Ecological Modelling 326 (2016) 152–161

Contents lists available at ScienceDirect

Ecological Modelling

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

Divide and conquer: Configuring submodels for valid and efficient analyses of complex simulation models

Iris Lorscheid *, Matthias Meyer

Institute of Management Accounting and Simulation, Hamburg University of Technology, Am Schwarzenberg-Campus 4, 21073 Hamburg, Germany

ARTICLE INFO

Article history:
Available online 23 December 2015

Keywords:
Sensitivity analysis
Design of experiments
Ecological theory
Computational modeling
Model analysis
Validation

ABSTRACT

Individual-based modeling is considered an important tool in ecology and other disciplines. A major challenge of individual-based modeling is that it addresses complex systems that include a large number of entities, hierarchical levels, and processes. To represent these, individual-based models (IBMs) usually comprise a large number of submodels. These submodels might be complex by themselves and interact with each other in many ways, which in turn can affect the overall system behavior in ways that are not always easy to understand. As a result, both the validity and credibility of IBMs can be limited. We here demonstrate how a cascaded design of simulation experiments (cDOE) may support the validity and efficiency of the analysis of IBMs and other ecological simulation models. We take a systematic approach that adopts a divide-and-conquer strategy. In a preparatory phase, submodels and their parameters are configured in "subexperiments". Consequently, the "top-level experiments" of the simulation model can assess the research questions in a more valid and efficient way. Our strategy thus supports the structural realism of individual-based models because both the behavior of their main components and the relationships between these components are explicitly addressed.

1. Introduction

Individual-based modeling (IBM) is considered an important approach both for advancing ecological application (Stillman et al., 2015) and theory (Huston et al., 1988; Railsback and Grimm, 2012). However, ecological IBM addresses complex systems which lead to models which usually are complex as well. The existence of submodels representing, for example, resource dynamics or the adaptive behavior of individual organisms, is a common characteristic of such models (Grimm and Railsback, 2005). These submodels can be complex by themselves and interact with each other in complex ways. The resulting system behavior can be difficult to understand for the modeler and even more so for outsiders (Van Nes and Scheffer, 2005). Models that are not well understood are likely to reduce the general acceptance of simulation as a method, as illustrated by the following statement: "With simulation it may be impossible to drill down to what assumptions are responsible for conclusions, to discern the causal connections between initial conditions and results, and simulation invites unsophisticated and sloppy research together with naive hocus-pocus about the magic of emergence." (Roughgarden, 2012, p. 8). Although such extreme points of view are getting less common, the underlying skepticism in individual-based and other types of ecological simulation models is still a healthy attitude (Augusiak et al., 2014) as it challenges simulation modelers to improve their methodology.

A more comprehensive understanding can only be partially achieved by standard analysis techniques such as local or global sensitivity analysis (Cariboni et al., 2007; Saltelli et al., 2008), especially in the context of complex IBMs. Our paper therefore introduces the cascaded design of experiments (cDOE) approach to address and eventually overcome these limitations. The approach consists of a sequence of simulation experiments that prepare and improve the investigation of the overall research questions addressed with a model. Dividing the model into components along the simulation process organizes the simulation model analysis and secures its systematics. Subexperiments are used to assess the model's components and their interdependencies. This contributes to a better understanding, (face) validation and verification of submodel behavior. Moreover, they help establish the configuration of parameter values and ranges for the submodels. We call these activities in the preparatory phase the "configuration of submodels".

Analyzing submodels separately has been suggested for IBMs before (Grimm and Railsback, 2005) and demonstrated in

* Corresponding author.
E-mail addresses: iris.lorscheid@tuhh.de (I. Lorscheid), matthias.meyer@tuhh.de (M. Meyer).

http://dx.doi.org/10.1016/j.ecolmodel.2015.11.013
0304-3800/

textbooks (Railsback and Grimm, 2012). Still, such analyses can be too ad hoc to increase overall understanding and validity, and many, if not most, IBMs are still published without any or only limited analyses of submodels and their interactions (Grimm and Berger, 2016). We therefore here embed the analysis of submodels in a more systematic framework, which is based on literature on design of experiments (DOE) and which we originally formulated for simulation models in the social sciences (Lorscheid et al., 2012). We hope that our framework will foster the structural realism of complex simulation models, given that the behavior of their main components and the relationships between those components are explicitly addressed. Over time, our approach may contribute to ecological theory because the similarities and differences of models can be better traced back to the behavior and interactions of their submodels (see also Grimm and Berger, 2016; Thiele and Grimm, 2015).

The paper is structured as follows. The next section reviews current suggestions for the analysis of complex simulation models in ecology. Section 3 introduces the idea of analyzing models through a systematically planned hierarchical sequence of designed simulation experiments. Section 4 illustrates the potential of the cDOE approach by applying it to a sample model. The final section concludes and presents avenues for future research. It should be noted that this article includes elements which correspond to tutorials, or guided tours, rather than research articles. We considered this form of presentation necessary because so far there is no culture in ecological modeling of carefully analyzing submodels before the full model.

