How to Read and Write TRACE Documentations

Volker Grimm & Amelie Schmolke[1]

Helmholtz Centre for Environmental Research – UFZ, Department of Ecological Modeling,
Permoserstr. 15, 04318 Leipzig, Germany

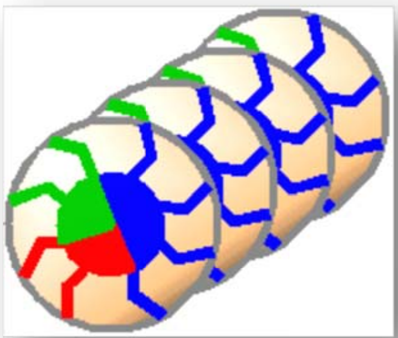
1st draft - 2011-03-04

Please send feedback, suggestions, or questions to: volker.grimm@ufz.de.
See http://cream-itn.eu/trace for updates of this document.

http://cream-itn.eu/

[1] Present address: 321 NW Harrison St, Pullman, WA 99163, USA, email: Amelie_schmolke@fastmail.net

CONTENTS



How to use this guideline

Schmolke et al. (2010b) propose a standard format for documenting ecological models that are developed to support environmental decision making, and their analyses: transparent and comprehensive ecological modeling (TRACE) documentation. The purpose of this standard format is to disclose all parts of the modeling process to scrutiny and make modeling itself more efficient and coherent. TRACE is completely described in Schmolke et al. (2010b) and ready to be used in modeling projects. However, a more detailed guidance will enhance the coherent application of TRACE, and is especially needed for clarification for two important groups of users: (1) beginners in ecological modeling and (2) decision makers who are going to use TRACE documentations for assessing models submitted for decision support. In the following, we provide a detailed guideline for TRACE documentations. All elements of TRACE, which correspond to the so-called "modeling cycle" (Fig. 1), are introduced by giving short answers to always the same set of questions. For every element, specific guidance for those who produce and those who read TRACE documentations is given. Additionally, we suggest ways to keep a "modeling notebook" and how to use it for producing TRACE documentations.

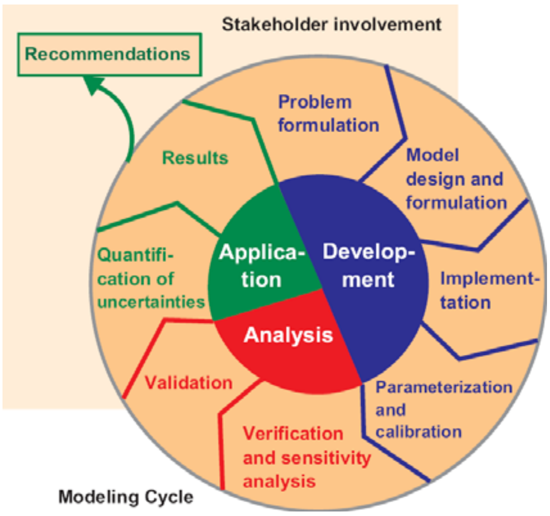


Figure 1. The "Modeling Cycle". The elements of the cycle correspond to the main task of any modeling projects. They therefore are used as main categories and section titles in TRACE documentations. The main parts are Model Development (blue), Model Testing and Analysis (red), and Model Application (green). The modelers' task is to implement, test, and analyze a model, but decision makers and stakeholders need to be involved in problem formulation, assessment of uncertainties and results, and formulation of recommendations (from Schmolke et al., 2010b).

In this guidance document, we are not repeating what has been said in Schmolke et al. (2010b), which should therefore be read first before reading and using this document:

> Schmolke A, Thorbek P, DeAngelis DL, Grimm V. 2010b. Ecological modeling supporting environmental decision making: a strategy for the future. *Trends in Ecology and Evolution* 25: 479-486. (Please contact Amelie Schmolke (amelie_schmolke@fastmail.net) or Volker Grimm (volker.grimm@ufz.de) for a pdf of the article.)

Modelers and model users: In the following, we will refer to "modelers" as those who formulate, implement, test, and analyze a model, and to "model users" as those who are going, in the widest sense, to use the insights gained from the model, i.e. stakeholders, decision makers, and experts of the species, natural resource, or environmental sector of concern. Some model users need to be directly involved in the modeling process, but once a model and its results exist, the model is also likely to be used by other users, who were not involved.

Maintaining modeling notebooks: In the article that introduces the TRACE framework, we emphasize that a routine day-to-day documentation of modeling projects is very important. Such a modeling notebook will mainly be written for the use of the modelers themselves or the modeling team as a reference. However, it should be written such that it can be used by a person who joins the modeling team at a later stage or takes over the project. Also, it is essential that the notebook can be consulted in case detailed questions or issues arise during model review and use. This compares to the status of the laboratory notebook as maintained for all laboratory experiments (Kanare 1985, Denny-Gouldson 2007).

The chronology of the modeling notebooks' assembly will not exactly follow the order of topics in the TRACE document because topics will be reiterated at different stages of model development. Nevertheless, the notebook should serve as a basis for the compilation of the final TRACE document. Thus, it is essential that all entries are clearly marked by the topic they are dealing with, the date of the entry, the version of the model or submodel they refer to, and names of files containing the results.

Some old-fashioned modelers may actually use a book (numbered blank sheets of paper bound together between a cover and a back) for their notes about the modeling project. If you are one of them (like one of the authors), we recommend using a separate book for each modeling project. Otherwise, diverse software can be used to write the notebook (Rabin and Pastorius, 2005). The simplest version is to use a word processing program which will allow searching for the topics. More advanced tools, such as Microsoft OneNote, already include functions specific to maintaining a project notebook of any kind, including links to other files, and the easy inclusion of tables and figures. A notebook written in html or xml, basically following the idea of a web blog, also has those advantages, and may be particularly helpful if several people need to have access to the notebook. No matter what software solution is chosen, separate files should be generated for each project, and the notebook file should be stored together with all other files relevant for the project, even after the completion of the project.

Preferably, and hopefully routinely in the future, TRACE documentations are assembled from modeling notebooks that document the modeling process. However, for models that were developed before TRACE was suggested in 2010, TRACE documentations can easily be assembled from existing documentations of model tests, etc. If no such documentations exist, new tests need to be run in retrospect and documented according to the format suggested by TRACE.

The structure of this guideline: This guideline will address each topic of the TRACE document as introduced in the article by Schmolke et al. (2010b; Box 1). For each point, we will provide explanations and definitions ("What is it?"), reason why it has to be included in a TRACE document ("Why is it important?"), and describe the content of the documentation including short examples ("How to 'trace' it?"). We highlight the benefits for the parties involved in the modeling project ("Benefits of TRACE documentation"), discuss what should be documented in the modeling notebook ("Modeling notebook and TRACE"), and give some general guidance how each part of the documentation can be assessed by model users or other reviewers ("Assessment criteria"). Additionally, we give background information and further reading ("Background and further reading").

Caveat: A complete TRACE documentation is necessary for a model to be acceptable for informing decisions, but not sufficient, i.e. a model is not per se suitable for its purpose because it comes with a TRACE documentation However, a model that is poorly designed, documented, or simply unsuitable for supporting the decisions to be made should be easy to recognize from the TRACE document.

Suggested reading: In addition to the numerous references given in the "Background" sections of this manual, and to general textbooks and monographs on modeling, we recommend downloading and reading the "Guidance on the development, evaluation, and application of environmental models", which was produced by the Council for Regulatory Environmental Modeling (CREM) of the US Environmental Protection Agency (EPA) (CREM, 2009[2]).

Box 1. TRACE (transparent and comprehensive ecological modeling) **documentation structure**. Note that here we introduced numbered section titles, whereas Schmolke et al. (2010b) only numbered the three main parts (from: Schmolke et al., 2010b).

1 MODEL DEVELOPMENT

1.1 Problem formulation: *Context* in which the model will be used, and the type of audience addressed; *specification of the question(s)* that should be answered with the model; statement of the *domain of applicability* of the model, including the extent of acceptable extrapolations; assessment of the *availability of knowledge and data*; specification of necessary *model outputs.*

1.2 Design and formulation: Description of the *conceptual model*; description and justification of the *modeling approach* used and of the *complexity*; *entities and processes represented* in the model; most important, the applied *assumptions* about the system.

1.3 Model description: Detailed *description of the actual model*, and how it has been *implemented* (programs, software platforms, scripts).

1.4 Parameterization: *List of all parameter values* used in the model, the *data sources*, and how the parameter values were obtained or calculated; *uncertainties* associated with each parameter.

1.5 Calibration: Documentation of the *data sets used* for calibration; *which parameters* were calibrated; what *optimization method* was used.

2 MODEL TESTING AND ANALYSIS

2.1 Verification: Assessment of whether the model is *working according to its specifications*; documentation of *what tests have been conducted.*

2.2 Sensitivity analysis: Exploration of the model behavior for *varying parameters*; documentation of which *parameter combinations* have been tested; *justification* of used parameter ranges and combinations.

2.3 Validation: Comparison of model or submodel outputs with *empirical data that were not used for parameterization or calibration*; documentation of *data sources*; what *parts (submodels) have been validated*; what *validation methods* were applied.

3 MODEL APPLICATION

3.1 Results: Outputs that are used to *inform decisions*; description of *simulation experiments (scenarios)* conducted; *statistics applied* to analyze model outputs.

3.2 Uncertainty analysis: *Uncertainties in model outputs* used for recommendations; description of *variance, noise, and bias in empirical data*; determination of *stochasticity* in the model; description of *model uncertainty* which can be assessed through application of different models or submodels; best- and worst-case scenarios.

3.3 Recommendation: Description of how *initial question(s) could be answered*; summary of *conclusions* drawn from model; clarification of *extrapolations* used (in time and space).

#

[2] http://www.epa.gov/crem/library/cred_guidance_0309.pdf

1 MODEL DEVELOPMENT

This part 1 of TRACE documentations should be written in a way that is accessible to model users who can be assumed to have a scientific background but no specific training in mathematical or simulation modeling. This is important because this part is the basis of assessing, and understanding, the remainder of the documentation. It is also decisive because model users, who formulated the original purpose of the model, need to be involved during iterative model development

1.1 Problem formulation

What is it?

Context in which the model will be used, and the type of audience addressed; *specification of the question(s)* that should be answered with the model; statement of the *domain of applicability* of the model, including the extent of acceptable extrapolations; assessment of the *availability of knowledge and data*; specification of necessary *model outputs*.[3]

Why is it important?

- Modelers need to have a clearly defined problem to be addressed with the model, otherwise they have no clear guideline for what elements and processes of the real system to explicitly represent in the model, and in what detail.

- Model users need to know, and be aware of, the model's purpose to understand why a model was designed in a certain way, to be able to assess whether or not this design is meaningful, and to be sure that they do not try using the model for other purposes than those it was designed for.

Models are simplified representations of real systems. They are designed to answer a specific question or solve a certain problem. This question or problem serves as a kind of filter: models include only those elements and processes in the model that are considered most relevant with regard to this question. Thus, without being clear and explicit about the model's purpose, it is impossible to decide whether or not a certain aspect of the real system should be included in the model. Likewise, without knowing for what purpose a model was developed, it is impossible to understand why it was designed in a certain way.

Formulating the question to be addressed with a model is therefore the first step in the modeling cycle. Iterating this step and re-formulating the model's purpose is quite common, because ecological systems and environmental problems are complex, thus being absolutely clear about the question that should be answered usually is an essential part of the problem to be solved.

Problem formulation is a task to be performed, iteratively, by modelers and model users together. Iterative participation of model users in this task is indispensable. In domains where models were not used so far, not only the modelers need to learn specific context of this domain, but also model users need to learn how models could help making better informed decisions. In such cases "iteration" thus might mean that several rounds of the modeling project are required to better understand what kind of model outputs, for example endpoints in ecological risk assessment of chemicals, are most useful.

[3]All text in blocks is taken verbatim from Schmolke et al. (2010b), box 1.

How to “trace” it?

Concise answers to the following questions should be provided:

Context: What is the intended field of application?

> Examples: “Higher tier regulatory risk assessment of a pesticide.” – “Decision support for ranking alternative management scenarios for threatened species.” – “Development of schemes for sustainable use of natural resources.”

Questions: What are the specific questions addressed with the model? What problem solving is the model supposed to support?

> Examples: How do sublethal effects of a certain pesticide on larval development affect winter survival of a honey bee colony? How does landscape structure affect the sensitivity of a population to pesticides? Which of three alternative options for supporting a threatened bird population (buying additional habitat, providing additional nesting sites, additional feeding in winter) is most effective in both in terms of increasing population viability and costs?

Domain of applicability: For which spatial and temporal scales is the model applicable? What kind of environmental conditions can the model take into account (only those for which model parameters were determined, or also new conditions)?

> Examples: “The model does not represent space and therefore describes population dynamics for regions that are small or homogeneous enough for spatial effects to be ignored, or for questions where spatial effects might be averaged out.” – “The model includes only empirical parameters, e.g. vital rates, which are based on mark-recapture data which were taken for a certain range of environmental conditions. The model can thus not be extrapolated to completely new conditions without further justification.” – “Key parameters of the model, e.g. daily growth and survival rates, are driven by lower level mechanisms, e.g. adaptive behavior, physiology, dynamic energy budgets, etc. The model can therefore be applied for a wide range of environmental conditions, including unprecedented ones.”

> *Availability of knowledge and data*: On what body of empirical information is the model based? How comprehensive, and how uncertain (variable) are the data?

> Examples: Own experiments or field studies, data from the literature, expert knowledge, theoretical principles, and insights from previous studies. – “Model parameters are based on only one comprehensive field study carried out in England over two years [Reference].” – “We compiled model parameters, such as habitat-specific population growth rate and carrying capacity, from the existing literature and unpublished reports. In Table [X], the range of reported values is listed as well as the reference value chosen for our model.”

