, as concluded from its confidence interval relatively narrow width. It can be seen that due to the inverse crime (in this case the experimental data were simulated with the approximated solution) the convergence and quality of the estimates improved, as shown Figs. 3 and 4, even without prior information for the parameter *C*. Nonetheless, the estimated confidence interval for *C* does not encompass the exact value, once again confirming the difficulties associated with the estimation of this parameter, as also demonstrated with the sensitivity analysis performed in Ref. [9].

In Case 3, The Markov chains were constructed with 50,000 states. These results illustrate that when prior information is considered available for the parameter *C*, much better estimations are achieved. Nonetheless, the results obtained still lead to some estimations that are relatively far from the expected values for some parameters, even if good convergence of the Markov chain is observed for all parameters, as shown in Figs. 5 and 6. For example,
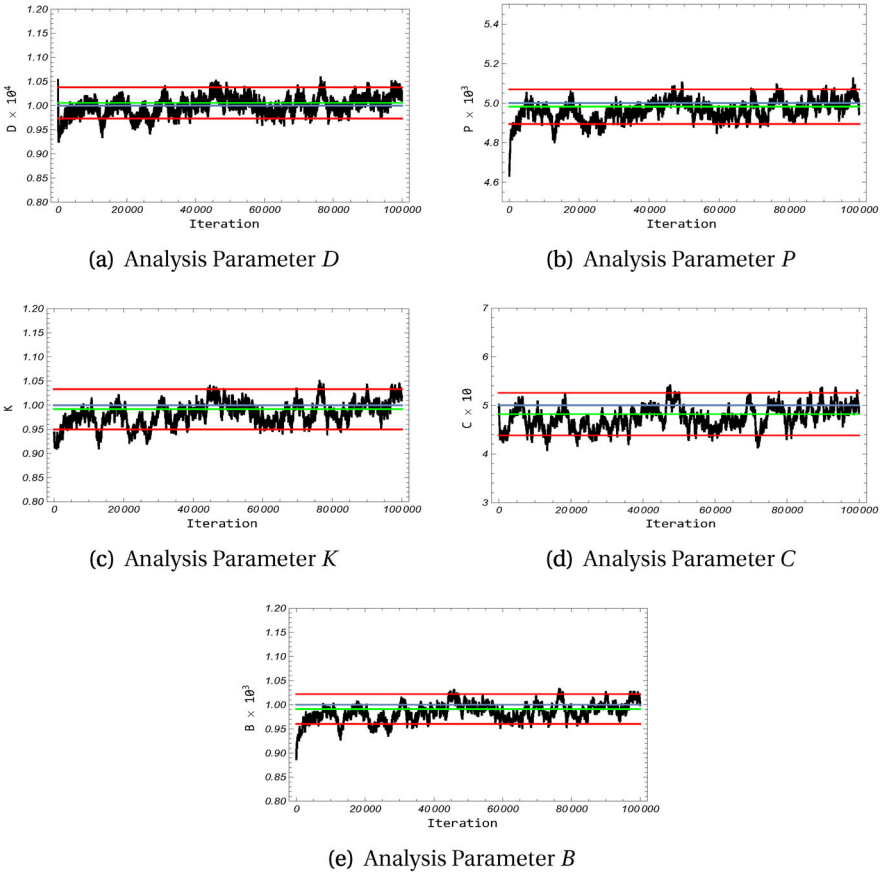
Int. J. Appl. Comput. Math (2021) 7:189

Page 21 of 25   **189**

(a) Analysis Parameter *D*

(b) Analysis Parameter *P*

(c) Analysis Parameter *K*

(d) Analysis Parameter *C*

(e) Analysis Parameter *B*

Fig. 13 Case 6—Markov Chain for each parameter (black), Mean (green), 95% confidence interval (red) and exact value (blue)

one may observe that the estimated confidence intervals for the parameters D, P and B do not encompass the exact values of such parameters. This is probably explained by the errors associated with the approximated solution. To demonstrate this, we finally consider in Case 4 the case with inverse crime and prior information available for the parameter C. In Case 4 the Markov chain was considered with 120,000 states, being the 20,000 first states discarded as the burn-in. Table 7 summarizes the results obtained. One may observe that for all parameters, the exact value lies inside the estimated 95% confidence intervals, demonstrating that the errors associated with the approximated solution, even if small, can have significant influence on the estimated parameters. In Fig. 7 it is shown that all the chains present fast convergence and the posterior density distribution, presented in Fig. 8, approaches a normal distribution.

