

Chapter 17

Designing, Formulating, and Communicating Agent-Based Models

1
2
3

Volker Grimm and Steven F. Railsback

4

Abstract Designing, formulating, and communicating agent-based models (ABMs) poses unique challenges, especially how to choose the right level of complexity and how to describe – and, even, think about – a model in a way that captures both its essential characteristics and its complete detail. Two techniques for dealing with such challenges have become established among agent-based modellers. The “ODD” (Overview, Design concepts, Details) protocol is a standard for describing ABMs in publications, but also provides design patterns for the model developer. ODD starts with an overview of what the model is and does, and then describes how the model implements ten “design concepts” that capture essential and unique characteristics of ABMs. Last come all the details needed to completely replicate the model. “Pattern-oriented modelling” (POM) is a set of strategies for using patterns observed in the systems to ensure that an ABM captures the right “essence” of the system. POM starts with identifying multiple patterns of behaviour in the real system and its agents that seem to capture the essential internal mechanisms for the problem being modelled. These patterns are then used to decide what kinds of entities, state variables, and processes need to be in the model; compare and test alternative “theory” for key agent behaviours; and filter potential parameter values to limit uncertainty. ODD and POM are important steps toward the acceptance of agent-based approaches as established, credible ways to do science.

5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24

V. Grimm (✉)
Department of Ecological Modelling, Helmholtz Center for Environmental Research,
Leipzig, Germany
e-mail: volker.grimm@ufz.de

S.F. Railsback
Department of Mathematics, Humboldt State University, Arcata CA, USA
e-mail: steven.railsback@humboldt.edu

25 **17.1 Introduction**

26 Agent-based models (ABMs) can be hard to communicate. They usually include
27 different types of agents and spatial units, each distinguished by a suite of state
28 variables, attributes, behaviours, and environmental processes. Model results can
29 depend strongly on how model entities are initialised, how processes and events are
30 scheduled, the data used to drive the simulated environment, and the details of how
31 submodels represent model processes. As a consequence, many, if not most, descrip-
32 tions of ABMs are lengthy but nevertheless incomplete.

33 Communication is, however, part and parcel of using ABMs as a scientific tool.
34 Models are the “Materials and Methods” used by modellers to obtain scientific
35 results. Thus, as with any other method used in science, the very feature that renders
36 an ABM “scientific” is replicability: that the same results can be obtained by peers
37 if they use exactly the same materials and methods. But most published ABMs
38 would be impossible, or at least challenging, to replicate. This situation cannot be
39 tolerated because it undermines the credibility of ABMs as a scientific method.
40 Moreover, because ABMs are increasingly built to support real-world decision
41 making, model structure and assumptions need to be transparent to decision makers;
42 otherwise the models are likely to be (justifiably) ignored, or used inappropriately.

43 Incomplete and inefficient communication is linked to a second challenge of
44 agent-based modelling: the lack of a framework for designing ABMs. Current prac-
45 tice is that most ABMs are developed from scratch and that the choice of model
46 structure and process representation is more or less *ad hoc*. Model design and for-
47 mulation often reflect the disciplinary background, experience, and personal prefer-
48 ences of the modeller more than general principles of model design that would lead
49 to similar model designs for similar problems.

50 Lack of efficiency and coherence in design and communication is to be expected
51 for an emerging scientific approach. But agent-based modelling is no longer in its
52 infancy: hundreds of ABMs have proven the potential of this approach, both in
53 theory and application, in a wide array of disciplines; many research projects and
54 programs dealing with complex systems include ABMs in their portfolio of
55 approaches; and the issues of standardisation, communication, and transparency
56 have increasingly been addressed over the last 10 years.

57 There are several examples of the drive for a general framework for ABMs.
58 Railsback (2001) listed design concepts that are important for designing ABMs, in
59 particular when adaptive behaviour is included. Grimm and Railsback (2005)
60 devoted their book on individual-based modelling in ecology to making this
61 approach more “coherent and efficient”. A series of workshops addressed the issue
62 of model replication (e.g., Hales et al. 2003). Standards for model development
63 (Richiardi et al. 2006) and communication were proposed (Grimm et al. 2006,
64 2010), and a community has been established with the mission to “foster the agent-
65 based modelling development, communication, and dissemination for research,
66 practice and education” (Janssen et al. 2008): the Open Agent Based Modelling
67 Consortium (www.openabm.org).

Standardisation is an indicator for a maturing approach: the approach has been used enough to understand its strengths and weaknesses, what decisions have to be made when using it, and how standards can make its use more efficient and coherent. Standards must of course avoid restricting the approach's flexibility and usefulness.

Here we present the ODD protocol, a standard for communicating and formulating ABMs; and pattern-oriented modelling (POM), a general strategy for tying model structure to multiple observed patterns to make model design and analysis less *ad hoc*. Both approaches were formulated in ecology but are relevant in any discipline using ABMs. Since both ODD and POM have been described in detail elsewhere, we focus here on what these approaches are, how they can be used, and what benefits they provide to the agent-based modeller.

17.2 The ODD Protocol 80

The ODD protocol was developed as a standard format for describing individual- and agent-based models (Grimm et al. 2006; Grimm and Railsback 2005). The acronym stands for the three blocks by which ODD's elements are grouped: Overview, Design concepts, Details. ODD is designed to be used for *all* ABMs, independent of their domain, purpose, complexity, and computer implementation. The main idea of ODD is to present ABMs in a hierarchical way: an overview of model structure and processes first, with details on processes last. Between overview and details is a checklist that explains how important general concepts for designing ABMs have been taken into account. This checklist ensures that modellers make important design decisions consciously and that readers understand how and why certain design decisions were made.