Model outputs: What main quantitative and qualitative output is the model going to deliver, and how can this output be used to make decisions?

> Examples: Risk of extinction over 100 years. Average population size over 10 years. Time to recovery after pulse exposure to a stressor. Average winter mortality. Ranking of management scenarios. Recommendation for mitigation, or risk management. Recommended stocking rates. Simple rules for adaptive management. Indicators for sustainable use. Understanding how certain landscape features serve as buffer mechanisms. Understanding why the population is more sensitive to stressors in certain landscapes, or certain seasons of a year.

Benefits of TRACE documentation

Modelers: A clearly defined, and iteratively refined, problem formulation is key to choosing an appropriate model structure, for making sure the problem to be addressed has been fully understood, and for making sure that model output directly can be used for supporting decisions. The detailed problem formulation gives the modeler insight in the actual problem at hand. Misunderstandings between modeler and model users will be limited, and efforts in model development will be reduced.

Model users: Only a comprehensive and clear problem formulation guarantees that the modeler will design and analyze the model in a way that makes model output relevant for supporting decisions. A clear formulation of the model objectives with the model user's participation assures the usefulness of the model for the problem at hand. In the same time, model users learn what the model may be able to provide, and in the following, will be able to judge if the model actually provides what it promises at the outset of the project.

Modeling notebook and TRACE

Problem formulation has to be done together with model users and, if required and possible, experts who know the ecological system to be represented. It is decisive for the success of any modeling project to write down, after every major discussion of this issue, an updated version of the problem formulation and to circulate it, including date, possibly the model version (number) on which the update of the problem formulation was based, and the main changes compared to the previous version. For the TRACE documentation, the formulation underlying the model version that has been actually used should be included, but also major insights that were gained through the modeling project regarding the problem formulation. This can help those model users who were not directly involved in the project to better understand why the problem was formulated in this way.

Assessment criteria

The only criterion for assessing whether or not, or to what degree a model can be used to support decisions is whether a complete and clear problem formulation exists. If not, the model cannot be used.

Background and further reading

It is recognized widely that the formulation of objectives is essential for any modeling project (Bart, 1995; Radomski and Goeman, 1996; Rykiel, 1996; Jackson et al., 2000; Nicolson et al., 2002; Pastorok et al., 2002; Bartell et al., 2003; Jakeman and Letcher, 2003; Landsberg, 2003; CREM, 2009; Prisley and Mortimer, 2004; Jakeman et al., 2006; Boorman et al., 2007; Janssen and van Ittersum, 2007). Boorman et al. (2007) list questions that have to be addressed at the beginning of a modeling project: "What is the management issue? What are the key variables and processes? What data are available? What model outputs are required? What solutions or understandings are expected?". In many cases, models are designed by scientists. However, problems arising in academia are usually quite different from management issues (Landsberg, 2003). It is a common pitfall to extend the scope of a model beyond the problem at hand (Jakeman et al., 2006).

A model can only be evaluated in the context of its objectives (Bart, 1995). Accordingly, clear and measurable objectives have to be defined (Radomski and Goeman, 1996; Pastorok et al., 2002). This includes the specification of the acceptable level of total uncertainty of the model results (Bart, 1995; Pastorok et al., 2002; CREM, 2009). Measurable results have to be based on, and evaluated in the context of empirical data available about the system (Bart, 1995; Boorman et al., 2007). Modeling objectives should state how model outputs will inform decision making (Beissinger and Westphal, 1998; Bartell et al., 2003; CREM, 2009; Boorman et al., 2007).

1.2 Design and formulation

What is it?

Description of the *conceptual model*; description and justification of the *modeling approach* used and of the *complexity*; *entities and processes represented* in the model; most important, the applied *assumptions* about the system.

Why is it important?

- Modelers need to assemble a conceptual model of the system to be represented. The conceptual model reflects the assumptions on how the system is organized and functions, and on what factors and processes are most important with regard to the purpose of the model. The conceptual model reflects the preliminary understanding of the system. If no such understanding exists, no further modeling is possible.

- Model users need to know, and understand, the conceptual model, which is the point of departure of the modeling process. It makes sure that more formal structures, which come later in the modeling process, were not chosen *ad hoc* but based on careful consideration of existing knowledge, data, and understanding.

How to "trace" it?

Concise answers to the following questions should be provided:

Conceptual model: What conceptual model is the starting point of the modeling process?

> This can be, for example, a graphical model, one or more assumptions on what the key factors and processes to be considered are, or an established theoretical concept.
>
> *Examples:* "We assume that the population is organized as a metapopulation, i.e. distributed across disjunct habitat patches, which have their own, only partly coupled population dynamics. We assume that local extinctions can be balanced by re-colonization from other patches." – "We assume that surviving the winter is the limiting factor of this population, therefore we focus on a detailed representation of this season." – "We assume that for the model's purpose it is sufficient to compare population growth rates, which are based on current, constant demographic parameters." – "We assume that response to multiple stressors such as food limitation, social stress, and toxicants, depends on how energy is allocated to the various tasks of individual organisms. We therefore base our model on Dynamic Budget Theory" (Fig. 2)

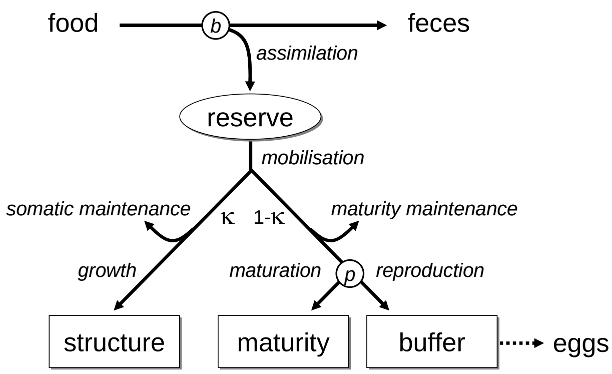


Figure 2. Example of a conceptual model depicted as a diagram. Schematic diagram of the primary state variables (boxes) and fluxes (italics) of the standard Dynamic Energy Budget model. The circles containing "b" and "p" denote maturity switches for birth and puberty (from Jager et al., 2004).

Modeling approach and complexity: What model type is used and why do you think it is appropriate for the purpose of the model? How complex is the model, approximately, in terms of number of entities, state variables, parameters, and processes? Why has no simpler, or more complex model been developed? Have such alternatives been explored?

> In population ecology, three broad classes of models can be distinguished: (1) unstructured models, where only population size, or total biomass, is considered; such models are usually formulated using ordinary differential (continuous time) or difference (discrete time steps) equations. (2) Structured models, where groups of individuals of different age, size, stage, or other characteristic are distinguished. If time is represented in discrete time steps, the models are in principle sets of couple difference equations which can, if they are linear with regard to the dependent variable, be explored via matrix algebra; therefore, such models are often referred to as matrix models. (3) Individual-based (or agent-based) models, where each individual is represented as independent, unique entity which is characterized by a set of state variables or attributes. Such models are simulated on computers and therefore implemented as computer programs. Note that these model types can overlap: individual-based models can include differential equation models as submodels; matrix models can include density dependence and environmental noise, which makes them similar to simple individual-based simulation models. Further aspects that may characterize a model are: spatial/non-spatial; deterministic/stochastic; population/community/ecosystem level.
>
> *Examples:* "We want to learn how characteristic differences in vital rates of soil organisms affect the species' sensitivity, in terms of changes in population growth rate, to pulse and chronic exposure of chemicals." – "We want to predict the response of the population to unprecedented changes in their habitat, for which no data exist. Therefore, we decided to base the model on behavioral and physiological mechanisms that mediate the response of organisms to changes in environmental variables such as temperature, salinity, resource level, or local intraspecific competition."

Model assumptions: What are the main assumptions about key factors and processes of the system and with regard to the purpose of the model? What are the most important simplifying assumption? How are these assumptions justified?

> Examples: "The model is non-spatial, which means that we ignore spatial effects, including habitat heterogeneity, local interactions, and migration. Our model thus is relevant only for homogeneous or small habitats, or for landscapes large enough that spatial effects can assumed to be averaged out." – "Although certainly all parameters can be assumed to vary from year to year, we chose to explicitly only represent variation in juvenile survival over the first 4 weeks of their life because this is known to strongly depend on weather, i.e. temperature and rainfall."

Benefits of TRACE documentation

Modelers: The conceptual model can be considered a first blue print of the model, which is necessary for the implementation, verification, and documentation of the model. The model design corresponds to the design of a scientific experiment. Careful consideration of the model's blue print and the design of experiments make the following tasks of the modeling cycle considerably more efficient. The conceptual model is essential for communicating the model to model users because the actual model may not be fully comprehensible for non-modelers or the assessment of the implementation may not be feasible within reasonable time limits.

Model users: Since the conceptual model should directly match the model implementation, model users can check if the model is actually representing what they expected. The conceptual model must not elude public understanding, but model users have to be able to assess the conceptual model.

Modeling notebook and TRACE

Design and formulation of the conceptual model is a highly iterative process. Transparent and comprehensive documentation of every major version allows to quickly identify misunderstandings of the problem, inconsistencies in model assumptions, and concerns about too complex, or too simple, aspects of the model. Quite a few model versions might need to be discussed before proceeding to the next task of the modeling cycle, the formal description of the model. The notebook should include all these major versions, including version numbers, and each new version should come with a summary of major changes and why there were considered essential. For TRACE documentations, only the final design and formulation will be documented, but the major insights gained during their iterative development should be included as well as it may document that, e.g., the conceptual model and model type were not chosen ad hoc but are based on thorough considerations and discussions.

Assessment criteria

A model can only be used for supporting environmental decision making if it is clear, logical, and convincing how the problem formulation led to the overall design and formulation of the conceptual model.

Background and further reading

Conceptual model: In a conceptual model, assumptions about the system are formalized, and thus, the formulation of a conceptual model is an integral part of the modeling process (Radomski and Goeman, 1996; Guisan and Zimmermann, 2000; Jackson et al., 2000; Nicolson et al., 2002; Pastorok et al., 2002; CREM, 2009; Jakeman et al., 2006). A conceptual model can clarify the relations between modules of the system (Pastorok et al., 2002). "Creating a good conceptual model forces an ecologist to formulate hypotheses, determine what data are available and what data are needed, and assess the degree of understanding about key components of the system" (Jackson et al., 2000). Conceptual models are usually written using arrows and boxes (Jackson et al., 2000), and may include the formulation of relationships in the form of mathematical equations (CREM, 2009). Radomski and Goeman (1996) refer to the conceptual model as a "cognitive map".

Model design subsumes the compilation of the conceptual model and its formalization. In the model design, the assumptions about the system as collected in the model objectives are laid out in a way that serves as a direct blue print for the model implementation. In this step, the decisions about the model type and model complexity are taken. Both decisions have to be guided by the model objectives and the available data. The model should not neglect processes that have been recognized as important, and should make use of data related to these processes. In the same time, no processes that have been deemed marginal should be included, or processes that cannot be supported by empirical data.

Choice of model approach: Modelers tend to use model approaches they are familiar with, which usually comprises a small subset of available models (Boorman et al., 2007; Jakeman et al., 2006). However, it is unlikely that a single approach is suitable for all purposes and across all available data (Ferrier and Guisan, 2006). Rather, the most appropriate approach for the goal of the study should be found (Ferrier and Guisan, 2006; Guisan and Zimmermann, 2000; Petersen et al., 2008). Accordingly, the choice of the model approach should be given ample attention (Beissinger and Westphal, 1998; Guisan and Zimmermann, 2000; Jackson et al., 2000; Bartell et al., 2003; CREM, 2009; Wiegand et al., 2003; Ferrier and Guisan, 2006; Boorman et al., 2007; Jongejans et al., 2008; Petersen et al., 2008; Schmolke et al., 2010a).

Choice of model complexity: The choice of the model approach is tightly linked to the appropriate level of complexity represented in the model. Increasing complexity of models is assumed to be linked to higher data needs and increasing uncertainty of model results due to the number of variables (Fulton et al., 2003; Jakeman et al., 2006). However, realism is emphasized to increase the relevance and usefulness of models (Bartell et al., 2003; Fulton et al., 2003; Mooij and DeAngelis, 2003). Some authors advertise the application of simple models (Wennergren et al., 1995; Beissinger and Westphal, 1998; Jakeman et al., 2006). In contrast, most authors recommend to find an optimal level of model complexity which includes as much complexity as needed in the specific context (Parker et al., 2002; Bartell et al., 2003; Fulton et al., 2003; Jakeman and Letcher, 2003; Mooij and DeAngelis, 2003; CREM, 2009; Wiegand et al., 2003; Jongejans et al., 2008; Nelson et al., 2008). Indeed, it is not a simple task to determine the optimal level of complexity. Fulton et al. (2003) state the lack of understanding how complexity influences model performance, but studies suggest a humped relationship (Fulton et al., 2003; Mooij and DeAngelis, 2003; "pattern-oriented modeling", Grimm et al., 2005).

1.3 Model description

What is it?

Detailed *description of the actual model*, and how it has been *implemented* (programs, software platforms, scripts).

Why is it important?

- Modelers need complete written descriptions of their model to communicate with model users, colleagues, journal reviewers, or academic advisors and to allow independent replication of the model and its analysis.

- Model users need to have a clear and comprehensive written description of the model to be able to understand, evaluate, and use the model.

How to "trace" it?

Model descriptions should include both, the factual description of the model, i.e. its equations, rules, and algorithms, and rationale underlying the model and its parts. Factual description and justification should be clearly separated from each other, so that readers easily can decide whether they only want to check the factual description, or also the explanation and justification of the chosen formulations.

For simulation models, following the structure of the so-called "ODD protocol" is recommended (Grimm et al., 2006; 2010). It was originally developed for describing individual- or agent-based models but has turned out to be suitable for simulation models in general. The acronym stands for: Overview, Design concepts, and Details (Table 2). Some of the design concepts are specific to individual-based models, which can simply be ignored for other model types.

