

MATHEMATICAL MODELS OF BIOLOGY

P. Haccou

Institute of Biology, Leiden University, The Netherlands

Keywords: modeling philosophy, function of models, evolution, ecology, population dynamics, evolutionary genetics, evolutionary ecology, optimal foraging, ideal free distribution, marginal value theorem, game theory, quantitative genetics, population genetics, branching processes, adaptive dynamics, ESS

Contents

1. About modeling
 - 1.1. What Are Models?
 - 1.2. Features of Models
 - 1.2.1. Manageability
 - 1.2.2. Generality
 - 1.2.3. Precision
 - 1.2.4. Realism
 - 1.3. Application of Models at Different Stages of a Research Program
 - 1.4. Modeling Philosophies
 2. Archetypical models of evolution and ecology
 - 2.1. Population Biology
 - 2.1.1. Non-Spatial Models
 - 2.1.2. Migration and Spatial Structure.
 - 2.1.3. Metapopulation Dynamics
 - 2.1.4. Demographic Stochasticity
 - 2.2. Evolutionary Genetics
 - 2.2.1. Population Genetics
 - 2.2.2. Quantitative Genetics
 - 2.3. Evolutionary Ecology
 - 2.3.1. Optimal Foraging Models and the Ideal Free Distribution
 - 2.3.2. Evolutionary Game Theory
 - 2.3.3. Adaptive Dynamics
 3. Conclusions
- Acknowledgements
Glossary
Bibliography
Biographical Sketch

Summary

Mathematical modeling is essential in contemporary biology. The requirements of models differ according to their goal and the phase of research at which they are applied. This should be taken into account when considering the merits of different models, but especially aspects relating to the research phase are often overlooked. Even though there is a wide variety of conceptual models almost all of them can be considered as generalizations that are derived from a small number of archetypical

models. In this chapter we give an overview of such basic models in evolution and ecology. Despite their simplicity, these models have proved useful in the study of many biological systems. In general, the bottom up approach, going from simple models with a high degree of abstraction to models with a higher complexity appears to be more fruitful than the reverse, top down approach. Examples of how models are applied to study phenomena in different fields of biology can be found in the remaining chapters under this topic.

1. About Modeling

1.1. What Are Models?

The concept of a model is derived from scale models, referring to simplified replicas of larger structures such as, for instance, buildings or ships. This implies a structural similarity between a model and its original. Thus models are simplified representations of certain aspects of our study systems. The degree of simplification may vary, and even might resemble a caricature, but it is always there.

The aspects of the research system that are considered in a model may differ, from underlying mechanisms (so-called mechanistic models), to global forms of results of these mechanisms (as in phenomenological models), or statistical properties of observations on certain aspects of a process (as in 'black box models', such as regression). Thus, the meaning of 'structural similarity' between a model and its original depends on the aspect of the system that is being modeled. In 'black box models', this similarity may seem to be absent, but it is not. Since these kinds of models are applied in data analysis, the aspect of the studied system that is modeled here concerns the empirical data. Thus in these models the aim is a structural similarity with the observations, such as their distributions, their relations with measured or manipulated variables, and their interdependencies.

The most important function of models is to order our thoughts. With models we formulate what we (think we) know about the world, and we may use them to perform thought experiments through 'what if' scenarios. Since this is an essential part of research, it can be argued that all scientists make models of their study system. Initial models are often verbal rather than mathematical. In simple cases verbal models may suffice. As things become more complicated, however, it is more and more difficult to keep track of verbal arguments and to check their consistency. Verbal models therefore involve a risk of overlooking important factors and/or introducing logical inconsistencies. Here, mathematics provides a powerful language that forces us to be logically consistent. Although a set of equations may seem daunting and complex, in fact in most instances it is much easier to check the logic of an argument from such a list than if it were formulated in ordinary language. The latter also takes up much more space!