Using ODD implies referring to this standard explicitly by stating: "The model description follows the ODD (Overview, Design concepts, Details) protocol for describing individual- and agent-based models (Grimm et al. 2006, 2010)" and then presenting the seven elements of ODD (Table 17.1) using exactly the given sequence and element labels. By using ODD always in this way, writers do not waste time figuring out how to organise their model description and readers avoid frustration by knowing exactly where to find information about a model. Moreover, understanding of a model is greatly facilitated by first providing an overview of its structure and processes. For most readers, this information is sufficient to grasp the overall rationale of the model and to understand its results. Readers hooked by the overview can go into more detail and see how certain processes have been represented. Finally, an ODD description should be complete enough to allow the model's replication.

For complex models, the Details and possibly also the Design concepts parts may be too long, so they can be moved to an electronic appendix, or the full ODD model description could be published in a separate report.

t1.1 **Table 17.1** The seven elements of the ODD protocol (See also the ODD webpages at www.ufz.de/oesatools/odd)

t1.3	ODD	ODD element	Questions to be answered	
t1.4	Overview	Purpose	What is the purpose of the model?	
t1.5		Entities, state variables, and scales	What kind of entities are in the model? By what state variables, or attributes, are these entities characterised? What are the temporal and spatial resolutions and extents of the model?	
t1.6				
t1.7				
t1.8		Process overview and scheduling	Which entities do what, in what order?	
t1.9			When are state variables updated? How is time modelled — as discrete steps or as a continuum over which both continuous processes and discrete events can occur?	
t1.10				
t1.11		Design concepts	Design concepts	There are ten design concepts (see text). How have these concepts been taken into account in the model's design?
t1.12				
t1.13				
t1.14	Details	Initialisation	What is the initial state of the model, i.e. at time $t=0$?	
t1.15		Input data	What input does the model use from external sources such as data files or other models to represent processes that change over time?	
t1.16				
t1.17				
t1.18		Submodels	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, tested, and parameterised?	
t1.19				
t1.20				
t1.21				
t1.22				
t1.23				
t1.24				
t1.25				
t1.26				
t1.27				

108 **17.2.1 Design Concepts**

109 The design concepts of ODD were first summarised by Railsback (2001), then
 110 discussed in detail in Grimm and Railsback (2005), and most recently updated by
 111 Grimm et al. (2010). The ten concepts are not actually needed to replicate a model,
 112 but they are important for communicating the essence of how (and why) an ABM
 113 has been designed—because traditional formats such as equations and diagrams
 114 cannot capture the essence of most ABMs. For example, the first design concept is
 115 “emergence”. Its discussion should explain whether key model results are imposed, or
 116 whether instead they emerge from agent behaviours and other processes – information
 117 that is the key to understanding the rationale of the model. Here “imposed” means
 118 that model rules and equations force the model to behave in a predictable way. For
 119 example, if a model is designed so that an event like a traffic jam occurs with a
 120 constant probability of 5%, then it is no surprise that on average the model produces
 121 traffic jams 5% of the time. If, however, traffic jams emerge from the behaviour and
 122 decisions of the agents and the roads they follow, there is no simple way to predict
 123 how often traffic jams occur: we have to run the model and see what emerges.

In many descriptions of ABMs, it is not entirely clear what the authors wanted to emerge and what they imposed. Similar “ad hoceries” often occur with the other design concepts. The following description of the concepts is adopted from Grimm et al. (2010):

- *Emergence.* What key outputs of the model are modelled as emerging from the adaptive behaviour of its agents? Are there other outputs that are more tightly imposed by model rules and hence less dependent on what individuals decide to do?
- *Adaptation.* What rules do agents have for changing behaviour in response to changes in themselves or their environment? Do these traits explicitly aim at increasing some measure of individual objectives or success? Or do they instead cause individuals to reproduce observed behaviours that are implicitly assumed to convey success?
- *Objectives.* If adaptive behaviour is represented as explicitly seeking some objective, what is the objective and how is it measured? Examples of “objectives” are “fitness” for organisms, “utility” for economic reward in social models, or simply “success”.
- *Learning.* Do individuals change their adaptive behaviour over time as a consequence of their experience? How?
- *Prediction.* To make decisions, model agents often need to predict future consequences of their alternatives. What internal models are used by the agents to estimate future conditions or consequences of their decisions? What “tacit” or hidden predictions are implied in these internal models?
- *Sensing.* What information (state variables of other model entities and themselves) can agents sense and consider in their adaptive decisions? Are the mechanisms by which agents obtain information modelled explicitly, or are agents simply assumed to “know” these variables?
- *Interaction.* What kinds of interactions among agents are in the model? Are there direct interactions, or are the interactions indirect, e.g. via competition for a mediating resource? How do agents interact with their environment?
- *Stochasticity.* What processes are modelled by assuming that they are random or partly random? Why is stochasticity used – to represent variability in a simple way, or to cause events or behaviours to occur with a specified frequency?
- *Collectives.* Are there aggregations of agents that affect, and are affected by, the agents? Examples include social groups, fish schools and bird flocks, human networks and organisations, or cells constituting an organ. Are collectives represented as emergent properties of the agents or as a separate kind of entity with its own state variables and traits?
- *Observation.* What data and patterns must be observed from the ABM for testing, understanding, and analyzing it, and how are they collected?

