The virtual ecologist approach: simulating data and observers

Damaris Zurell, Uta Berger, Juliano S. Cabral, Florian Jeltsch, Christine N. Meynard, Tamara Münkemüller, Nana Nehrbass, Jörn Pagel, Björn Reineking, Boris Schröder and Volker Grimm

Models permeate every field in ecology. They have become an indispensable tool for a wide range of tasks, including the understanding of mechanisms, capturing the processes behind the emergence of ecological phenomena, quantifying relationships between species presence or abundance and environmental conditions, and forecasting effects of changing environments on broad spatial and temporal scales (DeAngelis and Mooij 2005, Araújo and Rahbek 2006, Thüller et al. 2008).

There is, however, a further important field of application of ecological models that so far has not been thoroughly acknowledged in ecological research: evaluating methods for data sampling, analysis and modelling methods by means of virtual data. Here, the idea is to generate virtual data by simulating not only ecological processes, but also the sampling processes that are used to collect these data in reality and the methodological tools used to analyse them. We propose to call this the ‘virtual ecologist’ (VE) approach (see Glossary). The virtue of this approach is its ability to rigorously test method performance against a known truth. The VE approach is concerned with practical questions regarding ecological methods: Is a method able to identify patterns that we know exist (Grimm et al. 1999)? Can we infer the mechanisms underlying these patterns given a certain set of data (Tyre et al. 2001)? Can we correctly and reliably predict future events (Zurell et al. 2009)?

To evaluate methods of data collection, statistical analysis, and modelling we would ideally compare their outcome to reality. This would allow us to assess whether existing patterns were detected correctly, whether correct estimates of process rates were obtained, or whether the distribution of a species was predicted correctly. However, we have no privileged access to reality independent of and beyond field observations and analytical methods. The ability of field data to represent reality depends not only on the time interval and the spatial extent of observation but also on the disturbances.
the observation procedure might induce. We can never know the complete ‘truth’ because any knowledge about the real world is based on (limited) data, because the methods to derive and analyse real world data sets are subject to constraints and biases (Grimm et al. 1999, Halle and Halle 1999, Hirzel et al. 2001, Austin et al. 2006), and because amount of data is limited by time and costs. Many factors cannot be controlled: underlying environmental factors; historical factors such as disturbances, catastrophes, past land uses; and ecological processes such as competition, dispersal and diseases.

With the VE approach all relevant information can be obtained at all times in the virtual world which is taken as a surrogate of reality. We know, for example, the full movement path of model animals, or the exact location of all individuals or subpopulations at a given time. In the virtual reality, we can generate certain patterns a priori as well as biases introduced by the (virtual) observer.

The idea of generating virtual data to evaluate different methods is quite natural and not new. An early example for evaluating sampling methods is given by Stickel (1954). Stickel analysed the quality of mark–recapture data describing the dispersal of small mammals. For this, the author used as a virtual habitat a sheet of paper divided into grid cells. Some of the grid cells marked traps. Animal movement was simulated by random movements of a pencil. Based on the virtual capture data, movement indices were calculated and compared to those derived from the full trajectories of the pencil. By this the accuracy of different observational algorithms was evaluated.

In statistics it is quite common praxis to use high-quality data or artificially created, error-free data to qualify different sampling or modelling methods (Hirzel et al. 2001). For example, Fortin et al. (1989) subsampled a large, real vegetation data set of sugar-maple Acer saccharum in southwestern Québec, simulating three different types of sampling designs (random, systematic and systematic-cluster). This allowed them to evaluate the effects of these sampling designs and of different sampling efforts on the estimation of spatial structures as well as the sensitivity of different spatial analysis methods. Statistical ecologists also build replicate or simulated data sets with known properties to demonstrate the unbiasedness of new modelling methods they have developed or to show their superior efficiency in comparison to previous methods (Bolker 2008). Many introductory textbooks on statistics deal with such topics. Bolker (2008) recommends using simulated data as a ‘best-case scenario’ to test whether correct estimates of the parameters of an ecological system can be inferred from the data before proceeding to real data.

In this review, we identify two main fields of application for VE: (1) testing and improving sampling schemes and methods; (2) testing and comparing models. The first includes the evaluation of spatial and temporal sampling designs, and the assessment of sampling bias as well as the sensitivity of sampling methods to extrinsic conditions, trappability or observability (Halle and Halle 1999). For the latter, VE may help to assess whether a particular model fitted to the virtual data is principally capable of describing and predicting underlying patterns and processes. Also, contests can be arranged between competing models (Hanski 1999), and their application domain can be circumscribed theoretically (Hirzel et al. 2001). In this way, VE helps to select the most appropriate model for a given situation.

The primary aim of this review is to give the VE approach, which emerged and keeps emerging independently under different names in the literature, a common name and summarise its potential and current limitations. We want to introduce VE as a generic, rigorous and unifying approach that can be used as a common basis for testing methods of data collection and for testing modelling methods. First we will characterise the virtual ecologist approach and its elements in more detail. Secondly, we will review past uses of VE and list specific examples within the two above-mentioned main fields of application. We will thereby show that VE can be applied in a broad and diverse range of problems in ecology. Then we will discuss potential uses for empirical ecologists and ecological modellers, and give some practical guidelines which might help to design VE studies for given purposes. Finally, we will outline future directions and list specific research fields that we feel would benefit from VE.