> Examples: Appendix A includes a template for writing ODD model descriptions. Appendix B includes a list of publications including ODD model descriptions which which are clear and comprehensive and therefore can be taken as good examples for using ODD.

Even mathematically formulated models can benefit from adopting the ODD protocol, since mathematical equations do not necessarily inform us about the model's rationale. It should also be noted that nowadays most mathematically formulated models are not solved analytically, but numerically, which means they are in fact simulated. The software used for this should be specified, and the corresponding script files, for example for the software platforms MatLab, R, or Mathematica, should be provided.

ODD is a verbal format and despite existing detailed instruction for its use (Appendix A) some details might still be ambiguous. It is therefore important to also include, in electronic supplementary material, the program code that has been used to implement the model. In any case, the programming language and operation system has to be specified, including version number and distributor. Correspondence between ODD model description and computer code should be easy to check, via references to file names, HTML links, using the same names for variables and submodels in verbal description and programs. Preferably, also an executable version of the computer program implementing the model should be provided so that model users can check whether they can replicate reported results, and to perform, to better understand the model, own simulation experiments. It should not be required, though, that all files are made available that are needed to produce, e.g. by using a compiler, the executable problem because some of these files might be proprietary to the distributors of the software used.

Benefits of TRACE documentation

Modelers: The benefits of a clear and comprehensive model description are evident. Without it, communication and replication would be impossible.

Model users: The model, and its rationale, can be explored and scrutinized without needing to have been trained in mathematics, programming, or computer science.

Modeling notebook and TRACE

Since models, especially in early stages of model development, continuously change, not all exploratory modifications to equations and computer code can be documented. Preferably, exploratory modifications are performed on the time scale of hours and should finally lead to clearly defined new versions, or subversions, of the model. If exploratory tests last longer than a day, daily documentation is recommended, with summaries of the modifications explored and the insights gained. Every major change that leads to new simulation experiments and, accordingly, new insights, should be documented and explained. Obviously, every major model version that is analyzed needs to start with a written model description. For TRACE documentations, the actual version is taken from the modeling notebook. It will usually contain a mixture of submodels that did not change for a long time and more recent submodels.

Assessment criteria

The written model description is a central element of TRACE documentations. Model users need to have a means to understand and evaluate a model without being forced to try and understand computer code, or being assumed to have solid training in mathematics or computer science. The main acceptance criteria are: Is the model description transparent and comprehensive? Could the model be re-implemented based on the information provided? Are computer code and script files provided so that, if in doubt, they can be checked for implementation details? Is an executable program provided to re-run simulation experiments? Models which are not fully documented should not be accepted for supporting decision making.

Background and further reading

Grimm et al. (2006, 2010) introduce the ODD protocol and refer to further publications addressing standardization and replication. See also the guidance document produced by CREM (2009).

1.4 Parameterization

What is it?

List of all parameter values used in the model, the *data sources*, and how the parameter values were obtained or calculated; *uncertainties* associated with each parameter.

Parameters are the constants in the equations and algorithms that are used to represent the processes in a model. They determine the interactions and controls among model mechanisms, i.e. depending on parameter values, completely different system dynamics may occur.

“Parameterization” is the word modelers use for selecting values for a model’s parameters (Railsback and Grimm, 2011).

Why is it important?

- Modelers need to assign values to model parameters, which quantify relationships and processes in the model. Robustness of model predictions, and acceptance of the model for supporting decisions depends on thorough parameterization and its comprehensive documentation.

- Model users need to know to what degree the model is tied to empirical evidence. A look at the table(s), which is (are) usually used to document parameterization, informs about data sources for each parameter and about uncertainties. These tables thus indicate which parameters had to be determined indirectly, via calibration (see next element of TRACE), and which parameters have to be tested, and with how much effort, in sensitivity and uncertainty analyses (see part 2 of TRACE documentations).

How to “trace” it?

The following elements should be documented, if possible by using tables.

List of all parameter values: The parameter table(s) should include the symbol or name used to refer to the parameter, its meaning, its units, its chosen reference value, its chosen maximum and minimum values (i.e., its range), and existing *data sources*. These can be own experiments, which then have to be documented separately, as an appendix to the parameter table, or values reported in the literature.

In academic publications, it is appropriate to refer to data sources by citing other publications. This, however, is not sufficient for TRACE documentations. Model users cannot be assumed to get hold of all cited publications themselves and check in detail how parameter values were extracted from that publication. Rather, if values were directly taken, references to page numbers should be included. If values are deduced, by re-scaling, estimations, or extrapolations, from the publications, this has to be documented in detail. It can also be useful to indicate under what condition (region, environmental conditions, season, etc.) values reported in the literature were observed.

If parameters were determined via calibration, references to the corresponding sections in part 1.5 of the TRACE documentation should be included.

Uncertainties associated with each parameter should also be listed: e.g. the deviations in empirical measurements, potential biases of measurements, uncertainties due to derived parameters, etc.

Benefits of TRACE documentation

Modelers: Model parameterization is a decisive part of any model documentation. Keeping track, and documenting, the often considerable amount of work that went into parameterization will increase the credibility of the model.

Model users: A comprehensive parameter table informs about the character of the model: if it is based on many parameters that are related to low-level, “first principle” processes, which were determined directly and independently, then the model might be able to provide robust predictions for a wide range of conditions. If, on the other hand, many parameters, which refer to more aggregated processes, were quite uncertain and had to be narrowed down via calibration to observations, the model might still be useful but not easily be used for extrapolation to new conditions. If many parameters are just guessed, or cannot be directly measured at all, the model is likely to be more exploratory and theoretical and might therefore only be of limited use for supporting decisions.

Modeling notebook and TRACE

While a model is developed and parameterized, the following information should be saved in the modeling notebook (preferably in tables that later can directly be copied to TRACE documents):

- Every new parameter: in what equation or submodel does it occur, what does it mean, what are its units, what is the assumed range of values, what is the reference value used in model analysis, on what data are range and reference value based.
- Make sure to keep track of all data sources as detailed as possible: on what page of a certain reference are the data reported, what values are reported, under which conditions where they collected, did you need to re-scale data.
- Keep a list of "hidden parameters", i.e. parameters that you will not vary in the model. Add short explanations why you are not varying these parameters.
- Make notes of every parameter assignment that is not directly based on empirical data: how were values assigned? In many cases, "educated guesses" can be sufficient for the purpose of a model, or because that parameter is not very sensitive anyway, but on what "education", expert opinion, or rationale is that guess based?
- Clearly identify those parameters that were determined indirectly, via calibration. Include links to the corresponding sections in TRACE element 1.5, "Calibration".
- Uncertain parameters (ones that were just guessed or come with a high uncertainty in empirical data) should be marked for thorough sensitivity analysis (see TRACE element 2.2)

Assessment criteria

Undocumented model parameters, or parameterizations, should not be accepted for models that are to be used for supporting decision making. Parameters with high uncertainties should be subject to thorough sensitivity analysis (see TRACE element 2.2): if the model results are highly sensitive to small changes in the parameters, the model might not be very useful for decision support.

Background and further reading

"Model parameters are terms in the model that are fixed during a model run or simulation" (CREM, 2009). Parameter values are drawn from empirical data or knowledge about the system, i.e. a model is parameterized with available data about the system. Often, parameter values cannot be derived directly and exactly from data. In this case, parameters have to be estimated (Jackson et al., 2000). Calibration is a systematic way of estimating and adjusting model parameters to improve fit between model outputs and data (Rykiel, 1996; Pascual et al., 2003; Glaser and Bridges, 2007). The optimization of fit between model and data may be confounded with a measure of model quality, but model parameterization and calibration are part of the model development, and cannot be used for model evaluation (CREM, 2009).

1.5 Calibration

What is it?

Documentation of the *data sets used* for calibration; *which parameters* were calibrated; what *optimization method* was used.

Calibration is a special kind of parameterization: a few especially important parameters, which are very uncertain or were identified by sensitivity analyses (TRACE element 2.2), are „fine-tuned" to make model output match observations as close as possible. This fine-tuning can be important if quantitative predictions are important. An alternative term to „model calibration“ is „model fitting“. In simple cases, calibration can be done „manually“, by manually adjusting certain parameters and checking a quantitative measure of how well model output matches observation. More often, quite sophisticated optimization approaches are required, including gradient methods, genetic algorithms, neural networks, pattern-oriented (Monte-Carlo filtering), and – more recently – approximate Bayesian statistics (Hartig et al., *unpubl. manuscript*). Note that for more complex models, calibration should be done versus different sets of observations, or patterns, via multi-criteria assessment.

Why is it important?

- Modelers aiming at quantitative predictions need to calibrate their model. For example, predicted population growth rate of the brown bear re-invading the Alps from Slovenia was too uncertain to be useful for management recommendations. After pattern-oriented calibration of the model, predicted mean population growth rate was different to the non-calibrated prediction, and uncertainty in the prediction was considerably reduced (Wiegand et al., 2004).

- Model users need to know what parameters where determined via calibration, and what methods and data were used. This is important because in the extreme, a quite unrealistic model could be “forced”, via calibration, to reproduce certain observations. The predictive potential of such models would be very low, since they would not capture the true mechanisms generating observations and patterns, and overall structure and dynamics of the system in question. Moreover, the model might be valid only for situations similar to those in which data for calibration where obtained. Knowing details of calibration methods and data is thus critical for delineating the domain of application of a model.

How to “trace” it?

What parameters were calibrated? What optimization methods were used? What metrics and statistical methods were used to quantify differences between model output and observations? The rationale of the methods used should be briefly explained, and references to more detailed literature given. Modelers should discuss to what degree calibration interfered with the overall aim to make a model structurally realistic, and to let important dynamics and structure emerge from the interaction of model entities, instead of imposing them.

Benefits of TRACE documentation

Modelers: If quantitative predictions are critical to the usefulness of the model for supporting decisions, a lot of work will go into model calibration. If calibration is well documented, easy to grasp, and thoroughly discussed, model results will more readily be accepted, or used appropriately, i.e. by taking into account uncertainties regarding parameters and model structure.

Model users: Calibration is usually required to make model predictions quantitative. Having the chance to check all aspects of calibration in details is critical for assessing the robustness and structural realism of a model.

Modeling notebook and TRACE

Calibration can require a lot of resources for simulations and statistical analyses. If well documented, calibration methods used can more easily be re-used for new model versions, or new scenarios and parameter sets. Thus, any effort going into calibration should be documented in a way that would later on allow to re-run exactly the same calibration, which implies keeping the right model version, script files for statistical packages, and results files.

Assessment criteria

If calibration is not well documented, it remains unclear how robust the methods and observations using during calibration are, to what degree model results are imposed or emergent, and how likely it is that the model is structurally realistic, i.e. reproduces the right observations, or patterns, for the right reasons, i.e. the key generative processes in the real system.

Background and further reading

It is a widely held myth that the more complex a model, the more it has to rely on calibration. However, if a model is based on „first principles", i.e. mechanistic submodels that can be parameterized independently of the full model, even fairly complex ABMs have proven useful with little or no calibration of the full model (Pitt et al. 2003, Railsback et al. 2005, Goss-Custard et al. 2006). Nevertheless, calibration remains an important step in the development and analysis virtually any ecological model, especially when the model's purpose is to deliver quantitative predictions regarding specific systems.

2 MODEL TESTING AND ANALYSIS

This part 2 of TRACE documentations refers to the most technical part of the modeling cycle, which includes implementing the model, often as a computer program, testing the program, testing and verifying the model, and analyzing the sensitivity of the model to changes in parameters and initial values of state variables. Modelers should document all their work and tests here as described below, but they cannot be expected to explain all tests in all technical detail. Nevertheless, they should briefly describe the rationale of the methods used, include references to literature where the methods used are introduced and explained, and in particular briefly describe the purpose of each test, its main results, and what conclusions the results allow for. Model users should easily get an overview to what degree software and model have been tested so that they can assess the reliability and robustness of model predictions. For any submodel, they should easily be able to find the corresponding tests and possibly spot check them, if needed by invoking external modeling experts. Model users are not expected to know and understand all details of the methods used, but they should understand the test's rationale and the implications of test results for model reliability and robustness.

2.1 Verification

What is it?

Assessment of whether the model is *working according to its specifications*; documentation of *what tests have been conducted.*

Terminology regarding the issues of verification and validation is extremely diverse. Because for TRACE we have to decide for, and use consistently, a certain terminology, with "verification" we will refer to testing whether the model has been implemented according to its written specification. This includes, for simulation models, testing the software that was produced to implement the model, and testing whether submodels and the entire model behave not only as specified but also in a logical, consistent way. Verification thus also includes testing to what degree model output matches observed properties and dynamics of real systems (which usually requires calibration; see TRACE element 1.5).

In contrast, with "validation" we refer to devising and testing secondary, or independent model predictions, i.e. predictions of properties of the real system which were not taken into account at all for model parameterization and calibration (see TRACE element 2.3). Note that in TRACE terminology, we draw a line between "verification" as described here, and "validation", whereas many other authors put these two in the same category and refer to it as, e.g., "verification", "validation", "testing", or "corroboration". However, independent on how the term "verification" is used in various fields, or by the individual modeler: referring to "verification" in a TRACE document implies using it as described in this paragraph.

Why is it important?

- Modelers need to prove that they did their job right: translating a verbal model into equations and/or computer programs which behave exactly as intended, and making sure that the model reproduces one or more observations of real systems. Direct comparison to real systems is not required for theoretical, or demonstration, models, which are not, however, designed for supporting decisions but for providing general insights. Nevertheless, also theoretical model need to be verified regarding their implementation and regarding overall, generic features of real systems.

