

Data-adaptive reduction of process-based models

An der Fakultät für Mathematik und Naturwissenschaften
der Carl von Ossietzky Universität Oldenburg
zur Erlangung des Grades und Titels eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

angenommene Dissertation von

Dipl.-Geoökol. Knut Bernhardt

geboren am 03.01.1978

in Kassel

Erstgutachter: Prof. Dr. Kai W. Wirtz

Zweitgutachter: Prof. Dr. Wolfgang Ebenhö

Tag der Disputation: 15.02.2008

To Jan

*I can watch and can't take part
Where I end and where you start*

Thom Yorke

Abstract

Process-based models of environmental systems typically are very complex structures. This complexity arises from the attempt to describe the manifold natural processes and intricate biological interactions of environmental systems in mathematical terms. Because of their high number of state variables and parameters, the resulting complex models are difficult to calibrate and detailed model analysis is needed to extract the key governing processes within these complex structures.

Acknowledging these problems, the present thesis aims at finding a method to reduce complex process-based models. The main objectives for the development of the new method are its general applicability, its automated execution and its ability to construct reduced models which are interpretable in terms of system-specific mechanisms. The Mapping-based Genetic Reduction technique (MAGER) proposed in this thesis is a data-adaptive black-box procedure based on Genetic Programming which can be applied to ordinary differential equation models.

In the course of this thesis, the MAGER scheme is applied to three predator-prey and consumer-resource models of different dimensionality. It is found that even relatively simple models can be reduced further and the results show that a formal conformity of physical and biological oscillating systems exists. In addition, the reduction of the biological systems involves a change in description level. Instead of traditional density or traits variables, the new models incorporate descriptions of biological interactions which leads to the notion of Ecological Interaction Models (EIM) for this new model class. The uniformity of the results further points to the generality of the EIM descriptions.

As the MAGER scheme only depends on time series data and ignores former model structures, it is suggested that the method can also be applied to measured data or models from other scientific disciplines which offers many possibilities for further studies.

Zusammenfassung

Prozessbasierte Modelle von Umweltsystemen sind in der Regel sehr komplex. Diese Komplexität entsteht durch den Versuch, die mannigfaltigen natürlichen Prozesse und komplizierten biologischen Wechselwirkungen in mathematischen Beschreibungen abzubilden. Die resultierenden komplexen Modelle sind aufgrund ihrer großen Zahl an Zustandsvariablen und Parametern schwierig zu kalibrieren. Desweiteren sind detaillierte Modellstudien nötig, um die Schlüsselprozesse innerhalb dieser komplexen Strukturen aufzudecken.

Aufgrund dieser Probleme soll in der vorliegenden Arbeit eine Methode zur Reduktion prozessbasierter Modelle entwickelt werden. Die zentralen Zielsetzungen der Methodenentwicklung sind dabei die generelle Anwendbarkeit, die automatisierte Durchführung sowie die Herstellung reduzierter Modelle, die hinsichtlich systemspezifischer Mechanismen interpretiert werden können. In dieser Arbeit wird dazu die „Mapping-based Genetic Reduction“ Methode (MAGER) entwickelt. MAGER ist ein auf Genetischer Programmierung basierendes, datenadaptives Blackbox Verfahren zur Reduktion gewöhnlicher Differentialgleichungssysteme.

Die Anwendung des MAGER Verfahrens umfasst drei Räuber-Beute und Konsumenten-Ressourcen Modelle unterschiedlicher Dimensionalität. Es zeigt sich, dass selbst relativ einfache Modelle weiter reduziert werden können und die Ergebnisse weisen auf eine formale Übereinstimmung physikalischer und biologischer Schwingungssysteme hin. Die Reduktion der biologischen Systeme führt darüber hinaus zu einem Wechsel der Beschreibungsebene. Anstelle traditioneller Dichte- oder Eigenschaftsvariablen beinhalten die neuen Modelle direkte Beschreibungen der biologischen Wechselwirkungen. Der Begriff „Ecological Interaction Models“ (EIM) wird eingeführt, um diese neue Modellklasse zu beschreiben. Die Einheitlichkeit der Ergebnisse weist außerdem auf die Generalität der EIM Beschreibungen hin.

Da das MAGER Verfahren lediglich auf Zeitreihendaten basiert und ursprüngliche Modellstrukturen ignoriert werden, ist es naheliegend, dass die Methode auch auf Messdaten oder Modelle anderer wissenschaftlicher Disziplinen angewendet werden kann. Hieraus ergeben sich vielfältige Möglichkeiten für weitergehende Studien.

Contents

1	Introduction	1
1.1	The simplicity/complexity dilemma of process-based models	1
1.2	The MAGER approach and previous work	3
1.3	Structure of the thesis	4
2	Reduction of complex models	7
2.1	Introduction	7
2.2	A model of species competition	8
2.3	The Self-Organizing Map	10
2.3.1	Vector quantization of the dataset	11
2.4	Nonlinear projection	11
2.4.1	NLPCA of the SOM-filtered data	12
2.4.2	Interpretation of modes	12
2.5	Discussion	14
3	Finding alternatives and reduced formulations for process-based models	17
3.1	Introduction	17
3.2	State reduction	20
3.2.1	Temporal neighborhood	20
3.2.2	Principal Component Analysis	21
3.2.3	Principal Curves	22
3.3	Building an alternative model	22
3.4	Structure optimization	23
3.4.1	Fitness calculation	23
3.4.2	Genetic Programming	24

3.5	Parameter optimization	24
3.5.1	GA and local optimization	25
3.5.2	Linear-in-parameter models	26
3.6	Reconstruction of model dynamics	27
3.7	Model examples	28
3.7.1	The mathematical pendulum	28
3.7.2	Simple predator-prey model	29
3.8	Results	30
3.8.1	Mathematical pendulum	30
3.8.2	Predator-prey model	34
3.9	Discussion	37
3.10	Conclusion	41
4	Oscillation of structural properties in food web models	43
4.1	Introduction	43
4.2	Models used	45
4.2.1	Predator-prey model	45
4.2.2	NPZ model	46
4.3	Model reduction	46
4.3.1	State transformation and model learning	47
4.3.2	Reduction of the model systems	48
4.4	Results and Discussion	49
4.4.1	Predator-prey model	49
4.4.2	NPZ model	52
4.4.3	Ecological Interaction Models	54
4.4.4	Generality of the reduced models	55
4.5	Conclusion	57
5	Explicit formulation of indirect interactions. Part I: Reducing model complexity	61
5.1	Introduction	61
5.2	Model system and data	64
5.3	Data preparation	65
5.4	Linear dimensionality reduction	67
5.5	Nonlinear dimensionality reduction	67
5.6	Model learning and reconstruction	68
5.7	Model results and discussion	69
5.7.1	Results for the transient-free time series	69

5.7.2	Results for the complete time series	70
5.7.3	Performance of the reduction procedure	71
5.8	Conclusion	74
6	Explicit formulation of indirect interactions. Part II: Biological implications	77
6.1	Introduction	77
6.2	Method overview	79
6.3	Limit cycle dynamics	80
6.3.1	Effective consumer variables	80
6.3.2	Ecological Interaction Models	85
6.4	Modification by transients	87
6.4.1	Effective consumer variables	87
6.4.2	Ecological Interaction Models for the complete dynamics	89
6.5	Conclusion	90
7	Conclusion and outlook	93
8	Appendix	95
A	Appendix of Chapter 4	95
A.1	Tables	95
A.2	Normalization of the Rosenzweig-McArthur model	95
A.3	Derivation of the pendulum equation for (3.15)	97
B	Appendix of Chapter 5	98
B.1	Parameter settings of the competition model	98
B.2	Tables	98
B.3	Quality measures	99
C	Appendix of Chapter 6	100
C.1	Tables	100
	Bibliography	103
	Acknowledgements	115
	Curriculum Vitae	117

1.1 The simplicity/complexity dilemma of process-based models

The computational reproduction and prediction of environmental system behavior can easily be seen as one of today's greatest challenges in process-based modeling. The combination of biological, chemical and physical processes on different space and time scales as well as the manifold, often nonlinear, interactions between different biological units result in the overall complexity of environmental systems. Given this intertwined structure of natural phenomena and our inability to apprehend all of the processes involved, modeling accordingly is the attempt to abstract essential features and reduce the complexity of the real world. In the past, a large emphasis of ecological modeling has been put on deterministic descriptions in form of ordinary differential equations (ODE) because of their process-based interpretability. Attempts to incorporate all environmental processes of perceived importance in ODE descriptions naturally lead to an immense increase of model complexity. This tendency to "complexification" is continually fuelled by the ongoing gain in more detailed ecological understanding (see e.g. Anderson, 2005). Examples for complex environmental model systems are the European Regional Seas Ecosystem Model (ERSEM, Baretta et al., 1995) and the Gypsy Moth Life System Model (GMLSM, Sharov and Colbert, 1994) with tens or hundreds of state variables and model parameters.

As a downside, handling models of high complexity is fraught with many difficulties. The large number of parameters typically leads to overparameterization and underdetermination of the systems. Anderson (2005) discussed the impact of parameter uncertainties as well as poorly understood ecological processes of plankton functional types on marine biogeochemical modeling. He argued that these problems among others impede the incorporation of extra complexity beyond simple nutrient-phytoplankton-zooplankton-detritus (NPZD) models. In fact, relative to the amount of parameters, the available data for model verification are too sparse to reliably estimate the parameter values in these cases. Instead, agreement between model output and observations can be obtained with different sets of parameters. This, however, reduces the explanatory power of complex models as it allows for the proposition of multiple (often incompatible) theories about the data-generating processes (Matear, 1995; Friedrichs et al., 2006). Furthermore, parameter calibration is not the only aspect of complex systems which is subject to uncertainty. The notion of non-uniqueness has also been discussed for

the identification of model structures. This type of identification problem is known as the equifinality thesis in the context of hydrological modeling (see e.g. Beven, 2006). It states that equally acceptable fits to observational data are given by multiple complex models which can, thus, not be rejected easily. Support for this thesis can, for example, be derived from the results of a recent comparative model study assessing the usefulness of different ecological models to simulate measured biogeochemical data of the Joint Global Ocean Flux Study (JGOFS) (Friedrichs et al., 2006). All investigated models of this study showed similar low data fitting capabilities which were, most notably, independent of the respective model complexity.

Acknowledging the aforementioned problems, it is implied that huge models are expected to serve too many purposes simultaneously (Lee, 1973) and that approaches covering less process details and dynamic situations may be favorable. It is obvious that models of reduced complexity circumvent some of the problems inherent to complex descriptions. Simplified models with a small number of state variables and parameters are less prone to parameter uncertainty, overfitting and equifinality (Beven, 2006). In addition, they give a reduced, and hopefully more general, view of observed natural patterns and are thus more likely to provide insights into the underlying system's most important governing processes. The search for and preference of simple scientific explanations is in fact one possible verbalization of a famous scientific paradigm, known as the parsimony principle or Occam's razor. Though this epistemological rule of "not increasing a theory's complexity beyond need" is often attributed to the famous 14th century logician William of Ockham, its origin can be traced back to other sources in the Middle Ages and to ideas of Aristotle (Rodríguez-Fernández, 1999). However, simplicity and parsimony of models or theories do not necessarily describe the same concepts. Parsimony should rather be seen as *problem-adapted* simplicity and, as pointed out by Simon (2002), striving for simplicity in its one right eventually leads to theories which are too simple to provide useful system information. In fact, as pointed out by Logan (1994), complex systems need complex solutions. Thus, simplicity in the parsimonious sense could also be described as appropriate or problem-adapted complexity. Instead of trying to capture as many process details as possible on the one hand and oversimplification on the other hand, the model complexity should be adjusted according to the problem at hand, the available data and the purpose of interest (e.g. Sivakumar, 2007). Implications of this "balancing problem" have been investigated by a number of comparative model studies (Costanza and Sklar, 1985; Håkanson, 1995; Fulton et al., 2003) with similar results. In (Costanza and Sklar, 1985) a comparison of different ecosystem models with respect to model performance and complexity was performed. It was found that the effectiveness or explanatory power was highest for models with intermediate complexity. Thus, in this thesis, the term "model simplicity" is always used in the parsimonious sense which automatically implies a compromise between descriptive or predictive accuracy and the incorporation of a minimal set of dominant processes.

1.2 The MAGER approach and previous work

The obvious question arising from the introductory last section is: how can process-based models with the most appropriate level of complexity be found? In the present thesis, a new methodology of model building and transformation is developed in order to give a possible answer to this question. Simplified system descriptions can thereby either be derived in a bottom-up fashion in the course of model building or by top-down reduction of existing complex models. Because of its usability as a means to reduce process-based models to simpler structures, the new method is called *Mapping-based Genetic Reduction* (MAGER). Although it may also be applied for model building, only the model reduction (MR) potential of MAGER is presented in the framework of this thesis to simplify its introduction. The reduction of known model systems thereby facilitates the evaluation of the method's performance as it allows for an interpretation of the results in relation to the original system descriptions.

Four central properties of MAGER are responsible for its uniqueness in comparison with other MR schemes. These properties are (1) its independence on existing system knowledge or dynamic details, (2) the automated operation, (3) the high reduction potential and (4) the interpretability of the results in process-based terms. Some of these properties can also be found in earlier approaches but the main strength of the new method is their combination (see chapters 3 and 5 for detailed reviews on earlier works). Some well-known reduction approaches like model aggregation based on singular perturbation rely on the existence of specific dynamic properties of the systems (e.g. Auger and Poggiale, 1996b; Nayfeh, 1973). In the absence of these properties, e.g. long-term fixed point dynamics, these methods may produce reduced models with dynamics deviating from the original ones (Schaffer, 1981). Some black-box model identification approaches are only applicable for linear systems or need a priori specifications of the type of model nonlinearities (e.g. Young and Garnier, 2006; Phillips, 2000). Other approaches depend on detailed system knowledge and analysis and cannot be automated easily (e.g. Wirtz and Eckhardt, 1996; Van Nes and Scheffer, 2005; Raick et al., 2006). Finally, with strictly statistical or data-mining approaches, like neural networks or Krylov subspace methods, we end up with reduced system descriptions which are not interpretable in process-based terms (Carreira-Perpiñán, 1997; Fodor, 2002; Antoulas et al., 2001).