**The virtual ecologist approach**

The virtual ecologist approach requires four elements (Fig. 1): (a) the virtual ecological model, (b) the virtual sampling model, (c) (statistical) modelling and (d) evaluation. The virtual ecological model (a) represents the virtual
species and/or ecosystem, and includes key processes of the ecological system relevant to the question under study. Thus, the virtual ecological model may comprise a single or multiple species, single individuals or entire populations; it may be temporally and spatially implicit or explicit, fine-scaled or coarse-scaled; it may be governed by abiotic factors etc. The virtual sampling model (b) simulates the observation process. Data is collected from the virtual ecosystem (by a ‘virtual ecologist’) according to a sampling scheme mimicking the way the data would be collected by real ecologists in real ecosystems. (Statistical) Modelling (c) is used to draw inferences from the collected data. Examples include estimation of population size, identification of factors influencing species distribution or abundance, and estimation of process parameters. (Statistical) Modelling can also be used to predict the effects of ecological processes. Finally, the results are evaluated against ‘true’ simulated data (d). Essentially, the ‘virtual ecologist’ operates in the same way as an empirical ecologist (Fig. 1). However, in a VE study we have full access to all information created by the virtual ecological model which allows us to draw strong conclusions about our sampling and (statistical) modelling methods.

Different names have emerged throughout the literature for the very same approach: “artificial data” or “artificial species” (Austin et al. 2006, Meynard and Quinn 2007, Cabral and Schurr 2009), “virtual species” (Hirzel et al. 2001), “virtual ecologist” (Grimm et al. 1999, Tyre et al. 2001, Zurell et al. 2009), “simulated data” (Hanski 1999, Dormann et al. 2007), “virtual ecology” (Grimm et al. 1999, Nehrbaas et al. 2006), to name but a few. Of these, virtual ecologist approach seems to best capture the central idea that not only a virtual reality is created but that the sampling itself or the observer’s behaviour is also being simulated in a second model in a hierarchical way. The term virtual ecologist is thus not ambiguous in contrast to terms such as ‘virtual experiment’ or ‘virtual ecology’ which are also used for studies simply employing conceptual models for hypothesis testing where the effect of different scenarios on some system response is explored (Parysow and Gertner 1997, 1999). The current inconsistent terminology emphasises the importance to give the approach a common name which, we believe, will make it more visible and coherent.

In addition to various studies that we simply knew from regular scanning of the ecological literature, our overview of applications of the virtual ecologist approach is based on extensive literature searches carried out between autumn 2008 and spring 2009 using both the search engines <www.scirus.com> and <www.sciencedirect.com>. We used multiple keywords such as ‘virtual ecologist’, ‘virtual biologist’, ‘virtual experiment’, ‘virtual species’, ‘artificial species’, ‘artificial data’ and ‘simulated data’. Due to the lack of a general terminology, it is possible that we have not detected all studies that would have been relevant to our review of the VE approach. However, we are confident that we included a representative set of worked examples and of ecological research fields.

Both the virtual ecological model and the virtual sampling model can be of different complexities. Depending on how much process detail is put into these models the VE approach covers quite a broad range of scientific questions and applications. Generally, we can distinguish descriptive and mechanistic models representing the virtual species/ecosystem (see Glossary). In the same way, the virtual sampling model, i.e. the virtual ecologist, may be descriptive or mechanistic.

Throughout our literature survey, we found an approximately equal ratio between descriptive and mechanistic representations of the virtual ecological model (Table 1; 21 descriptive models vs 25 mechanistic models). In most studies that aimed at testing and improving sampling regimes (n = 14) the virtual ecosystem was simulated by means of mechanistic modelling (12). Within the second field of application, testing and comparing models, 19 out of 32 reviewed studies used descriptive models of the virtual ecosystem. The field of mechanistic modelling is vast and, thus, mechanistic modelling types employed in VE studies are manifold (Table 1). They range from grid-based models and patch network models (Hanski 1998) to individual-based models (Grimm 1999, Grimm and Railsback 2005).

Likewise, the virtual sampling model (Fig. 1, b) covers a wide range of complexities and model types. In most studies we reviewed within the two main fields of application, virtual sampling was modelled as simple subsampling from the full simulated data, and in rare cases virtual sampling was modelled probabilistically (Table 1; 37 out of 46 VE studies employed subsampling, eight of which carried out a full census; seven VE studies employed probabilistic sampling). Simple subsampling means that the virtual ecologist acts flawlessly according to a certain sampling design, makes no observational or measurement errors and does not interact with the virtual species in any way (Tyre et al. 2001). Probabilistic sampling includes e.g. probability of detection and regards observation as a stochastic process (Reese et al. 2005). For instance, even if the species is present, it may not be detected. Still the virtual sampling includes no interaction between virtual species and virtual ecologist. If the virtual ecosystem is based on a mechanistic model, direct feedbacks may be included between the models of virtual species and virtual sampling, such as observer induced individual escapes (Nott 1998, Berger et al. 1999).
### Table 1. Applications of the virtual ecologist approach. (ANN: artificial neural networks; ENFA: environmental niche factor analysis; GAM: generalised additive model; GARP: genetic algorithm of rule-set prediction; GLM: generalised linear model; IBM: individual-based model; IFM: incidence function model; MARS: multivariate adaptive regression splines; PVA: population viability analysis; SDM: species distribution model; SPOM: stochastic patch occupancy model)