- Model users need to ask the key question: Why should I care about predictions of a model? There might be software errors, logical errors, inconsistencies, and the model might reproduce the right observations for the wrong reasons (it has been "forced" to do so via calibration). Thus, model users need to see how thoroughly the model and its implementation were tested, and how well, with regard to one or more patterns observed in the real system, the model represents the real system and its functioning.

How to "trace" it?

Part 2 of the modeling cycle can be complex and require many different kinds of tests. Details on how this is done depends on the model type used, on the characteristics of the model, on the model's purpose, how much time the modeler can devote to this task, and on how well the modeler is trained in specific techniques, e.g. software testing (Railsback and Grimm, 2011; chapter 6).

We therefore recommend that modelers here use, according to their needs and constraints, their own outline where they introduce briefly the testing and verification approaches used (including references for further reading) and then describe all major tests, their results and conclusions, and finally briefly summarize what these results tell us.

For more complex models, formal methods of software testing may be required. Likewise, comparing model output to data can involve sophisticated methods of calibration, including Bayesian statistics. It is impossible to even summarize the vast literature on these techniques here. Modelers are expected to be aware of standard and more recent approaches and to use them, where appropriate. For TRACE documentations, it is important that modelers include a short summary of the rationales of the methods used, and of references to introductory and advanced literature on these methods. In particular, it is not sufficient to present just the result of a certain test, but also to include a brief discussion of to what degree the test allows to conclude that the tested program is error-free or that the tested submodels behave as intended.

Scientific publications and reports based on ecological modeling usually contain information on some tests of the model and its elements, but not of the software (an exception is the trout model "inSTREAM" of Railsback and co-workers). Although most modelers certainly perform numerous tests of all kinds, there is no culture of documenting and communicating these tests. Establishing such a culture is the declared purpose of TRACE documentations. In the CREAM project, a demonstration TRACE documentation of a very simple model will be available in Spring 2011; further TRACE documentations will be produced in the different individual CREAM projects (see CREAM website, http://cream-itn.eu).

The central technique for testing and analyzing simulation models (including mathematically formulated models that are numerically solved or simulated) is controlled simulation experiments.

Just as in real experiments, modelers define simplified scenarios where, e.g., all parameters but one are kept constant. The remaining free parameter is varied over a certain range and the results are interpreted and contrasted to the output of further simulation experiments. As for real experiments, the "design of experiments" (DOE) needs to be specified: What is the purpose of the experiment? What is the setting of the experiment (model version, parameter values, initial values of state variables), what "factors" (i.e., parameters or initial values) are varied, over what range, for what values (Campolongo et al., 2007)? How many repetitions are run, if the model is stochastic? What statistical analysis is used to evaluate the model output (significance of effects, uncertainty, sensitivity, etc.)

A very important element of model verification is multi-criteria assessment. The more observed patterns a model can reproduce simultaneously, the more likely does the model capture the functioning of the represented system. It can be easy to force a model, via calibration, to reproduce one pattern, e.g. a time series, but making a model reproducing two or more patterns narrows down the degrees of freedom regarding model structure and process formulation. Multi-criteria assessment is the core of a modeling strategy called "pattern-oriented modeling" (Grimm et al., 2005; Grimm and Railsback, 2005; Wiegand et al.; 2003; 2004). Note that most approaches used for verification also can be used for validation, the only difference being that for the validation, independent, new data and observations are used, which were not used for model parameterization and/or calibration or generally during model development in previous modeling cycles, including also verification.

Benefits of TRACE documentation

Modelers: Documenting verification makes modeling more efficient and more systematic. It allows keeping track of all tests performed. Documenting tests, including all script files and test protocols used, also allows re-using test designs. Systematic verification of software and model allows the modeler to always proceed from "firm ground": if a model behaves strangely after a new element has been added, not the entire model and software has to be re-checked. And, transparent and comprehensive documentation of tests for verification increases credibility of the model and its predictions.

Model users: Why should a model user trust a model? Because it is not only well-designed, completely documented and its design easy to understand, but also because the modeler invested a lot of time for testing and for documenting these tests in a format that can be reviewed by non-modelers.

Modeling notebook and TRACE

Like with model formulation, not every single model run, where one parameter or another has been modified, can or should be documented. Model development and testing includes heuristic interactions with the model, which is important to understand the model and how it works. Ideally, after some time these heuristic exercises lead to a simulation experiment where a certain piece of computer code of the model is thoroughly tested. These tests should be documented in the modeling notebook, including information on model version tested, corresponding files, methods used, specific software (e.g., statistics software) used etc. Documenting all these verification experiments, very often carried out with model versions that are soon updated, very much corresponds to keeping a laboratory notebook, or diary: describe the experiment, its purpose, summarize the results, and briefly summarize the lessons learned. For TRACE documentations, all tests available of the actual model version should be compiled. If the model is complex, it can be appropriate to summarize tests in a table and provide detailed information in an appendix or separate files.

Assessment criteria

The more detailed verification tests the modeler documents, the more confidence can be put on the reliability of the model as a tool for supporting decision making. The more complex a model, the more important is its verification. Most weight should be given to tests that are documented, usually in electronic files, in such detail that the model user can, usually via spot checks, see whether the right model version was tested and the tests themselves performed, described, and interpreted correctly.

Background and further reading

Regarding comparison to real systems, note that the purpose of verification is, first, determining "if the model is acceptable for its intended use, i.e., whether the model mimics the real world well enough for its stated purpose"; and second, "how much confidence to place in inferences about the real system that are based on model results" (Rykiel 1996[4]; Schmolke et al. 2010b). The quantifiers "well enough" and "how much" indicate that verification does not try to show whether or not a model is "true" or "false", since all models are based on simplifying assumptions and therefore cannot be true in an absolute sense. Rather, it depends on the purpose of the model, on the knowledge and understanding of the system represented, and on the "culture" of dealing with verification and validation (see next TRACE element), which can vary widely between different domains.

Verification is recognized by many authors as an important step in the model testing (Ferson, 1996; Rykiel, 1996; Oriade and Dillon, 1997; Guisan and Zimmermann, 2000; CREM, 2009; Prisley and Mortimer, 2004; Jakeman et al., 2006; Schmolke et al., 2010a).

2.2 Sensitivity analysis

What is it?

Exploration of the model behavior for *varying parameters*; documentation of which *parameter combinations* have been tested; *justification* of used parameter ranges and combinations.

Sensitivity analysis reveals how model outputs depend on model inputs, i.e. model parameters and initial values of state variables of the model entities, and also of changes in model structure. If variation of a certain input does not alter model output markedly, the model is not sensitive to this input, and accordingly, uncertainty in this input may not add much to the uncertainty in the output of the model. In contrast, if sensitivity is high, inputs have to be well-founded on empirical data because model outputs depend strongly on them. Moreover, sensitive parameters indicate which processes control the dynamics of the model system for the chosen scenario.

Why is it important?

- Modelers need to make sure they understand how their main model predictions emerge. Sensitivity analysis is an important tool for this because it helps identifying control mechanisms and tipping points in control. Sensitivity analysis can also guide the improvement of the model structure: parameters with low sensitivity can, for the time being, be ignored in further model analyses; parameters with high sensitivity indicate that it might be worthwhile investing time and replacing that parameter by a submodel, from which that parameter, instead of being imposed (Grimm and Railsback, 2005), emerges dynamically and driven by low-level processes (e.g., physiology, metabolism, adaptive behavior, disturbance regimes, etc.).

- Model users need to see results of sensitivity analysis to better understand a model and to see how well sensitive parameters are supported by data. They also can check the overall sensitivity of model output to variation in model input. If the majority of parameters is highly sensitive, the

[4] Note that Rykiel calls this „validation".

model might not be useful for supporting decisions. If only a few parameters are highly sensitive, but have been thoroughly explored, or are relatively certain, model predictions are more robust and reliable.

How to "trace" it?

The tests used should be briefly explained, including their purpose and references to the literature. Test results are usually summarized in tables and figures. These should be thoroughly described and the corresponding conclusions summarized. Should all parameters be included in this exercise? Most sensitivity analyses presented in the modeling literature only include a subset of all model parameters because modelers tend to "hide" parameters they never considered for being varied. There can be good reasons for this, for example because the modeler wants to focus on certain scenarios, but these reasons, and all "hidden" parameters should be documented nevertheless, preferably in a separate table or paragraph.

Justification for chosen ranges and reference values of parameters should be included, or references made to parameter tables included in part 1.4 of the TRACE document ("Parameterization").

Benefits of TRACE documentation

Modelers: Sensitivity analysis is a central, indispensable tool for model analysis. Documenting all sensitivity tests thus facilitates systematic and comprehensive analysis of a model, and helps convincing model users in trusting the main model predictions.

Model users: Sensitivity analyses allow to quickly grasping the relative importance of model processes and whether or not important processes are linked to particularly uncertain parameters. The rationale of such analyses is largely independent of the model type used.

Modeling notebook and TRACE

Like any other major test, all sensitivity analyses should be documented in the modeling notebook because such tests usually require considerable time for design, computation, and interpretation. Compiling this part 2.2 of the TRACE documentation of the final model, that was used for making predictions, is then straightforward, in particular if the notebook includes for every test a short (verbal) description of the tests purpose, main results, and conclusions. The TRACE document should include a summary discussion of model sensitivity (and robustness) and an explanation why, or to what degree, model users can or should rely on model predictions despite of the reported sensitivities.

Assessment criteria

Simulation models should not be used for supporting decision making if they are not supported by thorough and clear sensitivity analyses. However, no simple criteria exist for what level of sensitivity can be accepted. Sensitivity and uncertainty cannot be avoided in models, but at least models allow quantifying relative sensitivities and uncertainties, which is not possible for verbal, or mind models. Model users should keep in mind that all models are simplifications, so wholly exact predictions cannot be expected. Rather, models provide general insights into the relative importance of processes and they allow for "relative predictions", for example ranking of alternative management scenarios.

Background and further reading

Note that the purpose of sensitivity analyses is not to show that sensitivity is negligible for all parameters (this is virtually never the case) but to learn how many parameters, or processes, are sensitive, and which ones, and why. However, if too many parameters are highly sensitive, model predictions can be too uncertain to allow for any clear conclusions, unless the parameters are more or less certain.

In “Sensitivity experiments”, a single parameter is varied over its range and the response of one or more outputs (referred to as “observations” in the ODD protocol) determined and usually plotted. In “Local Sensitivity Analyses”, one parameter at a time is slightly “perturbed”, e.g. by +/- 5% of its value, and the resulting change in model output determined. The ratio of the relative change in model output to relative change in model input can be used to compare sensitivities of different parameters (many modifications of this simple scheme exist; Railsback and Grimm, 2011). As a result, we obtain a table of model parameters and inputs and their corresponding sensitivities (as an example, see Table 1). Such tables inform us which processes are, for the reference parameter set chosen and for not too large changes in parameters, the most important ones in controlling model output.

Table 1. Example of a local sensitivity analysis. The model represents a grasshopper metapopulation inhabiting a network of gravel bars in an Alpine river. Parameter values were increased by 5 to 20%; sensitivity S+ was calculated as the ratio of relative change in the population’s intrinsic mean time to extinction to the relative change in parameter value (after Stelter et al., 1997).

Process and Parameter	**Meaning of parameter**	**Reference value**	**Quality of knowledge**[1]	**Sensitivity S^+**
N	Number of gravel bars	30	5	2.92
Succession				
K_{max}	Maximum habitat capacity (# grasshoppers)	50	5	9.87
K_{min}	Minimum habitat capacity (# grasshoppers)	5	2 - 3	**22.0**
Shrub	Loss of habitat capacity per year (# grasshoppers)	1	4	**-18.9**
Delay	Time until new habitats can be colonized (years)	3	4 - 5	-1.54
Floods				
Flood	Probability per year of flood event	0.1	3	-3.67
Wash	Probability of being washed away during flood event	0.25	3	-0.85
Local extinction				
Extinct	Probability of extinction due to environmental noise	0.1	1	-2.07
My(*K*)	Probability of extinction due to demographic noise	1/K	1	
Colonization				
Migrate	Probability that a female disperses to another bar	0.1	3	3.64

[1] Estimated quality of the empirical knowledge used to set the parameter value: 5 means highly certain knowledge, 1 means low certainty.

Sensitivity analysis can also be more global. A very useful technique is using contour plots: one parameter is varied along the x-, a second one along the y-axis. For each combination of values explored, the corresponding model output is plotted. If such contour plots are arranged in a matrix, the influence of four parameters can be visualized (Fig. 3).

Beyond contour plots, specific protocols for global sensitivity analyses exist (Campolongo et al., 2007; Kleijnen and Sargant, 2000; Kramer-Schadt et al., 2009; Saltelli et al., 2008a; 2008b). They usually include a protocol for sampling the “parameter space” (the set of all possible combination of all parameters) and a statistical analysis (Analysis of Variance, Generalized Linear Models, etc.). However, even with global sensitivity analysis, a complete, comprehensive exploration of the behavior of simulation models which have more than, say, five parameters is usually impossible. Narrowing down the parameters’ range, via calibration, can therefore be necessary.

A vast literature on sensitivity analysis exists. Note, however, that sensitivity analysis is not a technical exercise that can automatically be run by computers. Rather, it is part of the scientific attempt to try and understand how the model works. They need to be augmented by exploring the sensitivity, or robustness, of model behavior to changes in model structure. Nicolson et al. (2002) emphasize that “thorough sensitivity analysis involves testing not only different parameter values but also the assumptions and the effect of alternative educated guesses at the underlying processes".