MAGER aims at combining the benefits of the aforementioned approaches and at eliminating some of their shortcomings. It is a black-box aggregation procedure which can be applied to measured or modeled time series data. The proposed MR scheme is composed of a mapping-based state variable reduction and a subsequent model learning step generating ordinary differential equation (ODE) models with simple structures. For all applications discussed in the next chapters, the new dynamical equations are able to reproduce the original time series to a large degree and can be interpreted in terms of the transformed state variables. This transformation aspect is also seen as the basis for the high reduction potential of the MAGER scheme. It is well known in scientific philosophy that the

measuring of simplicity depends on the chosen language (Sober, 2002). This effect can be formulated by quoting Goodman's "new riddle of induction" (Goodman, 1983). Consider a simple and a more complex proposition: (1) all emeralds are green and (2) all emeralds are green up to time t (which has yet to pass) and blue afterwards. We now define "grue" as green until t and blue afterwards and, accordingly, "bleen" as blue until t and green afterwards. Thus, the two statements can be rephrased as (1.1) all emeralds are grue until t and bleen after that and (2.1) all emeralds are grue. We see that the simpler proposition translates to a more complex one whereas the complex second statement is simplified, which is caused by the transformation of description. The same concept applies for MAGER: the simplification of models is supported by the adoption of a new "language" which is based on the transformed state variables. The benefits of a detachment from well-known description concepts have also been discussed in the context of latent variable models. For example, Malaeb et al. (2000) note that "thinking only in terms of directly observable variables confines our horizons and limits our assessment of complex systems". However, the reduction process of MAGER not only produces simple models but also accounts for parsimony as the tradeoff between simplicity and accuracy is an integral part of the model learning process. Because of the method being data-driven, the resulting models are most likely to provide insights into the dominant processes which are only derived from the information inherent to the available data.

1.3 Structure of the thesis

Chapters 2 to 6 consist of original publications and submitted manuscripts documenting the development process of MAGER from first ideas to biological applications. The development and implementation of the algorithms as well the analysis of the results have been carried out by the author of this thesis alone. The discussion of the results has been developed in collaboration with the co-author of chapters 4 to 6 who also provided the alternative EVA reduction in chapter 5 as well as fruitful additions to the individual introductory sections of the manuscripts. The analytical derivation of the oscillator equation in chapter 4 was also developed by both authors.

Chapter 2, which was published in the proceedings of the 2nd meeting of the International Environmental Modelling and Software Society (iEMSs), contains a preliminary application of the mapping-based dimensionality reduction step of MAGER. Starting with a ten-dimensional dataset obtained by parameter variation of a consumer-resource model, a nonlinear mapping algorithm performing a nonlinear principal component analysis (NLPCA) is used to reduce the dimensionality of the data. The results indicate that the first two nonlinear principal components capture the dominant dynamic aspects of the system, such as the separation of oscillatory states, and can eventually be taken as new "effective" state variables of reduced-form mechanistic models. In addition, this chapter introduces the Self-Organizing Map as a means to reduce the amount of data in order to speed up the nonlinear

mapping procedure. As the later studies are based on smaller datasets, this data-mining step is not used subsequently.

Chapter 3 gives a full account and technical details of the completed MAGER approach. The manuscript has been accepted by *Evolutionary Computation* and will be published in volume 16, issue 1 (Spring 2008) of this journal. Apart from the dimensionality reduction it also covers the Genetic Programming (GP) and parameter optimization approaches used for model learning. Extending the first results given in chapter 2, a comparison of linear and nonlinear mappings for state reduction is performed. The results show that nonlinear approaches must be used carefully for this task as they are likely to produce temporal mapping errors. The chapter also gives first indications that the well-known root-mean-square error (RMSE) is inappropriate for the calculation of data fitting capabilities in case of transient oscillatory dynamics. In addition to the RMSE, a simple criterion to discover oscillating systems approaching limit cycle dynamics is introduced. Finally, the MAGER scheme is used to transform a linear oscillator model (pendulum) and a two-dimensional predator-prey system. The results show the fundamental analogy between the two oscillatory systems which supports the notion of biological oscillators (Vandermeer, 2004; Gertsev et al., 2008).

The manuscript in chapter 4 has been submitted to *Theoretical Biology*. It extends the results of chapter 3. The oscillator transformation of the two-dimensional predator-prey model is compared with an approximate analytical derivation which offers an interpretation of the dynamic equations in terms of biological forces as introduced by Ginzburg (1986). This way, the interaction of biological populations is found to produce driving and damping forces in analogy to those observed in physical systems. In addition, the MAGER scheme is used to reduce a food-chain model with three trophic levels consisting of nutrients, phytoplankton and zooplankton (NPZ). The resulting simplified model again resembles a nonlinear oscillator. Furthermore, the model's two new state variables are interpreted as feeding limitations governing the phytoplankton/nutrient interactions on the one hand and the zooplankton/phytoplankton interactions on the other hand. Direct mutual inter-relations of these interaction variables thus capture indirect interactions between the original state variables which are, however, not formulated explicitly in the NPZ model.

Chapters 5 and 6 document the MAGER reduction results for an eight-dimensional consumer-resource (CR) competition model producing transient and limit cycle dynamics. The study has been split up into two parts discussing the reduction performance of MAGER for this model setup (chapter 5) and the corresponding biological implications (chapter 6). This two-part study will be submitted to *Theoretical Population Biology*. The CR model is reduced to two dimensions using a combination of the linear and nonlinear mapping approaches. This way, the temporal mapping errors found for the application of NLPCA in chapter 3 can be avoided. Chapter 5 further describes the reduction of the given CR model with two other reduction schemes based on the aggregation and omission of state variables. The study shows that MAGER has a much higher reduction performance compared to

these methods while its simplified models are comparable or better in terms of data reproduction. The data-adaptivity of MAGER and the transformation to interaction variables are proposed to be of major impact for this performance gain. The analysis in chapter 6 shows that the transformed state variables can be interpreted biologically. It turns out that these "effective consumer variables" describe shifts in community composition of the dominant interacting species. These shifts are thereby related to tradeoffs in resource requirements. Thus, the results of chapters 4 and 6 unify the predator-prey and consumer-resource model formulations in terms of processes leading to oscillatory dynamics. In all investigated cases, we find meta-level descriptions of biological networks featuring explicit formulations of direct or indirect biological interactions. From the perspective of the simulated datasets, the observed dynamics are independent from biological details mediating these interactions, such as shared nutrients or grazing formulations. The new class of *Ecological Interaction Models* (EIM) emerging from the application of the MAGER scheme substantiates the unifying notion of process-independent interactions of biological entities.

Finally, chapter 7 outlines limitations and possible extensions of MAGER. It closes with a discussion of future work and further applications of the new MR scheme.

Chapter 2

Reduction of complex models using data-mining and nonlinear projection techniques

Abstract

Complex models of environmental systems typically depend on a large amount of uncertain parameters. Therefore, they are often difficult to handle and do not provide an insight into effective modes of the underlying system's dynamics. Unlike earlier analytical attempts to find more effective model representations, we present a new combination of methods that only relies on data generated by complex, process-based models. These methods are taken from the field of data-mining and enable the recognition of patterns in measured or modeled data by unsupervised learning strategies. As these methods do not directly lead to a better understanding of the systems' driving processes, we suggest the linkage between pattern recognition and process identification by a multi-stage approach. In a first step, a large data-base was produced by a mechanistic model for species competition in a virtual ecosystem for a range of parameter settings. Using Vector Quantization and nonlinear projection techniques such as Self-Organizing Maps and nonlinear Principal Component Analysis, typical states of the complex model's dynamics as well as major pathways connecting these states were then identified. The visualization of the results points to the existence of nonlinear transformations of former model state variables and parameters to few effective variables. Effective variables built this way preserve most of the model's dynamic behavior, while they are nonetheless easier to use and require much less parameterization effort.

2.1 Introduction

Process-based models are widely used for modeling key mechanisms ruling ecosystem dynamics. The vast number of potentially relevant interactions and adaptations in ecosystems thereby increases the corresponding model complexity. Secondly, process identification is rarely unique, i.e. data can be reproduced with a variety of models or parameterizations (see e.g. Beven, 2001). Given this high complexity and the sparseness of data, parameter uncertainty is difficult to handle in these models.

An alternative way to reproduce and extrapolate data is the use of methods taken from the field of data-mining, such as Neural Networks (NN), clustering methods or (non-)linear projection techniques which are able to 'learn' distinct features of a dataset (Fodor, 2002). No knowledge of the underlying system is required using data-driven methods. New understanding of underlying key mechanisms

however, can not be gained and generic models with a large domain of applicability are not provided. The aim of this work is the construction of a new type of deterministic but reduced and efficient model from a classic complex mechanistic model by only using information contained in the data generated by the complex model. A nonlinear statistical analysis and reduction of this data should reflect the overall dynamics even for uncertain model parameterizations and should yield interpretable information on dominant internal modes. We propose a multi-step analysis using data-mining and nonlinear projection techniques to extract these modes or “effective variables” (Wirtz and Eckhardt, 1996). Those variables have been shown to successfully replace complex descriptions of adaptive processes in biological systems (e.g. Wirtz, 2002). Up to now they had to be built using intuitive modeling knowledge which is a major impediment for a broader use.

The recombination of the effective variables resulting in a reduced-form deterministic model consequently combines the benefits of the process-oriented as well as data-mining approaches. The existence of such a reduced representation is supported by the finding that even huge ecosystem models have a limited number of internal dynamic modes (Ebenhöh, 1996).

In principal, the proposed reduction scheme can be applied to any deterministic process-based model. In this study, we present the extraction of effective variables using a combination of Vector Quantization algorithms such as the Self-Organizing Map (Kohonen, 1997) and nonlinear Principal Component Analysis (Kramer, 1991).

We have chosen the reduction of a prominent model of species competition with rich dynamics including chaotic behavior (Huisman and Weissing, 1999) as a test case.

2.2 A model of species competition

The model analyzed in this study was proposed by Huisman and Weissing (1999). It describes competition for resources like phytoplankton species competing for nitrogen and phosphorus.

Consider n_P phytoplankton species and n_N nutrients. Let state variables \hat{P}_i and \hat{N}_j be the population abundance of species i and the availability of resource j respectively. The dynamics of species i follows

$$\frac{d\hat{P}_i}{dt} = \hat{P}_i \cdot (\mu_i - \omega_i) \quad i = 1, \dots, n_P \quad (2.1)$$

where ω_i are model parameters describing the mortality. The growth rate μ_i is controlled by the most limiting resource via a minimum of Monod functions with K_{ji} denoting the half-saturation constant for resource j of species i and g_i the maximal growth rate:

$$\mu_i = \min_v \left(\frac{g_i \hat{N}_v}{K_{vi} + \hat{N}_v} \right) \quad v = 1, \dots, n_N. \quad (2.2)$$

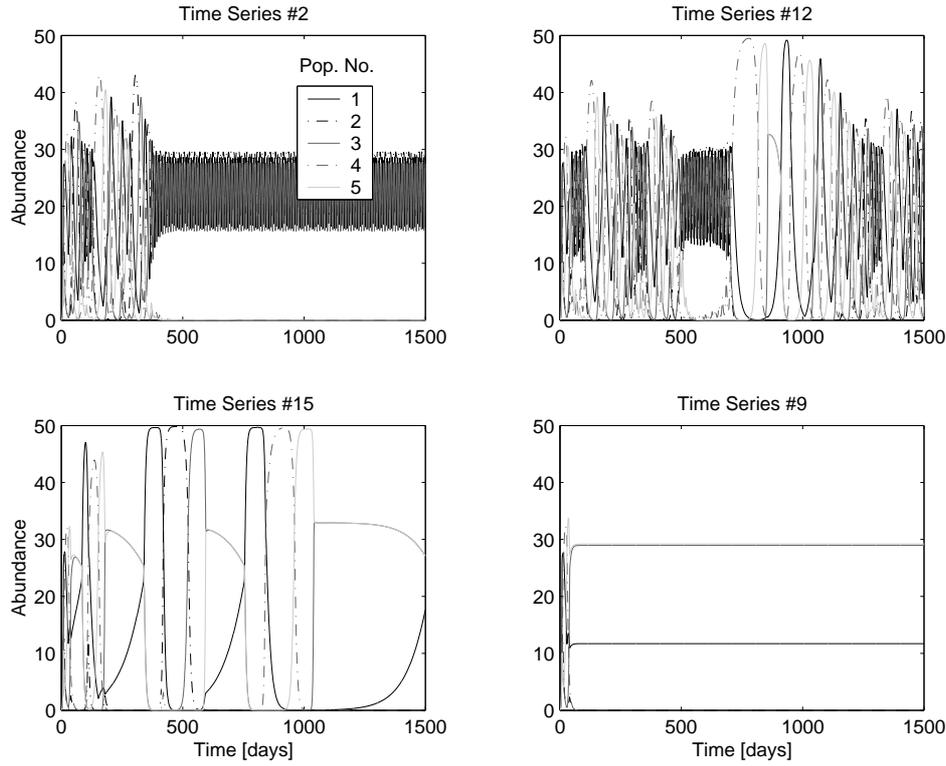


Figure 2.1: Time series of $\hat{P}_i (i = 1, \dots, 5)$ created by parameter variation of the test model. Parameter settings $\{k_{21}, k_{25}\}$ are $\{0.2, 0.325\}$ for time series #2, $\{0.275, 0.4\}$ for time series #12, $\{0.325, 0.35\}$ for time series #15 and $\{0.275, 0.275\}$ for time series #9.

The time evolution of the abiotic resource j is described as

$$\frac{d\hat{N}_j}{dt} = D \cdot (S_j - \hat{N}_j) - \sum_i c_{ij} \cdot \mu_i \cdot \hat{P}_i \quad j = 1, \dots, n_N \quad (2.3)$$

where D is a constant factor describing the nutrient turnover rate, S_j is the supply concentration and parameters c_{ij} quantify the content of nutrient j in species i .