2. Approaches to analyzing complex ecological models

Several approaches for analyzing and, thereby, better understanding simulation models exist (Grimm et al., 2006, 2010; Lorscheid et al., 2012; Thiele et al., 2014). An important approach is sensitivity analysis, which can be defined as "the study of how uncertainty in the output of a model (numerical or otherwise) can be apportioned to (. . .) uncertainty in the model input" (Saltelli et al., 2004). There are two types of sensitivity analysis: local and global (Cariboni et al., 2007). Local sensitivity analysis focuses on the effect of single model inputs varying over limited ranges, whereas global sensitivity analysis reveals the effect of several inputs, varying over their full range, and explores also possible interactions between inputs. A comprehensive review of the state of the art in sensitivity analysis is given by Saltelli et al. (2004) and a "cookbook" for applying a range of widely used techniques, based on existing R packages is provided by Thiele et al. (2014).

For sensitivity analysis, efficiency and computational costs are commonly raised issues (Campolongo et al., 2007; Cariboni et al., 2007; Marino et al., 2008; Thiele et al., 2014). Therefore, many existing techniques focus on enhanced efficient parameter sampling to make sensitivity analysis applicable to complex models. However, the understanding of complex simulation models is not only hampered by their high number of parameters but also by the many relations among submodels (Van Nes and Scheffer, 2005). Performing a more detailed assessment of submodels and their interactions is necessary both to understand model behavior (Brugnach, 2005) and to determine whether that behavior matches system behavior sufficiently well (Beck, 2002).

Brugnach (2005) addresses complex submodels of ecological models and proposes depicting the input–output relations of these submodels, which represent certain processes, to analyze complex process-based models. The proposed analysis explicitly screens for effects of complex interactions and their relationships on the overall simulation outcome.

In addition to these approaches, the literature on the design of experiment techniques (DOE) offers valuable techniques for analyzing complex simulation models. DOE was first applied in agriculture and biology, and its general potential for analyzing simulation models has been recognized (see, e.g., Antony, 2014; Law, 2015; Lorscheid et al., 2012; Saltelli et al., 2008). DOE can be defined as "the process of planning, designing and analyzing the experiment so that valid and objective conclusions can be drawn effectively and efficiently" (Antony, 2014, p. 8). Two aspects of this definition stand out. First, the technique supports the validity and reliability of the conclusions drawn from the results, e.g., by controlling for possible biases. Second, DOE offers solutions to overcome resource limitations caused by computational costs by reducing the range of possible experimental settings to the most important combinations. Adopting a social simulation perspective, Lorscheid et al. (2012) propose a standardized procedure to effectively integrate DOE principles into the analysis of complex simulation models. The establishment of standards for dealing with complex systems and tasks within a certain field indicate that methodology in this field is maturing beyond its pioneer phase, where methodological "ad-hockery" prevailed (Heine et al., 2005). Our framework is thus consistent with other initiatives for standardization such as the ODD protocol (Grimm et al., 2006, 2010) and TRACE (Grimm et al., 2014; Schmolke et al., 2010).

Whereas most approaches to analyzing complex models primarily focus on the efficiency aspect, the DOE approach developed by Lorscheid et al. (2012) also has the potential to address the validity aspect. They focus on conducting subexperiments to better understanding the behavior of submodels. However, they do not outline a process to tackle the possibly complex interactions between submodels. If this issue is not addressed effectively, it could threaten the soundness of the conclusions drawn from the simulation model analysis, therefore affecting the validity and reliability of results.

Overall, the literature review reveals that submodels of simulation models warrant a closer look and analysis. So far, only a few authors have explicitly considered existing model submodels in the analysis process. No systematic configuration of components and their parameters to prepare the simulation model for the experimental analysis has been developed. We therefore in the following suggest a cascaded setup to manage the internal complexity of models in a preparatory model-analysis phase.

3. Analyzing complex models by cDOE

Simulation methods can be seen as a complex experimental environment that can benefit from a preparatory phase. In line with lean management's philosophy of doing things right the first time (Crosby, 1979), we advocate a thorough planning and preparation of simulation experiments with both the submodels and the full model.

The goal of this configuration phase is to identify and focus on the (relevant) model components and the interplay between them to understand and validate their behavior. Submodels are configured by specifying their parameters to run the "right" experiments with respect to the purpose of the full model. Reducing the model analysis to reasonable and practicable parameter combinations will increase efficiency. Doing so also enables the detection and correction of possible implementation errors at an early stage of the analysis process, which again has a positive impact on efficiency because later adjustments and corrections may be more expensive. As indicated by the iceberg metaphor in Fig. 1, it is important to go beyond the immediately visible input–output relations of a model