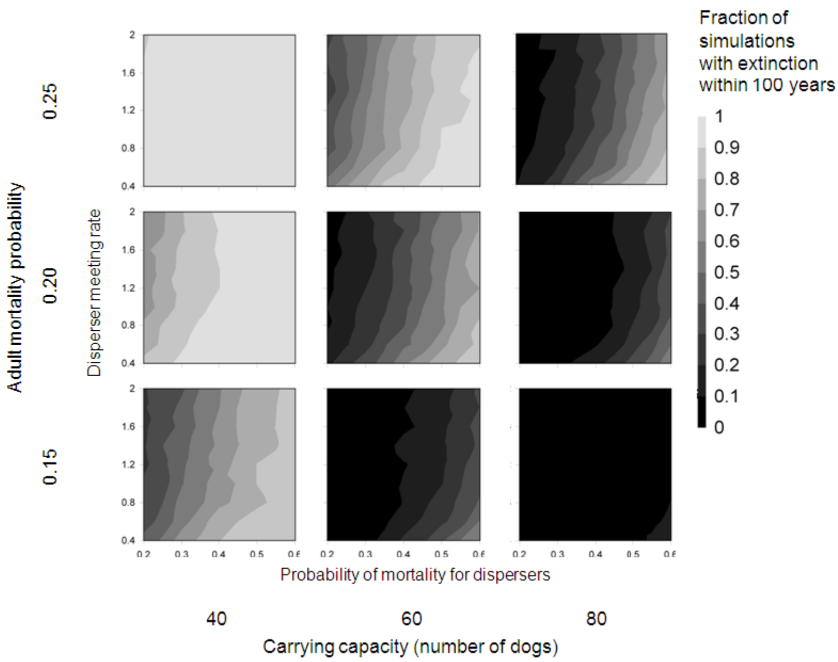


Figure 3. Example of using contour plots for sensitivity analysis. The fraction of extinctions within 100 years of 1,000 simulations is taken as the probability of extinction of a model of a small population of African wild dog (Gusset et al. 2009). Darker gray scales indicated higher risks of extinction. In each panel, two parameters are varied, as well as in each row and column. (After Gusset et al., 2009; Railsback and Grimm, 2011).

This is particularly important for simulation models, for example individual-based models, which are driven by one or a few key behaviors, for example foraging, habitat choice, competition with conspecifics, etc. We recommend testing alternative formulations of such submodels, usually including a simple “null model” and a more detailed formulation, to explore how robust the general insights gained with the model are with regard to such variations. Ideally, such tests will show that the chosen formulation is a good compromise between over-simplification (the null model leads to unrealistic behavior) and over-parameterization (the more detailed submodel does not lead to any new insight). Such tests of alternative formulations are part of virtually any modeling project, but so far they are usually not documented. However, if main tests are documented, model users learn that model design was not ad hoc, but based on thorough design decisions and tests. For a more detailed description of “Robustness Analysis” (exploring sensitivity to changes in model structure) and pattern-oriented model selection, see Grimm and Railsback (2005) and Railsback and Grimm (2011).

2.3 Validation

What is it?

Comparison of model or submodel outputs with *empirical data that were not used for parameterization or calibration*; documentation of *data sources*; what *parts (submodels) have been validated*; what *validation methods* were applied.

Validation is very similar to verification (TRACE element 2.1), the only difference being that for validation we use data sets that were not used for designing, parameterizing, and calibrating the model. With validation we test predictions devised from the model concerning properties of the real system that were not considered at all during model development and verification. Such predictions are therefore referred to as “secondary” or “independent” prediction. Note that in the current literature, such independent predictions are not very common.

Why is it important?

- Modelers want to make sure that the capture the internal organization of the real system they want to represent sufficiently well. If they can show that the model reproduces independent predictions, they can show that model predictions are robust and reliable and that the model can (cautiously) be extrapolated to new questions and scenarios.

- Model users can use validation as a very strong indicator that the model is reliable and robust (still within certain bounds).

How to "trace" it?

Virtually all what has been said on "tracing" verification also applies to validation, with the one difference that here it needs to be documented how certain patterns in the model were identified and how significant and robust these patterns are. Visualizations of model output and statistical methods play an important role here and should be documented. (It should be clear by now, that for any documented test, it is critical and mandatory to include information on model version and parameter values used, and the corresponding file names and where they are stored.) The relationship of the data sets used for validation to the ones used in model development, including verification, should be discussed.

Benefits of TRACE documentation

Modelers: Successful validation is the best a modeler can do to convince model users that they can put confidence in model predictions. For this, all elements of validation have to be convincing in themselves, which can be achieved by careful and thorough documentation.

Model users: A clear documentation of model validation considerably facilitates the assessment of the usefulness of a model for supporting decision making.

Modeling notebook and TRACE

See corresponding section TRACE element 2.1, "Verification".

Assessment criteria

Validation should be seen as an additional means to prove model reliability and robustness. For simple models, like matrix models, validation usually neither is possible nor necessary, because the purpose and scope of matrix models is different from, e.g., individual-based models (Grimm, 2010). The relevance of validation, but also the feasibility of model validation, increases with model complexity. Thus, the more complex a model, the more model users should base their assessment of a model on the degree to which it was validated.

Background and further reading

Validation is distinguished from verification because with the latter, modelers test whether the implemented model is working according to the specifications as determined by the conceptual model and the model description. In contrast, validation shows that the model can reproduce (to a sufficient extent) data sets that were disregarded during model development, i.e. that were not taken into account when parameterizing, calibrating, and verifying the model. Especially calibration implies the risk that a model is "forced" to reproduce the "right" observations for the wrong reasons. This risk can be reduced to a large degree during verification if more than one observed patterns is reproduced ("pattern-oriented modeling"), but still the best evidence that the internal organization of a real system has been captured sufficiently well is that the model correctly predicts system behaviors that the modeler did not consider and which therefore were not "forced" into the model. An example is the individual-based model of natural European beech forests "BEFORE" (Rademacher et al., 2004), which during development and testing only evaluated average properties of small spatial units, very much like in empirical surveys. Information about age, size, and distribution of individual canopy trees was never considered. Only after the first publication, including verification, was the model tested for patterns to be identified *in the model*. For example, it turned out that neighboring canopy trees had an average difference in age of about 60 years. This secondary prediction was confirmed by searching for corresponding information in the literature (Rademacher and Winter, 2003; Grimm et al., 2005).

The importance of using independent data for validation has been emphasized also by many others (Bart, 1995; Rykiel, 1996; Oriade and Dillon, 1997; Beissinger and Westphal, 1998; Brook et al., 2000; Lang et al., 2000; Wiegand et al., 2003; Holst et al., 2007; Jakeman et al., 2006; Glaser and Bridges, 2007; Janssen and van Ittersum, 2007), although some authors refer to "corroboration" (CREM, 2009) or "evaluation" (Guisan and Zimmermann, 2000; Prisley and Mortimer, 2004) instead. Rykiel (1996) resolves this problem presenting a comprehensive discussion of terminology and methods. Although validation is an essential step, validation results still are subject to interpretation (Glaser and Bridges, 2007). For instance, validation and calibration data sets may have been collected under very similar conditions, and thus, validation may not be very meaningful (Oriade and Dillon, 1997). Accordingly, the mere notion that a model has been validated is meaningless if no validation criteria are given (Rykiel, 1996). The power of validation is increased if secondary outputs are used for validation in addition to the main model results (Bart, 1995). "If the model is able to reproduce multiple patterns, each describing a different feature of the system, the risk that many different processes may have caused this particular combination of observed patterns is lower than when relying only on one pattern" (Wiegand et al., 2003).

3 MODEL APPLICATION

Part 1 of the modeling cycle, and TRACE documentation, focused on model formulation, Part 2 on model implementation and testing. In Part 3, the focus is on using the model for its original purpose: answering certain questions, or solving certain problems that would be difficult or impossible to answer by observing the real system or by real experiments. The model is taken as a – simplified – surrogate of the real system that can be experimentally manipulated as much as we want and observed from any angle and in any detail that is required.

3.1 Results

What is it?

Outputs that are used to *inform decisions*; description of *simulation experiments (scenarios)* conducted; *statistics applied* to analyze model outputs.

Generating results includes two aspects: getting results that help answering the original questions, and performing simulation experiments that help *understanding* how these answers emerge. Typical questions would be: How do changes in landscape structure affect population viability? How do sublethal effects of chemical stressors affect the ability of populations to cope with other stressors? Which of the available management options to increase the viability of a small population is likely to be most cost-effective?

Without understanding how these answers emerge, however, we would have to blindly trust predictions made by a "black box". This is unacceptable for decision making because the model might still, despite all the effort that went into its testing and analysis, contain implementation errors, unrealistic parameter values, inappropriate representation of processes and their interactions, and it might still miss key factors of the real system. Understanding means that we are able to clearly identify causal relationships between certain factors and scenarios, and model results. Understanding is, in contrast to correlative statistical models, a defining feature of mechanistic models. With simulation models it can be achieved to a large degree because in the model we can, in contrast to reality, perform any kind of controlled simulation experiment. In trying to understand a model, often unrealistic, simplified scenarios have to be explored, for example homogeneous landscapes, constant environments, identical individuals, random instead of adaptive behavior, etc. By analyzing such scenarios the relative importance of different elements of the models becomes clearer.

Why is it important?

- Modelers need to document "results", i.e. answers to the original question addressed with a model, because this is the very purpose of the model. They also need to demonstrate that they understand the model, otherwise confidence in model output will be limited.

- Model users want to know the "answers" to their original question so that they can use them to make better decisions. They also need to understand the causal relationships leading to these answers, because blind faith in model predictions cannot be the basis of decision making.

How to "trace" it?

This element of TRACE documentations partly overlaps with Part 2 because many simulation experiments that are performed to test the model, for example for sensitivity analysis, also significantly contribute to understanding how the model works. The main difference is that here the focus is on scenarios and output variables, or – as it is referred to in the ODD protocol – "observations" that were defined before the model existed, during problem and model formulation (Part 1). Of course, additional outputs, that help answering the model's question and understanding the model usually are developed during model analysis and should be documented as well. For example, in risk assessment of pesticides, regulatory "endpoints" for risk assessment at the population level are quite vague and need to be refined using well-tested models (Wang and Grimm, 2010).

Thus, this element of TRACE should be outlined along the list of questions to be addressed with the model that are listed in TRACE element 1.1, "Problem formulation". For each question, answers should be provided using figures, tables, and text, very much as in the results section of scientific publication. In addition, all main results should briefly but thoroughly be discussed, as in the discussion section of scientific publications: What do these results mean? What conclusions to they allow for? Is there supporting evidence for these conclusions in the literature? Are there opposing views or data? These discussions are important because model users cannot be expected to come up with their own conclusions: they need to know the modeler's interpretation of the results to be able to assess to what degree the model results should be taken into account in their decisions.

An important point is documenting evidence that the mechanisms leading to the main results are well understood. If, for example, mortality due to acute exposure to a certain pesticide leads to slightly lowered equilibrium population size in one type of landscape, but to continuous decline and extinction in another one: why is this so (Wang and Grimm, 2010)? Model users should ask the same question as software engineers who are testing their programs: "Is it a bug or a feature?". Thus, modelers should design and document controlled simulation experiments that clearly demonstrate, in a way that is assessable to non-modelers, how important model results emerge from the mechanisms and structure built into a model. Preferably, clear figures, tables, and – if applicable – statistical tests should be used for this purpose. For complex models, modelers should find and document simulation experiments that establish causal relationships between the behavior of the model's entities, for example individual organisms or small spatial units, and population-level dynamics. For this purpose, simplified scenarios, including smaller areas and fewer individuals, are very helpful (see chapter 22 in Railsback and Grimm, 2011).

Benefits of TRACE documentation

Modelers: Whereas part 2 of the modeling cycle often is not very well documented in scientific publications, “Results” are of course included because they are the very purpose of a model. Nevertheless, due to limitations of space often only main results are presented. In particular, simulation experiments that aim at providing insights into how the model works are often not included. Documenting such experiments, summarizing their results, and discussing in what way they help understanding the model and how the main “results” emerge will add considerably to the acceptance of the results by model users. For example, in wildlife epidemiology, to get model predictions regarding, e.g., rabies or classical swine fever accepted by decision makers it was not only critical that the models and their tests and analyses were well documented but also that model results could be well understood (Thulke and Grimm, 2010).

Model users: Understanding how the main results of a model emerge is critical for accepting a model as a tool for supporting decisions. TRACE documentations allow model users to know exactly where in the model documentation they can look for evidence that the model can be understood.

Modeling notebook and TRACE

Here, only analyses of the model should be included that are based on an established, well-tested, and fully parameterized version of the model (but possibly including simplified scenarios). Thus, after the model has sufficiently been tested and analyzed in Part 2 of the modeling cycle, its design is, for the time being, “frozen”, preferably by giving the model a version number, and answers to the original model question are sought for. Additional analyses for understanding this model version should be documented as well (and can be repeated for later versions). If important insights into the functioning of the model were obtained in Part 2, due references to those TRACE elements should be included here. (See also corresponding section for TRACE element 2.1, “Verification”.)

Assessment criteria

Do the result clearly include the scenarios and outputs that are listed in the problem formulation (TRACE element 1.1, “Problem formulation”)? Are the answers to the original questions clearly formulated? Are the mechanisms underlying these answers well understood? Is this understanding well documented? If the answer to any of these questions is “no”, the model should not be fully considered for supporting decisions. Note, however, that “well understood” is a relative term, which should not be taken as “full and comprehensive understanding” but as the establishment of causal relationships (“This is because..”) and evidence that this relationship exists.

Background and further reading

Virtually every monograph or textbook on ecological modeling includes a chapter on why it is important to understand a model, e.g. Haefner (2005), Railsback and Grimm (2011), Otto and Day (2007).

3.2 Uncertainty analysis

What is it?

Uncertainties in model outputs used for recommendations; description of *variance, noise, and bias in empirical data*; determination of *stochasticity* in the model; description of *model uncertainty* which can be assessed through application of different models or submodels; best- and worst-case scenarios.