For different choices of model parameters the system can be driven into attractors with different topologies containing fixed-point dynamics (no changes in species abundances for one or more species), limit cycles (fluctuating coexistence of species) or chaotic behavior. For further details on the parameter settings see (Huisman and Weissing, 1999, 2001b).

To keep the analysis simple, we numerically integrated (2.1) and (2.3) to produce 16 time series of 2000 points each for a model configuration with five species ($n_P = 5$) and three abiotic resources ($n_N = 3$) by varying only two of the half-saturation constants (k_{21} and k_{25}). The other model parameters (D , S_j , ω_i , g_i and c_{ij}) were kept at the fixed values used by Huisman and Weissing (2001b).

Time series modeled this way show all sorts of dynamics described above (see Figure 2.1).

2.3 The Self-Organizing Map

The Self-Organizing Map (SOM) algorithm was introduced by Kohonen (1997). It resembles a neural network variant consisting of topologically ordered nodes on a grid of predefined dimensionality.

A SOM is able to “learn” structures of high-dimensional input data vectors and to project them onto a lower-dimensional output space. It is therefore often used for Vector Quantization (VQ) where a reduced representation of complex datasets is built by replacing the data vectors with a smaller subset of so-called prototype vectors. Additionally, the existence of the typically two-dimensional output grid simplifies the visual inspection of the dataset and helps to identify patterns inherent to the data.

The algorithm transforms a dataset consisting of vectors $\mathbf{x}(t) = (x_1(t), x_2(t), \dots, x_n(t))^T \in \mathbb{R}^n$ with discrete-time coordinate $t = 0, 1, 2, \dots$, e.g. measurements of n variables over time. In this case, each $\mathbf{x}(t)$ is a ten-dimensional vector with entries $\hat{P}_i, \hat{N}_j, k_{21}$ and k_{25} ($n = n_P + n_N + 2$).

The SOM-network consists of a z -dimensional array of k nodes associated with prototype-vectors $\mathbf{m}_k \in \mathbb{R}^n$ with orthogonal or hexagonal neighborhood relationships between adjacent nodes.

The data vectors are iteratively compared with all \mathbf{m}_k by using Euclidean distances to find the best-matching node denoted by c . The updating procedure for prototype s then follows

$$\mathbf{m}_s(t+1) = \mathbf{m}_s(t) + h_{cs} \cdot [\mathbf{x}(t) - \mathbf{m}_s(t)], \quad (2.4)$$

where h_{cs} is a neighborhood function that asserts the convergence of the algorithm for $h_{cs} \rightarrow 0$ when $t \rightarrow \infty$. Mostly, the Gaussian function in dependence of $\|r_c - r_s\|$ is used, where $r_c \in \mathbb{R}^z$ and $r_s \in \mathbb{R}^z$ are the location vectors of nodes c and s . Additionally, h_{cs} is multiplied by the learning-rate factor $\alpha(t) \in [0, 1]$ that decreases monotonously over time to prevent the distortion of already ordered parts of the map at later time steps.

In measuring the quality of the SOM-mapping (see e.g. Bauer and Pawelzik, 1992; Villmann et al., 1997) a compromise has to be made between an optimized reproduction of the data vectors and the minimization of the topological distortion by neighborhood violations. In this work the SOM-Toolbox 2.0 package (Vesanto et al., 1999) was used that calculates the average quantization error and the topographic error (Kiviluoto, 1996). The best network of different map configurations was assumed to minimize the sum of these two measures. This procedure tends to find solutions overfitting the dataset but this drawback was accepted as the details of the VQ step were found to be of minor importance for the following analysis.

2.3.1 Vector quantization of the dataset

In advance of the analysis the data matrix was standardized by mean and standard deviation of the individual variables ($\hat{P}_i \rightarrow P_i$). To incorporate information about control parameters of the competition model into the learning procedure, the constant time series of k_{21} and k_{25} were added as additional variables to the training dataset.

SOM networks of different map configurations were trained and the best network with 50 x 50 prototype vectors was found to explain 96.2% of the data variance.

2.4 Nonlinear projection

Even though the SOM itself represents a kind of nonlinear projection technique it is not very well suited for the extraction of distinct modes of the underlying dynamics as the vectors spanning the SOM network can not be interpreted in terms of variable model entities. This limitation also exists for other unsupervised learning strategies that construct relevant topological constraints directly from the data (e.g. Martinetz and Schulten, 1991; Baraldi and Alpaydin, 2002). Hence, the need for finding “directions” along which features of the system vary continuously remains. A promising technique to extract these effective variables is nonlinear Principal Component Analysis (NLPCA) put forward by Kramer (1991).

The NLPCA relies on an autoassociative feed-forward neural network as depicted in Figure 2.2. It projects data vectors $\mathbf{x}(t)$ onto a so-called bottleneck layer u and compares the decoded vectors $\mathbf{x}'(t) = (x'_1(t), x'_2(t), \dots, x'_n(t))^T$ with the input data to minimize the cost function $J = \langle \|\mathbf{x}(t) - \mathbf{x}'(t)\|^2 \rangle$.

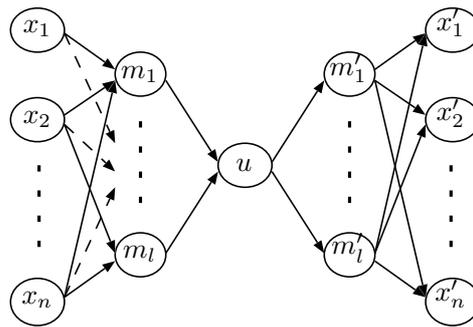


Figure 2.2: Example for an autoassociative neural network with l nodes m_1, \dots, m_l in the first and m'_1, \dots, m'_l in the second hidden layer.

The mappings $\mathbf{x} \rightarrow \mathbf{m}$ and $\mathbf{m}' \rightarrow \mathbf{x}'$ are typically performed by nonlinear transfer functions (e.g. the sigmoidal function), whereas mappings from and to the nonlinear principal component u use the

identity function. The number of nodes in the hidden layers of the network determines the approximation quality of the data.

Typical problems arising during neural networks training are overfitting and local minima in the cost function J . In our analysis we employ the NeuMATSA (Neuralnets for Multivariate and Time Series analysis) package (Hsieh, 2001) where multiple runs and penalty terms for the network weights smooth the nonlinear responses of the transfer functions to obtain results less sensitive to local minima.

Only by the data reduction of the preceding SOM analysis the NLPCA step is made applicable. Thus, an immense speed-up of the minimization of the neural networks' weights is gained and the already smoothed SOM representation additionally accounts for the avoidance of local minima in the cost function.

2.4.1 NLPCA of the SOM-filtered data

To prevent the NLPCA from overfitting, 20% of the SOM-filtered dataset were chosen randomly as test-dataset and ensembles of 25 runs were selected for configurations of nodes in the hidden layers ranging from one to five. The analysis was terminated when the quality of the mapping as quantified by the mean squared error (MSE) for the test set decreased subsequently to an initial rise.

After extraction of the first nonlinear PCA, further components were iteratively found by subtracting earlier solutions from the SOM dataset and by repeating the analysis using the residuals.

The first nonlinear mode found this way explained 61%, whereas the second and third mode accounted for 16.5 and 2.7% of the SOM networks' variance, respectively. Thus, the dataset can be assumed to be essentially two-dimensional and the first two nonlinear modes extracted by NLPCA can be interpreted as effective variables (EV) of the underlying model. Figure 2.3 shows the first two modes in state space $\{P_1, P_2\}$. Clearly, variation of the original model variables is constrained indicating implicit model trade-offs and the existence of a reduced EV representation.

2.4.2 Interpretation of modes

Figure 2.4 shows successive segments of an example time series projected into the EV space. The typical cycles with varying periods found in the dataset (see Figure 2.1) are clearly separated from each other and resemble the form of limit-cycles in the complex model's phase-space. Successive changes between these cycles illustrate the ability of the method to separate dynamic states.

To further investigate and interpret the effective nonlinear modes in terms of former model variables, the projected data was aggregated into bins of equal size with a minimum of five data points per class. Figure 2.5 shows the class distributions of the first mode.

The smooth course of the distributions together with the relatively small inner class variability even for

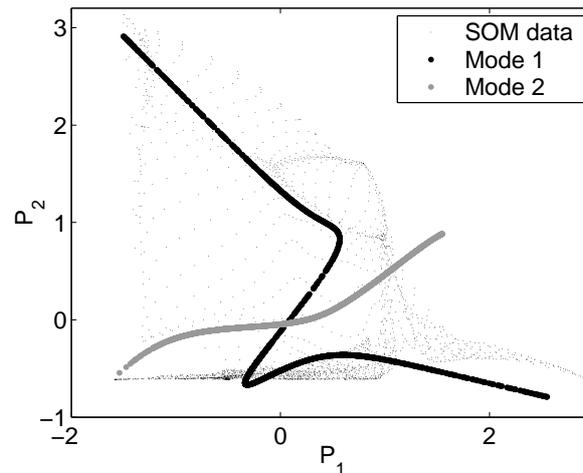


Figure 2.3: Nonlinear PCA modes 1-2 in a VQ state subspace with dots representing SOM-filtered model data.

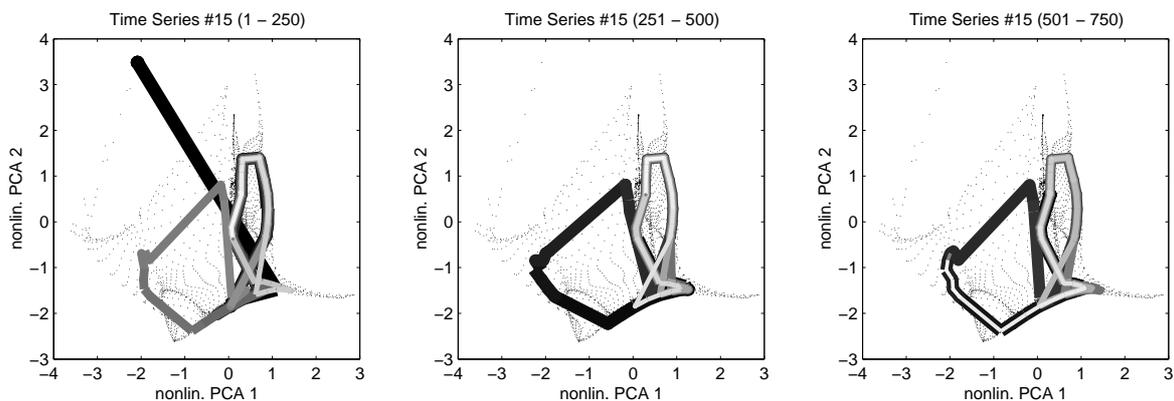


Figure 2.4: Plot of time series examples in EV space (spanned by the first two NLPCA modes). Shown are the first 750 time steps of series #12. Earlier time steps are drawn in thick dark and later ones in thin light lines.

densely covered bins (e.g. small positive values of the first nonlinear PCA) may provide a meaningful interpretation of the nonlinear modes. For example, a comparison of the chaotic transition (from small negative to small positive values of PCA 1 in Figures 2.4 and 2.5) for P_i provides an insight into the particular case when species coexist. Thus, this type of coexistence can be imagined as an occupation of “dynamically separated” niches.

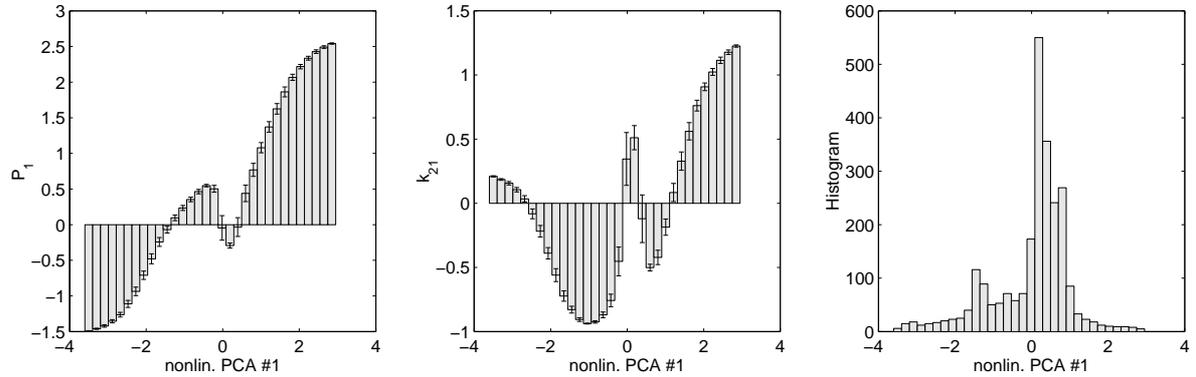


Figure 2.5: Aggregation of the projected data along the first nonlinear mode together with a histogram of bin occupancy.

2.5 Discussion

Improvements of the methodical parts of this work, as discussed in section 2.4 for the SOM algorithm, can be thought of. As outlined by Malthouse (1998), NLPCA solutions for the projection problem are only suboptimal and alternatives like the Principal Curves approach of Hastie and Stuetzle (1989), for example, can be tested as well. In this work however, the projection discrepancy does not constrain the usefulness of NLPCA as smooth solutions following mean features of the dataset are explicitly requested.

First outcomes of this work show that a combination of Vector Quantization and nonlinear projection can already provide valuable insights into the dynamics underlying process-oriented models. The extraction of relevant nonlinear modes describing a model on a higher or aggregated level is a first step towards effective variable models that are easier to use and better to interpret than their complex model equivalents.

The results shown here point to the existence of non-linear but nonetheless simple transformations of former model state variables and parameters to effective variables. The projections of quantized model data along the first two nonlinear modes, from which only a subset is presented in Figure 2.5, already support a piecewise linear transformation from the original space of model entities to new aggregated variables. In future studies, reduced-form models will be formulated using effective variables as provided by the approach put forward in this study. We will thereby rely on results and techniques presented here comprising (i) the simultaneous incorporation of model outcomes and varied model coefficients into the analysis, (ii) internal trade-offs between model variables for different attractors of the model dynamics and (iii) the smoothness of the projections of a small set of effective variables to the original model space. The extraction of these nonlinear transformations constitutes an analytical

means to interpret the nonlinear principal components in terms of simulated processes as an essential step towards a reduced-form representation of complex, mechanistic models.

