Journal of Environmental Radioactivity 162-163 (2016) 328–339

Contents lists available at ScienceDirect

Journal of Environmental Radioactivity

journal homepage: www.elsevier.com/locate/jenvrad

Identifiability of sorption parameters in stirred flow-through reactor experiments and their identification with a Bayesian approach

V. Nicoulaud-Gouin [a, *], L. Garcia-Sanchez [b], M. Giacalone [a], J.C. Attard [a], A. Martin-Garin [b], F.Y. Bois [c]

[a] Radionuclide Transfers in the Environment Research Laboratory (LRTE), IRSN, centre de Cadarache, bât. 159, BP 3, 13115, Saint-Paul-lez-Durance, France
[b] Laboratory of Biogeochemistry, Bioavailability and Transfers of Radionuclides (L2BT), IRSN, centre de Cadarache, bât. 183, BP 3, 13115, Saint-Paul-lez-Durance, France
[c] Models for Ecotoxicology and Toxicology Unit (METO), INERIS, Parc ALATA, BP 2, 60550, Verneuil-en-Halatte, France

ARTICLE INFO

Article history:
Received 1 April 2015
Received in revised form
7 June 2016
Accepted 8 June 2016

Keywords:
Bayesian inference
Sorption parameters
Identifiability
MCMC
Convergence monitoring
Equilibrium kinetic model

ABSTRACT

This paper addresses the methodological conditions —particularly experimental design and statistical inference— ensuring the identifiability of sorption parameters from breakthrough curves measured during stirred flow-through reactor experiments also known as continuous flow stirred-tank reactor (CSTR) experiments. The equilibrium-kinetic (EK) sorption model was selected as nonequilibrium parameterization embedding the K_d approach. Parameter identifiability was studied formally on the equations governing outlet concentrations. It was also studied numerically on 6 simulated CSTR experiments on a soil with known equilibrium-kinetic sorption parameters. EK sorption parameters can not be identified from a single breakthrough curve of a CSTR experiment, because $K_{d,1}$ and k^- were diagnosed collinear. For pairs of CSTR experiments, Bayesian inference allowed to select the correct models of sorption and error among sorption alternatives. Bayesian inference was conducted with SAMCAT software (Sensitivity Analysis and Markov Chain simulations Applied to Transfer models) which launched the simulations through the embedded simulation engine GNU-MCSim, and automated their configuration and post-processing. Experimental designs consisting in varying flow rates between experiments reaching equilibrium at contamination stage were found optimal, because they simultaneously gave accurate sorption parameters and predictions. Bayesian results were comparable to maximum likehood method but they avoided convergence problems, the marginal likelihood allowed to compare all models, and credible interval gave directly the uncertainty of sorption parameters θ. Although these findings are limited to the specific conditions studied here, in particular the considered sorption model, the chosen parameter values and error structure, they help in the conception and analysis of future CSTR experiments with radionuclides whose kinetic behaviour is suspected.

1. Introduction

Solute retention on solid phases such as soils —termed sorption (Sposito, 1984) — strongly governs the mobility and bioavailability of substances in the environment, and, in the case of radionuclides, their radiological impact.

In assessment level models, a very common approach is to represent sorption by a (i) kinetically fast, (ii) concentration independent and (iii) reversible reaction, with an equilibrium coefficient also called distribution coefficient and noted K_d (IAEA, 2010). Deviations from these assumptions in flowing conditions can have an influence on predictions such as leachable contamination and cleanup times (Bahr and Rubin, 1987). Two-site models offer an alternative parameterization of sorption that covers a wider range of behaviours. Their approach relaxes the equilibrium hypothesis (i) by decomposing sorption sites into two categories that describe the “fast“ and “slow“ sorption kinetics. Governing equations vary with authors (e.g. Cameron and Klute, 1977; Choi et al., 2007; van Genuchten et al., 1974; van Genuchten and Wagenet, 1989), but generally encompass the classical K_d approach. The equilibrium kinetic (EK) model proposes a nonequilibrium parameterization embedding the K_d approach

* Corresponding author.
E-mail address: valerie.nicoulaud_gouin@irsn.fr (V. Nicoulaud-Gouin).

http://dx.doi.org/10.1016/j.jenvrad.2016.06.008
0265-931X/

(Garcia-Sanchez et al., 2014).

In contrast to the batch reactor method, the stirred-flow technique (Carski and Sparks, 1985) allows the observations of kinetics in the sorption process of the soil-water system. A stirred flow-through reactor also known as continuous flow stirred-tank reactor (CSTR) experiment consists in injecting by an inlet port a reactive tracer through a stirred cell, containing a known mass m of solid and a known volume of solution V, and measuring at the outlet port its solution-concentration dynamics C_w with time (Fig. 1).

Useful quantities can then be derived directly by mass balance considerations from concentration data C_w, such as: sorbed and desorbed amount to study reaction reversibility, sorption rates in nonequilibrium and equilibrium conditions, and sorption isotherms (e.g. Guimaraes et al., 2015; Li, 2012; Limousin et al., 2007; Martin-Garin et al., 2003; Szenknect et al., 2005).

Sorption parameters θ, such as equilibrium constants and first-order reaction rates, cannot be calculated directly but can be estimated from outlet observations $C_w(t)$ by solving the inverse problem (e.g. Garcia-Sanchez et al., 2014; Martin-Garin et al., 2003):

$$C_w(t) = f(t;\theta) \tag{1}$$

for a set of observations at times $t \in \{t_1, t_2, \cdots, t_n\}$, where f is the retained reactive transport model for the solute in the reactor. However, CSTR experiments were not initially designed to calibrate sorption parameters, and the experimental design and model inference practice ensuring optimal parameters *identifiability* have not been fully addressed, notably for the EK model.