Uncertainty analysis (UA) looks at how uncertainty in parameter values (and initial values of state variables) affects the reliability of model results. UA provides a way to see, therefore, which results of a model we should or should not have much confidence in. The difference to sensitivity analysis is that there, we usually vary parameters uniformly within a freely chosen range to understand how different model outputs respond to this parameter variation. In UA, we take the *given* uncertainty in parameters, including their expected *probability distributions*, and explore the consequences for uncertainties in model outputs that are *relevant for answering the models questions* (see TRACE n element 3.1, “Results”).

Note, however, that UA can be uncertain by itself, because we might not know the right probability distribution. Moreover, even if main results turn out to be so uncertain that absolute predictions of the model make little sense, the model might still be very useful for relative predictions, i.e. for comparing alternative management options (e.g., risk management), and for getting general insights. This is the typical situation in models developed for „Population Viability analysis“ in conservation biology, where scarcity of data virtually always is immanent to the problem to be solved. Note also that uncertainties of single parameters do not necessarily multiply to total uncertainty because parameters represent processes that might interact to the effect that simple „error propagation“ is buffered (Wiegand et al., 2004). Note also that the relevance of UA for a certain model is defined during „Problem formulation“: if absolute predictions are sought for, UA is decisive; if ranking of management scenarios is the main purpose of the model, UA is much less important and sensitivity analysis might be sufficient.

Why is it important?

- Modelers need to test and document uncertainty of model predictions regarding the questions addressed because otherwise model users would have no clue how much they can rely on absolute, quantitative predictions.

- Model users need to be informed about uncertainties because too uncertain predictions would be useless for supporting decision making.

How to “trace” it?

The same applies that has been said about how to “trace” sensitivity analysis above (TRACE element 2.2): Tests used should be briefly explained. Test results are usually summarized in tables and figures. These should be thoroughly described and the corresponding conclusions summarized.

Should all parameters be included in this exercise? Most uncertainty analyses presented in the modeling literature only include a subset of all model parameters because modelers tend to “hide” parameters they never considered for being varied. There can be good reasons for this, for example because the modeler wants to focus on certain scenarios, but these reasons, and all “hidden” parameters should be documented nevertheless, preferably in a separate table or paragraph.

Justification for chosen ranges and probability distributions of parameters should be included. If uncertainties of some, or many, parameters are very high: what are the conclusions? Does the modeler nevertheless believe that model can be useful for supporting decisions? Why? Is error propagation buffered by process interaction? Has this been shown via global sensitivity analysis? Has the uncertainty in the ranking of management scenarios been tested?

Benefits of TRACE documentation

Modelers: In domains where decision makers are used to rely on absolute threshold values of risk indicators, like in regulatory risk assessment of chemicals, model users will expect clear statements regarding uncertainty in model output. Documenting this uncertainty, or explaining why for a certain problem uncertainty is less relevant, is therefore critical for getting model results accepted by model users.

Model users: If absolute, numerical predictions are a main purpose of a model, knowing uncertainty of this prediction to uncertainties in model inputs (parameters, initial values) is critical.

Modeling notebook and TRACE

For every major version of model, UA usually is the last task, after main model results were obtained. UA should be documented just like sensitivity analyses. In particular, conclusions from early UA should be documented, as they often can lead to changes in model structure and to investing additional resources into parameterization and calibration.

Assessment criteria

The TRACE documentation should include comprehensive UA if a declared purpose of the model is to make quantitative predictions. If reported uncertainty is high, modelers should provide reasons in what way the model still can support decision making. Note, however, that it is impossible to provide absolute margins for acceptable uncertainties.

Background and further reading

Railsback and Grimm (2011) summarize the steps of UA:

- "Identify the parameters to include in the analyses. The most uncertain and sensitive parameter should be included, but since UA can be require high effort and computation time, often only the most sensitive or uncertain parameters can be included.
- For each parameter included, define a distribution of its values that describes the uncertainty it is believed to have. Distributions are characterized by the type (continuous vs. discrete), shape (e.g., uniform, normal, log-normal), and parameters (e.g., minimum and maximum; mean and standard deviation) of a stochastic function from which values will be drawn.
- Run the model many times, each time drawing new random values of the parameters from their distributions. Each such run therefore has a unique combination of values drawn from the parameter space.
- Analyze the distribution of model results produced by all the random parameter combinations. "

In the literature on modeling supporting environmental decision making, UA is generally listed as a critical element (Bart, 1995; Radomski and Goeman, 1996; Brook et al., 2000; Guisan and Zimmermann, 2000; Lang et al., 2000; Pastorok et al., 2002; Harwood and Stokes, 2003; Jakeman and Letcher, 2003; CREM, 2009; Wiegand et al., 2003; Jager et al., 2005; Jakeman et al., 2006; Glaser and Bridges, 2007; Holst et al., 2007; Scheller and Mladenoff, 2007; Fuller et al., 2008). The assessment of uncertainties is considered essential for the judgment of the usefulness of model outputs (Radomski and Goeman, 1996; Harwood and Stokes, 2003; CREM, 2009; Jakeman and Letcher, 2003). High uncertainties imply a low value of the model for management or regulatory decision (Radomski and Goeman, 1996), but can reveal gaps in the knowledge or data about the system (Wiegand et al., 2003). Uncertainties in the model output originate from uncertainties in the input data and from the inherent uncertainty in the model assumptions and structure (Harwood and Stokes, 2003; CREM, 2009; Scheller and Mladenoff, 2007).

Input data is burdened with process stochasticity and measurement errors. The model itself can never be an exact representation of the real system, and thus, adds to the uncertainty in the model output. In addition, Harwood and Stokes (2003) argue that an implementation error has to be taken into account, i.e. inaccuracy in the application of the recommendations derived from the model results. It has been argued that uncertainties may not present a problem if scenario comparisons are conducted instead of direct predictions (Glaser and Bridges, 2007). Uncertainties influence the different scenarios in the same way, and thus, do not affect their relative performance. However, Fuller et al. (2008) claim that this assumption may not apply in all cases, but scenarios may be differently influenced by uncertainties. Accordingly, uncertainties have to be quantified whether the model output consists of direct quantitative predictions or recommendations derived from relative assessment of alternative scenarios (Fuller et al., 2008).

In the literature, it remains a controversial issue if ecological models can provide quantitative predictions that can directly be used in the formulation for policies or decisions (Brook et al., 2000; Nelson et al., 2008). Brook et al. (2000) argues that quantitative predictions facilitate model validation with empirical data. In contrast, most authors emphasize the role of models in increasing understanding, and they recommend the comparison of alternative scenarios (Radomski and Goeman, 1996; Beissinger and Westphal, 1998; Fulton et al., 2003; Harwood and Stokes, 2003; Jakeman and Letcher, 2003; Perry and Enright, 2006; Glaser and Bridges, 2007; Holst et al., 2007; Pretzsch, 2007; Scheller and Mladenoff, 2007).

3.3 Recommendation

What is it?

Description of how *initial question(s) could be answered*; summary of *conclusions* drawn from model; clarification of *extrapolations* used (in time and space).

"Results", as described in TRACE element 3.1, refer to scientific results, or answers, to the originally posed scientific question or problem. In contrast, "recommendation" refers to the initially posed management, or decision, problem. Recommendations are conclusions that are based on scientific results, but which – as any conclusion – have also to be based on judgment and interpretation. Ideally, conclusions are formulated by modelers and model users together.

Why is it important?

- Modelers have to present scientific results in a way that allows drawing conclusions that are directly relevant for management. All too often, however, modelers are not fully aware of the management context and therefore deliver results that cannot be used by decision makers. For example, if the environmental risk of a new pesticide has to be assessed for population of certain species, model results focusing on population growth rate are not directly suitable for assessing risk and therefore possibly not very useful.

- Model users usually do not have solid, or even any, training in ecological modeling. If confronted only with scientific results, they may thus not be able to deduce possible consequences for their decision making. Thus, explicit formulation of management recommendations, which are based on scientific results, should be included in TRACE documentation.

How to "trace" it?

What are the recommendations regarding to management problem addressed? How are they based in the models scientific results? Are these recommendations directly relevant to the decisions to be made, or do they only refer to "proxies", i.e. outputs that are only vaguely related to the outputs which are of direct interest to model users? How are uncertainties taken into account in these recommendations?

Benefits of TRACE documentation

Explicit formulations of recommendations for decision making ensure that the transfer from scientific results to conclusions directly relevant for decision making has been tried.

Modeling notebook and TRACE

Ideally, recommendations are formulated together by modelers and model users. Modelers should start with, and document, first drafts of their recommendations which then should be discussed with model users. Bridging the gap between scientifically relevant model output and conclusions that are relevant for management can be a major challenge by itself. For example, for regulatory risk assessment of chemicals it is still unclear what population level metrics, or endpoints, should be used to assess “adverse effects” on populations: population growth rate, effects on average abundance, time to recovery, extinction risk, etc. (Wang and Grimm, 2010).

Assessment criteria

TRACE documentation should include explicit formulations of recommendations.

References

Bart, J. 1995. Acceptance criteria for using individual-based models to make management decisions. Ecol.Appl. **5**:411-420.

Bartell, S. M., R. A. Pastorok, H. R. Akcakaya, H. Regan, S. Ferson, and C. Mackay. 2003. Realism and relevance of ecological models used in chemical risk assessment. Human and Ecological Risk Assessment **9**:907-938.

Beissinger, S. R., and M. I. Westphal. 1998. On the use of demographic models of population viability in endangered species management. J.Wildl.Manage. **62**:821-841.

Boorman, D., R. Williams, M. Hutchins, E. Penning, S. Groot, and J. Icke. 2007. A model selection protocol to support the use of models for water management. Hydrol.Earth Syst.Sci. **11**:634-646.

Brook, B. W., J. J. O'Grady, A. P. Chapman, M. A. Burgman, H. R. Akcakaya, and R. Frankham. 2000. Predictive accuracy of population viability analysis in conservation biology. Nature **404**:385-387.

Campolongo, F., J. Cariboni, A. Saltelli. 2007. An effective screening design for sensitivity analysis of large models. Environmental Modeling & Software **22**: 1509-1518.

Denny-Gouldson, P. 2007. Meeting the specialized requirements of the biologist through a biology-focused electronic laboratory notebook. American Laboratory **39**:26-28.

Ferrier, S., and A. Guisan. 2006. Spatial modeling of biodiversity at the community level. J.Appl.Ecol. **43**:393-404.

Ferson, S. 1996. Automated quality assurance checks on model structure in ecological risk assessments. Human and Ecological Risk Assessment **2**:558-569.

Fuller, M. M., L. J. Gross, S. M. Duke-Sylvester, and M. Palmer. 2008. Testing the robustness of management decisions to uncertainty: Everglades restoration scenarios. Ecol.Appl. **18**:711-723.

Fulton, E. A., A. D. M. Smith, and C. R. Johnson. 2003. Effect of complexity on marine ecosystem models. Marine Ecology-Progress Series **253**:1-16.

Glaser, D., and T. S. Bridges. 2007. Separating the wheat from the chaff: the effective use of mathematical models as decision tools. Integrated Environmental Assessment and Management **3**:442-449.

Council for Regulatory Environmental Modeling (CREM). 2009. Guidance on the development, evaluation, and application of environmental models. U.S. Environmental Protection Agency, EPA/100/K-09/003 (March 2009), Washington DC. http://www.epa.gov/crem/library/cred_guidance_0309.pdf

Goss-Custard, J., Burton, N., Clark, N., Ferns, P., McGrorty, S., Reading, C., Rehfisch, M., Stillman, R., Townend, I., West, A. and Worrall, D. 2006. Test of a behavior-based individual-based model: response of shorebird mortality to habitat loss. Ecological Applications: **16**, 2215-2222.

Grimm, V., and S. F. Railsback. 2005. *Individual-Based Modeling and Ecology*. Princeton University Press, Princeton.

Grimm, V., E. Revilla, U. Berger, F. Jeltsch, W.M. Mooij, S.F. Railsback, H.-H. Thulke, J. Weiner, T. Wiegand, and D. L. DeAngelis DL . 2005 Pattern-oriented modeling of agent-based complex systems: lessons from ecology. Science **310**: 987-991.

Grimm, V., U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. Heinz, G. Huse, A. Huth, J. U. Jepsen, C. Jørgensen, W. M. Mooij, B. Müller, G. Pe'er, C. Piou, S. F. Railsback, A. M. Robbins, M. M. Robbins, E. Rossmanith, N. Rüger, E. Strand, S. Souissi, R. A. Stillman, R. Vabø, U. Visser, and D. L. DeAngelis. 2006. A standard protocol for describing individual-based and agent-based models. Ecological Modeling **198**:115-126.

Grimm, V. 2010. Short introduction to ecological modeling. In: Thorbek, P., V. Forbes, F. Heimbach, U. Hommen, H.-H. Thulke, P. J. van den Brink, j. Wogram, V. Grimm(eds). *Ecological models for regulatory risk assessments of pesticides: developing a strategy for the future*. Pensacola and Boca Raton (FL): Society of Environmental Toxicology and Chemistry (SETAC) and CRC Press, pp. 15-26.

Grimm, V., U. Berger, D. L DeAngelis, G. Polhill, J. Giske, S. F. Railsback. 2010. The ODD protocol: a review and first update. Ecological Modeling **221**: 2760-2768.

Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecological Modeling **135**:147-186.

Gusset, M., O. Jakoby, M.S. Müller, M.J. Somers, R. Slotow, and V. Grimm. 2009. Dogs on the catwalk: modeling the re-introduction of endangered wild dogs in South Africa. Biological Conservation **142**: 2774-2781.

Haefner J. W. 2005. *Modeling Biological Systems*. Springer, New York.

Harwood, J., and K. Stokes. 2003. Coping with uncertainty in ecological advice: lessons from fisheries. Trends Ecol.Evol. **18**:617-622.

Holst, N., I. A. Rasmussen, and L. Bastiaans. 2007. Field weed population dynamics: a review of model approaches and applications. Weed Research **47**:1-14.