ODD model descriptions should describe how these design concepts were taken into account. It is possible to leave out some of the concepts if they are not included in the ABMs design at all (e.g., if adaptive behaviour is simple, then Objectives, Learning, and Prediction may be irrelevant). Four concepts are relevant to almost every ABM: emergence, interaction, stochasticity, and observation.

168 **17.2.2 Examples**

169 As a simple example, we apply the ODD protocol to the “Segregation” model
170 published as part of NetLogo (Wilensky 1997), which was inspired by the work of
171 T. Schelling (e.g., Schelling 1971).

172 *Purpose.* The model addresses segregation of households in cities: why members
173 of two different groups tend to separate into different neighbourhoods. The model
174 explores the relationship between how tolerant individuals are of the opposite group
175 and how segregated neighbourhoods are, when individuals express intolerance by
176 moving.

177 *Entities, state variables, and scales.* The model entities include mobile agents that
178 represent households, and square grid cells that represent houses. Households are
179 characterised by their location (which house they occupy) and their colour, which is
180 either blue or red. Households also have a state variable “happy?”, a boolean vari-
181 able set to false if the household is intolerant of its neighbours. The grid cells make
182 up a square 51 × 51 cells in extent, with no depiction of roads or other spaces between
183 them. The space is “toroidal”: if an agent moves off one edge of the space, it reap-
184 pears on the opposite edge; and the neighbours of a household on one edge of the
185 space include those on the opposite edge.

186 The length of a time step is unspecified but represents the time in which a house-
187 hold would decide whether to move. The number of time steps in a model run is an
188 emergent outcome: the model runs until no households are intolerant of their neigh-
189 bours and want to move (see submodel *move*).

190 *Process overview and scheduling.* The following actions are executed, in this order,
191 once per time step.

- 192 • If no households are intolerant of their neighbours (“happy?” is true for all
193 households), then the model stops.
- 194 • The households who are intolerant of their neighbours (“happy?” is false) execute
195 the submodel *move*. The order in which these households execute is randomly
196 shuffled each time step.
- 197 • All households update their “happy?” variable (see submodel *update*).
- 198 • Outputs for system-level results are updated.

199 *Design concepts.* The key outcomes of the model are segregation patterns, espe-
200 cially how strongly segregated the entire system is; these outcomes *emerge* from
201 how tolerant households are to unlike neighbours. The households’ *adaptive behav-*
202 *iour* is to move when their *objective* – to live in a neighbourhood with the fraction
203 of unlike neighbours below their tolerance threshold – is not met. The behaviour
204 does not involve learning, or prediction other than the implicit prediction that mov-
205 ing will lead to a neighbourhood where the tolerance objective is met. Households
206 *sense* the colour of other households on the eight surrounding grid cells. *Stochasticity*
207 is used in only two ways: to initialise the model so that it starts unsegregated; and to
208 determine the new location of households when they move, because modelling the

details of the movement is beyond the scope of this model. *Observations* include a visual display of which colour household is on each grid cell, and two numerical results: the mean (over all households) percent of neighbours of similar colour, and the percent of households with “happy?” false.

Initialisation. A user-chosen number (typically, 2000) of households are initialised. They are each placed on a random empty grid cell and given a colour randomly, with equal probability of red and blue.

Input data. The model does not use input from external models or data files.

Submodels. The submodel *move* is performed by individual households if they are not tolerant of their neighbours. The household chooses a direction randomly from a uniform real distribution between 0° and 360°, then moves forward a distance drawn randomly from a uniform real distribution of 0–10 grid cell widths. If there is already a household on the grid cell at this new location, the household repeats the move. If the new grid cell is empty, the household moves to its centre.

The submodel *update* is conducted by all households, to determine whether they tolerate their neighbourhood. The tolerance of households is determined by a parameter “%-similar-wanted”, the value of which ranges from 0 to 100 and applies to all households. A household’s “neighbours” are all households on the eight surrounding patches. The household’s variable “happy?” is set to false unless the number of neighbours with the household’s colour is greater than or equal to “%-similar-wanted” divided by 100 and multiplied by the number of neighbours.

Using ODD for such simple models may look overdone. However, the ODD model description has a clear hierarchical structure, different elements of the model are easy to find, and all information required for implementation is provided. Moreover, the design concepts provide some explanation of why the model was designed as it was.

An important benefit of ODD is that any ABM can be described in exactly the same format as the segregation model. ODD thus provides a unifying format. Consequently, using ODD to rewrite model descriptions made it much easier to compare three different models of land use/land cover changes (Polhill et al. 2008a): SLUDGE (Parker 1999; Parker and Meretsky 2004), SOME (Brown et al. 2004, 2005), and FEARLUS (Polhill et al. 2001; Gotts et al. 2003) with the ELMM extension (Polhill et al. 2005, 2008b).

Just identifying the entities and state variables of the three models made it easy to see the conceptual similarities of these models, but also their differences in complexity, resolution, and scope (Table 17.2). Obviously, SLUDGE and SOME are conceptually very simple and similar, whereas FEARLUS+ELMM can represent quite complex landscapes and scenarios. FEARLUS+ELMM could, however, probably be mapped to the two simpler models by choosing simplified parameterisations and initialisations. Model-to-model comparison, and thus transfer of questions, model designs, and results, is greatly facilitated by using ODD. This applies also to comparing process overviews and schedules, and design concepts (Polhill et al. 2008a).

