Parameter estimation in a nonlinear dynamic model of an aquatic ecosystem with meta-heuristic optimization

Katerina Tashkova [a,*], Jurij Šilc [a], Nataša Atanasova [b,c], Sašo Džeroski [d]

[a] Computer Systems Department, Jožef Stefan Institute, Jamova cesta 39, Ljubljana, Slovenia
[b] Center for Marine and Environmental Research (CIMA), University of Algarve, Faro, Portugal
[c] International Center for Coastal Ecohydrology, Faro, Portugal
[d] Department of Knowledge Technologies, Jožef Stefan Institute, Jamova cesta 39, Ljubljana, Slovenia

ABSTRACT

Parameter estimation in dynamic models of ecosystems is essentially an optimization task. Due to the characteristics of ecosystems and typical models thereof, such as non-linearity, high dimensionality, and low quantity and quality of observed data, this optimization task can be very hard for traditional (derivative-based or local) optimization methods. This calls for the use of advanced meta-heuristic approaches, such as evolutionary or swarm-based methods.

In this paper, we conduct an empirical comparison of four meta-heuristic optimization methods, and one local optimization method as a baseline, on a representative task of parameter estimation in a nonlinear dynamic model of an aquatic ecosystem. The five methods compared are the differential ant-stigmergy algorithm (DASA) and its continuous variant (CDASA), particle swarm optimization (PSO), differential evolution (DE) and algorithm 717 (A717). We use synthetic data, both without and with different levels of noise, as well as real measurements from Lake Bled. We also consider two different simulation approaches: teacher forcing, which makes supervised predictions one (small) time step ahead, and full (multistep) simulation, which makes predictions based on the history predictions for longer time periods.

The meta-heuristic global optimization methods for parameter estimation are clearly superior and should be preferred over local optimization methods. While the differences in performance between the different methods within the class of meta-heuristics are not significant across all conditions, differential evolution yields the best results in terms of quality of the reconstructed system dynamics as well as speed of convergence. While the use of teacher forcing simulation makes parameter estimation much faster, the use of full simulation produces much better parameter estimates from real measured data.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Parameter estimation, also known as parameter fitting, model calibration or the inverse problem, is a key step in the mathematical modeling of dynamic systems from measured data (Janssen and Heuberger, 1995). Given a model structure and measured data, its goal is to estimate the model parameters in order to minimize the distance between the measurements and model predictions. Essentially, parameter estimation is an optimization task, that for the complex dynamics of ecological models can turn into a hard problem for traditional (derivative-based or local) optimization methods, calling for the use of advanced meta-heuristic approaches, such as evolutionary or swarm intelligence methods. Typically, ecosystem models are nonlinear and have many parameters, the measurements are sparse and imperfect due to noise, and the studied system can often be only partially observed. All of these constraints can lead to identifiability problems, i.e., the inability to uniquely identify the unknown model parameters, making parameter estimation an even harder optimization task (Marsili-Libelli, 1992; Omlin and Reichert, 1999; Marsili-Libelli et al., 2003).

There are two broad classes of approaches to the parameter estimation task: *frequentist* ("classical") and *Bayesian* (probabilistic) estimation (Rice, 2007; Samaniego, 2010). The most representative approach of the first class is the *maximum-likelihood* estimation, according to which the most likely parameter values are the ones that maximize the probability (likelihood) of observing the given data. A special case of maximum-likelihood estimation, based on the assumption of independent and normally distributed errors in the experimental data, leads to the well known *least-squares* estimation.

For maximum-likelihood estimation, we do not need any extrinsic information about the parameters. In contrast, the Bayesian

* Corresponding author.
E-mail addresses: katerina.taskova@ijs.si (K. Tashkova), jurij.silc@ijs.si (J. Šilc), naatanasova@ualg.pt (N. Atanasova), saso.dzeroski@ijs.si (S. Džeroski).

estimation approach treats the parameters to be estimated as random variables with a prior distribution: this distribution represents the knowledge about the parameter values before taking the data into account. Both approaches have shown benefits for specific applications across different domains, e.g., while Bayesian approaches explicitly treat the uncertainty of a parameter value by providing as a result its distribution rather than a single point estimate, frequentist approaches can be used for high-dimensional models. Note, however, that we cannot argue in favor of one class of approaches against the other in a general manner (Samaniego, 2010).

Related work in the domain of ecological modeling includes numerous applications involving representative methods from both classes. For probabilistic parameter estimation, we refer the reader to the work by Omlin and Reichert (1999), Dowd and Mayer (2003), Qian et al. (2003), and Jones et al. (2010). Here, we will focus on the parameter estimation task from the frequentist point of view, using least-squares estimation.

Classical approaches used for parameter estimation in ecological models include local optimization methods, such as derivative-based methods (e.g., the quasi-Newton method used for calibration of a phytoplankton model in the work by Mocenni et al. (2008)) and direct-search methods (e.g., parameter estimation of well-known ecological models with a flexible method based on polyhedron search by Marsili-Libelli (1992). These methods converge very fast to the optimum, if the search is started from a good initial point (in a close neighborhood of the optimum). However, they can only guarantee local convergence, as they do not have a mechanism to escape from a local optimum.