Another important use of models is to function as an idealization of the world. Models force us to make a choice concerning which aspects of 'the real world' we include into our description and which we choose to ignore (for the moment). This is true for verbal as well as mathematical models. It is usually only noted more quickly in mathematical

models. The latter is yet another advantage: mathematical formulations reveal implicit assumptions and thus force us to be explicit. Thus mathematical formulations enforce consistency of assumptions and make clear what the range of validity of conclusions derived from them is.

1.2. Features of Models

Modeling involves a choice what to include and what to leave out. This choice is first of all limited by what we know about a system. It is simply impossible to observe all possible aspects of reality, and we are thus necessarily constrained. But, besides that, we make choices concerning which aspects of what we know about a system to put into a model. This implies that there is a multiplicity of possible models for each single study system. As a consequence there is an ongoing discussion in biology about the general use of models, and the application and merits of specific models.

Levins pointed out that several different models might be useful for the same study system, depending on the modeling goal. He distinguished three ultimate goals of modeling: understanding, predicting and modifying nature, and argues that achieving one might require different models than achieving another. One of the great merits of Levins' paper is that it lists four features of models: manageability, generality, precision, and realism. Levins believed that there are trade offs between these, that is, that they cannot be maximized with one single model, and, thus, depending on the goal a choice should be made which aspects are valued the most and which should be sacrificed. This paper has led to an ongoing debate about whether these features could (and should) be used as a basis for model selection and whether trade offs exist or not. We will not go into that discussion here, but instead focus on these four features and their implications for modeling.

1.2.1. Manageability

The most manageable models are completely tractable. This means model equations can be solved analytically, and different types of model dynamics can be related explicitly to parameters or combinations thereof. The advantage of complete analytical tractability is that it provides explicit expressions for the dependencies between (combinations of) values of model parameters and model results. Thus, it is clear where results come from. Furthermore, we can completely specify the boundary conditions for which predictions of the model are valid. Manageability also relates to being able to examine robustness of a model with respect to its assumptions. In the ideal case we are able to examine effects of different types of generalizations of a model and make statements on how the results depend on model assumptions, which assumptions are crucial for its main predictions and which are not. Note that a decreased manageability of models automatically implies a decreased knowledge of the boundary conditions and robustness of their predictions.

In less tractable models it is not possible to completely characterize the dynamics of a model. Yet, sometimes still a more-or-less complete picture of the behavior may be obtained by approximation techniques such as linearization, separation of time scales, or pair approximations, combined with numerical determination of, for instance, regions of attraction and repulsion.

In very complex models even this is usually impossible. Then it depends on whether visualization or pattern recognition leads to a reduction of the dimensionality of the state space. If, however, no complete characterization is possible mathematically, there are no such ‘summary statistics’ of a model’s behavior. Then we are left to try and grasp complex dynamics in a high dimensional state space and we are limited by our own cognitive abilities. However, in many instances what we considered complex a hundred years ago is considered basic teaching material right now. Thus, apparently manageability is changing. As more advanced analytical and computational techniques become available, more and more complex models become manageable.

1.2.2. Generality

Models that make few assumptions (are parsimonious) are often stated to be more general than those with many. Especially in the literature on conceptual models statements connecting simplicity of models with generality are often made. This is, however, not necessarily valid, since all models contain assumptions about conditions that do NOT occur. These assumptions are not explicitly stated, simply because the set of conditions that do not occur is infinitely large.

For instance, a model that explicitly assumes two conditions, A and B (denoted model 1) appears to be less general than one that only assumes A (model 2). However, if model 2 implicitly assumes that B does NOT occur this is not necessarily true: it depends on how large is the intersection of sets A and B compared to that of A and NOT B. A truly more general model than model 1 is, for instance, a model that assumes A and (B or C or D). These considerations are usually not made explicit in the literature, but should be given more attention. This implies that robustness of a model’s predictions with respect to changes in its assumptions should be examined. The more robust models will be the more general ones. For instance, models that assume general functional forms (e.g. increasing, convex) are truly more general than those that assume specific functional equations. An example is Levins’ research on effects of changing environments in relation to fitness landscapes.