Jackson, L. J., A. S. Trebitz, and K. L. Cottingham. 2000. An introduction to the practice of ecological modeling. BioScience **50**:694-706.

Jager, T., T. Crommentuijn, C.A.M. van Gestel, and S.A.L.M. Kooijman. 2004. Simultaneous modeling of multiple endpoints in life-cycle toxicity tests. Environ. Sci. Technol. **38**:2894–900

Jager, H. I., A. W. King, N. H. Schumaker, T. L. Ashwood, and B. L. Jackson. 2005. Spatial uncertainty analysis of population models. Ecological Modeling **185**:13-27.

Jakeman, A. J., and R. A. Letcher. 2003. Integrated assessment and modeling: features, principles and examples for catchment management. Environ.Model.Softw. **18**:491-501.

Jakeman, A. J., R. A. Letcher, and J. P. Norton. 2006. Ten iterative steps in development and evaluation of environmental models. Environ.Model.Softw. **21**:602-614.

Janssen, S., and M. K. van Ittersum. 2007. Assessing farm innovations and responses to policies: A review of bio-economic farm models. Agric.Syst. **94**:622-636.

Jongejans, E., O. Skarpaas, and K. Shea. 2008. Dispersal, demography and spatial population models for conservation and control management. Perspect.Plant Ecol.Evol.Syst. **9**:153-170.

Kanare H. M. 1985. Writing the Laboratory Notebook. American Chemical Society, Washington, D.C.

Kleijnen J. P. C. , and R. G. Sargent. 2000. A methodology for _tting and validating metamodels in simulation. European Journal of Operational Research **120**: 14-29.

Kramer-Schadt, S., M. Fernándéz, V. Grimm, and H.-H. Thulke. 2009. Individual variation in infectiousness explains long-term disease persistence in wildlife populations. Oikos **118**: 199-208.

Landsberg, J. 2003. Modeling forest ecosystems: state of the art, challenges, and future directions. Can.J.Forest Res. **33**:385-397.

Lang, Railsback & Associates. 2000. Instream flow assessment methods: guidance for evaluating instream flow needs in hydropower licensing. EPRI (Electric Power Research Institute, Inc.), Palo Alto, CA: 2000. 1000554.

Mooij, W. M., and D. L. DeAngelis. 2003. Uncertainty in spatially explicit animal dispersal models. Ecol.Appl. **13**:794-805.

Nelson, W. A., A. Potapov, M. A. Lewis, and A. E. Hundsdörfer. 2008. Balancing ecological complexity in predictive models: a reassessment of risk models in the mountain pine beetle system. J.Appl.Ecol. **45**:248-257.

Nicolson, C. R., A. M. Starfield, G. P. Kofinas, and J. A. Kruse. 2002. Ten heuristics for interdisciplinary modeling projects. Ecosystems **5**:376-384.

Oriade, C. A., and C. R. Dillon. 1997. Developments in biophysical and bioeconomic simulation of agricultural systems: a review. Agric.Econ. **17**:45-58.

Otto S. P., and T. Day. 2007. A Biologist's Guide to Mathematical Modeling in Ecology and Evolution. Princeton University Press, Princeton, NJ.

Parker, P., R. Letcher, A. Jakeman, M. B. Beck, G. Harris, R. M. Argent, M. Hare, C. Pahl-Wostl, A. Voinov, M. Janssen, P. Sullivan, M. Scoccimarro, A. Friend, M. Sonnenshein, D. Baker, L. Matejicek, D. Odulaja, P. Deadman, K. Lim, G. Larocque, P. Tarikhi, C. Fletcher, A. Put, T. Maxwell, A. Charles, H. Breeze, N. Nakatani, S. Mudgal, W. Naito, O. Osidele, I. Eriksson, U. Kautsky, E. Kautsky, B. Naeslund, L. Kumblad, R. Park, S. Maltagliati, P. Girardin, A. Rizzoli, D. Mauriello, R. Hoch, D. Pelletier, J. Reilly, R. Olafsdottir, and S. Bin. 2002. Progress in integrated assessment and modeling. Environ.Model.Softw. **17**:209-217.

Pastorok RA, Bartell SM, Ferson S. 2002. Ecological Modeling in Risk Assessment: Chemical Effects on Populations, Ecosystems, and Landscapes.Lewis, Boca Raton, FL, USA.

Perry, G. L. W., and N. J. Enright. 2006. Spatial modeling of vegetation change in dynamic landscapes: a review of methods and applications. Progr.Phys.Geogr. **30**:47-72.

Petersen, J. H., D. L. DeAngelis, and C. Paukert. 2008. An overview of methods for developing bioenergetic and life history models for rare and endangered species. Transactions of the American Fisheries Society **137**:244-253.

Pitt, W. C., P. W. Box, and F. Knowlton. 2003. An individual-based model of canid populations: modelling territoriality and social structure. Ecological Modelling **166**:109-121.

Pretzsch, H. 2007. Analysing and modeling forest stand dynamics for practical application - an european review and perspective -. Eurasian Journal of Forest Research **10**:1-17.
Prisley, S. P., and M. J. Mortimer. 2004. A synthesis of literature on evaluation of models for policy applications, with implications for forest carbon accounting. Forest Ecol.Manag. **198**:89-103.
Rabin, R., and E. Pastorius. 2005. Electronic laboratory notebook systems. Genetic Engineering News **25**.
Rademacher, C., C. Neuert, V. Grundmann, C. Wissel, and V. Grimm V. 2004. Reconstructing spatiotemporal dynamics of central European natural beech forests: the rule-based model BEFORE. Forest Ecology and Management **194**: 349-368.
Rademacher, C., and S. Winter. 2003. Totholz im Buchen-Urwald: Generische Vorhersagen des Simulationsmodels BEFORE-CWD zur Menge, räumlichen Verteilung und Verfügbarkeit. Forstwissenschaftliches Centralblatt **122**:337-357.
Radomski, P. J., and T. J. Goeman. 1996. Decision making and modeling in freshwater sport-fisheries management. Fisheries **21**:14-21.
Railsback, S., B. Harvey, J. Hayes, and K. LaGory. 2005. Tests of a theory for diel variation in salmonid feeding activity and habitat use. Ecology **86**: 947-959.
Railsback, S.F., and V. Grimm V. 2011 *Agent-based and Individual-based Modeling: A Practical Introduction*. Princeton University Press, Princeton, N.J. (to appear October 2011)
Rykiel, E. J. 1996. Testing ecological models: the meaning of validation. Ecological Modeling **90**:229-244.
Saltelli, A., K. Chan, and E. M. Scott (eds.) 2008a. *Sensitivity analysis*. Wiley-Interscience, New York.
Saltelli, A., M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola (eds.) 2008b. *Global sensitivity analysis: The primer*. Wiley-Interscience, New York.
Scheller, R. M., and D. J. Mladenoff. 2007. An ecological classification of forest landscape simulation models: tools and strategies for understanding broad-scale forested ecosystems. Landscape Ecol. **22**:491-505.
Schmolke, A., P. Thorbek, P. Chapman, and V. Grimm. 2010a. Ecological modeling and pesticide risk assessment: a review of current modeling practice. Environmental Toxicology and Chemistry **29**: 1006-1012.
Schmolke, A., P. Thorbek, D.L. DeAngelis, and V. Grimm. 2010b. Ecological modeling supporting environmental decision making: a strategy for the future. *Trends in Ecology and Evolution* 25: 479-486.
Stelter, C., M. Reich, V. Grimm, and C. Wissel. 1997. Modeling persistence in dynamic landscapes: lessons from a metapopulation of the grasshopper *Bryodema tuberculata*. Journal of Animal Ecology **66**:508-518.
Thulke H-H, Grimm V. 2010. Ecological models supporting management of wildlife diseases. In: Thorbek, P., V. Forbes, F. Heimbach, U. Hommen, H.-H. Thulke, P. J. van den Brink, j. Wogram, V. Grimm(eds). *Ecological models for regulatory risk assessments of pesticides: developing a strategy for the future*. Pensacola and Boca Raton (FL): Society of Environmental Toxicology and Chemistry (SETAC) and CRC Press, pp. 67-76.
Wang, M., and V. Grimm. 2007. Home range dynamics and population regulation: an individual-based model of the common shrew. Ecological Modeling **205**: 397-409.
Wang, M., and V. Grimm. 2010. Population models in pesticide risk assessment: lessons for assessing population-level effects, recovery, and alternative exposure scenarios from modeling a small mammal. Environmental Toxicology and Chemistry **29**: 1292-1300
Wennergren, U., M. Ruckelshaus, and P. Kareiva. 1995. The promise and limitations of spatial models in conservation biology. Oikos **74**:349-356.
Wiegand, T., F. Jeltsch, I. Hanski, and V. Grimm. 2003. Using pattern-oriented modeling for revealing hidden information: a key for reconciling ecological theory and application. Oikos **100**:209-222.
Wiegand, T., E. Revilla, and F. Knauer. 2004. Dealing with uncertainty in spatially explicit population models. Biodiversity and Conservation **13**: 53-78.

Appendix A: Template for writing ODD model descriptions

This is a supplement to: Grimm V, Berger U, DeAngelis DL, Polhill JG, Giske J, Railsback SF. 2010. The ODD protocol: a review and first update. Ecological Modeling **221**: 2760-2768.

This supplement can be used as a template for writing ODD model descriptions of individual-/agen-based or any other type of structured simulation model. It contains Section 3 of the publication. After reading the explanations and typing the answers to the question, ODD users should have a clear and complete ODD model description of their individual- or agent-based models. Questions and explanations (and underlines) should, of course, be deleted then.

ODD Template

[5]The model description follows the ODD (Overview, Design concepts, Details) protocol for describing individual- and agent-based models (Grimm et al. 2006, 2010).

1. Purpose

Question: What is the purpose of the model?

Answer: ...

Explanation: Every model has to start from a clear question, problem, or hypothesis. Therefore, ODD starts with a concise summary of the overall objective(s) for which the model was developed. Do not describe anything about how the model works here, only what it is to be used for. We encourage authors to use this paragraph independently of any presentation of the purpose in the introduction of their article, since the ODD protocol should be complete and understandable by itself and not only in connection with the whole publication (as it is also the case for figures, tables and their legends). If one of the purposes of a model is to expand from basic principles to richer representation of real-world scenarios, this should be stated explicitly.

2. Entities, state variables, and scales

Questions: What kinds of entities are in the model? By what state variables, or attributes, are these entities characterized? What are the temporal and spatial resolutions and extents of the model?

Answer: ...

Explanation: An entity is a distinct or separate object or actor that behaves as a unit and may interact with other entities or be affected by external environmental factors. Its current state is characterized by its state variables or attributes. A state variable or attribute is a variable that distinguishes an entity from other entities of the same type or category, or traces how the entity changes over time. Examples are weight, sex, age, hormone level, social rank, spatial coordinates or which grid cell the entity is in, model parameters characterizing different types of agents (e.g., species), and behavioral strategies. The entities of an ABM are thus characterized by a set, or vector (Chambers, 1993; Huse et al., 2002), of attributes, which can contain both numerical variables and references to behavioral strategies.

One way to define entities and state variables is the following: if you want (as modelers often do) to stop the model and save it in its current state, so it can be re-started later in exactly the same state, what kinds of information must you save?

If state variables have units, they should be provided. State variables can change in the course of time (e.g. weight) or remain constant (e.g. sex, species-specific parameters, location of a non-mobile entity). State variables should be low level or elementary in the sense that they cannot be calculated from other state variables. For example, if farmers are represented by grid cells which have certain spatial coordinates, the distance of a farmer to a certain service centre would not be a state variable because it can be calculated from the farmer’s and service centre’s positions.

Most ABMs include the following types of entities:

Agents/individuals. A model can have different types of agents; for example, wolves and sheep, and even different sub-types within the same type, for example different functional types of plants or different life stages of animals. Examples of types of agents include the following: organisms, humans, or institutions. Example state variables include: identity number (i.e., even if all other state variables would be the same, the agent would still maintain a unique identity), age, sex, location (which may just be the grid cell it occupies instead of coordinates), size, weight, energy reserves, signals of fitness, type of land use, political opinion, cell type,

[5] References are given in Grimm et al. (2010).

species-specific parameters describing, for example, growth rate and maximum age, memory (e.g., list of friends or quality of sites visited the previous 20 time steps), behavioral strategy, etc.

Spatial units (e.g., grid cells). Example state variables include the following: location, a list of agents in the cell, and descriptors of environmental conditions (elevation, vegetation cover, soil type, etc.) represented by the cell. In some ABMs, grid cells are used to represent agents: the state and behavior of trees, businesses, etc., that can be modeled as characteristics of a cell. Some overlap of roles can occur. For example, a grid cell may be an entity with its own variables (e.g., soil moisture content, soil nutrient concentration, etc., for a terrestrial cell), but may also function as a location, and hence an attribute, of an organism.

Environment. While spatial units often represent environmental conditions that vary over space, this entity refers to the overall environment, or forces that drive the behavior and dynamics of all agents or grid cells. Examples of environmental variables are temperature, rainfall, market price and demand, fishing pressure, and tax regulations.

Collectives. Groups of agents can have their own behaviors, so that it can make sense to distinguish them as entities; for example, social groups of animals, households of human agents, or organs consisting of cells. A collective is usually characterized by the list of its agents, and by specific actions that are only performed by the collective, not by their constitutive entities.

In describing spatial and temporal scales and extents (the amount of space and time represented in a simulation), it is important to specify what the model's units represent in reality. For example: "One time step represents one year and simulations were run for 100 years. One grid cell represents 1 ha and the model landscape comprised 1,000 x 1,000 ha; i.e., 10,000 square kilometers".