Parameters are not identifiable if there is not a unique solution θ to the inverse problem (1) (Cobelli and DiStefano, 1980) which might result in the parameters's estimated values to be very imprecise or physically unsound. Some formal and numerical analyses allow to check parameter identifiability, but they have not yet been applied to sorption parameter estimation from CSTR experiments.

The problem of parameter identification can stem from the mathematical structure of the model $f(t, \theta)$ and its *structural identifiability* (Cobelli and DiStefano, 1980), i.e. the existence and unicity of the solution f^{-1}, transforming observable breakthrough curves $C_w(t)$ into solution points θ in the parameters space:

$$\theta = f^{-1}[C_w(t)] \tag{2}$$

Diagnosing existence and unicity is difficult in general, but is well-developed in the particular case of solutions of ordinary differential equations (e.g. Kaplansky, 1976; Denis-Vidal and Joly-Blanchard, 2004; Ritt, 1950).

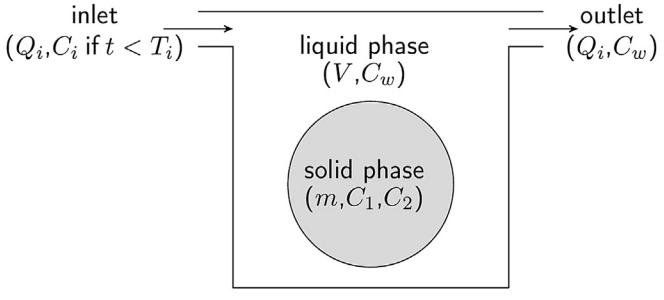

Fig. 1. Physical model and main physical variables characterizing stirred flow-through reactor experiments. Solution is injected by the inlet port at flow rate Q_i with contaminant concentration C_i during contamination stage ($t < T_i$). Contaminant experiences sorption with the solid phase in the chamber (associated Type-1 and Type-2 concentrations are C_1 and C_2). The reactor chamber is well-mixed and has homogenous contaminant concentration in solution C_w that is monitored at the oulet port.

Inaccurate parameter calibration can also stem from poor experimental information, due to observation errors and inappropriate experimental design. These problems, known as *practical identifiability*, can be studied on the sensitivity matrix or design matrix X defined by (Raue et al., 2009):

$$X = \left[\frac{\partial f(t_i;\theta)}{\partial \theta_j}\right] \tag{3}$$

The linear approximation of the inverse problem (1) is ill-conditioned if X is not of full rank, and estimations are highly inaccurate if X exhibits multi-collinearity (Seber and Wild, 1989).

Identification procedures are generally based on a statistical model for the observations:

$$C_w(t) = f(t;\theta) + \varepsilon \tag{4}$$

where ε is the random error unexplained by the model f.

In the frequentist nonlinear regression framework, θ is assumed to have a true value θ_0. This theory underlies most applications, notably for sorption parameter identification (e.g. Bajracharya and Barry, 1997; Fesch et al., 1998), although it can give rise to difficulties (Pinheiro and Bates, 2000; Seber and Wild, 1989). Numerically, parameter estimates are taken as absolute optimum of an objective function (generally the likelihood or the sum of squared deviates) in the parameter space. They are searched by iterative algorithms that can fail to converge or only find a local optimum. Statistically, parameter estimators are asymptotically unbiased and normally distributed, but their confidence intervals can be extremely misleading at finite sample sizes, and resampling techniques like the bootstrap are recommended to estimate their uncertainty (Seber and Wild, 1989).

In the Bayesian framework, θ is assumed to be random. It has a prior probability distribution $P(\theta)$, reflecting the existing knowledge before the calibration experiment, that can be updated by the observed data C_w through application of the Bayes' rule (Jeffreys, 1961):

$$P(\theta|C_w) \propto P(C_w|\theta) \cdot P(\theta) \tag{5}$$

The Bayesian approach is gaining audience in the field of radioecology (Beresford et al., 2014; Hosseini et al., 2013; Sy et al., 2015), including sorption parameter identification (Nordén, 2010), because its output directly quantifies parameter uncertainty and makes use of prior knowledge. However, its identification equation is a complex problem of numerical integration. The posterior distribution $P(\theta|C_w)$ can be identified numerically by Markov chain Monte Carlo algorithms (like Gibbs or Metropolis-Hastings) that are proved to simulate random samples from the target probability distribution when they reach convergence. The application of these methods requires a particular attention on the existing numerical and graphical convergence tests (Brooks and Roberts, 1998; Cowles and Carlin, 1996), that remain an ongoing topic of research. These algorithms were initially implemented for explicit models, e.g. in the WinBugs and RJAGS softwares (Lunn et al., 2000, 2009). More recently, GNU-MCSim software (Bois and Maszle, 2011) has extended these algorithms to ordinary differential equations models, but without implementation of convergence tests.

The objectives of this work were to check sorption parameters identifiability in CSTR experiments, and particularly to evaluate the abilities of the Bayesian approach to select a sorption model among alternatives and estimate its parameters. It also aimed to define experimental designs with optimal identifiability properties. For these purposes, outlet concentrations were simulated for 6 CSTR