Table 17.2 Comparison of the entities (*bold*) and their state variables (*italicized lists*) of three land use model (see text for references). For a full ODD description, in addition the precise number and type of state variables needs to be specified, for example for “biophysical characteristics” or “aesthetic quality.” (See also Polhill et al. 2008a)

	SOME		FEARLUS+ELMM	
SLUDGE				
Land-scape cell	- Productivity	- Aesthetic quality	Land parcel	- Biophysical characteristics
	- Externalities	- Coordinates		- Land use (of the past)
	- Output price for land use 0			- Yield (of most recent land use)
	- Coordinates			- Owning land manager
Land owner			Land manager	
	- Land use	- Relative importance		- List of owned land parcels
	- Coordinates	- of three different aspects		- Land manager's subpopulation
		- of landscape cell		- Land managers market
		- Utility		- Account
		- Coordinates		- Aspiration threshold
				- Memory size
				- Land offer threshold
				- Bidding strategy
				- Selection strategy
Market location			Subpopulation	- Land manager list
	- Coordinates	- Coordinates		- Incomer offer price distribution
				- Imitative probability distribution
				- Aspiration threshold distribution
				- Land offer threshold distribution
				- Bidding strategy distribution
Land-scape			Land market	- Auction type
	- Dimensions			- Parcels for sale list
				- Bid list
			Environment	- Spatial topology
				- External conditions
				- Array of probabilities of changes in external conditions
				- Break-even threshold
				- Possible land uses

t2.1
t2.2
t2.3
t2.4
t2.5
t2.6
t2.7
t2.8
t2.9
t2.10
t2.11
t2.12
t2.13
t2.14
t2.15
t2.16
t2.17
t2.18
t2.19
t2.20
t2.21
t2.22
t2.23
t2.24
t2.25
t2.26
t2.27
t2.28
t2.29
t2.30
t2.31
t2.32

17.2.3 ODD as Design Patterns for Formulating ABMs

252

Since 2006, ODD has been used for over 70 individual- and agent-based models. ODD has mainly been used by ecologists, but recently applications are appearing in other domains as well, e.g., microbiology (Hellweger et al. 2008), socio-ecology (Guzy et al. 2008), biomedical research (Galvão and Miranda 2009), and anthropology (Premo and Hublin 2009). After reviewing most of these applications of ODD and summarising the questions and comments from many ODD users, Grimm et al. (2010) revised ODD slightly but completely rewrote its description and explanation to make its use easier.

253
254
255
256
257
258
259
260

Reviewing these applications of ODD revealed unexpected benefits, including that ODD provides a general format for thinking about and designing ABMs. ODD makes us ask questions about model structure and design in a hierarchical way: structure and overview first, processes and details later. This facilitates translating the often confusing and vague conceptual models are the starting point of any modelling process (Grimm and Railsback 2005) into model formulations. Scientists with no background in agent-based modelling found ODD helpful for understanding what would be involved to implement an ABM of their system and problem. Likewise, communication between students who develop ABMs and their supervisors is made much more efficient by using ODD. ODD also facilitates communication between different domains, which usually have their own styles of modelling and communicating models.

261
262
263
264
265
266
267
268
269
270
271
272

Last but not least, ODD corresponds to “design pattern” in software engineering (Gamma et al. 1994), which describe “recurring solutions to common problems in software design” (Wikipedia 2009). In agent-based modelling, ODD is a design pattern for formulating new models, and for re-formulating existing models.

273
274
275
276

17.3 Patterns and Stylized Facts

277

While ODD has proven very important and useful, it provides only a general structure for formulating ABMs. It cannot, by itself, prevent “ad hocery” in the choice of entities and state variables. ABMs can still be too simple or too complex to be useful for inferences about how their real counterparts are working. We need a strategy for increasing the chance that our models explain the internal organisation of real complex systems by guiding our choices of what to include in a model and in what detail.

278
279
280
281
282
283
284

Such a strategy emerged in ecology over the last 15 years: pattern-oriented modelling (POM). It is based on the notion that complex systems provide indicators of their internal organisation: patterns, signals, or regularities that cannot be explained by random processes and therefore call for an explanation (Heine et al. 2007). For example, if vegetation in semi-arid regions shows a characteristic banded structure (Fig. 17.1), we can take this pattern as an indicator of how vegetation is affected by,

285
286
287
288
289
290



Fig. 17.1 Banded vegetation pattern in semi-arid regions produced by a simple grid-based model (Thiery et al. 1995; model reimplementation courtesy of U. Berger). The 100×100 cells each have the size of a tree. *Dark cells* are covered by woody vegetation, *white cells* are not. Implicitly, the landscape is assumed to have a gentle slope so that runoff from rain flows from *top* to *bottom*

291 and affects, the distribution of water. Our task thus is to “decode” the patterns
292 observed in the real systems. We do this inversely by trying to find models that
293 produce the same patterns.

294 This approach is certainly not new but just the basic idea of science, or infer-
295 ence (Platt 1964). Patterns were key to revealing the internal organisation of atoms
296 and the universe (spectral patterns), identifying quasars (periodic signals), and
297 proposing asteroids as the cause of mass extinctions (unusual Iridium concentra-
298 tions in geologic strata). However, one important and often neglected point in the
299 science of agent-based complex systems is that a single pattern is rarely enough
300 to decode the internal organisation. Many different models are, for example, able
301 to reproduce cycles in the population dynamics of small mammals (Czárán 1998).
302 How can we select the right model or falsify models that produce cycles for the
303 wrong reasons?