Before proceeding to the inverse problem solution with the AEM approach, some considerations must be made about the characterization of the modeling error ϵ. To generate the samples, the average of the parameters obtained in Case 4 were disturbed by a random noise with zero mean and standard deviation of 5% of the mean, each of these parameters was
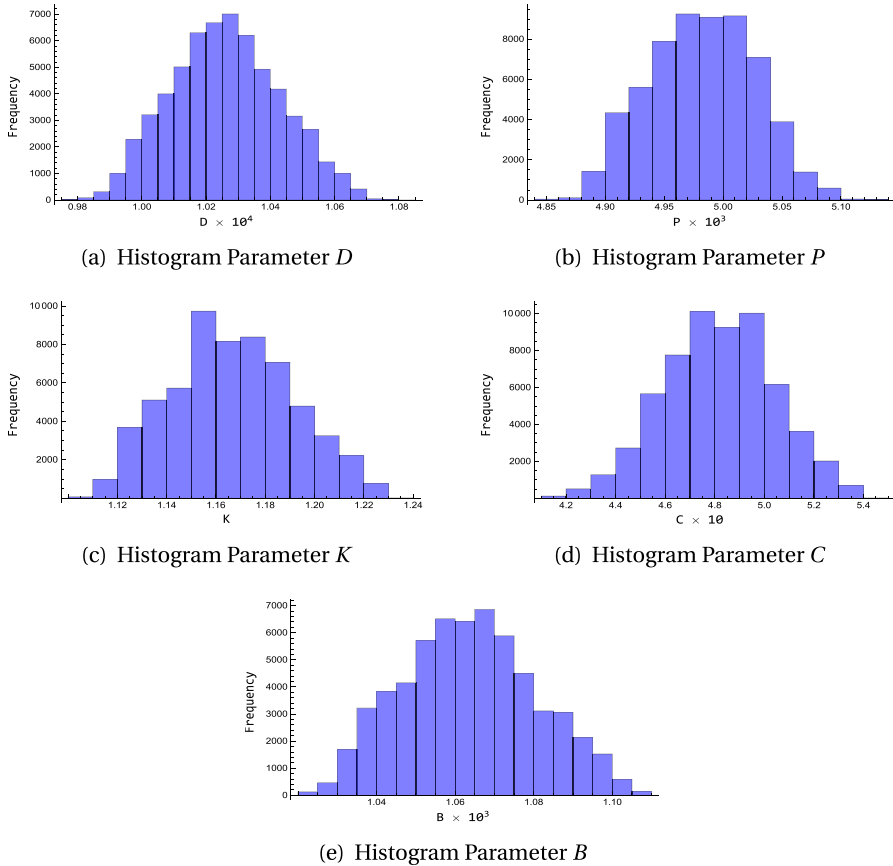
(a) Histogram Parameter D

(b) Histogram Parameter P

(c) Histogram Parameter K

(d) Histogram Parameter C

(e) Histogram Parameter B

Fig. 14 Case 6—posterior distribution

used to generate a measure of the approximation error. For the calculation of the mean and the covariance matrix, 15,000 samples were generated, which proved to be sufficient when considering the convergence of the mean and the trace of the covariance matrix, as can be seen in the Figs. 9 and 10.

In Case 5, no prior information was used for the parameter C, and the Markov chains were constructed with 150,000 states. Convergence was achieved and all parameters were estimated with reasonable values, near the exact ones, with exception of parameter C, as expected due to the sensitivity issues already discussed here. These results are summarized in Table 8 and Figs. (11) and (12).