Chapter 3

Finding alternatives and reduced formulations for process-based models

Abstract

This paper addresses the problem of model complexity commonly arising in constructing and using process-based models with intricate interactions. Apart from complex process details the dynamical behavior of such systems is often limited to a discrete number of typical states. Thus, models reproducing the system's processes in all details are often too complex and over-parameterized. In order to reduce simulation times and to get a better impression of the important mechanisms, simplified formulations are desirable.

In this work a data-adaptive model reduction scheme automatically building simple models from complex ones is proposed. The method can be applied to the transformation and reduction of systems of ordinary differential equations. It consists of a multi-step approach using a low-dimensional projection of the model data followed by a Genetic Programming/Genetic Algorithm hybrid to evolve new model systems. As the resulting models again consist of differential equations, their process-based interpretation in terms of new state variables becomes possible.

Transformations of two simple models with oscillatory dynamics, simulating a mathematical pendulum and predator-prey interactions respectively, serve as introductory examples of the method's application. The resulting equations of forces indicate the predator-prey system's equivalence to a nonlinear oscillator. In contrast to the simple pendulum it contains driving and damping forces that produce a stable limit cycle.

3.1 Introduction

Mathematical models are important tools for reproducing, understanding and predicting the behavior of physical and environmental systems. However, by incorporating many processes to reproduce complex interactions and dynamics, their usefulness is limited by the growing amount of state variables and adjustable parameters. Datasets from biology or biogeochemistry, for instance, are often too sparse for a satisfactory verification of complex models. Furthermore, many model parameters are not measurable directly by existing monitoring techniques. Models often are over-parameterized and may accurately fit the data with different parameter sets, thereby masking relevant key processes of the underlying system (Beven, 2001). These aspects of mathematical models are part of the concept

of model complexity. A major guideline in model building, based on Occam's razor, is the preference of a simple model if it is equally well suited to reproduce the data.

The principal aim of this study is the development of a complexity reduction technique for process-based models. In this context, model complexity is understood as a combination of the number of state variables and parameters on the one hand and nonlinear interactions of the states on the other hand. A model's complexity may be reduced in the course of model building by including essential processes and states only. This, however, can be a very difficult task as model formulation must be done on a high level of abstraction and a thorough knowledge of the system's relevant key processes is essential. Alternatively, models may be derived from existing complex ones by simplification (see Van Nes and Scheffer, 2005).

A short overview of current model simplification approaches is given before the method developed in this study is described. The models concerned here are systems of ordinary differential equations (ODE) that are widely used to represent dynamical systems in many fields of science.

Two methodically different groups of model reduction methods will be discussed. Approaches of the first group may be termed heuristic as their application typically requires a certain amount of user interaction. They include a detailed analysis and assessment of the interactions and feedbacks between model components in order to identify key processes of the system. The reduction implies changes of the model structure and must typically be done by domain experts. An overview of different approaches was given by Van Nes and Scheffer (2005).

One possible way of finding key parameters of a system, for instance, is to measure the sensitivity of the model's results towards a variation of its parameters. Brooks et al. (2001), for example, used a sensitivity and structure analysis to simplify a model for wheat yield prediction. The simplification of model parts may thereby involve the deletion or replacement of process descriptions and state variables, e.g. with simplified mathematical terms or constants (Cox et al., 2006). These methods primarily aim at increasing the scientific knowledge about the underlying system.

Model reduction methods of the second group are solely based on mathematical concepts. They are applied to shorten the simulation times of high-dimensional ODE systems by reducing their order. The most prominent approaches are projection-based (de Villemagne and Skelton, 1987). The systems are projected onto a lower-dimensional subspace and the model equations are solved for the substituted projected states. Methods of this group may be categorized according to the (non-)linearity of the model equations and the basis vectors spanning the low-dimensional subspace. The theory is best established for linear systems, see Antoulas et al. (2001) for an overview. Basis selection methods for these systems include Krylov subspace methods (Grimme, 1997) and techniques using Singular Value Decomposition, also known as Principal Component Analysis (PCA) or Proper Orthogonal Decomposition (Lall et al., 2003; Rowley et al., 2004), among other techniques.

Nonlinear model reduction in this context is rather challenging. There are still only a small number of

established methods, most of which are extensions of the aforementioned linear techniques. The most simple ones are based on linearization or reduced-order series expansion of the system's nonlinearities using Taylor or Volterra series (see Phillips, 2000). These low-order approximations, however, are only possible for weakly nonlinear systems. Other approaches use multiple linearizations about selected states along a trajectory (Rewieński, 2003) or nonlinear extensions of the Galerkin method (Matthies and Meyer, 2003).

The motivation for the development of a new model reduction technique in this study is based on the benefits and drawbacks of the aforementioned methods. The simplified models of the first group are interpretable in terms of the original state variables but their derivation can hardly be done automatically as it involves a detailed analysis and knowledge of the system's processes. Automatic model reduction is performed by members of the second group. The original interpretation of the states, however, is typically lost in the course of the reduction process.

The proposed approach is intended to combine the benefits of both groups: it generically produces alternative model formulations which are open to process-based interpretation. In the course of the method, new state variables and model structures are produced. The new model states thereby are transformed versions of the former ones and low-dimensional ODE systems are constructed to describe the dynamics of these states. The new systems can finally be interpreted in terms of the new state variables.

Using model simulation data, the approach starts similar to some projection-based reduction schemes with a transformation of the data to a lower dimensional structure. The projection may be linear or nonlinear depending on the problem at hand. Some examples of projection methods are discussed in section 3.2. A Genetic Programming (GP) scheme in combination with a Genetic Algorithm (GA) is used to build new ODE systems based on the projected data. The reduction potential of the proposed scheme is given by the dimensionality of the system's projected state space. The projection preserves the model dynamics up to a certain error and varies from system to system.

The following sections give an outline of the proposed method. Two simple model examples, consisting of two-dimensional first order differential equations, demonstrate its application. The examples show two different basic realizations of oscillatory dynamics and are transformed into one-dimensional ODEs of second order. This conversion of the ODEs in equations of forces is, in fact, no real model reduction as the dimensionality of the systems is not changed. The resulting formulations, however, make a comparative interpretation of both systems possible and expose the central terms responsible for the realization of the dynamics.

3.2 State reduction

The model reduction approach can be formally stated as follows. Starting from the ODE system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t), \quad \mathbf{x} \in \mathbb{R}^m, \quad (3.1)$$

with time t and state vector \mathbf{x} , the reduction scheme approximates (3.1) with a model of dimension $l < m$:

$$\dot{\boldsymbol{\xi}} = \mathbf{g}(\boldsymbol{\xi}, t), \quad \boldsymbol{\xi} \in \mathbb{R}^l. \quad (3.2)$$

Note, that in general $\mathbf{g}(\cdot) \neq \mathbf{f}(\cdot)$.

As the method should be applicable with only a minor knowledge of the model itself, it is based solely on model-generated data. Therefore, similar to the method of snapshots (Sirovich, 1987), a data matrix $\mathbf{X} = [\mathbf{x}(t_0), \dots, \mathbf{x}(t_{n-1})]^T \in \mathbb{R}^{n \times m}$ was built from a simulation of (3.1) using the state information $\mathbf{x}(t_i)$ of n time steps $t_i, i = 0, \dots, n-1$. The state reduction was performed by a mapping $\mathbf{X} \rightarrow \boldsymbol{\Xi}$ onto a low-dimensional representation $\boldsymbol{\Xi} = [\boldsymbol{\xi}(t_0), \dots, \boldsymbol{\xi}(t_{n-1})]^T \in \mathbb{R}^{n \times l}$ of \mathbf{X} .

Many variants of dimension reduction techniques have been proposed earlier in different contexts, e.g. Principal Component Analysis (Fukunaga and Koontz, 1970) with nonlinear variants (Nonlinear PCA (Kramer, 1991), Principal Curves (PCurve; Hastie and Stuetzle, 1989), and Kernel PCA (Schölkopf et al., 1998)), Self-organizing Maps (Kohonen, 1997) or Isomap (Tenenbaum et al., 2000) among others. Conceptually, these methods provide a topology-preserving mapping, i.e. local neighborhood relationships between data points are maintained by the mapping. A short introduction to the applied methods, PCA and PCurve, is given in sections 3.2.2 and 3.2.3.

3.2.1 Temporal neighborhood

A major drawback of the methods mentioned above is the neglect of a dataset's temporal information, as they are based on the data distribution alone. The mapping learned by a dimension-reduction technique therefore may contain jumps when points close to each other in time are mapped to different regions of the reduced structure.

Taking this into account, a measure is proposed which quantifies the temporal continuity of the learned mapping. This mean smoothness error (*MSmE*) follows the concept of false nearest neighbors in the context of attractor reconstruction from time series via embedding (Kennel et al., 1992). The *MSmE* penalizes differences between the temporal behavior of the original and mapped datasets using the normalized distances $\tilde{x}(t_i)$ and $\tilde{\xi}(t_i)$ between consecutive elements of the time series. The measure is

defined as

$$MSmE = \frac{1}{n-1} \sum_i \left| \log_2 \left(\frac{1 + \tilde{x}(t_i)}{1 + \tilde{\xi}(t_i)} \right) \right|, \quad (3.3)$$

$$\tilde{x}(t_i) = \frac{\|\mathbf{x}(t_{i+1}) - \mathbf{x}(t_i)\|}{\|\mathbf{x}_{max} - \mathbf{x}_{min}\|},$$

for $i = 0, \dots, n-2$. Here, \mathbf{x}_{max} and \mathbf{x}_{min} denote the maximum and minimum of the state variable vector for all time steps, $\tilde{\xi}(t_i)$ is defined in the same manner as $\tilde{x}(t_i)$.

According to this definition, the worst mapping is realized when all pairs of consecutive points are separated maximally in one space while being equal in the other. The expression in brackets then has a value of 2 or 1/2, depending on where the maximal separation occurred, and $MSmE = 1$. An optimal mapping with a value of zero for $MSmE$ is reached when the relative changes between subsequent points in the original and mapped space are the same for all time steps. The performance of the measure is demonstrated for the nonlinear mapping of model system (3.15) in section 3.8.2.

3.2.2 Principal Component Analysis

The most commonly used method of dimension reduction is PCA. It is applied to find orthogonal linear combinations of the variables which are optimal linear approximations of the dataset in the least squares' sense. These so-called principal components are extracted from \mathbf{X} using an orthogonal decomposition of its covariance matrix $\Sigma = \frac{1}{n} \mathbf{X}^T \mathbf{X} \in \mathbb{R}^{m \times m}$ according to

$$\Sigma = \mathbf{U} \Lambda \mathbf{U}^T, \quad (3.4)$$

where $\mathbf{U} \in \mathbb{R}^{m \times m}$ contains the orthogonal eigenvectors as column vectors and Λ is a diagonal matrix of the ordered eigenvalues $\lambda_1 \leq \dots \leq \lambda_m$, i.e. $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_m)$. The principal components $\{s_i\}_{i=1}^m$ can then be found by mapping the data onto the eigenvectors of the covariance matrix according to

$$\mathbf{S} = \mathbf{X} \mathbf{U}^T, \quad (3.5)$$

with $\mathbf{S} = [s_1, \dots, s_m] \in \mathbb{R}^{n \times m}$.

A dimension reduction of \mathbf{X} can be achieved if only a subset of the s_i is used for further analysis. The first principal component with the highest eigenvalue thereby is the linear regression line that explains the main part of the data variance. For dimension reduction, the first k principal components can be retained to account for a certain percentage of the total variance (e.g. 95%).

As the mean values of the variables contain no information about a specific data vector they were removed from the datasets beforehand.

3.2.3 Principal Curves

The first few linear principal components usually fail to explain the main part of a dataset's variability if it is caused by nonlinear dependencies between the system's variables. In this case, nonlinear extensions of PCA may provide better dimension reduction capabilities. The Polygonal Line Algorithm (PLG) of Kégl et al. (2000), which is based on Principal Curves (Hastie and Stuetzle, 1989), was used in this work. It approximates the dataset with a polygonal line consisting of linearly connected vertices.

In the PLG a curvature penalty coefficient for the learned structure constrains the length and nonlinearity of the curve. As this is the main parameter that controls the shape of the principal curve, it was varied in a wide range to compare different mappings. The PLG is initialized with the first principal component and subsequently adds new vertices to the curve that are moved to minimize the mean distance to the points mapped onto it. The algorithm is stopped when the decrease in error drops below a certain threshold or a maximum number of vertices has been generated. A sufficiently continuous curve can be produced with the terminating condition coefficient set to a high value as it controls the maximum number of vertices used to construct the polygonal line. Additionally, the length penalty was set to zero to ensure the coverage of the data at the beginning and end points of the curve. Table 3.1 shows the parameter settings used.

curvature penalty coeff.	0.1 – 0.9
length penalty coeff.	0.0
terminating cond. coeff.	99.0
optimization threshold	$3 \cdot 10^{-3}$

Table 3.1: Parameter settings of the PLG.

3.3 Building an alternative model

State space reduction is the basis of the proposed model reduction scheme, but a further step is needed to arrive at a simple mathematical model of the reduced states' dynamics. A prominent way of finding functional descriptions for given datasets was followed here as provided by Genetic Programming (Koza, 1992). GP is one of many optimization schemes based on the principles of evolution, namely diversification, inheritance and survival of the fittest. In GP, structures representing functional forms are randomly initialized from a set of monomial functional terms and transformed using evolutionary inspired operators to optimize a given fitness measure in the course of the algorithm. Furthermore, as the models produced may contain adjustable parameters, their values have to be optimized in order to calculate the fitness of each model. Because of the diversity of the generated models, this is a very

demanding task for any optimization scheme. In this work a hybrid method combining a Genetic Algorithm (GA) and a gradient based optimizer was applied, making use of the benefits of both.