Some models can be considered as limiting cases of more complex models, which makes them very useful in a general setting. For instance, in most population models, exponential growth occurs at densities close to zero. Furthermore, in many cases model dynamics close to equilibria can also be described accurately by a model with exponential growth or decline.

1.2.3. Precision

Precision relates to the details up to which a model specifies a system. Therefore there might indeed be a strong trade off between precision and generality: a model that very precisely describes a specific biological system can not at the same time be very general, since different systems will always differ in some of their details. If this is true we do not have to consider precision as a separate aspect of models.

However, the author believes that there are some additional remarks to be made in this context. Models can be very detailed in that they specify individual behavior, spatial

structure etcetera, or they may consider the effects of collective behavior of groups of individuals. The latter type of models are not necessarily more general: they assume that characteristics of different individuals within the same group are equal.

Models can also differ in the degree of detail in which they consider mechanisms underlying certain phenomena. So-called mechanistic models contain equations for those phenomena that are explicitly derived from mechanisms, whereas phenomenological models use equations that are assumed to describe the phenomena reasonably well. An example of a well-known phenomenological model for population growth is the logistic growth model. In this model it is assumed that population growth goes to zero as population density is small as well as when it approaches the carrying capacity of the population. Above the carrying capacity, population density declines. These are reasonable assumptions for many populations. To describe the functional relation between population density and growth the model uses a quadratic function, which is the simplest polynomial function with the required properties. Whereas this model is formulated on the basis of considerations concerning interactions between individuals, however, this functional relationship is not derived from mechanisms of those interactions. Note, however, that the quadratic function can also be considered as a second order approximation of other, more mechanistically based functional relations. From this perspective the logistic model is an approximation that captures the most important features (or those that are deemed so) of more detailed population growth models.

1.2.4. Realism

Realism has to do with the structural similarity that there is supposed to be between models and the aspects of the underlying phenomenon that they are meant to describe. Model assumptions should therefore at least have some degree of realism. There are certain boundary conditions concerning realism, that every biologist would agree with, imposed by physical and biological laws. For instance, there is no '*generatio spontanea*', that is, at zero population densities no individuals should be born. Also models should respect mass and energy conservation laws. Further constraints are imposed by the fact that certain quantities, such as height, length, weight and others cannot become negative. Although this may seem obvious, these constraints are not always respected. This occurs for instance in models that incorporate statistical distributions with negative support, such as the Gaussian. And this may not even be such a bad assumption, if the mean of the distribution is sufficiently large compared to the variance, so that the probability of negative values is negligibly low. However, one needs to be aware of these boundary conditions and explicitly consider them when formulating a model. Otherwise, very unrealistic models may result.

Within these constraints, different degrees of realism are possible, but only up to a point. Models that are very far from reality are useless, since they cannot provide any relevant information about the studied system, but models with assumptions that are only partially true might give useful insights. For example most inheritable traits are determined by several loci rather than one. Yet single-locus models have generated much insight in population genetics. However, a model that assumes that traits of an individual of the next generation is totally determined by the current population trait

distribution, regardless of its parental genotypes, would not be considered useful at all.

Although most biologists are able to judge quite easily whether or not a model is unrealistic in a useful way, it is difficult to describe what exactly makes the difference. One important factor is certainly the possibility of embedding the model within another, more realistic one, without totally having to change its structure. Another point is the importance of specific assumptions for the main model predictions. Assumptions that crucially determine important results deserve to be looked at quite critically, and they usually are hotly debated.

Note that, since the meaning of structural similarity depends on the modeled aspect, realism also relates to that. For instance, a model of ecological interactions between species can be grossly unrealistic about its assumptions about the physiology of individuals. Such assumptions are not explicitly stated, but most models imply, for instance, that individuals are equal in all aspects relating to physiological traits or states. As long as those individual features are not important for the studied ecological interactions, however, they may be ignored.

