



Calibration and evaluation of individual-based models using Approximate Bayesian Computation



Elske van der Vaart^{a,*}, Mark A. Beaumont^{b,c}, Alice S.A. Johnston^a, Richard M. Sibly^a

^a School of Biological Sciences, University of Reading, Harborne Building, University of Reading, Whiteknights, Reading, Berkshire, RG6 6AS, United Kingdom

^b School of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, United Kingdom

^c School of Biological Sciences, University of Bristol, Woodland Road, Bristol, BS8 1UG, United Kingdom

ARTICLE INFO

Article history:

Received 16 February 2015

Received in revised form 13 May 2015

Accepted 17 May 2015

Keywords:

Approximate Bayesian Computation

Parameter estimation

Model selection

Individual-based models

Population dynamics

ABSTRACT

This paper investigates the feasibility of using Approximate Bayesian Computation (ABC) to calibrate and evaluate complex individual-based models (IBMs). As ABC evolves, various versions are emerging, but here we only explore the most accessible version, rejection-ABC. Rejection-ABC involves running models a large number of times, with parameters drawn randomly from their prior distributions, and then retaining the simulations closest to the observations. Although well-established in some fields, whether ABC will work with ecological IBMs is still uncertain.

Rejection-ABC was applied to an existing 14-parameter earthworm energy budget IBM for which the available data consist of body mass growth and cocoon production in four experiments. ABC was able to narrow the posterior distributions of seven parameters, estimating credible intervals for each. ABC's accepted values produced slightly better fits than literature values do. The accuracy of the analysis was assessed using cross-validation and coverage, currently the best-available tests. Of the seven unnarrowed parameters, ABC revealed that three were correlated with other parameters, while the remaining four were found to be not estimable given the data available.

It is often desirable to compare models to see whether all component modules are necessary. Here, we used ABC model selection to compare the full model with a simplified version which removed the earthworm's movement and much of the energy budget. We are able to show that inclusion of the energy budget is necessary for a good fit to the data. We show how our methodology can inform future modelling cycles, and briefly discuss how more advanced versions of ABC may be applicable to IBMs. We conclude that ABC has the potential to represent uncertainty in model structure, parameters and predictions, and to embed the often complex process of optimising an IBM's structure and parameters within an established statistical framework, thereby making the process more transparent and objective.

© 2015 Published by Elsevier B.V.

1. Introduction

Animal populations consist of autonomous, adaptive individuals, all figuring out their own ways of achieving their goals. From these activities of individuals emerge population consequences, such as spatial distributions, social structures and population dynamics. For many questions, both theoretical and applied, scientific knowledge exists at the level of the individuals or the population, but not both; in these cases, individual or agent-based models, here referred to as IBMs, can bridge the gap (DeAngelis

and Mooij, 2005). In IBMs, the interactions among individuals and between them and their surroundings are explicitly simulated, and all individuals have their own characteristics and make their decisions accordingly. This makes IBMs uniquely suited to exploring the effects of individual decisions on collective behaviour and to predicting how populations will change across time and space (Grimm and Railsback, 2005).

Examples of theoretical questions studied with IBMs include how ants build their nests (Buhl et al., 2005), how starlings coordinate their flocks (Hildenbrandt et al., 2010) and how macaques establish their relationships (Hemelrijk and Puga-Gonzalez, 2012). In practical applications, IBMs have been used to estimate the impact of development on coastal birds (Stillman and Goss-Custard, 2010), to aid in the management of fish stocks (Hartman and Kitchell, 2008), and to assess the effects of pesticides on non-target organisms (Schmolke et al., 2010). In all these cases, model

* Corresponding author. Tel.: +44 1183785467.

E-mail addresses: e.e.vandervaart@reading.ac.uk (E. van der Vaart), m.beaumont@bristol.ac.uk (M.A. Beaumont), a.s.a.johnston@pgr.reading.ac.uk (A.S.A. Johnston), r.m.sibly@reading.ac.uk (R.M. Sibly).

processes are fit to some data. However, due to the inherent complexity of IBMs, this process is often complicated, and the resulting outcome is often difficult to evaluate (Augusiak et al., 2014).

Currently, most IBMs are, implicitly or explicitly, built and evaluated using ‘pattern-oriented modelling’ (POM). This approach is essentially a protocol to be followed (Grimm and Railsback, 2012; Grimm et al., 2005). It specifies how multiple patterns observed in the real world should be used to iteratively design, select and parameterise IBMs, with each pattern serving as a ‘filter’ that rejects unsuitable model versions or parameterisations. Although the method has worked well in practice, a future goal is ideally to embed IBMs within mainstream statistical modelling and prediction. The Bayesian framework offers a comprehensive and well-trodden approach to this, but standard Monte Carlo methods for Bayesian inference are computationally intractable for all but the most carefully structured models (Keith and Spring, 2013). Conversely, ‘Approximate Bayesian Computation’, or ABC, enables approximate Bayesian inference for models of almost arbitrary complexity (Beaumont, 2010; Csilléry et al., 2010).