<table>
<thead>
<tr>
<th>System modelled</th>
<th>Issues addressed</th>
<th>Virtual ecological model</th>
<th>Virtual sampling model</th>
<th>Main conclusions</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testing and improving sampling schemes and methods</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plants</td>
<td>Testing and improving sampling schemes and methods</td>
<td>Lattice model</td>
<td>Subsampling, seed trap designs</td>
<td>Transsect and sector placement of traps performed best for estimating dispersal kernels. In cases of anisotropic dispersal with isotropy unknown to observer, anulli and grid arrays performed better.</td>
<td>Skarpaas et al. 2005</td>
</tr>
<tr>
<td>Wildlife</td>
<td>Evaluating sampling designs and sizes for species distribution modelling.</td>
<td>Descriptive model / SDM</td>
<td>Subsampling</td>
<td>Regular and equal-stratified sampling strategies were most accurate and robust. Greater sample sizes were advantageous. Guide to improve sample designs.</td>
<td>Hirzel and Guisan 2002</td>
</tr>
<tr>
<td>Seabirds (marbled murrelets)</td>
<td>Evaluating transect-layouts and sampling frequencies to detect population declines.</td>
<td>Lattice model</td>
<td>Subsampling, transect designs</td>
<td>Stratified and unstratified zigzags, and ten 8-km transects placed at random had high power to detect population trend, produced unbiased population estimates, and were logistically feasible.</td>
<td>Rachowicz et al. 2006</td>
</tr>
<tr>
<td>Coleoptera (darkling beetles)</td>
<td>Effect of habitat-specific movements of individuals on capture rates and population size estimates in relation to trap geometry.</td>
<td>IBM</td>
<td>Subsampling, pitfall traps</td>
<td>Rectangular trap arrangements have the highest capture probability when animals relatively sedentary and occupy territories or home ranges. Different geometries useful when individuals are transient to trap area and populations have open spatial structure.</td>
<td>Crist and Wiens 1995</td>
</tr>
<tr>
<td>Ground-dwelling arthropods (carabid beetles)</td>
<td>Evaluation of cross-shaped trap arrangements for estimating population density.</td>
<td>IBM</td>
<td>Subsampling, pitfall traps</td>
<td>Method is promising, especially for more mobile species, and is worth testing in the field.</td>
<td>Perner and Schueler 2004</td>
</tr>
<tr>
<td>Micro-algal grazers (Gastropods)</td>
<td>Evaluation of sampling designs for species of different foraging strategies.</td>
<td>IBM</td>
<td>Subsampling, confinement experiments</td>
<td>Mixing of data derived from field-enclosure experiments and from quadrat-based methods seems to be ill-advised as biases arise especially for intelligent, decision-making organisms.</td>
<td>Mac Nally 2001</td>
</tr>
<tr>
<td>Trees</td>
<td>Comparison of structurally different models to analyse tree growth-mortality relationships.</td>
<td>IBM</td>
<td>Subsampling, tree-ring data, repeated forest inventories</td>
<td>Flexible statistical approaches were superior to less flexible models only for large sample sizes. Study provides theoretical basis for sound estimation of growth-mortality models, and guidelines for efficient sampling schemes in real forests.</td>
<td>Wunder et al. 2008</td>
</tr>
<tr>
<td>Animals (grasshoppers)</td>
<td>Optimal observational interval for estimation of mobility of species and suitability of different mobility measures.</td>
<td>IBM</td>
<td>Probabilistic + Feedback, mark–recapture</td>
<td>Difference between observation and real movements of animals increases with less frequent surveys and with animal mobility. Daily surveys should only be done if species are not disturbed easily.</td>
<td>Berger et al. 1999</td>
</tr>
<tr>
<td>Sampling bias</td>
<td>Probability of sampling negative growth in dependence of time since invasion.</td>
<td>IBM</td>
<td>Subsampling, permanent plots</td>
<td>Probability of sampling negative growth increases with time since first invasion. Populations stagnate in size when maximum of local invasive potential reached.</td>
<td>Nehrbass et al. 2006</td>
</tr>
<tr>
<td>Wildlife</td>
<td>Quantifying SDM parameter bias conditional on detection errors.</td>
<td>Descriptive model / SDM + stochasticity</td>
<td>Probabilistic</td>
<td>Estimating and correct for non-detection error requires multiple sampling occasions. Estimating relationships between probability of detection and habitat covariates to identify patches with need for higher sampling effort.</td>
<td>Gu and Swihart 2004</td>
</tr>
<tr>
<td>Wildlife</td>
<td>Comparing methods to account for residual spatial autocorrelation in species distribution modelling.</td>
<td>Descriptive model / SDM</td>
<td>Subsampling</td>
<td>Spatial autocorrelation in response and environmental variables disturbs classical tests of significance while spatial autocorrelation in a single variable has no effect.</td>
<td>Legendre et al. 2002</td>
</tr>
<tr>
<td>Wildlife</td>
<td>Comparison of threshold criteria for a wide range of sample sizes and prevalences.</td>
<td>Descriptive model / SDM</td>
<td>Subsampling</td>
<td>Sensitivity–specificity difference minimizer and sensitivity–specificity sum maximizer criteria produced the most accurate predictions. However, in all cases, the threshold value chosen and the research goals that determined its choice must be stated.</td>
<td>Jiménez-Valverde and Lobo 2007</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>System modelled</th>
<th>Issues addressed</th>
<th>Virtual ecological model</th>
<th>Virtual sampling model</th>
<th>Main conclusions</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wildlife</td>
<td>Comparison of regularisation methods for SDMs.</td>
<td>Descriptive model / SDM</td>
<td>Subsampling</td>
<td>No regularization method performed best under all circumstances. Variable selection should be used with caution. Ridge and lasso are risk-averse model strategies, preferably esp. for small sample sizes.</td>
<td>Reineking and Schröder 2006</td>
</tr>
<tr>
<td>Forests</td>
<td>Comparison of modelling techniques for the broad-scale mapping of forest characteristics.</td>
<td>Descriptive model / SDM</td>
<td>Subsampling</td>
<td>MARS and ANN performed best within VE, but much smaller differences were seen with real data because of noise or possible lack of nonlinear relationships between response and predictor variables.</td>
<td>Moisen and Frescino 2002</td>
</tr>
<tr>
<td>Grasslands</td>
<td>Comparison of modelling techniques for predicting ecosystem attributes.</td>
<td>Descriptive model / SDM</td>
<td>Subsampling</td>
<td>For the six traits analysed, ANNs were able to make better predictions than regression models.</td>
<td>Paruelo and Tomasel 1997</td>
</tr>
<tr>
<td>Wildlife</td>
<td>Compare performance of SDM algorithms regarding underlying response shapes, direct and indirect predictors.</td>
<td>Descriptive model / SDM</td>
<td>Subsampling</td>
<td>Ecological knowledge and statistical skills of the analysts were more important than the method used.</td>
<td>Austin et al. 2006</td>
</tr>
<tr>
<td>Wildlife</td>
<td>Compare performance of SDM algorithms conditional on prevalence, sample size, selection procedure.</td>
<td>Descriptive model / SDM + stochasticity</td>
<td>Subsampling</td>
<td>Recommend the use of GAM or GLM over classification trees or GARP. SDMs for species with low prevalence can be improved through targeted sampling.</td>
<td>Meynard and Quinn 2007</td>
</tr>
<tr>
<td>Wildlife</td>
<td>Compare performance of SDM algorithms conditional on colonisation history.</td>
<td>Descriptive model / SDM + stochasticity</td>
<td>Subsampling</td>
<td>GLM was badly affected in the case of the spreading species but produced slightly better results than ENFA when the species was overabundant; at equilibrium, both methods produced equivalent results.</td>
<td>Hirzel et al. 2001</td>
</tr>
<tr>
<td>Arboreal marsupials (greater glider)</td>
<td>Suitability of SDMs for identifying source habitats.</td>
<td>IBM</td>
<td>Subsampling</td>
<td>SDMs based on logistic regressions measure the ability of species to reach colonize habitat, not their death/ birth rates.</td>
<td>Tyre et al. 2001</td>
</tr>
<tr>
<td>Cerrado vegetation (savanna)</td>
<td>Performance of SDMs coupled with spatial eigenvector mapping under range expansion.</td>
<td>Cellular automata</td>
<td>Subsampling</td>
<td>Mechanisms that generate range cohesion and determine species' distribution under climate changes can be captured by spatial modelling.</td>
<td>de Marco et al. 2008</td>
</tr>
<tr>
<td>Fish (stream trout)</td>
<td>Usefulness of SDMs for assessing the fitness potential provided by habitat and for predicting population responses to habitat alteration.</td>
<td>IBM</td>
<td>Subsampling</td>
<td>Little can be inferred about the fitness value of habitat from observed habitat selection. Recommend that SDMs be supplemented with mechanistic approaches.</td>
<td>Railsback et al. 2003</td>
</tr>
<tr>
<td>Wildlife (arthropods)</td>
<td>Effects of transient dynamics and ecological properties and processes on the prediction accuracy of SDMs under climate change.</td>
<td>Lattice model</td>
<td>Subsampling</td>
<td>Different range dynamics lead to different prediction accuracies of SDMs under climate change. Study pinpoints relevant processes which should be incorporated into SDMs.</td>
<td>Zurell et al. 2009</td>
</tr>
<tr>
<td>Descriptive community assembly models</td>
<td>Evaluate phylogenetic community metrics and their statistical power to detect phylogenetic patterns formed by ecological (competition, habitat filtering, or neutral processes) and trait evolution processes (conserved and convergent traits).</td>
<td>IBM</td>
<td>Full census</td>
<td>Very few tests gave consistent type I error rates over a range of different conditions. Most tests reject the null hypothesis (that only neutral processes structured spatially the local community) too often when the randomization algorithm broke down a structure in the original data set. Tests often showed better conformance when applied to a single study site rather than to multi-study sites.</td>
<td>Hardy 2008</td>
</tr>
</tbody>
</table>