3. Process overview and scheduling

Questions: Who (i.e., what entity) does what, and in what order? When are state variables updated? How is time modeled, as discrete steps or as a continuum over which both continuous processes and discrete events can occur? Except for very simple schedules, one should use pseudo-code to describe the schedule in every detail, so that the model can be re-implemented from this code. Ideally, the pseudo-code corresponds fully to the actual code used in the program implementing the ABM.

Answer: ...

Explanation: The "does what?" in the first question refers to the model's processes. In this ODD element only the self-explanatory names of the model's processes should be listed: 'update habitat', 'move', 'grow', 'buy', 'update plots', etc. These names are then the titles of the submodels that are described in the last ODD element, 'Submodels'. Processes are performed either by one of the model's entities (for example: 'move'), or by a higher-level controller that does things such as updating plots or writing output to files. To handle such higher-level processes, ABM software platforms like Swarm (Minar et al., 1996) and NetLogo (Wilensky, 1999) include the concept of the 'Model', or 'Observer', itself; that is, a controller object that performs such processes.

By "in what order?" we refer to both the order in which the different processes are executed and the order in which a process is performed by a set of agents. For example, feeding may be a process executed by all the animal agents in a model, but we must also specify the order in which the individual animals feed; that is, whether they feed in random order, or fixed order, or size-sorted order. Differences in such ordering can have a very large effect on model outputs (Bigbee et al., 2006; Caron-Lormier et al., 2008).

The question of when variables are updated includes the question of whether a state variable is immediately assigned a new value as soon as that value is calculated by a process (asynchronous updating), or whether the new value is stored until all agents have executed the process, and then all are updated at once (synchronous updating). Most ABMs represent time simply by using time steps: assuming that time moves forward in chunks. But time can be represented in other ways (Grimm and Railsback, 2005, Chapter 5). Defining a model's schedule includes stating how time is modeled, if it is not clear from the 'Entities, State Variables, and Scales' element.

4. Design concepts

Questions: There are eleven design concepts. Most of these were discussed extensively by Railsback (2001) and Grimm and Railsback (2005; Chapter. 5), and are summarized here via the following questions:

Basic principles. Which general concepts, theories, hypotheses, or modeling approaches are underlying the model's design? Explain the relationship between these basic principles, the complexity expanded in this model, and the purpose of the study. How were they taken into account? Are they used at the level of submodels (e.g., decisions on land use, or foraging theory), or is their scope the system level (e.g., intermediate disturbance hypotheses)? Will the model provide insights about the basic principles themselves, i.e. their scope, their usefulness in real-world scenarios, validation, or modification (Grimm, 1999)? Does the model use new, or previously developed, theory for agent traits from which system dynamics emerge (e.g., 'individual-based theory' as described by Grimm and Railsback [2005; Grimm et al., 2005])?

Answer: ...

Emergence. What key results or outputs of the model are modeled as emerging from the adaptive traits, or *behaviors*, of individuals? In other words, *what* model results are expected to vary in complex and perhaps unpredictable ways when particular characteristics of individuals or their environment change? Are there other results that are more tightly imposed by model rules and hence less dependent on what individuals do, and hence 'built in' rather than emergent results?

Answer: ...

Adaptation. What adaptive traits do the individuals have? What rules do they have for making decisions or changing behavior in response to changes in themselves or their environment? Do these traits explicitly seek to increase some measure of individual success regarding its objectives (e.g., "move to the cell providing fastest growth rate", where growth is assumed to be an indicator of success; see the next concept)? Or do they instead simply cause individuals to reproduce observed behaviors (e.g., "go uphill 70% of the time") that are implicitly assumed to indirectly convey success or fitness?

Answer: ...

Objectives. If adaptive traits explicitly act to increase some measure of the individual's success at meeting some objective, what exactly is that objective and how is it measured? When individuals make decisions by ranking alternatives, what criteria do they use? Some synonyms for 'objectives' are 'fitness' for organisms assumed to have adaptive traits evolved to provide reproductive success, 'utility' for economic reward in social models or simply 'success criteria'. (Note that the objective of such agents as members of a team, social insects, organs—e.g., leaves—of an organism, or cells in a tissue, may not refer to themselves but to the team, colony or organism of which they are a part.)

Answer: ...

Learning. Many individuals or agents (but also organizations and institutions) change their adaptive traits over time as a consequence of their experience? If so, how?

Answer: ...

Prediction. Prediction is fundamental to successful decision-making; if an agent's adaptive traits or learning procedures are based on estimating future consequences of decisions, how do agents predict the future conditions (either environmental or internal) they will experience? If appropriate, what internal models are agents assumed to use to estimate future conditions or consequences of their decisions? What tacit or hidden predictions are implied in these internal model assumptions?

Answer: ...

Sensing. What internal and environmental state variables are individuals assumed to sense and consider in their decisions? What state variables of which other individuals and entities can an individual perceive; for example, signals that another individual may intentionally or unintentionally send? Sensing is often assumed to be local, but can happen through networks or can even be assumed to be global (e.g., a forager on one site sensing the resource levels of all other sites it could move to). If agents sense each other through social networks, is the structure of the network imposed or emergent? Are the mechanisms by which agents obtain information modeled explicitly, or are individuals simply assumed to know these variables?

Answer: ...

Interaction. What kinds of interactions among agents are assumed? Are there direct interactions in *which* individuals encounter and affect others, or are interactions indirect, e.g., via competition for a mediating resource? If the interactions involve communication, how are such communications represented?

Answer: ...

Stochasticity. What processes are modeled by assuming they are random or partly random? Is stochasticity used, for example, to reproduce variability in processes for which it is unimportant to *model* the actual causes of the variability? Is it used to cause model events or behaviors to occur with a specified frequency?

Answer: ...

Collectives. Do the individuals form or belong to aggregations that affect, and are affected by, the individuals? Such collectives can be an important intermediate level of organization in an ABM; examples include social groups, fish schools and bird flocks, and human networks and organizations. How are collectives represented? Is a particular collective an emergent property of the *individuals*, such as a flock of birds that assembles as a result of individual behaviors, or is the collective simply a definition by the modeler, such as the set of individuals with certain properties, defined as a separate *kind* of entity with its own state variables and traits?

Answer: ...

Observation. What data are collected from the ABM for testing, understanding, and analyzing it, and how and when are they collected? Are all output data freely used, or are only certain data *sampled* and used, to imitate what can be observed in an empirical study ("Virtual Ecologist" approach; Zurell et al., 2010)?

Answer: ...

Explanation: The 'Design concepts' element of the ODD protocol does not describe the model *per se*; i.e., it is not needed to replicate a model. However, these design concepts tend to be characteristic of ABMs, though certainly not exclusively. They may also be crucial to interpreting the output of a model, and they are not described well via traditional model description techniques such as equations and flow charts. Therefore, they are included in ODD as a kind of checklist to make sure that important model design decisions are made consciously and that readers are aware of these decisions (Railsback, 2001; Grimm and Railsback, 2005). For example, almost all ABMs include some kinds of adaptive traits, but if these traits do not use an explicit objective measure the 'Objectives' and perhaps 'Prediction' concepts are not relevant (though many ABMs include hidden or implicit predictions). Also, many ABMs do not include learning or collectives. *Unused concepts can be omitted in the ODD description*.

There might be important concepts underlying the design of an ABM that are not included in the ODD protocol. If authors feel that it is important to understand a certain new concept to understand the design of their model, they should give it a short name, clearly announce it as a design concept not included in the ODD protocol, and present it at the end of the Design concepts element.

5. Initialization

Questions: What is the initial state of the model world, i.e., at time t = 0 of a simulation run? In detail, how many entities of what type are there initially, and what are the exact values of their state variables (or how were they set stochastically)? Is initialization always the same, or is it allowed to vary among simulations? Are the initial values chosen arbitrarily or based on data? References to those data should be provided.

Answer: ...

Explanation: Model results cannot be accurately replicated unless the initial conditions are known. Different models, and different analyses using the same model, can of course depend quite differently on initial conditions. Sometimes the purpose of a model is to analyze consequences of its initial state, and other times modelers try hard to minimize the effect of initial conditions on results.

6. Input data

Question: Does the model use input from external sources such as data files or other models to represent processes that change over time?

Answer: ...

Explanation: In models of real systems, dynamics are often driven in part by a time series of environmental variables, sometimes called external forcings; for example annual rainfall in semi-arid savannas (Jeltsch et al., 1996). "Driven" means that one or more state variables or processes are affected by how these environmental variables change over time, but these environmental variables are not themselves affected by the internal variables of the model. For example, rainfall may affect the soil moisture variable of grid cells and, therefore, how the recruitment and growth of trees change. Often it makes sense to use observed time series of environmental variables so that their statistical qualities (mean, variability, temporal autocorrelation, etc.) are realistic. Alternatively, external models can be used to generate input, e.g. a rainfall time series (Eisinger and Wiegand, 2008). Obviously, to replicate an ABM, any such input has to be specified and the data or models provided, if possible. (Publication of input data for some social simulations can be constrained by confidentiality considerations.) If a model does not use external data, this element should nevertheless be included, using the statement: "The model does not use input data to represent time-varying processes." Note that 'Input data' does *not* refer to parameter values or initial values of state variables.

7. Submodels

Questions: What, in detail, are the submodels that represent the processes listed in 'Process overview and scheduling'? What are the model parameters, their dimensions, and reference values? How were submodels designed or chosen, and how were they parameterized and then tested?

Answer: ...

Explanation: The submodels are presented in detail and completely. The factual description of the submodel, i.e., equation(s) and algorithms, should come first and be clearly separated from additional information. From what previous model this submodel was taken or whether a new submodel was formulated, and why, can be explained. If parameterization is not discussed outside the ODD description, it can be included here. The parameter definitions, units, and values used (if relevant) should be presented in tables.

Any description of an ABM and its submodels will seem ad hoc and lack credibility if there is no justification for why and how formulations were chosen or how new formulations were designed and tested. Because agent-based modeling is new and lacks a firm foundation of theory and established methods, we expect ODD descriptions to include appropriate levels of explanation and justification for the design decisions they illustrate, though this should not interfere with the primary aim of giving a concise and readable account of the model. Justification can be very brief in the Overview and Design concepts sections, but the complete description of submodels is likely to include references to relevant literature, as well as independent implementation, testing, calibration, and analysis of submodels.

ODD-based model descriptions consist of the seven elements described above; however, in most cases it will be necessary to have a simulation experiments or model analysis section following the model description (see Discussion).

Appendix B: Example model description using the ODD protocol

The following list of publications has been identified in the Review by Grimm et al. (2010) as including good examples of the use of the ODD protocol for describing individual- or agent-based models:

Banitz T, Huth A, Grimm V, Johst K (2008) Clumped versus scattered: how does the spatial correlation of disturbance events affect biodiversity? *Theoretical Ecology* 1: 231-240

Beaudouin R, Monod G, Ginot V (2008) Selecting parameters for calibration via sensitivity analysis: An indlividual-based model of mosquitofish population dynamics. *Ecological Modelling* 218: 29-48

Charles S, Subtil F, Kielbassa J, Pont D (2008) An individual-based model to describe a bullhead population dynamics including temperature variations. *Ecological Modelling* 215: 377-392

Dur G, Souissi S, Devreker D, Ginot V, Schmitt FG, Hwang JS (2009) An individual-based model to study the reproduction of egg bearing copepods: Application to Eurytemora affinis (Copepoda Calanoida) from the Seine estuary, France. *Ecological Modelling* 220: 1073-1089

Giacomini HC, De Marco P, Petrere M (2009) Exploring community assembly through an individual-based model for trophic interactions. *Ecological Modelling* 220: 23-39

Gusset M, Jakoby O, Müller MS, Somers MJ, Slotow R, Grimm V (2009) Dogs on the catwalk: Modelling re-introduction and translocation of endangered wild dogs in South Africa. *Biological Conservation* 142: 2774-2781

Hellweger FL (2008) The role of inter-generation memory in diel phytoplankton division patterns. *Ecological Modelling* 212: 382-396

Jovani R, Grimm V (2008) Breeding synchrony of colonial birds: from local stress to global harmony. *Proceedings of the Royal Society B-Biological Sciences* 275: 1557-1563

Kristiansen T, Lough RG, Werner FE, Broughton EA, Buckley LJ (2009) Individual-based modeling of feeding ecology and prey selection of larval cod on Georges Bank. *Marine Ecology-Progress Series* 376: 227-243

Le Fur J, Simon P (2009) A new hypothesis concerning the nature of small pelagic fish clusters An individual-based modelling study of Sardinella aurita dynamics off West Africa. *Ecological Modelling* 220: 1291-1304

Meyer KM, Wiegand K, Ward D, Moustakas A (2007) SATCHMO: A spatial simulation model of growth, competition, and mortality in cycling savanna patches. *Ecological Modelling* 209: 377-391

Meyer KM, Vos M, Mooij WM, Hol WHG, Termorshuizen AJ, Vet LEM, van der Putten WH (2009) Quantifying the impact of above- and belowground higher trophic levels on plant and herbivore performance by modeling. *Oikos* 118: 981-990

Pagel J, Fritzsch K, Biedermann R, Schröder B (2008) Annual plants under cyclic disturbance regimes: Better understanding through model aggregation. *Ecological Applications* 18: 2000-2015

Piou P, Berger U, Hildenbrandt H, Grimm V, Diele K, D'Lima C (2007) Simulating cryptic movements of a mangrove crab: recovery phenomena after small scale fishery. *Ecological Modelling* 205: 110-122

Strand E, Huse G (2007) Vertical migration in adult Atlantic cod (Gadus morhua). *Canadian Journal of Fisheries and Aquatic Sciences* 64: 1747-1760

van Nes EH, Noordhuis R, Lammens EHHR, Portieje R, Reeze B, Peeters ETM (2008) Modelling the effects of diving ducks on zebra mussels Dreissena polymorpha in lakes. *Ecological Modelling* 211: 481-490