3.4 Structure optimization

The basic structure of the GP algorithm used in this study follows the approach put forward by Cao et al. (2000) with some major differences concerning the fitness calculation and parameter optimization (see below). Target systems to evolve are ODEs of the form

$$\dot{\xi}_i = f_i(\xi_1, \dots, \xi_l, c_1, \dots, c_j), \quad (3.6)$$

with $i = 1, \dots, l$, where ξ_i denotes state variable i and $\{c_g\}_{g=1}^j$ are free parameters. Each system of differential equations is represented as a set of l GP-trees $\{T_i\}_{i=1}^l$ with a maximum tree depth d_m . The GP-trees encode the right-hand sides of (3.6), see Figure 3.6. Starting with a randomly generated initial population of $n_{p,GP}$ members, their structure was evolved using GP to maximize a certain fitness measure or, respectively, minimize an error measure (see section 3.4.1). Furthermore, as the fitness value of a given structure naturally depends on the exact values of its free parameters, a parameter estimation/optimization was carried out for each member once a new generation had been built (see section 3.5).

3.4.1 Fitness calculation

The calculation of a GP population member's fitness is one of the most time consuming steps, as it involves the numerical integration of each system. As the type and complexity of the ODE system to solve are not known beforehand, algorithms applicable here must be fast, precise and robust. To account for the possibility of stiff differential equations, a fast variable step size/variable order algorithm (CVODE) was used that is part of the SUNDIALS package (Hindmarsh et al., 2005). The error measure quantifying the goodness-of-fit for an ODE system was the mean of the root mean square error (*RMSE*) of the simulated and original time series for all state variables:

$$RMSE = \frac{1}{l} \sum_{i=1}^l \sqrt{\frac{1}{n} \sum_{j=1}^n \left(\hat{\xi}_i(t_j) - \xi_i(t_j) \right)^2}, \quad (3.7)$$

where $\hat{\xi}$ denotes the simulated states, l being the number of state variables and n the number of time steps. As the error was averaged over all state variables they should have similar ranges or must be scaled appropriately before calculating the *RMSE*.

Additionally, as a minimal mathematical description of the dynamics was desired, the number of nodes in all trees served as a measure of complexity. To account for the tradeoff between structural simplicity and goodness-of-fit, a simple multi-objective optimization scheme was used: all members were

ranked according to both measures (*RMSE* and number of nodes) individually and the population was sorted according to the sum of ranks. This scheme is a variant of the basic aggregating optimization approach (see Coello Coello, 2000, and references therein). Independence of the objectives' values was achieved by using the individual ranks .

3.4.2 Genetic Programming

Every GP tree consisted of several linked nodes representing either a basic arithmetical operator, a numerical constant parameter or one of the state variables ξ_i (Table 3.2). In every new generation an elitism strategy was used to prevent the loss of the best $n_{c,GP}$ members found in earlier generations by copying them to the new population. Additionally, $n_{r,GP}$ of the new members were created randomly to introduce 'new genetic material'. The remaining members of the new population were created by crossover and mutation of existing members. The selection of the members to cross or mutate was carried out using a tournament selection scheme with $n_{t,GP}$ members. Crossover and mutation were then applied with equal probability and produced one new GP member each. For crossover, the first two best members of the tournament were used and random switching of subtrees was performed for each pair of trees T_i of both members. The base trees were thereby chosen randomly from the first or second member. If a newly generated tree was larger than allowed according to the maximum depth d_m the corresponding base tree was left unaltered. For mutation, the tree and mutation point in this tree were selected randomly and a new subtree was created at this point. The crossover and mutation points were thereby selected from the equally weighted levels of the GP tree, similar to the depth-fair selection scheme of Kessler and Haynes (1999).

Finally, the structure of the newly generated members was compared to all population members built so far. In order to make this structural comparison possible, the new members were normalized by moving terminal constant nodes to the left hand side of binary nodes encoding commutable operations (i.e. + and *). When a member's structure was already present in the population, this member was deleted and a new one was produced using the selection procedure described above to prevent an accumulation of members having the same structure.

The GP algorithm was stopped when a maximum number of generations, $g_{m,GP}$, was reached or no fitness improvement could be achieved in the last $g_{i,GP}$ generations. Specifications of the GP algorithm are listed in Table 3.2.

3.5 Parameter optimization

Many different strategies can be used for optimization of parameters, among which are local gradient-based methods, like quasi-newton algorithms, or random and evolutionary search methods, like Simulated Annealing, Evolutionary Algorithms or GA (see e.g. Pham and Karaboga, 2000). In this work,

Binary functions		+, -, *, /	
Terminals		$const., \xi_i$	
$n_{p,GP}$	50	d_m	3,4,5
$g_{m,GP}$	500,1000	$g_{i,GP}$	250,500
$n_{t,GP}$	4	$n_{c,GP}$	2
$n_{r,GP}$	5		

Table 3.2: Specifications of the GP module.

a hybrid of global and local search methods was applied. There are multiple variations of such methods (see Goldberg and Voessner, 1999). As evolutionary search strategies are able to quickly find promising regions of the search space, but are slow to reach an optimum, local gradient-based search strategies can be enacted to improve the overall performance of the search process.

Another alternative for parameter optimization is to constrain the ODE structure to incorporate linear parameters only. For such a case, numerical estimation can be implemented instead of evolutionary search strategies. This numerical estimation is outlined in section 3.5.2.

3.5.1 GA and local optimization

For the hybrid optimization of general ODE structures, the global part of the optimization procedure was carried out with a Genetic Algorithm. A local search strategy was applied to a subset of the best population members every s_{loc} generation steps (see Table 3.3 for the control parameters).

Initially, a GA population with $n_{p,GA}$ members was created for an individual GP members' n_{par} parameters by randomly initializing the parameter values in a specified range $[p_{min}; p_{max}]$. Note, however, that this parameter range was chosen arbitrarily as no constraints for the GP parameters could be given beforehand. The local optimizer, as well as the crossover and mutation operations of the GA, may produce values outside this range.

After the calculation of all members' fitness values according to (3.7) a new population was created. Similar to the GP module, copies of the best $n_{c,GA}$ members of the last generation and $n_{r,GA}$ random members were included in the new population and the remaining GA objects were created by crossover and mutation using again a tournament selection scheme with $n_{t,GA}$ members. Because of the hybrid global/local optimization scheme used, crossover and mutation operations of the GA mainly served to explore the parameter space and were not needed to fine-tune a given parameter value. Crossover and mutation operators were therefore applied with equal probability to the selected members. Furthermore, the mutation scheme used by Cao et al. (2000), which is based on the Breeder Genetic Algorithm (Mühlenbein and Schlierkamp-Voosen, 1993), was altered to produce values with a higher dispersion about the original parameter value as follows. One of the GA members' parameters

$v_m, m = 1, \dots, n_{p,GA}$ was selected randomly and its value was changed according to

$$v_m^* = v_m \pm 0.3 \cdot 2^{-j} \cdot (p_{max} - p_{min}) \quad (3.8)$$

where v_m^* denotes the new parameter value and the integer j was chosen randomly in the range of 0 to 3. The crossover scheme used was the BLX-crossover with an α -value of 0.5 (Eshelman and Schaffer, 1992) which produces new parameter values $v_{m,1}^*, v_{m,2}^*$ from parameters $v_{m,1}$ and $v_{m,2}$ of two GA members according to

$$v_{m,1}^* = (1 - \mu) \cdot v_{m,1} + \mu \cdot v_{m,2} \quad (3.9)$$

$$v_{m,2}^* = \mu \cdot v_{m,1} + (1 - \mu) \cdot v_{m,2} \quad (3.10)$$

with uniform random value $\mu \in [-\alpha, 1 + \alpha]$.

A local optimization procedure was applied to the best n_{loc} members of the GA population when the fitness value of the best member could not be improved in the last s_{loc} generations. The scheme used for local optimization was a limited-memory quasi-Newton method for unconstrained optimization (L-BFGS-B; Zhu et al., 1997). n_{loc} was thereby set to a value higher than $n_{c,GA}$ to ensure that new GA objects which had not been copied from the last generation were optimized.

Parameter optimization was terminated when a maximum of $g_{m,GA}$ generations was reached or fitness values did not improve in the last $g_{i,GA}$ generations.

$n_{p,GA}$	30	$g_{m,GA}$	50
$g_{i,GA}$	30	$n_{t,GA}$	4
$n_{c,GA}$	2	$n_{r,GA}$	5
p_{min}	-10	p_{max}	+10
n_{loc}	$2 \cdot n_{c,GA}$	s_{loc}	10

Table 3.3: GA and local optimization specifications.

3.5.2 Linear-in-parameter models

If the models to learn are constrained to contain linear parameters only according to

$$\dot{\xi}_i = \sum_{g=1}^j c_g \cdot f_g(\xi_1, \dots, \xi_l), \quad i = 1, \dots, l \quad (3.11)$$

parameters c_g can be estimated directly by solving a least-mean-squares problem (LMS; Ando et al., 2002). A special class of ODE models of this type, namely models with polynomial structure, have been used often in system identification applications (Smirnov et al., 2002; Aguierre et al., 2001).

Because of these restrictions, however, the learned model structures may be relatively complex compared to models with nonlinear parameters, as more (higher order) polynomial terms are needed to approximate more complicated dynamics. For datasets that can be approximated with simple ODE models, however, this procedure offers a fast alternative to the time-consuming GA/local optimizer approach for the parameter optimization problem. The example calculations in sections 3.8.1 and 3.8.2 were carried out for linear- and nonlinear-parameter models to compare the different results.

3.6 Reconstruction of model dynamics

The *RMSE* alone may not be sufficient to decide about the ability of a model to reproduce a given time series. The simplest approximation for example, the mean value, is also a good model of a given dataset. The constructed models should rather be able to reproduce the *dynamics* of the time series $\xi_1(t)$. Focusing on oscillatory regimes we provide a simple way of capturing different types of dynamics by computing the change in amplitude of the n_{max} local maxima of the time series:

$$\Delta max_i = max_i - max_{i-1}, \quad (3.12)$$

with $i = 1, \dots, n_{max}$, where max_i indicates a local maximum, i.e. a point of $\xi_1(t)$ where the slope changes its sign from positive to negative. Typical temporal courses of Δmax_i and the associated time series are shown in Figure 3.1.

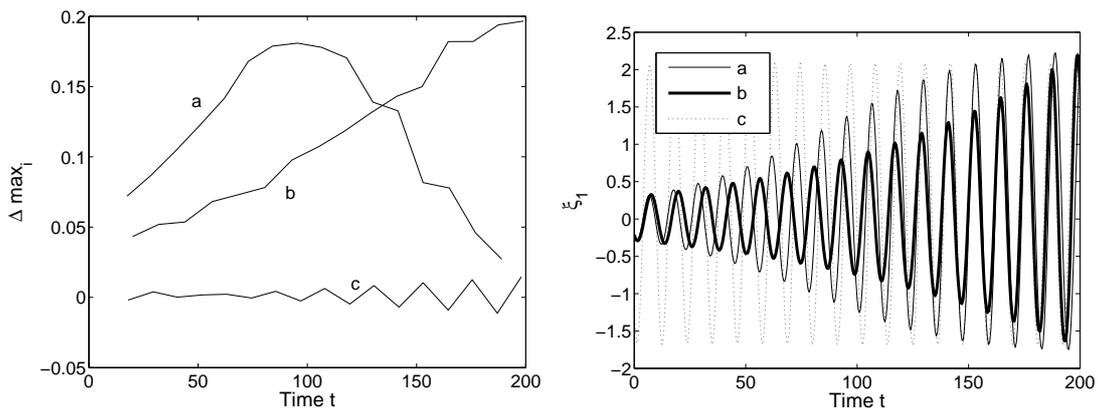


Figure 3.1: Typical examples of the change in amplitude of local maxima Δmax_i over time with corresponding time series.

3.7 Model examples

To demonstrate the proposed methodology, the reduction scheme was applied to two simple two-dimensional models: the model of a mathematical pendulum and a simple predator-prey model. While these systems have different origins, they share common features of their dynamics. They can be seen as the most basic model realizations of oscillatory behavior in a physical and biological system respectively. In analogy to the restoring gravitational force acting on the axis of the pendulum, the oscillations of the biological system are thereby caused by the predation pressure and prey abundance. The simple structure of both models and their similar dynamical behavior provide a good basis for the interpretation of the reduced models as the common principles underlying the oscillatory dynamics can be made transparent by the reduction process.

Both systems were reduced to one new state. However, as a one-dimensional model is not able to produce oscillations, the learned ODE models were taken to be of second order,

$$\dot{\xi}_1 = \xi_2, \quad \dot{\xi}_2 = f(\xi_1, \xi_2), \quad (3.13)$$

where ξ_1 indicates the newly built state variable and ξ_2 its derivative. This in fact means that the systems still were two-dimensional and the reduction of the models becomes a transformation.

3.7.1 The mathematical pendulum

Consider the idealized mathematical model of a pendulum: a point mass, connected to a rigid massless axis whose motion is not affected by friction. The temporal dynamics of the angle α between this pendulum and the vertical axis can be described with a second order differential equation:

$$\ddot{\alpha} = -\frac{g}{l} \cdot \sin(\alpha), \quad (3.14)$$

where g denotes the gravitational acceleration and l is the length of the pendulum's axis. A two-dimensional system was derived from (3.14) by considering observables of the model, namely its amplitude in vertical (y) and horizontal (x) direction,

$$x = l \cdot \sin(\alpha), \quad y = l \cdot \cos(\alpha).$$

The length of the pendulum and the starting angle were arbitrarily chosen to be $l = 3m$, $\alpha(t_0) = 3$ ($\approx 172^\circ$) and g was set to $9.81m/s^2$, resulting in the time series depicted in Figure 3.2. Though the physical units of the parameters and state variables are given here, the transformed models will be used in dimensionless forms. This can be justified by the fact that the projected state variables have other physical meanings than the original ones.

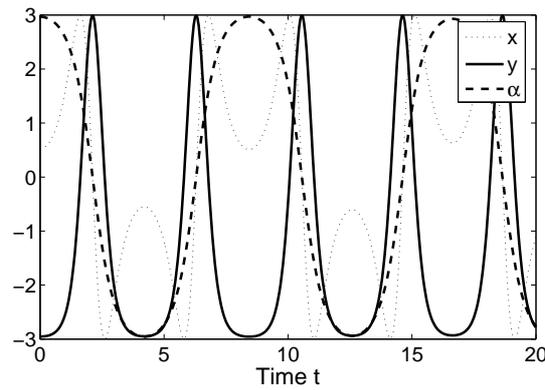


Figure 3.2: Time series of the pendulum (equations (3.14) and (3.15)).