**Table 1. (Continued)**
Natural communities

**Descriptive model**  
**Full census**

Patterns due to competition are better detected by nearest-relative tests, patterns due to habitat filtering are better detected with total community relatedness tests. Statistical power strongly depends on the size of the local community relative to the regional pool with larger pool sizes increasing power for habitat filtering patterns but decreasing power for competition patterns.

*Kraft et al. 2007*

Natural communities

**Descriptive model**  
**Full census**

A null model that preserves matrix row and column totals has lower type I and type II error probabilities than a null model that relaxes row and column totals (which is used in the popular nestedness temperature calculator).


Dynamic (meta-)population models

**Animals (butterflies)**

Introduce model of individual capture histories that allows to measure rates of migration and survival in metapopulations.

*Hanski et al. 2000*

**Animals (butterflies)**

Evaluation of an improved method for parameter estimation of IFMs for a range of varying data scenarios.

*Moilanen 1999*

**Wildlife**

Studying the effect of different types of error in data used to parameterise IFMs.

*Moilanen 2002*

**Wildlife**

Comparing the susceptibility of different SPOMs to predicting spurious trends in metapopulation size.

*Moilanen 2000*

**Animals (grasshoppers)**

Investigate data requirements of IFM, and compare two different SPOMs.

*Hilker et al. 2006*

**Wildlife**

Examine how robust five PVA models are to neglected effects of species interactions.

*Sabo and Gerber 2007*

**Wildlife**

Assessment of the reliability of predictions for extinction risks from PVA conditional on available data.