To solve the inverse problem in Case 6, prior information was used for parameter C, with chains of 100,000 states using 30,000 burn-in states for all parameters except for parameter P which used 40,000, as indicated by the convergence test from Heidel [26]. After removing the burn-in states, the statistical quantities were calculated and presented in the Table 9. In Fig. 13 it is shown that all the chains converge considering few states and the posterior density distribution, presented in Fig. 14, follows a normal distribution. One may observe that the use of the low cost solution with the AEM approach yielded very good estimations for all parameters considered, within relatively small and reliable confidence intervals. This result
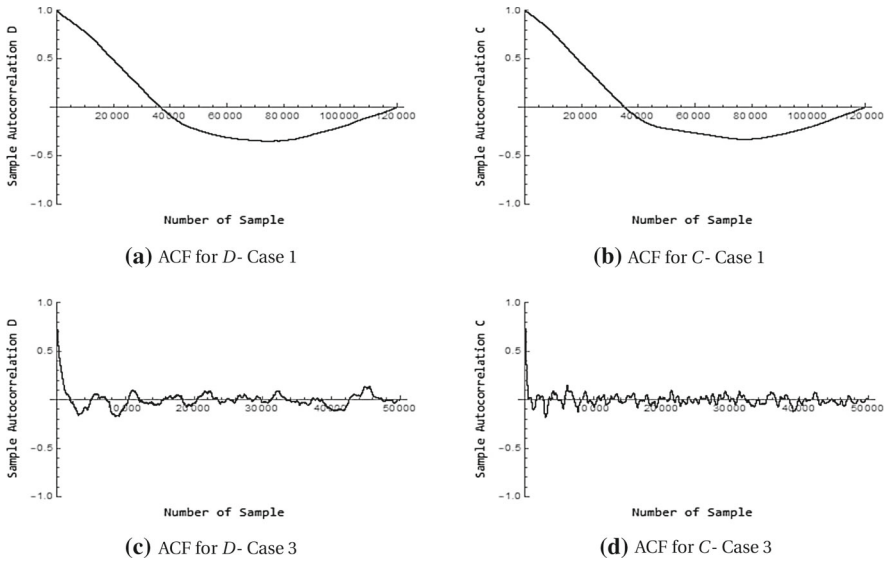
(**a**) ACF for *D*- Case 1　　　　　　　　　　(**b**) ACF for *C*- Case 1

(**c**) ACF for *D*- Case 3　　　　　　　　　　(**d**) ACF for *C*- Case 3

Fig. 15 Autocorrelation functions of the Markov chains regarding the parameters C and D in Cases 1 and 3

(**a**) ACF for *D*- Case 5　　　　　　　　　　(**b**) ACF for *C*- Case 5

(**c**) ACF for *D*- Case 6　　　　　　　　　　(**d**) ACF for *C*- Case 6
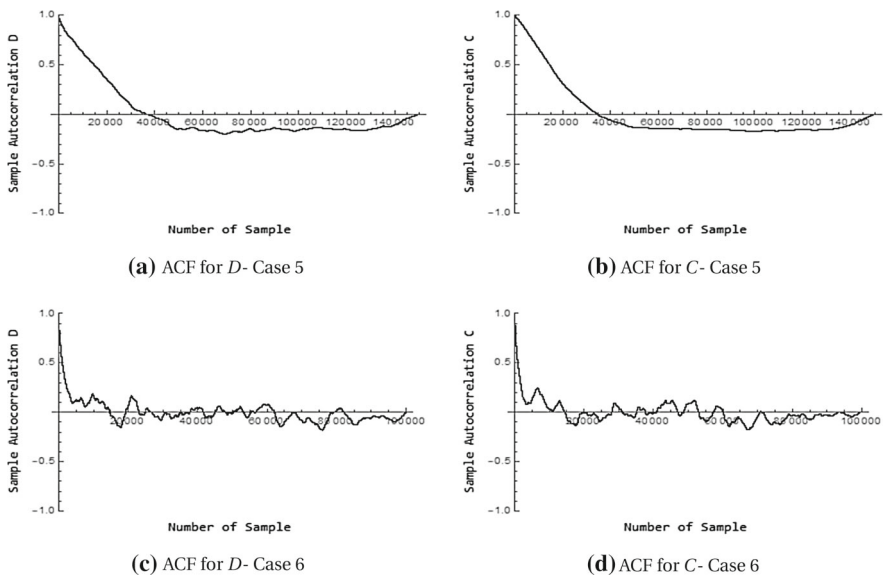
Fig. 16 Autocorrelation functions of the Markov chains regarding the parameters C and D in Cases 5 and 6