3.7.2 Simple predator-prey model

The simple predator-prey model chosen was the dimensionless Rosenzweig-McArthur model describing the dynamics of prey and predator abundances in the following way (Rosenzweig and McArthur, 1963):

$$\begin{aligned} \dot{x}_1 &= (1 - b_1 \cdot x_1) \cdot x_1 - \frac{x_1}{1 + b_2 \cdot x_1} \cdot x_2 \\ \dot{x}_2 &= b_3 \cdot \left(\frac{x_1}{1 + b_2 \cdot x_1} - 1 \right) \cdot x_2. \end{aligned} \quad (3.15)$$

Here, the interdependency between predator x_2 and prey x_1 is modeled via Holling-II terms (Holling, 1959) where parameters b_1 , b_2 and b_3 determine the temporal course of the system, e.g. the convergence towards a stable limit cycle or fixed point in state space. A snapshot of the model's limit cycle dynamics was generated using the parameter settings of Table 3.4 (see Figure 3.3).

$[x_1, x_2]_{t=0}$	[2, 1.5]
normalized inverse capacity b_1	0.15
normalized inverse half saturation for predation b_2	0.5
normalized predator mortality b_3	1

Table 3.4: Simulation settings for the simple predator-prey model (3.15).

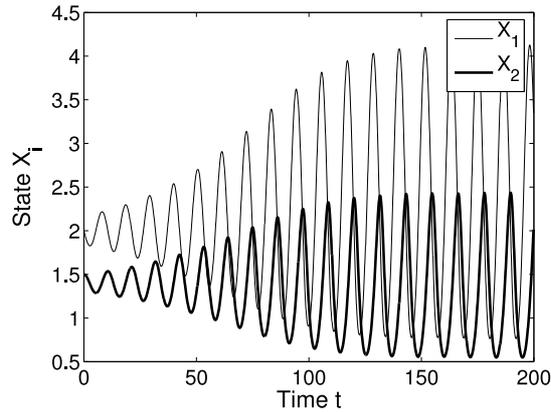


Figure 3.3: Time series of model (3.15) using the simulation settings of Table 3.4.

3.8 Results

3.8.1 Mathematical pendulum

State reduction

In order to study the effect of nonlinear versus linear mapping, both the reduction to the first linear principal component and a nonlinear principal curve were calculated. Settings for the PLG can be found in Table 3.1. Results for different settings of the curvature penalty coefficient were very similar. Only for values larger than $\rho \approx 0.65$ was the curvature constraint too high to allow the principal curve to follow the data distribution. For these values, the principal curve resembled the principal component (the starting condition of the PLG) instead. Figure 3.4 shows the resulting linear and nonlinear line structures in state space.

Note that a reduction to the first principal component retains only 54% of the data variance. Therefore, the reduction clearly is not justified as a lot of information about the system contained in the data matrix is lost. As the example is only used to demonstrate the multi-step method, however, this problem will be ignored.

The mapping from the dataset \mathbf{X} to its low-dimensional counterpart $\mathbf{\Xi}$ was achieved by finding the points on the linear or nonlinear structures which are closest to the data points in the least-squares sense. Figure 3.5 shows the time series resulting from the linear and nonlinear mappings.

Model learning

The GP-algorithm, using the settings from Table 3.2, was run repeatedly for the reduced data taken from the PCA ($\mathbf{\Xi}_{PCA}$) and PCurve ($\mathbf{\Xi}_{PCurve}$) mappings. Two modes of operation were carried out: for linear-in-parameters (LIP) models to be learned, the maximal depth of the GP-trees d_m was varied

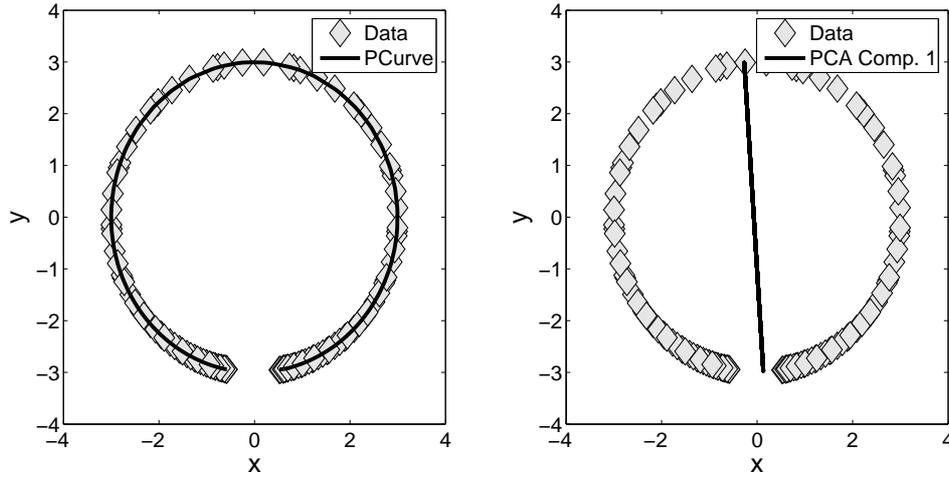


Figure 3.4: Principal curve and first principal component for the time series data of the pendulum in the original 2-dimensional state space.

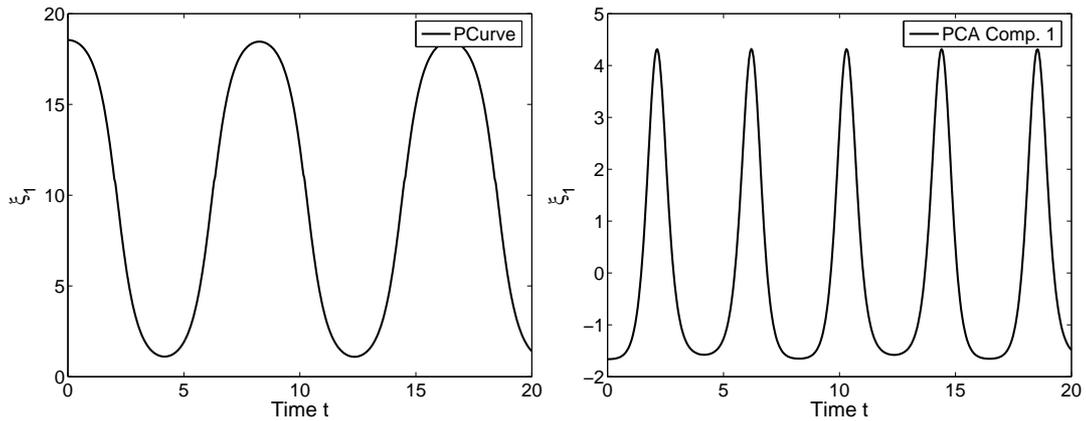


Figure 3.5: Time series of the new state variable ξ_1 resulting from mappings onto the line structures in Figure 3.4.

from 3 to 4, thereby increasing the maximal complexity of the learned ODE structures. For each of these configurations, 50 runs with $g_{m,GP} = 1000$ and $g_{i,GP} = 500$ were carried out. The parameter values were estimated directly from the time series data using the LMS method (section 3.5.2).

The GP algorithm was also run to evolve models with nonlinear parameters (NLIP). In these cases the hybrid GA/local scheme was used for parameter optimization. As this kind of optimization is very time consuming, d_m was again varied from 3 to 4 (20 runs each) but the GP was stopped earlier when the number of generations reached a maximum of $g_{m,GP} = 500$ or $g_{i,GP} = 250$.

In order to analyze an ensemble of models showing different tradeoffs between fitness and complexity, the best 10 members of the final populations for all runs were used. The best model in terms of

RMSE alone was always found to be part of this collection.

PCA results - The GP-tree of the best model found, with a good compromise between fitness and size, is depicted in Figure 3.6. Written in mathematical notation it follows

$$\begin{aligned}\dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= c_1 \xi_1^2 + c_2 \xi_1 + c_3\end{aligned}\tag{3.16}$$

with

$$c_1 = -3.1434, \quad c_2 = 2.0, \quad c_3 = 12.4023.$$

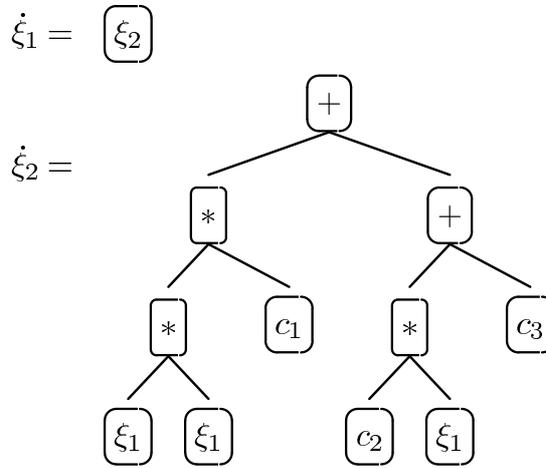


Figure 3.6: GP-tree of model (3.16).

Figure 3.7 shows the resulting simulated time series. This model structure was found in all modes of operation but the parameter optimization results turned out to be very different. As a rule, the hybrid optimization scheme gave much better results than the direct numerical estimation in LIP mode. This can partly be attributed to the strong sensitivity of the LMS optimization to errors in the numerical estimation of the second derivative. The parameter values of (3.16) were taken from the best NLIP run with $d_m = 3$.

PCurve results - The reduction of the two-dimensional pendulum using the nonlinear principal curve is much more viable than the linear reduction as the variance of the original system is explained completely. Figure 3.8 shows RMS errors and the number of nodes of the resulting systems. As expected, the mean fitness of the best members increases with higher complexity of the trees (higher values of d_m). Additionally, we can see that the best NLIP models found are significantly less complex than the LIP models while being comparable or better in terms of RMS error, though this is mainly

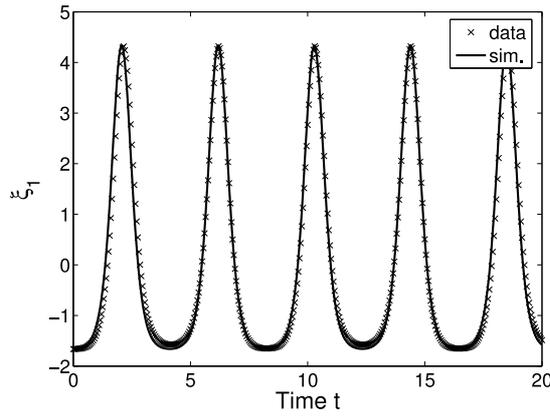


Figure 3.7: Simulated time series of the best model found for the PCA reduction of the pendulum data.

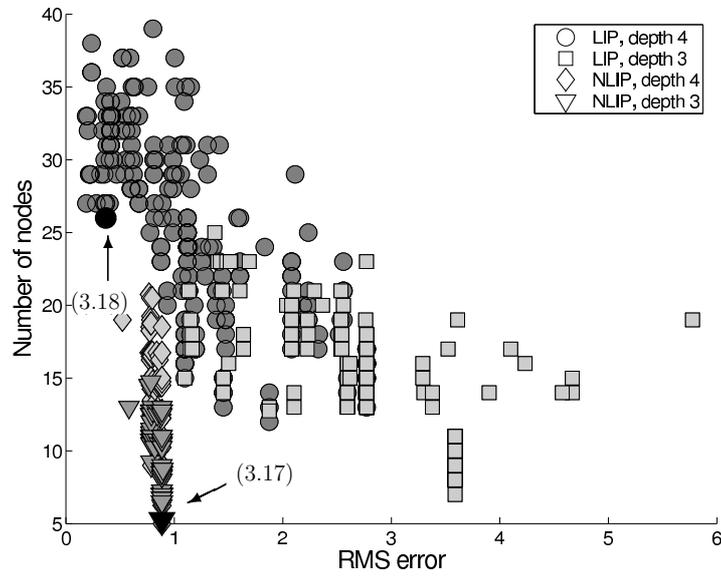


Figure 3.8: RMS errors and number of nodes of generated ODE systems for Ξ_{PCurve} of the pendulum. Arrows indicate models (3.17) and (3.18).

caused by differences between the parameter optimization schemes (see PCA results). The most simple NLIP model with $d_m = 3$ and $RMSE =$ can be written as

$$\begin{aligned}\dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= -c_1 \xi_1 + c_2,\end{aligned}\tag{3.17}$$

with

$$c_1 = 0.5817, \quad c_2 = 5.5242.$$

The resulting time series is shown in Figure 3.9(a). Models with a structure following (3.17) were also found in NLIP mode with $d_m = 4$ as well as in LIP mode with $d_m = 3$. For example, the smallest LIP model in Figure 3.8 is of this structure but its $RMSE$ is much higher than the error of model (3.17) as the optimization scheme failed to fit the linear parameters.

The more complex model results with even better fits all contained nonlinear terms of ξ_1 and ξ_2 like the rather complex best LIP model with $d_m = 4$ (Figure 3.9(b)):

$$\begin{aligned}\dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= c_1 \xi_2^2 \xi_1^2 + c_2 \frac{\xi_2^2}{\xi_1} + c_3 \xi_1 + c_4\end{aligned}\quad (3.18)$$

with

$$\begin{aligned}c_1 &= -0.001, & c_2 &= 0.9603 \\ c_3 &= -0.2667, & c_4 &= 2.1719.\end{aligned}$$

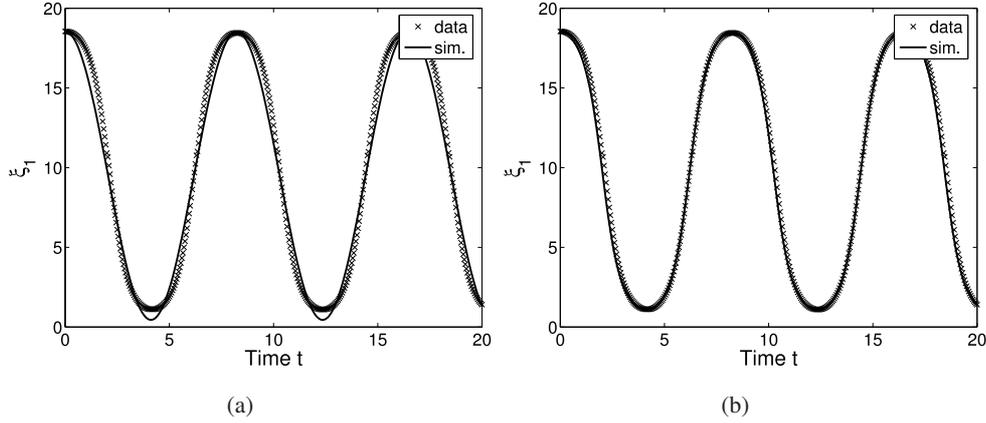


Figure 3.9: Time series of two models found for the reduced pendulum: (a) model (3.17), $RMSE = 0.89$ and (b) model (3.18), $RMSE = 0.37$.