*McCarthy et al. 2003*

**Plants (Proteaceae)**

Introduce likelihood framework for estimating demographic models for range dynamics of woody plants.

*Cabral and Schurr 2009*
Past use of VE

Testing and improving sampling schemes and methods

In many field studies, ecologists obtain data that are known to be biased. Nevertheless, such data may provide valuable information particularly in cases where the ratio of measured variables between ecological systems is of interest. Knowledge about the error range of each variable is essential, as it might differ depending on the particular observation scenario. An increasing number of studies already optimise the error ranges of their chosen observation scenario by a virtual or theoretical comparison of optional scenarios beforehand (Table 1). In the following we chose three of these studies to illustrate the range of potential fields of application.

Entomologists frequently use mark-recapture methods to monitor the position of grasshoppers or ground beetles in order to understand their behaviour and mobility depending on habitat quality, intra-daily variable climatic conditions, or interactions with con-specific and other animals. Based on the resulting data on positions at different times, various mobility variables are calculated, for example the mean daily movement, maximal distance between two locations an individual was captured, or mean activity radius. These indices may be biased and their quality may differ depending on the particular observation scheme, sample size, edge effects, spatial discretisation among others (Berger et al. 1999).

It seems reasonable to assume that the quality of mobility variables increases with the frequency of observations. However, too frequent or dense observations will disturb the individuals and might artificially increase their activity (Fig. 2). It is thus necessary to optimise the observation scenario related to the minimisation of the observation error and, simultaneously, to minimise the disturbance effect by the observer. The VE approach was used for this optimisation (Berger et al. 1999). The ‘virtual ecologist’ samples the data according to the observation schemes applied in the field and disturbance effects on grasshopper are included in the model. The comparison of the ‘real’ mobility variables (obtained in the virtual world) with the sampled variables provides a quality assessment of the various variables depending on the particular survey method and allows to rank their suitability.

The VE approach can also be used for assessing the compatibility of different sampling methods across spatial scales (Mac Nally 2001). Mac Nally asks whether comparing experimental units of different size may cause scaling artefacts. He tests the ability of the two most common methods to estimate the strength of interaction between competing species, enclosures and quadrate- or transect-based techniques, and whether information from the two sources can be mixed, which often is done for parameterising so-called community matrix models (Wootton 1995). In his simulation model, Wootton (1995) describes three types of foragers (mimicking micro-algal grazers on rocky shores) which are distinguished by their foraging strategy (‘random walkers’, ‘homing’, ‘searcher’). Mac Nally (2001) found that for foragers that apply a more ‘intelligent’ foraging strategy, including dynamic decision-making capabilities, the mixing of data from field-enclosure experiments and quadrate-based methods is ill-advised because the error of these two methods scales differently with the size of the sampling plot.

A third example is related to tree-mortality relationships. Tree mortality is a key process in forest dynamics. In many cases, tree death is preceded by periods of slow growth, and many forest succession models incorporate growth-mortality relationships. Few studies, however, quantify the growth-mortality relationship from empirical data. One question concerns the accuracy of growth-mortality models that are based on tree-ring data, forest inventory data or a combination of both. Wunder et al. (2008) address this question with a VE approach. An individual-based virtual forest model included growth, mortality, snag standing time and regeneration of trees. The forest was subjected to alternative sampling regimes (tree-coring, forest inventories). Growth-mortality relationships were estimated with statistical models of varying flexibility, and were compared to the a priori specified relationships. Highest accuracies were found for tree-ring based models, which require only a small sample size (60 dead trees). High model accuracies were also found for forest inventory-based models, starting at sample sizes of 500 trees. Overall, the study provided guidelines for efficient sampling schemes in real forests.

Testing and comparing models

Within this field of application we can compare the efficiency of different modelling approaches including algorithmic choices, or the effects of different model structures and complexities. We distinguish different classes of problems that can be unified conceptually or technically: first, we list examples of VE studies testing and comparing species distribution models (see Glossary), followed by studies that tested descriptive models in the context of community assembly theory. Finally, we present studies that used VE to test statistical modelling frameworks to parameterise dynamic population models of differing complexity.

Figure 2. Movement of one exemplary individual over a 100 day period; (a) undisturbed and (b) influenced by an observer’s motion during daily surveys (after Berger et al. 1999).
Species distribution models
Species distribution models are commonly used to characterise suitable environmental conditions for a species by relating incidence data to environmental variables (Guisan and Zimmermann 2000). The resulting species–habitat relationship can be extrapolated in space and time to identify the spatial distribution of potentially suitable habitats. Steps in species distribution modelling involve data acquisition, selection of modelling algorithm, model calibration including selection of important predictor variables and parameters, creation of habitat suitability maps, and model evaluation. VE studies usually focused on specific steps of this model building procedure.

Several VE studies tested and compared the performance of alternative modelling algorithms (Hirzel et al. 2001, Legendre et al. 2002, Moisen and Frescino 2002, Tyre et al. 2003, Reese et al. 2005, Austin et al. 2006, Dormann et al. 2007, Meynard and Quinn 2007) conditional on e.g. response shapes, direct and indirect predictor variables, prevalence, sample size, spatial autocorrelation, or colonisation history. Reineking and Schröder (2006) compared regularisation and variable selection methods for model calibration. Other studies tested different threshold criteria (Jiménez-Valverde and Lobo 2007) or the use of favourability functions (Real et al. 2006, Albert and Thuiller 2008) to convert the species distribution model output to maps of presence or absence.