304 The basic and simple idea of POM is to use multiple patterns, observed at differ-
305 ent scales and hierarchical levels, as multiple criteria for selecting among alternative
306 model structures, submodel formulations, and parameter sets. One single pattern
307 might be relatively “weak” and contain little information so that still, say, half of the
308 models considered would reproduce it. But by adding a second pattern, we possibly
309 reduce the degrees of freedom in model structure. Thus, every single pattern serves

as a filter. Again, this method of inference is old and regularly used by detectives who use motives, alibis, witness statements, and evidence to filter possible suspects. Or, in the famous story of the discovery of the structure of DNA (Watson 1968), X-ray diffraction patterns indicated a spiral structure, but many such structures were compatible with this patterns. Only after two additional patterns, Chargaff's rule and the geometry of purine and pyridimine bases, were taken into account was the real structure identified.

In social sciences and in particular in economics, the notion of "stylised facts" (after Kaldor 1961) corresponds to what we here mean by patterns. In the modelling literature of these domains, stylised facts often seem to be disregarded as too vague to use for rigorous inferences. The point of POM is, however, to increase inferential power by *combining* multiple vague, weak, or qualitative patterns or stylised facts (see also the discussion of 'middle range models' by Gilbert (2008, pp. 41–44)).

17.4 Pattern-Oriented Modelling 323

In POM, multiple patterns are used for three elements: designing model structure (what entities, state variables, and processes to include); selecting among alternative submodels; and determining parameter values.

17.4.1 Patterns for Model Structure 327

The key question of POM's first element is: what do experts who know the system well consider to be characteristic features, or essentials, of the system? For example, in growing cohorts of trees, competition among neighbouring trees increases as the trees grow. Initial differences in size and the trees' spatial distribution cause some trees to be more suppressed by competition so they die, which is called "self-thinning" in ecology. The resources used by the trees that die (space, light, nutrients) become available to survivors, who can grow further until competition is again strong enough that more trees die.

A characteristic pattern of this process is that the relationship between average tree biomass and tree density follows a power law: a straight line on a log-log graph. Moreover, the slope of this line is similar for virtually all plant species and environments. Until about 2000 it was believed that the slope has to be, for theoretical reasons, $-3/2$ but a new theory predicts a slope of $-4/3$ (Enquist et al. 1998).

The consequence of this pattern for modelling self-thinning was that models focussed on tree density and average biomass as state variables. Different models, however, can each produce a power law with the right slope. So, what other patterns characterise self-thinning? This question is not fully answered yet, but it seems that the spatial distribution of trees is quite regular during the entire process, which indicates that neighbour competition, depending on local tree density, is important.

347 To potentially explain this regular spatial distribution, a model should be individual-
348 based so that space and local interactions can be represented. Trees might be repre-
349 sented just by their biomass, which corresponds to a certain “zone of influence”
350 over which they interact with neighbours (Berger and Hildenbrandt 2000). No further
351 patterns seem to be known yet that would suggest including further state variables
352 of the trees, e.g. height, crown shape, etc. Interestingly, power laws are also found
353 in the size distribution of cities, where probably similar mechanisms of neighbour
354 competition are the underlying mechanism (Batty 2005).

355 To use POM to design a model structure, we identify multiple patterns, typically
356 two to five, which seem to characterise the system and problem being modelled.
357 Making it possible for the model to reproduce each pattern requires adding some
358 state variables and processes. These make the model more complex but also rich
359 enough in structure to be tested against not only one, but multiple patterns. The
360 models are also “mechanistically rich” (DeAngelis and Mooij 2003) so that they can
361 be checked for patterns that were not used for model formulation and calibration.

362 An example of a mechanistically rich model is the natural beech forest ABM of
363 Rademacher et al. (2004), who found characteristic patterns in the age structure and
364 spatial distribution of canopy trees in model results. These patterns were not at all
365 considered during model formulation and verification, which were driven by other
366 characteristic patterns, but agreed very well with observations (Rademacher et al.
367 2001). This discovery indicated that the model captures essentials of the system’s
368 internal organisation instead of just being fit to observations. One generic feature of
369 structurally realistic models is that they reproduce observed patterns at different
370 hierarchical levels simultaneously. They capture essential behaviours of both the
371 system and its agents.

372 Quite a few ABMs have been oriented towards reproducing sets of observed
373 patterns, but usually this is not stated explicitly. POM means to use multiple patterns
374 systematically and explicitly.

375 **17.4.2 Patterns for Selecting Submodels**

376 ABMs usually comprise a set of submodels that describe the model entities’ behav-
377 iour. The segregation model described above, for example, has two submodels,
378 *move* and *update*. Formulating submodels includes many degrees of freedom. The
379 decision of whether and where to move could, instead of being very simple and
380 partly random, have been a complex algorithm using knowledge about distant loca-
381 tions and prediction of future changes in neighbourhoods. Submodels, like entire
382 models, must compromise between being simple and unrealistic vs. more realistic
383 and harder to understand. How much detail, and what detail, must we have to solve
384 a particular problem?