Derivative-based methods use the derivatives of the objective function (which is optimized), while the direct-search methods are derivative-free, which makes them generally easy to apply. The problem with the first is that they can fail if the objective function is discontinuous (as well as for discontinuous derivatives of the objective function), non-smooth, multi-modal or ill-conditioned, while the disadvantage of the second is that they become less efficient for high-dimensional problems. Therefore, it is recommended to use global optimization approaches that are more robust regarding the dimensionality and the landscape characteristics of the search space.

One general classification of global optimization approaches is into deterministic (exact) and stochastic (probabilistic). The deterministic methods (e.g., branch and bound, interior-point, cutting planes, etc.; see Horst et al., 1995; Horst and Tuy, 1996) can locate the global optima and assure their optimality, but there is no guarantee that they can solve any type of global optimization problem in finite time. Stochastic methods, on the other hand, rely on probabilistic search rules to find good solutions (Törn et al., 1999). They can locate the neighborhood of the global optima relatively fast, but their efficiency comes at the cost of global optimality (which cannot be guaranteed) and computational effort.

In the last two decades, special attention has been paid to meta-heuristics: these are general-purpose algorithms that can find acceptable solutions in reasonable time in both complex and large search domains. Most meta-heuristics are inspired by natural processes such as evolution (e.g., evolutionary algorithms (Fogel, 2000)), social behavior of biological organisms (e.g., ant colony optimization (Dorigo and Stützle, 2004), particle swarm optimization (Kennedy and Eberhart, 1995)), and controlled cooling associated with physical process (e.g., simulated annealing (Kirkpatrick et al., 1983)).

Methods from both classes are used for parameter estimation of ecological and environmental models, with evident interest especially in the recent research on meta-heuristic optimization. An example application of a deterministic approach is the parameter estimation of a groundwater model with Lipschitzian global optimization (Russell Finley et al., 1998). There are numerous applications of genetic algorithms, e.g., calibration of water quality models (Mulligan and Brown, 1998), calibration of phytoplankton dynamics for lake Kasumigaura in Japan (Whigham and Recknagel, 2001), and calibration of an ecosystem model of lake Kinneret in Israel (Gilboa et al., 2009). There are also some applications of genetic programming to the calibration of lake ecosystems models (Cao et al., 2008). Particle swarm optimization has been used for the calibration of water quality models (Afshar et al., in press). Simulated annealing has been applied to the calibration of marine ecosystem models (Matear, 1995). Some comparisons of global optimization methods have been also performed in this context, e.g., calibration of an oceanic biogeochemical model (Athias et al., 2000) and calibration of a hydrological model (Zhang et al., 2009).

Here we address the task of parameter estimation in an ecological model of Lake Bled in Slovenia. So far, Lake Bled has been modeled using simple theoretical models (Rismal et al., 1997), machine learning approaches (Kompare et al., 1997), and automated discovery of the structure and parameters of a model of its dynamics (Atanasova et al., 2006). The considered modeling approaches indicate that Lake Bled is a complex ecosystem requiring appropriate models for describing its behavior. The model structure we consider was discovered in a previous study (Atanasova et al., 2006) with the automated modeling tool LAGRAMGE 2.0 (Todorovski and Džeroski, 2006) from measured data. The model includes three ordinary differential equations (ODEs) for three ecological variables, i.e., dissolved phosphorus, total phytoplankton concentration, and the concentration of a zooplankton species *Daphnia hyalina*, that describe the dynamics of the food web in Lake Bled. The model was calibrated with a limited amount of measured data. Due to the ecosystem's complexity, the estimated parameters explain the calibration data well, but fail to provide accurate predictions for unseen data.

One of the reasons for the low quality of system dynamics reconstruction by the model at hand is the use of a local optimization method for parameter estimation. Namely, LAGRAMGE 2.0 uses the derivative-based local search algorithm 717 (Bunch et al., 1993). Moreover, LAGRAMGE 2.0 simulates the considered ODE models using the so-called teacher forcing simulation (derived from the teacher forcing approach for training neural networks, as introduced by Williams and Zisper (1989)). Teacher forcing simulation uses the measured values of the system variables at a given time point to calculate the system response at the next time point, unlike full numerical ODE integration (based on the one-step Runge-Kutta or advanced methods, such as multistep predictor-corrector adaptive-step integrators), where only the initial values of the system variables are needed for the model simulation over longer periods of time.

To improve the quality of the reconstructed system dynamics and to facilitate the automated discovery of appropriate model structures, we propose to use meta-heuristic methods for parameter estimation in the ecological model at hand. In this context, our study includes an experimental comparison of four meta-heuristic optimization methods: the differential ant-stigmergy algorithm, a recently developed meta-heuristic method for global optimization (Korošec, 2006; Korošec et al., in press) inspired by the pheromone-based communication of ants; its conceptual sibling, the continuous differential ant-stigmergy algorithm (Korošec, and Šilc, 2011); particle swarm optimization, another bio-inspired meta-heuristic based on the idea of swarm intelligences (Kennedy and Eberhart, 1995); and differential evolution, a well-known meta-heuristic method for global optimization based on the concept of natural evolution (Storn and Price, 1995, 1997). These are used to address the task of parameter estimation in the food web model of Lake Bled in the context of the least-squares estimation framework. Their performance in terms of the quality of