All these studies focused on the methods’ ability to correctly reproduce the current distribution pattern of the virtual species. Simple descriptive models were used to create these patterns. Only few studies were concerned with the processes behind those distribution patterns, and simulated the virtual ecosystem and driving processes by means of mechanistic modelling (Tyre et al. 2001, Railsback et al. 2003, de Marco et al. 2008, Zurell et al. 2009).

Tyre et al. (2001) examined whether species distribution models are capable of identifying source habitats with high birth rates and low death rates and, thus, whether demographic processes can be inferred from simple distribution patterns. De Marco et al. (2008) evaluated the performance of SDMs coupled with spatial eigenvector mapping under range expansion. Railsback et al. (2003) and Zurell et al. (2009) assessed whether species distribution models are able to project species distribution into the future when species undergo transient dynamics due to environmental change. Species distribution models are increasingly used to project shifts in species distributions for different scenarios of climate change (Thomas et al. 2004, Thuiller 2004) and land use change (Pompe et al. 2008). Since the future is unknown, these expected distributional changes are difficult to evaluate, and the use of species distribution models for global change projections remains hotly debated (Dormann 2007).

Zurell et al. (2009) utilised VE to explore the performance of species distribution models under climate change scenarios, and tested the effects of transient dynamics and ecological processes on projection accuracies. To accomplish this, they created a virtual ecosystem by means of mechanistic modelling that included three species, a butterfly, a host plant and a predator, and incorporated species-specific properties and processes such as ecological niche width, dispersal and reproduction, interspecific ecological processes such as competition and predation, environmental stochasticity, and climate change. Virtually sampled data were used to calibrate species distribution models; then, future potential species distribution was projected and evaluated against the simulated ‘true’ distribution of the virtual species. With the VE approach, Zurell et al. (2009) were able to show that the performance of species distribution models for climate change projections strongly depends on the dispersal ability of the species and the extinction rate at the trailing edge of range shifts. Furthermore, their results indicated that species distribution models were useful tools in most of their tested situations. Zurell et al. (2009) were the first to rigorously assess the potential impacts of such factors like dispersal, demographic processes and biotic interactions on global change projections. Nevertheless, they also point out, that their study only scratched the surface of what could be done by using VE with mechanistic models of the virtual ecosystems to test species distribution models. In the future, the complicating effects of several other factors could be explored with this approach such as changing biotic interactions under environmental change, the effects of changing disturbance regimes, local ecological adaptation or the evolution of species niches.

Descriptive community assembly models
Several studies on community structure and assembly rules utilised the virtual ecologist approach. Local communities can be considered as a subset of the larger regional pool of potential community members. Numerous processes (including niche differentiation, environmental filtering, limited dispersal, niche conservatism and convergence) contribute to the formation of the local community from the regional species pool by fostering some species and excluding others. From certain patterns in distributional data, underlying community processes can be inferred by employing different metrics that characterise the community structure and by testing these for significant deviations from the null hypothesis (e.g. the community is locally neutral). Therefore, the question is twofold. First, do different processes result in different patterns of phenotypic, genotypic and trait diversity? Second, do the metrics and null models successfully distinguish between different patterns? The VE approach has been mainly used to address the second question, i.e. to test the performance of different metrics and null models in identifying non-random patterns in biodiversity distribution data.

Here, artificial communities that result from any of the proposed processes are created, for example by using simple filtering algorithms (Fig. 3). For instance, limiting similarity has been modeled by the stepwise exclusion of species with the lowest trait based Euclidean distances to other species while neutrality was modelled by random exclusion (Kraft et al. 2007). Then different metrics and null models are applied and their performance at distinguishing patterns created by different community processes is assessed. Patterns tested have considered nestedness (Fischer and Lindenmayer 2002, Greve and Chown 2006, Higgins et al. 2006, Ulrich and Gotelli 2007a, 2007b) and trait, phylogenetic and species diversity (Kraft et al. 2007).

Hardy (2008) studied how phylogenetic community metrics and null models perform in identifying neutral processes by using an individual-based model to represent the
virtual ecosystem. In contrast to Kraft et al. (2007), he found inflated type I error rates for some null model tests. Hardy argues that the difference in results are due to differences in the structure of the virtual ecological model, Kraft et al.’s (2007) model being much simpler (based on simple algorithms and neglecting individual differences, abundances, the influence of dispersal limitation, and the influence of community size variation). However, Hardy only simulated a neutral community. It would be interesting to see, what happens to the performance of the different indices and null models when applied to a range of distributional patterns generated not by simple filtering algorithms but by mechanistic models.

Dynamic (meta-)population models

The VE approach has also achieved prominence for models of population dynamics, whenever these are parameterised from data. A class of models which has been extensively explored with VE are metapopulation models or stochastic patch occupancy models (SPOMs, Hanski 1999, Hanski et al. 2000). SPOMs describe metapopulation dynamics in a patch network by rates of local extinction and colonisation and are parameterised either from recorded turnover events or spatial data on patch occupancy. For the latter, Moilanen (1999) presents an improved technique for parameter estimation based on maximising the likelihood of observed transitions in patch occupancy. By evaluating the new method with a VE approach, Moilanen (1999) demonstrates that parameter estimates were generally more accurate than those produced by the original method. In a similar study, the new method showed to be less susceptible to the prediction of spurious trend in metapopulation size than other methods (e.g. logistic regression of turnover rates), especially when only snapshot data from two years is used (Moilanen 2000). While both these studies used exact data, Moilanen (2002) imposed error on the virtual measurements of both patch area and patch occupancy and simulated oversight of patches during survey in order to study the effect of different error types on parameter estimation and predictions and, thus, to guide survey efforts accordingly. Extending the VE approach further by using an IBM for the ecological simulation enabled Hilker et al. (2006) to compare the performance and data needs of a patch-based SPOM against a grid-based analogue.

