Ecological Complexity 22 (2015) 178–191

Contents lists available at ScienceDirect

Ecological Complexity

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

Context-oriented model validation of individual-based models in ecology: A hierarchically structured approach to validate qualitative, compositional and quantitative characteristics

Andreas Kubicek[a], Fred Jopp[b,c], Broder Breckling[d,e], Christoph Lange[f], Hauke Reuter[a,g,*]

[a] Leibniz Center for Tropical Marine Ecology (ZMT), 28359 Bremen, Germany
[b] Systems Biology, Justus Liebig University, 35392 Giessen, Germany
[c] Department of Biology, University of Miami, P. O. Box 249118, Coral Gables, FL 33124, USA
[d] Center for Environmental Research and Sustainable Technology, Dept. General and Theoretical Ecology, University of Bremen, 28359 Bremen, Germany
[e] Landscape Ecology, University of Vechta, 49377 Vechta, Germany
[f] Bruker Daltonik GmbH, 28359 Bremen, Germany
[g] University of Bremen, 28359 Bremen, Germany

ARTICLE INFO

Article history:
Received 30 May 2014
Received in revised form 31 March 2015
Accepted 31 March 2015

Keywords:
Context-oriented validation
Calibration
Parameterisation
Individual-based model
Agent-based model
Hierarchically structured validation

ABSTRACT

Validation constitutes a vital process in model development and application, as it ensures the applicability of a model for the intended purposes and trustworthy results within the range of model assumptions. Commonly, independent empirical data sets are statistically compared with the generated model results, which is an adequate approach for models which operate on a single hierarchical level, such as most equation-based models. Individual-based models (IBM) can operate on different organisational levels synchronously and have an inherent complex and variable interaction structure for many applications. Thus a plain comparison of data congruity on the result levels might leave too many questions unanswered. However, a more comprehensive assessment of model validity can require additional investigations which encompass also qualitative and structural relationships.

Here we describe a *hierarchically structured validation* which is oriented towards the investigated context of the model and allows organising the validation process in close relation to the different hierarchical levels which are covered in the model. The context oriented organisation protocol for validation includes the following steps: (1) assessing the different model levels separately, then, (2) applying a set of different techniques such as visual inspection, statistical comparison, involvement of experts, aggregation of data on higher integration levels and experimental validation.

The context oriented approach accounts for the specificity of individual-based models – i.e., the dynamic self-organisation of model outcomes from biologically underpinned individual interactions without an inherent determination of properties on higher hierarchical levels – and extends the potential of the validation process qualitatively, as it allows to assess complex structural and causal relations and multi-level feedback processes of the developed models.

Contents

* *Corresponding author at:* Leibniz Center for Tropical Marine Ecology (ZMT), Ecological Modelling, Fahrenheitstr. 6, 28359 Bremen, Germany. Tel.: +49 4212380058.
E-mail address: hauke.reuter@zmt-bremen.de (H. Reuter).

http://dx.doi.org/10.1016/j.ecocom.2015.03.005
1476-945X/

1. Introduction

Validation is an essential part of model development because it ensures applicability of a model for the intended purpose and assesses how robust and reliable model results are. It reflects to which extent a model is rational and fulfils its objectives (Hamilton, 1991) and, furthermore, it should provide information on the model's range of validity and its constraints (Jakeman et al., 2006; Saltelli and Annoni, 2010). In a strict sense, validation is twofold consisting of part (a) informing how potential outcomes relate to a priori assumptions and (b) constitutes the last step in warranting the correctness and applicability of a model before it is applied to its intended purposes. Most of the broad spectra of validation approaches in ecological modelling have been tailored to equation-based models (Power, 1993; Rykiel, 1996) or describe general aspects of quality assurance (Oreskes et al., 1994; Janssen and Heuberger, 1995; Sargent, 1998; Troitzsch, 2004; Jakeman et al., 2006; Marks, 2007). There is also a growing literature on the evaluation of environmental models with an emphasis on management and decision support (e.g., McIntosh et al., 2011; Bennett et al., 2013; Filatova et al., 2013). In ecological modelling there is a long discussion on validation procedures with a focus on different aspects on the model evaluation process leading to some ambiguity in the term (Augusiak et al., 2014). Thus how 'validation' is performed requires a clear definition of targets and methods in each specific case (Jakeman et al., 2006; Aumann, 2011; Bennett et al., 2013). For our purpose the general definition of validation by Hamilton (1991) referring to the model objectives is sufficient as it allows to include the majority of approaches. In a specific sense model validation focuses on the comparison of model results with known patterns and processes to ensure that the model's behaviour is consistent with the constitutive behaviour of the investigated system.

In general, validation for most equation-based models constitutes a single level comparison because processes are generally described on the same hierarchical level as the results. For example, the Lotka-Volterra equations (Lotka, 1925; Volterra, 1926) describe population growth and population interactions (predator-prey in its basic form) to simulate population dynamics. In this case validation only makes sense on the integration level of the population. However, this procedure is not transferable to other modelling approaches, such as individual-based modelling (IBM, Huston et al., 1988; Judson, 1994) for which this relation between described model processes and model generated output is not always valid. An additional analysis of internal model generated patterns (Pattern Oriented Modelling) may increase reliability in model results (Grimm et al., 1996; Wiegand et al., 2003) in comparison to the analysis of single level results.

Individual-based models (IBM) have specific characteristics and therefore allow for specifically designed extended approaches for validation. In general, IBMs consist of rule-based systems that represent basic organisation (often individual organisms) which, in turn, can be combined with more or less distinguished mathematical equations. An inherent hierarchy of model processes emerges self-organised from interactions of the specified components or entities on higher integration levels resulting from lower level interactions (Breckling et al., 2006). Thus, model processes often span over several integration levels (e.g., individual life-history, population dynamics and community development) and exhibit across-level feedbacks requiring a specific validation on each of the levels (Reuter et al., 2005; Topping et al., 2012). In this regard, IBMs represent a more complex setting according to the represented context which in many cases requires specific considerations which do not necessarily conform to a type of situation which can be strictly standardised. Due to the fact that modelled processes and components are relatively close to the represented biological processes (e.g., physiological reactions, intra/inter-specific interactions and behavioural rules) most of the underlying assumptions are directly accessible to validation routines. In differential equations hidden implicit assumptions are not accessible for rigorous testing during the validation process. These assumptions refer for example to linear mortality rates proportional to population size, the existence of infinitely small populations (sometimes smaller than one individual), or the assumption of spatially homogeneous dispersal. This is usually not the case in IBM due to the explicit representation of individuals with finite life times. Causal relationships implemented in the model can be assessed directly, which allows for potential conclusions on the driving forces in natural systems. In other words, IBM allows to extend validation procedures by including the specific structural context, multi-level analyses as well as feedback processes between different model levels, to cope with inherent variability in a model's structure.

Compared to ecological IBMs, agent-based model (ABM) validation, is more advanced in the social sciences. Here the discussion on validation is an ongoing issue (Moss, 2008), where different approaches for validation and checking the accuracy of model representations are discussed for their context and appropriateness (e.g., Küppers and Lenhard, 2005; Qudrat-Ullah, 2005; Windrum et al., 2007; Filatova et al., 2013). ABM and IBM are quite similar approaches because they are based on similar modelling paradigms (Reuter et al., 2008). Troitzsch (2004) distinguishes between different types of model validity relating to replication (how well the model matches data), prediction (how well a model matches data before they are acquired) and structure (reflects the way the observed systems produces its dynamics). Moss (2008) differentiates ABM validation approaches in two ways: those that closely relate empirical procedures in data