385 The idea of POM for selecting submodels is to use the entire ABM as a virtual
386 laboratory in which we separately test alternative submodels: which submodel, or
387 theory (Grimm and Railsback 2005), is able to reproduce multiple observed patterns

simultaneously? If more than one submodel can do so, which is the most simple one? 388
 For example, Railsback and Harvey (2002) distilled from the literature six broad pat- 389
 terns in the habitat selection behaviour of trout; their ABM was designed so these 390
 patterns *could* emerge but were not *forced* to emerge. Three alternative habitat selec- 391
 tion submodels were tested, and only one could reproduce all six patterns. 392

Contrasting alternative submodels or theories by their ability to reproduce sets of 393
 patterns corresponds, in principle, to more formal model selection algorithms from 394
 information theory, which are used to select among alternative simple descriptive 395
 models (Piou et al. 2009). 396

One important advantage of contrasting alternative submodels is also that it helps 397
 communicate that “the model can be wrong” (J. Goss-Custard personal communi- 398
 cation) and thus counter the widespread notion that complex ABMs cannot be 399
 wrong because you can fit them to any result you want. The segregation model, for 400
 example, does *not* reproduce segregation if the decision to move is entirely random. 401
 Likewise, Janssen et al. (2009) found that two ‘naïve’ models of behaviour in a 402
 common dilemma game, random walk and greedy agents, did not reproduce all four 403
 patterns identified in controlled laboratory experiments. 404

Of course, this kind of high-level, pattern-oriented submodel selection cannot be 405
 applied to all submodels of an ABM but only to those that represent key behaviours; 406
 examples from ecology are habitat selection (Railsback and Harvey 2002) and for- 407
 aging strategy (Goss-Custard et al. 2006). 408

17.4.3 Patterns for Parameterisation 409

Pattern-oriented models are usually of moderate complexity with typically 10–20 410
 parameters. Ideally, all parameter values could be determined directly from informa- 411
 tion about the agents being modelled, but this is virtually never the case. Estimating 412
 parameter values from expert knowledge is often sufficient to get the qualitative behav- 413
 iour of the model right, because experts often know much more than can be expressed 414
 by hard numbers. One important benefit of POM is that this qualitative knowledge can 415
 be included via empirical if-then rules and tested against multiple patterns. 416

Nevertheless, if models are to be used to make inferences about the real world 417
 and support decision-making, “guestimated” parameterisations can be too uncertain 418
 For example, Wiegand et al. (2004) developed a spatially explicit ABM of 419
 brown bears re-invading the Alps from Slovenia. Uncertainty in the predicted popu- 420
 lation growth rate in a certain area was very high (Fig. 17.2). Wiegand et al. there- 421
 fore used multiple patterns in spatial distributions and census time series to reduce 422
 uncertainty in parameter values. 423

They created a large number of parameter sets by sampling values of the uncertain 424
 parameters from their assumed ranges. Then, each observed pattern was used as 425
 a filter: parameter sets leading to model output which did not reproduce the pattern, 426
 according to quantitative criteria, were discarded. Some patterns could be repro- 427
 duced by many parameter sets, others contained more information and reduced 428

this figure will be printed in b/w

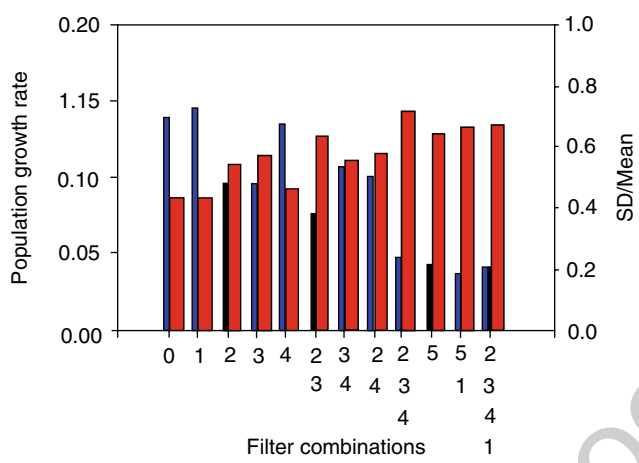


Fig. 17.2 Population growth rate (*red bars*) and its coefficient of variation (*black bars*) of a brown bear population as predicted by the model of Wiegand et al. (2004). Each pair of bars represents one set of model parameterisations, as (*left to right*) more patterns are used as “filters”: parameterisations that do not reproduce the pattern are no longer used. Uncertainty decreases greatly as more patterns are used

429 the set of possible parameterisation more. In combination, however, only a few
 430 (typically 10–30) parameter sets fulfilled all patterns. Using only these “filtered”
 431 parameterisations strongly reduced uncertainty in the model output (Fig. 17.2).

432 Again, this technique of “Monte Carlo filtering” is not new and is used in other
 433 domains under the name “inverse modelling”, but Wiegand et al. (2004) first dem-
 434 onstrated its power for parameterising ABMs and linked it to the other two elements
 435 of POM. Inverse parameterisation, like the entire POM strategy, is independent of
 436 its origin in ecology (Grimm et al. 2005) and applicable to ABMs in general (e.g.
 437 Janssen et al. 2009).

438 17.5 Discussion

439 ABMs are fundamentally different from the traditional models of many fields, espe-
 440 cially from models based on system-level equations. We now have enough experi-
 441 ence to know that the processes of designing, building, and doing science with
 442 ABMs is different from traditional approaches. Traditional system-level models are
 443 designed and described using established “languages” such as differential equations
 444 and flow diagrams, and have their complexity determined mainly by mathematical
 445 tractability. With ABMs, we need new standards and strategies. This chapter
 446 describes two such standards and strategies that have already proven very useful.