3.8.2 Predator-prey model

State reduction

Unlike the results of the pendulum, the form of the principal curve for the data distribution of the predator-prey model strongly depends on the value of the curvature penalty coefficient ρ (Figures 3.10 and 3.11).

We plainly see the effect of the learning algorithm being unaware of the temporal dynamics, as discussed in section 3.3, in the mapping for $\rho = 0.1$. On the one hand, this principal curve closely follows the data distribution as the low value of ρ permits a highly curved structure. On the other hand, however, the mapping results in a time series showing sharp transitions between subsequent time steps

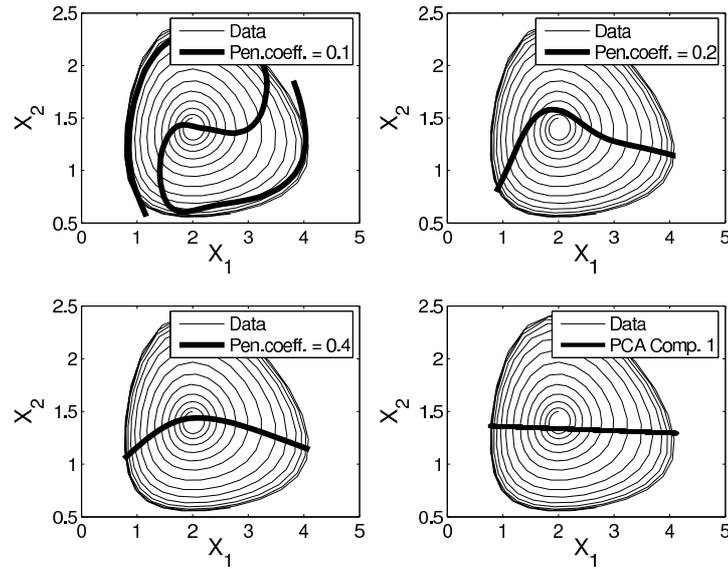


Figure 3.10: First principal component and principal curves for the predator-prey model with different settings of the curvature penalty coefficient ρ .

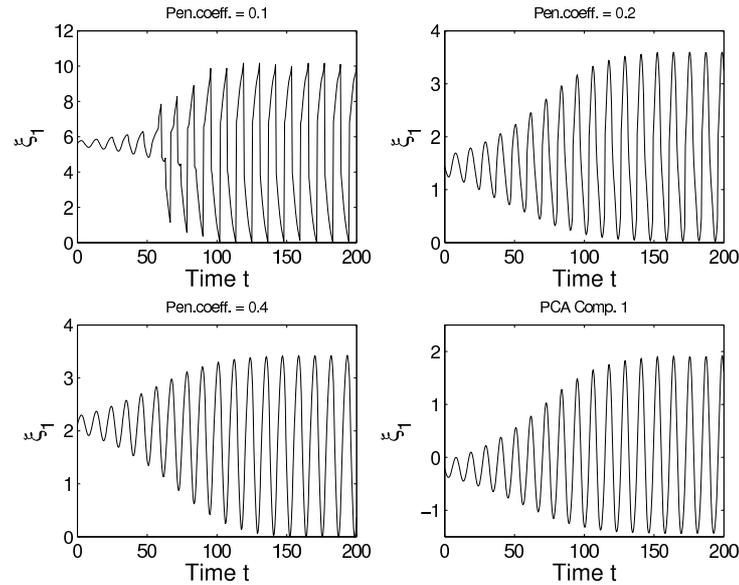


Figure 3.11: ξ_1 for structures in Figure 3.10.

which can not be reproduced by a meaningful model. A quantification of the temporal mapping errors can be achieved with the $MSmE$ measure from section 3.3 (Figure 3.12). Apart from the lowest values of ρ , where a pronounced drop in $RMSE$ and an increase in $MSmE$ occurs, the time series for the principal curves show virtually no temporal distortion and their $RMSE$ and $MSmE$ values are very similar to those of the PCA. Because of these similarities, the time series of the simplest

mapping, i.e. the PCA result, was chosen as input data for the model construction step.

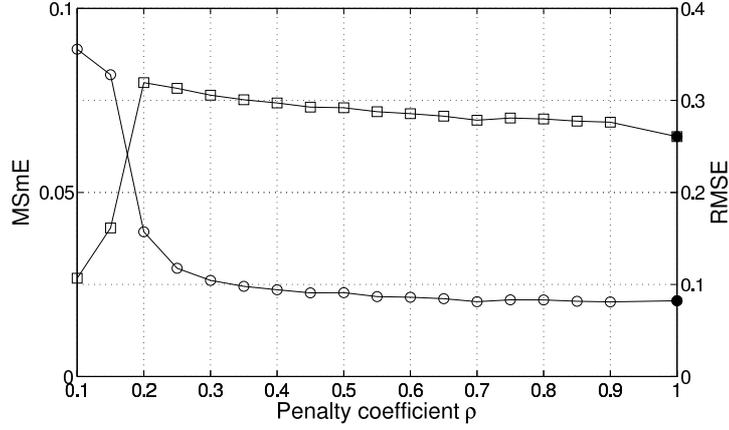


Figure 3.12: Mean smoothness (MSmE, ○) and mean root-mean-square error (RMSE, □) for different settings of the curvature penalty parameter ρ . PCA results are marked with filled symbols.

Model learning

The GP-algorithm was run repeatedly for the reduced data taken from the PCA mapping (Ξ_{PCA-1}). The modes of operation follow those of the pendulum but additional runs using $d_m = 5$ were carried out for the LIP models and in the NLIP case d_m was set to 3 (10 runs) and 5 (5 runs). Again, the best 10 members of the final populations for all runs were used for further analysis and Figure 3.13 shows RMS errors and number of nodes of the generated systems. Once more, the differences in tradeoff between fitness and size for LIP and NLIP mode can be seen: NLIP models are typically smaller and have a lower error value than the LIP models. As outlined in section 3.6, the RMS error alone is not able to distinguish qualitatively correct and incorrect models and the course of relative change in local maxima was used to narrow the selection of best models. The transient dynamics we are interested in are characterized by a distinct maximum in Δmax_i followed by a decrease towards the end of the time series. Models showing this temporal characteristics are drawn with filled symbols in Figure 3.13. It can be seen that no simple LIP models with $d_m = 3$ showing the correct transient behavior could be found. Thus, the advantage of NLIP over LIP models in this case can not be attributed to the parameter optimization scheme alone but is also based on the optimization of the structure. In fact, the two NLIP models with $d_m = 3$, which were able to reproduce the desired dynamics, contain a nonlinear parameter. The difference in size between these two systems is caused by redundant nodes only, and they can be expressed as

$$\begin{aligned} \dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= \frac{-c_1 \xi_1 + c_2 \xi_2}{c_3 - \xi_1 \xi_2}, \end{aligned} \quad (3.19)$$

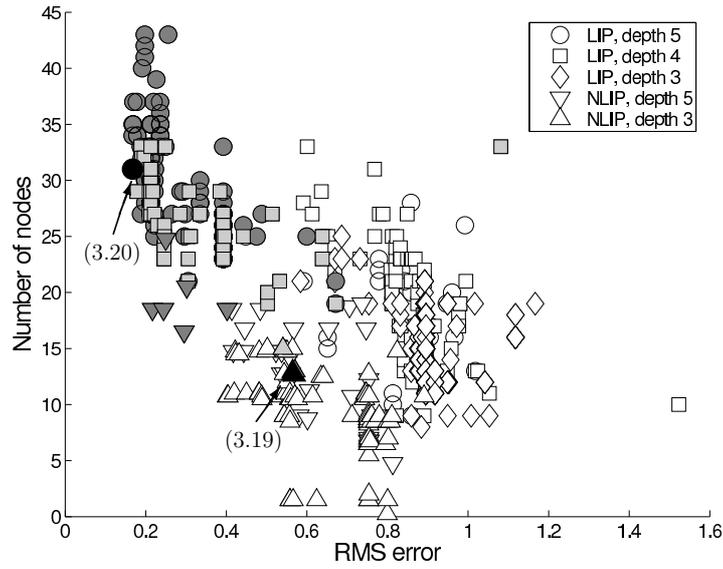


Figure 3.13: RMSE and number of nodes of generated ODE systems for Ξ_{PCA-1} . Models that were able to reproduce the transient dynamics are plotted with filled symbols and arrows indicate models (3.19) and (3.20).

with

$$c_1 = 2.8296, \quad c_2 = 0.2581, \quad c_3 = 9.0,$$

and $RMSE = 0.54$ (Figure 3.14). As with the results of the pendulum, more complex models with better fitness values again contain mixtures of higher order nonlinear terms of ξ_1 and ξ_2 . The best LIP model with $d_m = 5$ follows

$$\begin{aligned} \dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= -\xi_1(c_1 - c_2\xi_1^2) + \xi_2(c_3 - c_4\xi_1^2) + c_5\xi_1^2\xi_2^2, \end{aligned} \quad (3.20)$$

with

$$\begin{aligned} c_1 &= 0.3327, \quad c_2 = 0.0174, \quad c_3 = 0.0423 \\ c_4 &= 0.0574, \quad c_5 = 0.0461 \end{aligned}$$

and $RMSE = 0.17$ (Figure 3.14).

3.9 Discussion

The control parameters of evolutionary learning schemes must be tuned to the problem at hand in order to provide optimal results. Parameter tuning of the hybrid GP/GA module, however, is very difficult. On the one hand, the fitness calculations of the GP members, due to repeated numerical integrations

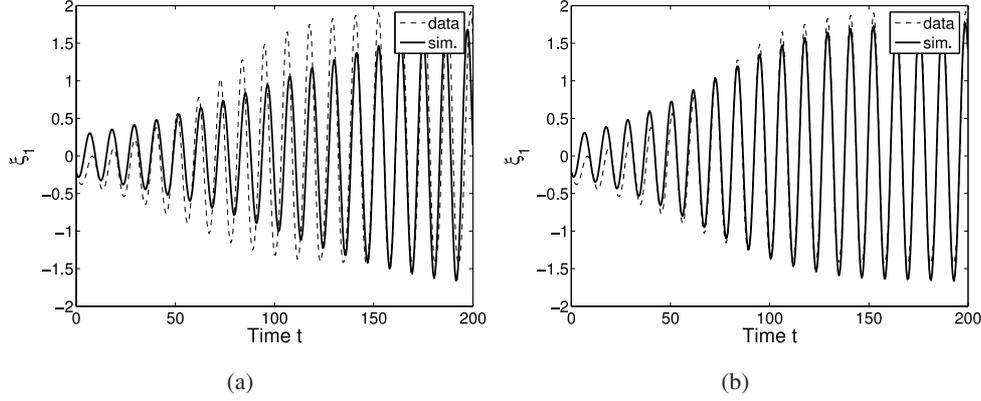


Figure 3.14: Time series of two models found for the reduced predator-prey model: (a) model (3.19), $RMSE = 0.54$ and (b) model (3.20) with $RMSE = 0.17$.

of the models, result in long operation times and inhibit the use of many different parameter settings. On the other hand, the tuning of GA parameters is also intractable because of the diversity of the constructed GP models. The choice of GP and GA parameter settings used (Tables 3.2 and 3.3) must therefore rather be seen as a first guess based on computing time considerations. Concerning the GA module alone, the incorporation of the local optimizer generally turned out to enhance the optimizer's performance most drastically. In comparison, the effect of different control parameter settings on the method's performance was negligible.

The simplified models built by the reduction method can typically not be derived from the original models analytically. One example of this is found in the PCA reduction of the pendulum, equation (3.16), which describes a simple nonlinear oscillator. This model can not be derived from the original equations (3.14) and (3.15), e.g. by approximating the principal component with the vertical axis y and further substitution of α .

In some cases, however, an analytical derivation of the learned models is possible if approximations of the original model are used. The derivation of model (3.17), the PCurve reduction result of the pendulum, is a good example. The dynamics of ξ_1 for the PCurve mapping of the pendulum resembles the dynamics of α and the learned model should therefore just be equal to (3.14). However, as no $\sin(\cdot)$ function was used in the GP, the algorithm could only approximate this function with a series expansion and the result would be a very complicated model with higher-order terms. Alternatively, equation (3.17) can be interpreted as a small angle approximation of the pendulum equation (3.14) with $\sin(x) \approx x$ and a linear transformation of the result. In order to match the frequency and amplitude of the original, the transformed pendulum's axis \hat{l} must be longer than l . As the values of ξ_1 were chosen arbitrarily, ξ_1 can be written as a linearly transformed version of α ,

$$\xi_1 = \frac{\alpha - \alpha_o}{\tilde{\alpha}} \quad (3.21)$$

with scale $\tilde{\alpha}$ and offset α_o . Using a new pendulum length \hat{l} and (3.21) to substitute α in the approximation of (3.14), $\ddot{\alpha} \approx -g/\hat{l} \cdot \alpha$, we get

$$\ddot{\xi}_1 \approx -a \cdot \xi_1 + b, \quad (3.22)$$

with $a = c_1 = g/\hat{l}$ and $b = c_2 = -(g \cdot \alpha_o)/(\hat{l} \cdot \tilde{\alpha})$. Thus, with an appropriate new pendulum length, mathematical viable small angle approximations of (3.14) can be found that match the data and the constant c_2 in (3.17) can be attributed to the linear transformation of α . Using the parameters of (3.17), this length is $\hat{l} = g/c_1 \approx 5.6 \cdot l$ and with an initial angle of e.g. $\xi_1(t_0) = 0.1$ ($\approx 5.7^\circ$), we get $\alpha_o = 3.0319$ and $\tilde{\alpha} = -0.3193$.