1.3. Application of Models at Different Stages of a Research Program

In the discussion about models and their merits it is usually not explicitly mentioned at what stage of a research program models are applied. Yet this is an important distinction. At the conceptual stage of research models are used to further develop theory. At the other extreme lies the practical stage, where models are used to closely examine real-life phenomena.

Models used at the conceptual stage are called ‘strategic’, or ‘conceptual’ models. At this stage the behavior of models rather than concrete biological systems is studied, and results are limited to the world of the model. As Kokko formulates it: Conceptual models are used for doing thought experiments of the form: ‘all else being equal’ what is the effect of a certain factor or process? She argues that in this respect they are comparable to experiments, where effects of confounding factors are also eliminated as far as possible. However, whereas such experiments are aimed at finding real life phenomena, the theoretical research is aimed at finding out what would happen if the world was like the model assumes. Thus, conceptual models are thinking aids rather than investigations of natural phenomena. For instance, we can gain insight into the question why sexual reproduction with two sexes evolved, by looking at what would happen if there were three or more types.

At the practical stage models are used for purposes such as inference from empirical data, making quantitative predictions concerning future system dynamics, or effects of studied factors, and guiding management decisions. These models comprise ‘statistical’ models, ‘computer models’, etc. Their aim is to make accurate statements about the studied system given the circumstances in the real world.

Which features of models are considered the most important depends on the stage of the research. For theory development it is necessary that relations between model results and parameter combinations are sufficiently clear, as are interdependencies between

variables etc. Therefore, manageability of a model is important. Also generality of results is usually valued highly here. So conclusions should be robust and not depend on specific details of a system. Therefore, approximate models are preferred over those that make highly specific assumptions.

With respect to realism, models used at this research stage usually incorporate global characteristics that can be found in several empirical systems, but they may be unrealistic with respect to details for specific systems. For instance, the marginal value theorem for optimal foraging is based on a model where food is encountered in patches, with a constant patch encounter rate, and a concave within-patch gain function. It is assumed that environmental depletion occurs much slower than the time scale of the foraging behavior. There are many species with clumped resource distributions, where patches are more or less randomly distributed in space. This model provides a useful starting point for developing theory concerning the behavior of all those species. However, it is highly unrealistic in assuming complete information, no effects of satiation, no intra patch competition, etc. Therefore it would not be wise to apply it directly to an empirical research system.

At the practical research stage, generality is considered less important, since the aim is to get precise insight in a specific system. However, we usually do want results to be robust for certain types of variation. For instance, it would not be very useful if a new model is needed for each different experimental setup. Manageability might not be very important either, provided that sufficiently accurate parameter estimates are available. However, if complex models are used, whose behavior can only be explored numerically, it is good to be aware of the fact that the validity of the conclusions is constrained by the examined range of parameter values. Manageability thus becomes especially important in situations where we do not know parameter values and cannot estimate them accurately. In such situations the sensitivity of model results to variation of unknown parameters should be examined explicitly.

The aim of modeling at this stage is to mimic behavior of the studied real-life system as accurately as possible given the available information. This does not necessarily mean that models should contain much mechanistic detail. Rather, it is important that they are in agreement with what is known about the data structure, e.g. statistical distributions, and interdependence relations between variables. Criteria for assessing models are the model's fit in relation to the number of parameters (Akaike criterion), and consistency with constraints, such as positivity of measures like weights, lengths, times. An important distinction at this point is whether models are used for inference or prediction (management goals). In the latter case more mechanistic models are preferable, since they need to be able to make predictions outside the range of observed measurements.

From the previous it will be clear that not all models are meant to be tested or fitted to empirical data: it depends on their goal. This is sometimes confused, partly due to the fact that mechanistic models with added error structure are sometimes used as statistical models in the practical research phase. Confusion may arise due to the fact that these models resemble conceptual models. In general one should not try to fit a strongly simplified conceptual model to data, since it is too far removed from the data. The same holds the other way round: the outcome of statistical models usually does not

immediately provide major conceptual insights. Rather, global trends apparent in these outcomes are used to adjust conceptual models, and then their consequences for the theory are examined.