Another field of population modelling studied by VE experiments is population viability analysis (PVA). For example, McCarthy et al. (2003) assessed absolute and relative predictions of extinction risks for a total of 160 parameter scenarios using the stochastic Ricker model. To scrutinise common assumptions of single-species PVA, Sabo and Gerber (2007) simulated time series of population abundance with a stochastic stage-based predator–prey model. Both demographic PVA models and time-series PVA methods were tested for the effect of neglected species interactions on predictions of quasi-extinction risk for the prey.

A more challenging task is the parameterisation of spatially explicit demographic models from species' count data. For the development and verification of parameterisation techniques the VE approach can be an (in-)valuable tool. An example was performed by Cabral and Schurr (2009) using hybrid models of species distribution (Fig. 4). The authors aimed to parameterise both the mechanistic demographic model, which simulated the range dynamics of a species within its suitable habitat, and the observation model, which incorporated sampling error of the survey data set used for parameterisation. With a selected combination of demographic and observation parameter values, they simulated virtual data in five different fractal landscapes. Using these virtual survey data, they assessed whether the applied parameterisation framework was able to recover the underlying parameters. Although the fitted parameter values could vary around the correct values, the median values over the five different landscapes were strikingly close to the correct values, confirming the suitability of the parameterisation technique.

Discussion

The VE approach provides an important, unifying framework to test sampling methods as well as statistical analysis and modelling methods (Hilker et al. 2006). More and new methods are constantly appearing in ecology, especially as more computer power becomes available. These methods need to be tested rigorously and continuously before applying them to real data. VE is an intuitive and powerful method to do so. It has been used in ecology for a long time without being properly recognised or acknowledged. We think that VE deserves a more prominent place in the ecological toolbox.

VE is particularly suitable for synthesising our mechanistic understanding of factors influencing our study results: system-immanent properties and processes such as animal movement, methodological aspects such as observer behaviour and analysis tools as well as interactions of both. The VE models can incorporate an increasing level of complexity that allows the separation of different factors, and it can be carried out at spatial and temporal scales that would be impossible to tackle in reality.
The behaviour of individual ecologists can be simulated in particular situations and, thus, potential problems arising during data sampling can be extensively explored: limited access to certain areas (e.g. lack of roads, steep slopes); spatial autocorrelation in the samples and in the way ecologists move; interactions with the observation target; varying detection probabilities among other factors. Specific problems can be isolated and thereby better understood. A simulation can help to optimise resources and get an idea of the necessary sampling effort for a desired level of accuracy, given site access, budget constraints, sampling bias, and current knowledge of the system. This becomes particularly important when we are about to spend a large budget in surveying a large area, for example.

VE allows to compare alternative methods and thereby to theoretically circumscribe their application domain. The most appropriate model for any situation can be selected, i.e. the best modelling approach for a given data set, and crucial data needs for the application of more complex descriptive or even mechanistic models may be identified (Hirzel et al. 2001). This has to be seen different from and is more sophisticated than model selection techniques. In model selection the fit of potential models to the data is assessed and models are then ranked according to their predictive power (Burnham and Anderson 2002). For instance, Gotelli et al. (2009) recently proposed a modelling strategy that employs parametric bootstrapping to assess the fit of simulation models and to rank competing models according to their ability to explain large-scale diversity patterns. At first sight, this sounds very similar to VE. However, in model selection the goodness of fit of alternative models can only be evaluated on the given data which might be limited and biased. In contrast, VE allows the models to be evaluated against known (virtual) truth. Thus, in a VE study the question is not about how well the model fits the data but how well the model represents (virtual) reality and under which circumstances it does this.

**Limitations**

Beside the merits of the virtual ecologist approach, modellers must be aware of possible limitations of VE, which are actually more related to the models used or to the simulation design than with VE itself. Foremost, the benefit of VE depends on the quality of the ecological model, and ignores whatever complexity is not covered by the model. Models by definition simplify; the real world is much more complicated, and conclusions drawn from the virtual data sets might be limited. Wunder et al. (2008) point out that when using VE to identify necessary sample sizes to achieve a desired level of accuracy, these values constitute only lower bounds as they were estimated under the controlled conditions of the virtual reality. In the model of Berger et al. (1999), grasshoppers moved according to a random walk. Deviations from this movement behaviour might lead to a different ranking of the observation errors. However, different movement modes can be implemented and tested in the model, as in the example of Mac Nally (2001). Generally, VE is better at discrediting
methods than at corroborating them. If a method fails in the virtual world, chances are that it fails in the real world as well, unless the method’s deficits fortuitously counterbalance the virtual world’s biases. However, if a method works well in the virtual world, this does not guarantee that it works in the real world as well.

In addition, models are prone to errors, and we should never put blind faith in our models (Wissel 1992); this also holds for VE. Numerous limitations can be hidden in the modelling process: uncertainty in input data, in underlying model assumptions, in parameters, and bugs in the simulation program itself (Grimm et al. 1999). Thus, as any other tool, VE needs to be used consciously and cautiously, and it should continuously be scrutinised.

Sometimes, the VE approach may seem a bit circular. For example, Hirzel et al. (2001) sampled from the same statistical modelling type, a logistic regression model that they aimed to test. However, even if one samples from the same (statistical) model, running VE is worthwhile. If the tested method is not able to recover the underlying model, then it will not be worth to further develop this particular method.