One key advantage of ABC, compared with other Bayesian methods, is that it is not necessary to analytically express how the likelihood of the data depends on the model parameters. Instead, ABC approximates these likelihoods by running models a large number of times, with parameters drawn randomly from their prior distributions, and then retaining the simulations closest to the observations. In this way, ABC provides a systematic way of assessing the support that different model versions and parameterisations receive from the available data, given some prior beliefs about how likely they are. Thus, for individual-based modellers, ABC has the potential to complement POM by making its rejections of unsuitable model versions and parameterisations more transparent and statistically rigorous. In addition, because ABC approximates full posterior parameter distributions, it provides a concise overview of the uncertainty in a model’s parameter values, which can then be propagated into a model’s predictions. Especially for IBMs that are then used in practical ecological decision making, this is an important feature.

Although the potential benefit of using ABC with ecological IBMs has been noted (Sibly et al., 2013; Thiele et al., 2014; Topping et al., 2012), whether it will work in practice is still uncertain. The original development of ABC was within population genetics (Beaumont et al., 2002; Pritchard et al., 1999; Tavaré et al., 1997), and in its basic form, rejection-ABC, it has yet to be applied to an IBM with more than two parameters (Sottoriva and Tavaré, 2010). In this paper, we apply rejection-ABC to a fairly typical ecological example, a 14-parameter IBM fitted to four existing experiments. In this IBM, Johnston et al. (2014) simulated the dynamic energy budgets of individual earthworms as they forage, grow and reproduce. We use rejection-ABC both to parameterise the model and to compare it to a simpler, possibly better model. In this way, we aim to introduce ABC to a wider audience, and to show that even simple implementations of ABC can provide surprising insights. In particular, we demonstrate how standard elements of an ABC analysis can be used to inform future modelling cycles. We consider the potential of more advanced versions of ABC in the context of the specific challenges posed by IBMs in the Discussion. A gentle introduction to the use of ABC with ecological IBMs together with a primer on building energy budget models from first principles will be available in van der Vaart et al. (submitted).

2. Material and methods

The simplest version of ABC, rejection-ABC, originally described by Pritchard et al. (1999), can be summarised as follows (Csilléry et al., 2010): First, for each of the parameters of each model, a

reasonable prior distribution is chosen. Statistically, for simplicity, we assume that each parameter has an independent prior distribution. Then, parameter values are sampled from these prior distributions, a large number of times, and the model is run for each of these samples, yielding some output. This output is compared with the empirical data. Some number of the runs that give the output closest to the empirical data are then ‘accepted’ as being ‘close enough’. The accepted simulations now provide a sample of the posterior distributions of the model’s parameters given the data. The same simulations can also be used for model comparison; in this case, the ratio in which different models are retained gives the relative probability that each model is correct. In the rest of this section, we first describe the empirical data available and the earthworm IBM; then, we give a detailed description of the ABC procedures used. All simulation results, the earthworm IBM and the ABC code have been deposited in a figshare repository (van der Vaart et al., 2015a,b), along with a brief guide to their use.

2.1. The empirical data

For our ABC analyses, we used the same empirical data that Johnston et al. (2014) originally used to assess the model’s fit. This empirical data consists of the growth and reproduction data for *Eisenia fetida* earthworms in different laboratory setups (Gunadi et al., 2002; Gunadi and Edwards, 2003; Reinecke and Viljoen, 1990). In each case, five to ten earthworms were placed in small containers and supplied with cattle manure for food, under various feeding schedules (see Fig. 4 and Table S1). The earthworms were weighed and all cocoons were removed and counted at regular intervals. These procedures were replicated in our simulations; we assumed that the weighing of earthworms entailed a randomisation of their position, but no homogenisation of the substrate. The mean individual body masses and total cocoon numbers so obtained are referred to as the summary statistics. In total, 160 summary statistics were used.

2.2. The individual-based model

For parameter estimation, we used Johnston et al.’s (2014) IBM of the earthworm *Eisenia fetida*. For model selection, this IBM was compared with a simplified version of itself, which is described in Section 3.2. The model is implemented in NetLogo, a programming platform designed specifically for IBMs (Wilensky, 1999). Using the IBM to simulate all the available empirical data takes approximately half a second on a 3.4 GHz i7 iMac. Each time step, the earthworms in the IBM forage independently and allocate the acquired energy to maintenance, growth and cocoon production in a fixed order of priority (Fig. 1a). These priorities are represented by algorithms and equations derived from fundamental physiological ecology (Sibly et al., 2013). Each time step corresponds to a day. The model’s parameters are in Table 1.

Eisenia fetida is predominantly surface-dwelling, and its environment becomes patchy when foraging earthworms deplete the food in their vicinity. To approximate this situation, we mapped the environment in two dimensions and divided it into four rectangular patches, with food being homogeneously distributed within each patch and variation between patches. The total area simulated in this manner is 0.0144 m (see the Supplementary Information). Whenever the experimental setup calls for food to be added or removed, the amount is distributed equally across all four patches.

Each earthworm starts the simulation as a juvenile, with mass M_j , until it reaches mass M_p , when it becomes an adult. Every daily time step, it moves around randomly, with speed s , and ingests food. This movement is the only stochastic aspect of the model. An earthworm acquires energy according to Eq. (1), where X is the food density and M is the earthworm’s current mass. This is

Download English Version:

<https://daneshyari.com/en/article/6296538>

Download Persian Version:

<https://daneshyari.com/article/6296538>

[Daneshyari.com](https://daneshyari.com)