Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2020-0178

Vol. **29**, No. 5, pp. 1583-1622
May 2021

COMPUTATIONAL SOFTWARE

DAFI: An Open-Source Framework for Ensemble-Based Data Assimilation and Field Inversion

Carlos A. Michelén Ströfer, Xin-Lei Zhang, Heng Xiao*

Kevin T. Crofton Department of Aerospace and Ocean Engineering, Virginia Tech, Blacksburg, VA 24061, USA.

Received 9 September 2020; Accepted (in revised version) 7 December 2020

Abstract. In many areas of science and engineering, it is a common task to infer physical fields from sparse observations. This paper presents the DAFI code intended as a flexible framework for two broad classes of such inverse problems: data assimilation and field inversion. DAFI generalizes these diverse problems into a general formulation and solves it with ensemble Kalman filters, a family of ensemble-based, derivative-free, Bayesian methods. This Bayesian approach has the added advantage of providing built-in uncertainty quantification. Moreover, the code provides tools for performing common tasks related to random fields, as well as I/O utilities for integration with the open-source finite volume tool OpenFOAM. The code capabilities are showcased through several test cases including state and parameter estimation for the Lorenz dynamic system, field inversion for the diffusion equations, and uncertainty quantification. The object-oriented nature of the code allows for easily interchanging different solution methods and different physics problems. It provides a simple interface for the users to supply their domain-specific physics models. Finally, the code can be used as a test-bed for new ensemble-based data assimilation and field inversion methods.

AMS subject classifications: 35R30, 76M21, 60-04

Key words: Data assimilation, inverse modeling, random fields, ensemble Kalman filter, Bayesian inference.

Program summary

Program title: DAFI

Nature of problem: This software performs ensemble-based, derivative-free, Bayesian inference of physical fields from sparse observations.

*Corresponding author. *Email addresses:* `cmich@vt.edu` (C. Michelén Ströfer), `xinleizhang@vt.edu` (X.-L. Zhang), `hengxiao@vt.edu` (H. Xiao)

Software licence: Apache-2.0

CiCP scientific software URL: `https://github.com/xiaoh/DAFI`

Programming language(s): Python

Computer platform: Any

Operating system: Any

Compilers:

RAM:

External routines/libraries:

Running time:

Restrictions:

Supplementary material and references:

Additional Comments: Runs on any system with Python and NumPy. Running time and computational requirements depend on specific problem being solved.

1 Introduction

Inverse problems in physical systems take many forms, and two broad classes—data assimilation and field inversion—are considered here. *Data assimilation* [1] refers to a class of inverse problems where a dynamic model is available and time-dependent observations are used to infer some property of a dynamic system. An example of a data assimilation problem is inferring the temperature field of a heated solid at the current time by using both sparse observations of the temperature (i.e. at a few locations) and a model forecast of the entire field. The model forecast could be obtained from propagating the temperature field at an earlier time using the diffusion equation. Here, *Field inversion* problems refer to a class of inverse problems where two sets of fields are related to each other through a forward model and observations of the output fields are used to infer the input fields. With the heat diffusivity example, a field inversion problem is inferring the material diffusivity from steady-state temperature measurements, where the heat diffusion equation is the forward model relating a diffusivity field to a temperature field. The approach taken here is to formulate both data assimilation and field inversion problems within a general framework of inverse problems and solve them by using ensemble Kalman filtering methods [2]. This is possible since field inversion problems can be recast as artificial dynamics problem and solved iteratively by using data assimilation procedures [3]. The main intended application for our code, which we named DAFI, is solving field inversion problems described by partial differential equations (PDE), a common type of problem in science and engineering. The code has several features that reflect this emphasis on PDE-based inversion problems, including: (i) ensemble-based solution approaches which are non-intrusive, requiring no gradients from and no code