These remarks do not imply that conceptual models should not in some way be confronted with experiments. In some cases empirical tests of conceptual evolutionary models have led to remarkable insights. Such tests usually concern the assumptions of models. Most insight is provided if empirical results are not consistent with the predictions of a model, since this teaches us something about the studied system. Namely, that for this system the model assumptions are incorrect, or that there is a fundamental disagreement between the conditions that are assumed in the model and the experimental setup. If, however, the results are consistent, the model might still be wrong. So it is good practice to go on testing model predictions until a disagreement is found, and thus to explore the boundaries for which a model still gives the right predictions.

Even if empirical research is directed at testing predictions from a specific conceptual model, different models are usually needed at different stages of research. Strategic models at most provide qualitative predictions about the system, concerning for instance trends, inequalities, presence or absence of a specific phenomenon in relation to certain circumstances, or coarse dependencies between variables, such as predictions of the form: if x increases then y increases. Statistical models translate these predictions into quantitative, testable statements, such as for example: y increases log-linearly with x . There may be a very loose connection between the conceptual and statistical models that are used to study a specific system. An example is given by the use of optimization models in behavioral ecology. Predictions from those models concern the relationship between environment, such as spatial resource distribution, and foraging experience, such as the encounter rate with prey in a patch, and behavioral decisions, such as patch residence times. However, these models cannot be applied directly to examine data since they are completely deterministic and do not take inter- and intra-individual variation in behavior into account. To test their predictions in practice, therefore, variables representing intra patch experiences are incorporated in statistical models for patch leaving tendencies, and their effects are estimated and tested. Results of the tests are interpreted in the context of the optimization model.

-
-
-

TO ACCESS ALL THE 27 PAGES OF THIS CHAPTER,
Visit: <http://www.eolss.net/Eolss-sampleAllChapter.aspx>

Bibliography

Balding, D.J., Bishop, M., Cannings, C. (Editors) (2007) *Handbook of Statistical Genetics* John Wiley & Sons; 3rd Edition. [An extensive overview of all main genetic models and statistical techniques for

analyzing genetic data].

Burger, R. 2000 *The Mathematical Theory of Selection, Recombination, and Mutation*, John Wiley & Sons, Ltd., Chichester. [A mathematical treatment of population genetic models].

Caswell, H. (2001) *Matrix population models - Construction, Analysis, and Interpretation* 2nd ed. Sinauer Associates, Inc. Sunderland, Massachusetts. [An extensive overview of matrix models in biology].

DeAngelis, D.L. and Gross, L.J. (eds.) (1992) *Individual-based Models and Approaches in Ecology*. CRC Press Inc. [An introduction to individual-based modeling].

Ellner, S.P. and Guckenheimer, J. (2006) *Dynamic models in biology*. Princeton University Press. [An introduction to modeling with examples from many different biological fields].

Haccou, P., Jagers, P., Vatutin, V.A. (2005) *Branching Processes - Variation, Growth, and Extinction of Populations*. Cambridge University Press. [An overview of branching processes, their analysis and applications in biology].

Hanski, I. (2002) *Metapopulation ecology*, Oxford Series in Ecology and Evolution, Oxford University Press. [Introduction to metapopulation modeling].

Hilborn, R. and Mangel, M. (1997) *The Ecological Detective, confronting models with data*. Princeton University Press. [Describes how to perform statistical analyses of data based on ecological models with error terms].

Houston, A.I. and McNamara, J.M. (1999) *Models of Adaptive Behavior - An approach based on state*. Cambridge University Press, Cambridge [An overview of state dependent optimal foraging models and dynamical games and their analysis].

Jagers, P. (1975) *Branching Processes with Biological Applications*, John Wiley and Sons, Ltd. London. [An overview of branching processes, their analysis and applications in biology].

Kimmel, M., and Axelrod, D.E. (2002) *Branching Processes in Biology* Springer-Verlag, New York. [An overview of branching processes, their analysis and applications in biology, with the emphasis on molecular biology].