confirms the importance of taking into account the error model when employing approximated solutions or surrogate models, especially in nonlinear models such as the case of most population dynamics formulations.

In order to better observe the influence of the prior information and the AEM approach in the Markov chains, Figs. 15 and 16 present the autocorrelation functions regarding the Markov chains obtained for the parameters with lower sensitivity (parameters C and D)

[9], for Cases 1 and 3 (in Fig. 15) and Cases 5 and 6 (in Fig. 16). The autocorrelation functions show that considering prior information for C leads to a significant improvement in the convergence behavior, with a remarkable decrease in the number of states needed to obtain an independent sample, for both parameters here analyzed, C and D, even with prior information only adopted for C.

Concluding Remarks

This work addressed the direct and inverse problem formulation and solution to tackle the diffusive population problem with impulsive culling sites. The direct problem was formulated using the reaction-diffusion equation with logistic birth and mortality rates and impulsive culling. The solution of this problem followed the formalism of the Generalized Integral Transform Technique (GITT) employing a computationally low cost solution, since the inverse methodology would require several calculations of the direct problem. The inverse problem was formulated within the Bayesian framework and the Markov Chain Monte Carlo method, implemented with the Metropolis-Hastings algorithm, which was employed to sample the posterior probability density function. The approximate solution of the direct problem was used to reduce the computational cost, but the error due to this type of approximation generated estimates and/or confidence regions incompatible with the values used in the simulation. To handle this issue, the Approximation Error Model (AEM) approach was employed, leading to reliable results, as presented in Case 6. This work also confirmed the sensitivity analysis performed in Ref. [16], now within the Bayesian framework, demonstrating that one of the parameters appearing in the logistic-like growth rate model employed in this formulation yields low sensitivity and is unlikely to be estimated without the use of prior information. It should be highlighted that the methodology here developed, combining a low cost numerical-analytical solution with the Approximation Error Model approach, can be efficiently employed for the solution of inverse problems within the Bayesian framework, despite of the sampling methodology considered, including newly developed MCMC methods.

Declarations

Competing interest The authors declare that they have no competing interests.