447 The ODD standard addresses the need for a way to describe both the essential
 448 characteristics and the full details of an ABM. The value and success of ODD is

illustrated by how widely and rapidly it has been adopted. Benefits of ODD include 449
 allowing readers to quickly get an overview of a model's most important character- 450
 istics, providing sufficient detail for replication, and providing the modeller with a 451
 framework for thinking about and designing an ABM from the start. 452

The POM strategy addresses three needs in the process of designing and using an 453
 ABM. First, it helps with the fundamental model design problem: determining what 454
 entities, variables, and processes must be in a model, so it is complex enough but not 455
 unnecessarily complex. Second, it provides a way to develop and test theory: rules 456
 for how individual agents behave that are shown, via pattern-oriented analysis, to 457
 reproduce essential dynamics of the system. Finally, POM can be used to filter 458
 potential parameter values in an efficient and rigorous way. 459

Both ODD and POM seem “messier” than the standard techniques of traditional 460
 modelling, but that is simply a reflection that ABMs are messier than simpler models: 461
 if we want to model complex systems, we must use more complex models. But 462
 the messiness of ABMs is exactly why we need techniques like ODD and POM: to 463
 help us cope with the complexity. We can view these techniques as evidence of the 464
 establishment and acceptance of ABMs as a way of doing science. 465

References

Batty, M. (2005). *Cities and complexity: Understanding cities through cellular automata, agent- 467
 based models, and fractals*. Boston: The MIT Press. 468

Berger, U., & Hildenbrandt, H. (2000). A new approach to spatially explicit modelling of forest 469
 dynamics: Spacing, ageing and neighbourhood competition of mangrove trees. *Ecological 470
 Modelling, 132*, 287–302. 471

Brown, D. G., Page, S. E., Riolo, R., & Rand, W. (2004). Agent-based and analytical modeling 472
 to evaluate the effectiveness of greenbelts. *Environmental Modelling and Software, 19*, 473
 1097–1109. 474

Brown, D. G., Page, S. E., Riolo, R., Zellner, M., & Rand, W. (2005). Path dependence and the 475
 validation of agent-based spatial models of land use. *International Journal of Geographic 476
 Information Systems, 19*, 153–174. 477

Czárán, T. (1998). *Spatiotemporal models of population and community dynamics*. London: 478
 Chapman & Hall. 479

DeAngelis, D. L., & Mooij, W. M. (2003). In praise of mechanistically-rich models. In C. D. Canham, 480
 J. J. Cole, & W. K. Lauenroth (Eds.), *Models in ecosystem science* (pp. 63–82). Princeton: 481
 Princeton University Press. 482

Enquist, B. J., Brown, J. H., & West, G. B. (1998). Allometric scaling of plant energetics and popu- 483
 lation density. *Nature, 395*, 163–165. 484

Galvão, V., & Miranda, J. G. V. (2009). Modeling the Chagas' disease after stem cell transplanta- 485
 tion. *Physica A, 388*, 1747–1754. 486

Gamma, E., Helm, R., Johnson, R., & Vlissides, J. (1994). *Design patterns: Elements of reusable 487
 object-oriented software*. Reading: Addison-Wesley. 488

Gilbert, N. (2008). *Agent-based models. Quantitative applications in the social sciences*. Thousand 489
 Oaks: Sage. 490

Goss-Custard, J., Burton, N. H. K., Clark, N. A., Ferns, P. N., McGrorty, S., Reading, C. J., 491
 Rehfish, M. M., Stillman, R. A., Townend, I., West, A. D., & Worrall, D. H. (2006). Test of a 492
 behavior-based individual-based model: Response of shorebird mortality to habitat loss. 493
Ecological Applications, 16, 2215–2222. 494