Kokko, H. (2007). *Modelling for Field Biologists and Other Interesting People*. Cambridge University Press. [An introduction to mathematical modeling in biology].

Lande, R., Steinar, E., Saether, B. (2003). *Stochastic Population Dynamics in Ecology and Conservation*, 212 pp. New York, USA: Oxford University Press Inc. [An overview of population dynamic models with stochastic elements].

Levins, R. (1966) The strategy of model building in population biology. *American Scientist* 54: 421-431 [An influential paper on modeling strategies in biology].

Mangel, M. and Clark, C.W. (1988) *Dynamic Modeling in Behavioral Ecology*. Princeton University Press, Princeton [An introduction to state dependent optimality modeling].

May, M. and McLean, A. (2007) *Theoretical Ecology, principles and applications*. Oxford University Press [An introduction to ecological modeling].

Maynard Smith, J. (1982) *Evolution and the Theory of Games*, Cambridge University Press, Cambridge [An introduction to evolutionary game theory].

Metz, J.A.J. and Diekmann, O. (eds.) (1986) *The Dynamics of Physiologically Structured Populations*. Springer Lecture Notes in Biomathematics 68, Springer- Verlag Berlin Heidelberg [An edited volume on modeling and analysis of structured population dynamics].

Mode, C.J. (1971) *Multitype Branching Processes - Theory and Applications*, American Elsevier Publishing Company, Inc. New York [An overview of mathematical results on multitype branching processes].

Odling-Smee, F.J., Laland, K.N., Feldman, M.W. (2003) *Niche Construction - The neglected process in evolution*. Princeton University Press, Princeton [Introduction to niche construction theory, with mathematical models].

Odenbaugh, J. (2006) The strategy of "The strategy of model building in population biology" *Biology & Philosophy* 21: 607-621 [Reaction to Levins' paper on model building].

Orzack, S.H. (2005) Discussion: What, if anything, is 'The strategy of model building in population biology?' A comment on Levins (1966) and Odenbaugh (2003). *Philosophy of Science*, 72, 479-485 [Reaction to Levins' paper on model building].

Otto, S.P. and Day, T. (2007) *A biologists' guide to mathematical modeling in ecology and evolution*. Princeton University Press [Introduction to modeling].

Okubo, A. and Levin, (2001) S.A. *Diffusion and Ecological Problems - Modern Perspectives* 2nd ed. Springer-Verlag, New York Inc. [An introduction to modeling ecological phenomena with diffusion processes].

Rice, S.H. (2004) *Evolutionary Theory - Mathematical and Conceptual Foundations*, Sinauer Associates, Inc. Sunderland, Massachusetts U.S.A. [An overview of mathematical models in evolutionary genetics].

Shigesada, N. and Kawasako, K. (1997) *Biological Invasions: Theory and Practice*, Oxford Series in Ecology and Evolution, Oxford University Press. [An overview of diffusion models of biological invasions and their applications.]

Tuljapurkar, S. (1990) *Population Dynamics in Variable Environments*, Lecture Notes in Biomathematics 85, Springer-Verlag Berlin Heidelberg [An introduction to and overview of the analysis of matrix population models in stochastic environments].

Weisberg, M. (2006) Forty years of 'The Strategy': Levins on model building and idealization. *Biology & Philosophy* 2: 623-645. [A discussion of Levins' seminal paper on modeling strategies].

Biographical Sketch

Patsy Haccou is Associate Professor of Theoretical Biology at the Institute of Biology of Leiden University, The Netherlands. She was Chalmers University 150th Anniversary Visiting Professor in Gothenburg, Sweden in 2005, and is Chair of the Dutch society for Theoretical Biology (NVTB) since 2005. Her general field of expertise is mathematical biology with a specialization in modeling stochastic aspects of evolutionary and ecological processes. Her current research interests concern the effects of competition on the evolution of resource exploitation strategies, and the dynamics of invasion processes in various contexts.