The role of mechanistic models

Following the famous words of Albert Einstein one should make the models “as simple as possible, but not simpler”. In good modelling practice this means that both the virtual ecological model and the virtual sampling model should be no more complex than is necessary to answer the scientific question. Of course, this also requires a clear definition of the problem and the target underlying the VE study.

If the scope of the VE study is to assess whether a pattern may be correctly identified by a particular sampling method or correctly predicted by a model then, in most cases, a descriptive model of the virtual ecosystem will be adequate. In contrast, if the scope is to test whether a specific sampling method is able to identify, or a model is able to predict, for example, certain spatial and temporal dynamics or process rates, then a more mechanistic model of the virtual ecosystem is needed in which the processes are simulated in a ‘structurally realistic’ way (Fig. 5; Wiegand et al. 2003, Grimm et al. 2005). Also, the decision whether the virtual sampling model should be descriptive or mechanistic should be driven by the scope of the VE study; that is questions like: should observer errors or biases be included; are there interactions between the observer and the species (Fig. 5)?

Nevertheless, we want to emphasise that a contemporary shift towards generating virtual species/ecosystem and observer from mechanistic models can qualitatively enhance the potential of the VE approach. Mechanistic models can account more realistically for complexity in both ecological and observational processes, including possible interactions. Specific problems or aspects of ecological systems can be incorporated. Data are still controlled, but potentially behave in a non-trivial manner. The exercise becomes one that is equally about understanding complex dynamics and optimising the way we can study them empirically by using mechanistic, ‘close to nature’ simulation models. In mechanistic models of virtual species/ecosystems one has to take care of complicating effects such as coloured noise, stochasticity, and deterministic chaos. We can thus test whether our method under study is working even in the face of such complex inherent interactions, and hence delineate the method’s application domain more accurately.

Individual-based models (IBMs) are the most general mechanistic models as the emergence of metapopulation dynamics is the result of individual interactions in a landscape mosaic (Grimm 1999, Hilker et al. 2006). IBMs differ from descriptive models or mechanistic models on a more aggregated, metapopulation level, in that the ‘true’ values of the population-level parameters we try to estimate are not necessarily known, but rather are an emergent property (Hilker et al. 2006). The parameters can be estimated, however, in the IBM with arbitrary precision because we can produce as many replicates as required (at least if sufficient computer power is available). The efforts of such a complex IBM might be justified if the field study is a non-repeatable project; if a wide-spread sampling method is to be evaluated; or if we want to test how mechanistic models on a more aggregated, metapopulation level converge to more complex (virtual) reality (Hilker et al. 2006). A full-fledged VE approach with the virtual species modelled by means of individual-based modelling and explicit interactions between virtual observer and virtual species (Berger et al. 1999) may be useful in survey planning of highly mobile and sensitive species.

Figure 5. Decision tree which methods to use for the virtual ecosystem and the virtual sampling model for which purposes (IBM: individual-based model).
Future directions

We have shown various applications and research fields where the virtual ecologist approach has been successfully employed, and has proven itself as a practical and worthwhile tool. As pointed out throughout this review, the approach is not yet fully explored and many more aspects of ecological surveys and modelling tasks can be addressed with VE.

The ecological community holds enormous stocks of data collected, for example, in herbaria; by voluntary or hobby ornithologists, entomologists; nature conservationists; PhD students etc. Sometimes trust in this data is rather limited because of suspected bias in survey design or observer behaviour. For instance, volunteers monitoring butterflies will often preferentially visit places where they expect to find the most enigmatic and interesting species. Conversely, places where observers do not expect to find many species are likely to not be monitored properly or only very short visits will be paid to such places. Through such unequal observer effort fallacious absences (and also presence) might be induced with unknown effects for subsequent data analyses. Here, VE could help to assess potential effects rigorously and to assess sampling bias if information on the observer effort is available; the data could then be corrected by these estimated values. However, we want to stress that VE is no panacea for flawed survey designs. It can merely be a way to salvage at least some of the information in the data.

Another important research field for which VE holds great potential for the future is global change research. Railsback et al. (2003), Cabral and Schurr (2009) and Zurell et al. (2009) show that VE can help to evaluate models which are intended to project species distributions into the future for different scenarios of environmental change. The effects of many other factors potentially complicating global change projections could be explored with VE: changing biotic interactions or spatially dependent biotic interactions that only take place at the edges of species distributions, behavioural adaptation, evolutionary effects, invasions, climatic extremes or catastrophic events. The VE approach would also allow to assess projections that address the effects of climate change or land use change for individual species with particular species-environment relationships, or to integrate species with different functional characteristics into assessing the effects of global change in whole communities or ecosystems.

In addition to these potential future directions, the virtual ecologist approach could, if it were used more routinely in the future, have more general and perhaps even more important benefits: it could foster the integration of theoretical and empirical work. Empiricists are often unaware of the potentials and limitations of ecological models, and the same holds for theoreticians regarding field work and sampling methods. Working together on the development of sampling methods, designs and efforts by using the VE approach could help overcome this mutual ignorance. It could help practitioners to better plan their work. It could help modellers to increase the practical value of their work. It could also stimulate work that goes far beyond sampling methods. While trying to test sampling methods, new and interesting ecological models and even theories might emerge; and while trying to use existing models for testing sampling methods, ecological models might become more realistic in structure and lead to new questions for empirical research. Also, field work could be oriented more directly towards data needs modellers have for specific modelling tasks. Looking at their models from the perspective of optimising empirical work might make work of theoreticians more valuable, and also it might help to better understand the system.

Acknowledgements – We would like to thank J.A.F. Diniz-Filho and T.F.L.V.B. Rangel for valuable comments.

References