- 495 Gotts, N. M., Polhill, J. G., & Law, A. N. R. (2003). Aspiration levels in a land use simulation.
496 *Cybernetics and Systems*, 34, 663–683.
- 497 Grimm, V., & Railsback, S. F. (2005). *Individual-based modeling and ecology*. Princeton: Princeton
498 University Press.
- 499 Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., Thulke, H.-H.,
500 Weiner, J., Wiegand, T., & DeAngelis, D. L. (2005). Pattern-oriented modeling of agent-based
501 complex systems: Lessons from ecology. *Science*, 310, 987–991.
- 502 Grimm, V., Berger, U., Bastiansen, F., et al. (2006). A standard protocol for describing individual-
503 based and agent-based models. *Ecological Modelling*, 198, 115–126.
- 504 Grimm, V., Berger, U., DeAngelis, D. L., Polhill, G., Giske, J., & Railsback, S. F. (2010). The
505 ODD protocol for describing individual-based and agent-based models: A first update.
506 *Ecological Modelling*, 221(23), 2760–2768. doi:10.1016/j.ecolmodel.2010.08.019.
- 507 Guzy, M.R., Smith, C.L., Bolte, J. P., Hulse, D. W., & Gregory, S. V. (2008). Policy research using
508 agent-based modeling to assess future impacts of urban expansion into farmlands and forests.
509 *Ecology and Society*, 13, 37. [online] URL: <http://www.ecologyandsociety.org/vol13/iss1/art37/>
- 510 Hales, D., Rouchier, J., Edmonds, B. (2003). Model-to-model analysis. *Journal of Artificial*
511 *Societies and Social Simulation*, 6, URL: <http://jasss.soc.surrey.ac.uk/6-4/5.html>
- 512 Heine, B.-O., Meyer, M., & Strangfeld, O. (2007). Das Konzept der stilisierten Fakten zur Messung
513 und Bewertung wissenschaftlichen Fortschritts. *Die Betriebswirtschaft*, 67, 583–608.
- 514 Hellweger, F. L., Kravchuk, E. S., Novotny, V., & Gladyshev, M. I. (2008). Agent-based modeling
515 of the complex life cycle of a cyanobacterium (*Anabaena*) in a shallow reservoir. *Limnology*
516 *and Oceanography*, 53, 1227–1241.
- 517 Janssen, M. A., Alessa, L. N., Barton, M., Bergin, S., & Lee, A. (2008). Towards a community
518 framework for agent-based modelling. *Journal of Artificial Societies and Social Simulation*,
519 11(2), 6. <http://jasss.soc.surrey.ac.uk/11/2/6.html>.
- 520 Janssen, M. A., Radtke, N. P., & Lee, A. (2009). Pattern-oriented modeling of common dilemma
521 experiments. *Adaptive Behavior*, 17(6), 508–523. doi:10.1177/1059712309342488.
- 522 Kaldor, N. (1961). Capital accumulation and economic growth. In F. A. Lutz & D. C. Hague
523 (Eds.), *The theory of capital* (pp. 177–222). London: Macmillan.
- 524 Parker, D. C. (1999). Landscape outcomes in a model of edge-effect externalities: A computational
525 economics approach. *SFI Publication 99-07-051*. Santa Fe: Santa Fe Institute. <http://www.santafe.edu/sfi/publications/Working-Papers/99-07-051.pdf>
- 526 Parker, D. C., & Meretsky, V. (2004). Measuring pattern outcomes in an agent-based model of
527 edge-effect externalities using spatial metrics. *Agriculture, Ecosystems and Environment*, 101,
528 233–250.
- 529 Piou, C., Berger, U., & Grimm, V. (2009). Proposing an information criterion for individual-based
530 models developed in a pattern-oriented framework. *Ecological Modelling*, 220, 1957–1967.
- 531 Platt, J. R. (1964). Strong inference. *Science*, 146, 347–353.
- 532 Polhill, J. G., Gotts, N. M., & Law, A. N. R. (2001). Imitative versus nonimitative strategies in a
533 land use simulation. *Cybernetics and Systems*, 32, 285–307.
- 534 Polhill, J. G., Parker, D. C., & Gotts, N. M. (2005). Introducing land markets to and agent based
535 models of land use change. In K. G. Troitzsch (Ed.), *Representing social reality: Pre-proceedings*
536 *of the third conference of the European Social Simulation Association* (pp. 150–157). Koblenz:
537 Dietmar Fölbach.
- 538 Polhill, J. G., Brown, D. G., & Grimm, V. (2008a). Using the ODD protocol for describing three
539 agent-based social simulation models of land use change. *Journal of Artificial Societies and*
540 *Social Simulation*, 11 URL: <http://jasss.soc.surrey.ac.uk/11-2/3.html>
- 541 Polhill, J. G., Parker, D. C., & Gotts, N. M. (2008b). Effects of land markets on competition
542 between innovators and imitators in land use: Results from FEARLUS-ELMM. In C. Hernandez,
543 K. Troitzsch, & B. Edmonds (Eds.), *Social simulation technologies: Advances and new discov-*
544 *eries* (pp. 81–97). New York: IGI Global.
- 545 Premo, L. S., & Hublin, J.-J. (2009). Culture, population structure, and low genetic diversity in
546 Pleistocene hominins. *PNAS*, 106, 33–37.
- 547

- Rademacher, C., Neuert, C., Grundmann, V., et al. (2001). Was charakterisiert Buchenurwälder? 548
 Untersuchungen der Altersstruktur des Kronendachs und der räumlichen Verteilung der 549
 Baumriesen mit Hilfe des Simulationsmodells BEFORE. *Forstwiss Centralbl*, 120, 288–302. 550
- Rademacher, C., Neuert, C., Grundmann, V., et al. (2004). Reconstructing spatiotemporal dynam- 551
 ics of Central European natural beech forests: The rule-based forest model BEFORE. *Forest 552*
Ecology and Management, 194, 349–368. 553
- Railsback, S. F. (2001). Concepts from complex adaptive systems as a framework for individual- 554
 based modelling. *Ecological Modelling*, 139, 47–62. 555
- Railsback, S. F., & Harvey, B. C. (2002). Analysis of habitat-selection rules using an individual- 556
 based model. *Ecology*, 83, 1817–1830. 557
- Richiardi, M., Leombruni, R., Saam, N., & Sonnessa, M. (2006). A common protocol for agent- 558
 based social simulation. *Journal of Artificial Societies and Social Simulation*, 9, URL: [http:// 559](http://jasss.soc.surrey.ac.uk/9/1/15.html)
jasss.soc.surrey.ac.uk/9/1/15.html 560
- Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology*, 1, 561
 143–186. 562
- Thiery, J. M., D’Herbes, J. M., & Valentin, C. (1995). A model simulating the genesis of banded 563
 vegetation patterns in Niger. *Journal of Ecology*, 83, 497–507. 564
- Watson, J. (1968). *The double helix: A personal account of the discovery of the structure of DNA*. 565
 New York: Atheneum. 566
- Wiegand, T., Revilla, E., & Knauer, F. (2004). Dealing with uncertainty in spatially explicit popu- 567
 lation models. *Biodiversity and Conservation*, 13, 53–78. 568
- Wikipedia (2009). Software design patterns. July 24, 2009. URL: [http://en.wikipedia.org/wiki/ 569](http://en.wikipedia.org/wiki/Software_design_patterns)
[Software_design_patterns](http://en.wikipedia.org/wiki/Software_design_patterns) 570
- Wilensky, U. (1997). NetLogo segregation model. Center for Connected Learning and Computer- 571
 Based Modeling, Northwestern University, Evanston. [http://ccl.northwestern.edu/netlogo/ 572](http://ccl.northwestern.edu/netlogo/models/Segregation)
[models/Segregation](http://ccl.northwestern.edu/netlogo/models/Segregation) 573