References

1. Malthus, T.R.: An essay on the principle of population (Printed for J. Johnson, in St. Pauls Church-Yard, London) (1798)
2. Verhulst, P.F.: Notice surlaloique la population poursuitdans son accroissement. Correspondance (1838)
3. Skellam, J.G.: Random dispersal in theoretical populations. Biometrika **38**(1/2), 196–218 (1951)
4. Okubo, A.: Diffusion and Ecological Problems: Mathematical Models. Biomathematics Texts. Springer, Berlin (1980)
5. Okubo, A., Levin, S.A.: Diffusion and Ecological Problems: Modern Perspectives, vol. 14. Springer, Berlin (2001)
6. Simons, R.R.L., Gourley, S.A.: Extinction criteria in stage-structured population models with impulsive culling. SIAM J. Appl. Math. **66**(6), 1853–1870 (2006)
7. White, R.E.: Populations with impulsive culling: control and identification. Int. J. Comput. Math. **86**(12), 2143–2164 (2009)
8. Mocenni, C., Madeo, D., Sparacino, E.: Linear least squares parameter estimation of nonlinear reaction diffusion equations. Math. Comput. Simul. **81**(10), 2244–2257 (2011)
9. Knupp, D.C., Sacco, W.F., Silva Neto, A.J.: Direct and inverse analysis of diffusive logistic population evolution with time delay and impulsive culling via integral transforms and hybrid optimization. Appl. Math. Comput. **250**, 105–120 (2015)
10. Anguelov, R., Dufourd, C., Dumont, Y.: Simulations and parameter estimation of a trap-insect model using a finite element approach. Math. Comput. Simul. **133**, 47–75 (2017)
11. Moura Neto, F.D., Silva Neto, A.J.: An Introduction to Inverse Problems with Applications. Springer, Berlin (2013)
12. Kaipio, J.P., Somersalo, E.: Computational and statistical methods for inverse problems. Appl. Math. Sci. **160**, 2004 (2004)
13. Teixeira, J.S., Stutz, L.T., Knupp, D.C., Silva Neto, A.J.: Structural damage identification via time domain response and Markov chain monte Carlo method. Inverse Probl. Sci. Eng. **25**(6), 909–935 (2017)
14. Mirsian, S., Khodadadian, A., Hedayati, M., Ajdad, A.M., Kalantarinejad, R., Heitzinger, C.: A new method for selective functionalization of silicon nanowire sensors and Bayesian inversion for its parameters. Biosens. Bioelectron. **142**, 111527 (2019)
15. Khodadadian, A., Noii, N., Parvizi, M., Abbaszadeh, M., Wick, T., Heitzinger, C.: A Bayesian estimation method for variational phase-field fracture problems. Comput. Mech. **66**(4), 827–849 (2020)
16. Khodadadian, A., Stadlbauer, B., Heitzinger, C.: Bayesian inversion for nanowire field-effect sensors. J. Comput. Electron. **19**(1), 147–159 (2020)
17. Noii, N., Khodadadian, A., Wick, T.: Bayesian inversion for anisotropic hydraulic phase-field fracture. arXiv preprint arXiv:2007.16038 (2020)
18. Noii, N., Khodadadian, A., Ulloa, J., Aldakheel, F., Wick, T., Francois, S., Wriggers, P.: Bayesian inversion for unified ductile phase-field fracture. arXiv preprint arXiv:2104.11114 (2021)
19. Zhao, H., Chen, B., Li, S., Li, Z., Zhu, C.: Updating the models and uncertainty of mechanical parameters for rock tunnels using Bayesian inference. Geosci. Front. **12**(5), 101198 (2021)
20. Lamiena, B., Orlande, H.R.B.: Approximation error model to account for convective effects in liquids characterized by the line heat source probe. In Proceedings of the 4th Inverse Problems, Design and Optimization Symposium, Albi, France (2013)
21. Cotta, R.M.: Integral Transforms in Computational Heat and Fluid Flow. CRC Press, Boca Raton (1993)
22. Sphaier, L.A., Cotta, R.M., Naveira Cotta, C., Quaresma, J.N.N.: The unit algorithm for solving one-dimensional convection–diffusion problems via integral transforms. Int. Commun. Heat Mass Transf. **38**(5), 565–571 (2011)
23. Cotta, R.M., Knupp, D.C., Naveira Cotta, C.P.: Analytical Heat and Fluid Flow in Microchannels and Microsystems. Springer, Berlin (2016)
24. Gamerman, D., Lopes, H.F.: Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference. CRC Press, Boca Raton (2006)
25. Nissinen, A., Heikkinen, L.M., Kaipio, J.P.: The bayesian approximation error approach for electrical impedance tomography-experimental results. Meas. Sci. Technol. **19**(1), 015501 (2007)
26. Heidelberger, P., Welch, P.D.: A spectral method for confidence interval generation and run length control in simulations. Commun. ACM **24**(4), 233–245 (1981)