These two model results also exemplify the advantage the state reduction using a nonlinear mapping may have over the reduction based on a linear mapping. The pendulum equation's nonlinearity was captured by the principal curve and the resulting model (3.17) is a simple linear oscillator. The linearly reduced model (3.16) is slightly more complex due to a nonlinear term. On the other hand it was shown that a nonlinear state reduction of time series data does not necessarily produce reasonable results as it may introduce temporal disruptions of the data. The MSmE was introduced as a sensible measure of these mapping errors and in combination with the RMSE it provides a good means to compare mappings of different nonlinearity.

Regarding the system identification step alone, the combination of GP with the hybrid parameter optimization scheme and no restrictions concerning the use of nonlinear parameters turned out to be superior to the common linear-in-parameter model identification approach. The higher flexibility of the structure optimization together with the successful parameter optimization scheme resulted in smaller and better models. This could be seen most clearly for the reduced predator-prey data as no simple LIP model could be found at all in this case.

As stated above, the model examples can be seen as basic realizations of oscillatory dynamics in physical and biological contexts, respectively. These examples were chosen to demonstrate that the reduction scheme as the transformation of both systems to the same structure, namely a second-order oscillator equation, may reveal common properties and differences of the system's dynamics. For the pendulum, on the one hand, the most simple models are composed of polynomial equations incorporating the reduced state only. It can easily be shown that fixed points of ODE systems having this structure are only marginally stable, i.e. every starting point in the state space is part of a limit cycle and no transient behavior can be found. This limitation, however, can be removed by the incorporation of nonlinear terms in the reduced predator-prey models. Model (3.19) has only one fixed point at $(x_1^*, x_2^*) = (0, 0)$. A local stability analysis of this point gives the necessary conditions for the existence of a limit cycle:

$$\frac{c_2}{c_3} > 0 \quad (3.23)$$

$$4 \cdot c_1 c_3 > c_2^2. \quad (3.24)$$

The first condition (3.23) determines the stability of (x_1^*, x_2^*) : for a limit cycle to exist the fixed point has to be unstable and the condition is met when c_2 and c_3 have the same sign. The conformance with the second inequality (3.24) is responsible for the oscillatory dynamics. Figure 3.15 shows a stability diagram of the model. The system shows oscillatory dynamics above the parabola. Below it,

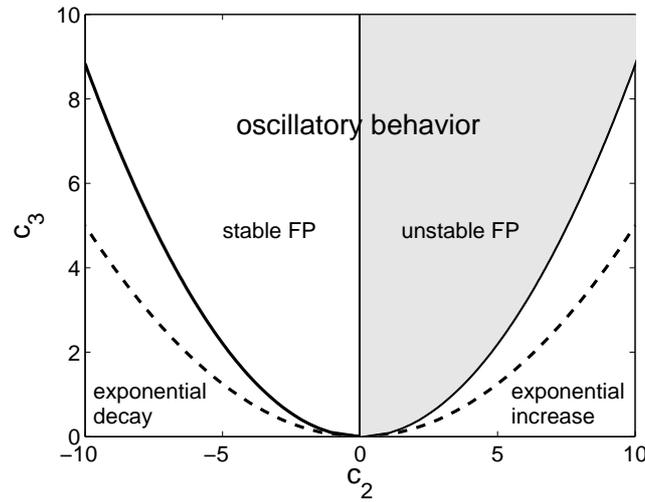


Figure 3.15: Stability diagram of model (3.19) for two different settings of c_1 . The solid parabola indicates the limit curve of the oscillatory regime for $c_1 = 2.8296$ and the parameter regime of the limit cycle is shaded in gray. The dashed curve shows the oscillatory limit for $c_1 = 5$.

the system's solutions move towards or away from the fixed point at an exponential rate, depending on the stability of the fixed point. For negative values of c_2 the fixed point is stable and the trajectories approach it in an oscillatory or exponential manner. Equations (3.23) and (3.24) also show that - different from c_2 and c_3 - the stability of the fixed point does not depend on parameter c_1 . All three parameters, however, affect the oscillatory behavior of the time series and a higher value of c_1 promotes the existence of oscillations (see the dashed line in Figure 3.15). In this respect, the parameters of the reduced model can also be dynamically linked to the parameters of the original model (3.15). An analysis of this model reveals that the stability of the system's non-trivial fixed point solely depends on two of the three parameters, namely b_1 and b_2 . The third parameter, b_3 , has an influence on the oscillatory regime only and high values of b_3 again promote the existence of a limit cycle. The parameters of models (3.15) and (3.19) therefore have very similar implications for their system's dynamical behavior.

Physically spoken, the second order differential equation can be seen as an equation of forces. Without the second additive term of (3.19), i.e. for $c_2 = 0$, shown as a vertical line in Figure 3.15, the system is still able to produce oscillations, but the limit cycle coincides with the marginally stable fixed point.

The resulting time series show damped oscillations that approach the fixed point asymptotically. The second term of (3.19), being proportional to the impulse ξ_2 , can therefore be interpreted as a driving force that increases the impulse and causes the trajectories to move away from the fixed point. The negative first term of (3.19) can be seen as a restoring force, as present in the small angle approximation of the pendulum, $\ddot{\xi}_1 = -a \cdot \xi_1$, with an additional nonlinear damping term in the denominator. Without this damping the system's solutions would explode because of the acting driving force. The existence of the stable limit cycle therefore is the result of a balance of both forces.

Finally, it should be noted that similar forces are present in the more complex models. The first two terms in (3.20), for example, can again be seen as nonlinearly corrected restoring and driving forces. As the corrections are quadratic in ξ_1 and therefore positive, however, this results in an additional damping term proportional to the impulse ξ_2 which is governed by c_3 . Just as the nonlinear denominator of (3.19), this term is responsible for the existence of the stable limit cycle. The last term of this model is an additional nonlinear driving factor. Thus, equation (3.20) essentially is just a more complex description of a balance of forces as given by (3.19).

3.10 Conclusion

The interpretable transformation and reduction of model systems is a methodically very demanding task. In this respect, the proposed method of this work, a combination of dimension reduction and evolutionary learning algorithms, was found to be a promising step towards an automatic reduction scheme producing simplified and interpretable model results. As the structure of the resulting transformed or reduced models is independent from the original formulation, alternative and potentially more general interpretations of a model's internal interactions are made possible.

This was demonstrated for the transformation of two simple oscillatory systems taken from physical and biological contexts. The present analysis thereby showed that the dimensionality of the original model formulations must not necessarily be high in order to get insights into relevant processes governing the dynamics. The transformation of the predator-prey model in this study revealed the system's fundamental analogy with a nonlinear oscillator. This result supports the concept of biological oscillators, originating in consumer-resource interactions, being the basic building blocks of general food web models (Vandermeer, 2004). In this respect, the nonlinear predator-prey oscillator found can be seen as a prototype for limit cycle dynamics of more complex biological systems.

For higher-dimensional models, the reduction ability of the presented approach naturally depends on the system's dynamics being confined to a low-dimensional subspace. In the biological context, the existence of such low-dimensional dynamics has for example been reported for lynx and hare limit cycle dynamics in a complex boreal food web (Stenseth et al., 1997). Other studies analyzing the dynamics of the complex European Regional Seas Ecosystem Model (ERSEM) revealed that the dynamics

of this model was confined to a small part of the state space (Wirtz and Wiltshire, 2005) and that some of the model's unexpected dynamic effects could be explained with a simple two-dimensional model (Kohlmeier and Ebenhöf, 1995). In a follow-up paper it will be shown that the limit cycle dynamics of more complex biological population models can readily be captured by reduced two-dimensional models (Bernhardt and Wirtz, 2007). The reduction process thereby offers new ways of interpreting the dynamic picture of biological interactions. Sustained limit cycle oscillations have further been reported for many biological population models as well as measured time series datasets (e.g. Turchin, 2003; Huisman and Weissing, 2001b). It may be proposed that the long-term dynamics in many of these cases is also intrinsically low-dimensional, offering fields of application for the model reduction mechanism.

Chapter 4

Oscillation of structural properties in food web models

Abstract

Simple models of biological food webs with a small number of state variables are, despite their simplicity, important building blocks of more complicated integrated models. Apart from the classical formulations using density-based state variables, the importance of trait-related effects on population dynamics has been pointed out in recent years. However, there still exist ambiguities about the choice of relevant traits as well as the corresponding dynamic equations. In addition, the relative importance of mutual direct and indirect effects of density- and trait-based variables is subject to ongoing research.

In this paper, we show that the dynamics of food webs can be described more efficiently when the model formulation is based on interaction variables incorporating the interdependencies of functional groups and resources. We apply a data-adaptive reduction method to build simple interaction models from time series data of a predator-prey and a NPZ model. In this framework, linear PCA is used to extract the new interaction variables describing structural food web properties which are not given explicitly in the original formulations. These variables can be interpreted as feeding limitations within the trophic subsystems of the food webs. Concurrently, the corresponding indirect interactions between the system's trophic levels are translated to direct interactions of the feeding limitations. Analogies with a physical oscillator further help to identify general biological forces controlling complex food web dynamics.

4.1 Introduction

Simple mathematical models describing the interactions of biological populations with a small number of state variables have been used and discussed by modelers for a long time and are still widely in use today. Based on the work of Lotka and Volterra in the 1920s and 1930s, most current model formulations of aquatic ecosystems are of the NPZ or NPZD type, in which one or two density variables for nutrients, phytoplankton, zooplankton and detritus form the main model structure (Palmer and Totterdell, 2001). Despite of their simplicity corresponding to an omission of many processes affecting food web dynamics, these models are still useful to reproduce bulk system properties and events such as mean seasonal patterns, phytoplankton blooms or trophic cascades (e.g. Schmitz, 1993; Franks, 2002). Aiming at improving their accuracy and prediction capabilities, the development of

much more complex models has progressed in the last decade. This increase in complexity is, for example, based on a larger number of functional types or additional structural information in the planktonic compartments (Flynn, 2003; Le Quéré et al., 2005). However, the incorporation of more state variables and parameters to represent a more detailed population structure leads to a number of problems. First of all, the scarcity of available process data strongly limits the verification of complex models. The corresponding large number of model functions and parameters, which can often not be further refined using existing quantitative information on the underlying biological processes, are typically under-determined (Anderson, 2005). As a consequence, one may conclude that classical ecosystem modeling is trapped in a simplicity/complexity dilemma and neither of the modeling strategies appears to give vital benefit. This notion has been substantiated recently in the US-JGOFS testbed project. In this study, a collection of NPZ-type models was compared with respect to their ability to simulate observed biogeochemical cycling using data assimilation techniques. It was found that the overall ability to fit a data set that was not used for calibration before was relatively weak for all models. More notably, this ability did not depend on model complexity (Friedrichs et al., 2006). Because of the problems associated with complex descriptions and the fact that increased complexity not necessarily leads to better models, it is still sensible to use simple models derived from the Lotka-Volterra (LV) formulation. These models include growth, mortality and feeding terms which directly determine changes of population density. However, direct physical interaction, i.e. consumption of biomass, is only one form of predator-prey interaction. Some of the dynamic patterns in real systems can be related to predation-induced changes in species traits such as diet and habitat selection, feeding time or food quality (e.g. Abrams, 1995; Abrams et al., 1996; Relyea and Yurewicz, 2002). The notion of behavior- or trait-related aspects of population dynamics has been used in individual-based models for a long time as so-called *i*-states (DeAngelis and Rose, 1992) and attempts have been made to reconcile the population- and individual-based approaches (e.g. Fahse et al., 1998). In this respect, the density-related state variables of classical population models must be regarded as a description of the community mean of functionally identical individuals. The indirect effects of species traits on population composition and dynamics have been found to be of substantial importance (Beckerman et al., 1997; Preisser et al., 2005) but they are typically masked by measuring the net effect of predator/prey density changes (Křivan and Schmitz, 2004). Based on these results, it has been proposed in some studies to consider both, density- and trait-related variables for model-building (e.g. Wirtz and Eckhardt, 1996; Fussmann et al., 2005). Similar to the procedure of the Adaptive Dynamics approach (Dieckmann and Law, 1996), the dynamics of traits in these studies have been linked to derivatives of the growth function with respect to traits. However, the selection of the most important adaptive traits is fraught with ambiguities which limits the general applicability of trait variables in food web models. As the relevant net effect of both, direct and indirect density- and trait-based interactions, is reflected in the overall population dynamics it can be argued that the distinction between density- and

trait-related aspects of population models is artificial to some extent. If this is the case we may ask whether it is possible to find alternative, more problem (or data) adapted sets of key state variables. These new variables would then provide a basis for new types of models explaining the main features of the dynamics by minimal means. The concept of minimal complexity in model formulations, known as Occam's razor, is one of the major aims of modeling (Turchin, 2003). More importantly, reduced model structures could point to principles underlying food web dynamics which have been obscured by the domination of NPZ-type models.

In this study, we introduce a method deriving new biological state variables within the context of model reduction (MR) techniques. As the oscillatory behavior of Lotka-Volterra models can be seen as a reference case in ecology, we will focus on two of the most simple food chain models yielding predator-prey cycles in order to demonstrate the model reduction procedure. Namely, we will use a predator-prey model consisting of two state variables and a slightly more complicated NPZ chemostat model with an additional trophic layer.

So far, the spectrum of MR methods ranges from the formal aggregation of state variables (Iwasa et al., 1989; Auger and Poggiale, 1996a) to approaches where model formulations are restructured using model analysis and detailed knowledge of the underlying system (Raïck et al., 2006). We apply the Mapping-based Genetic Reduction approach (MAGER), a new MR scheme constructing simplified mathematical descriptions of biological system dynamics (Bernhardt, 2007). Unlike the methods mentioned above, the new scheme is a data-adaptive black-box procedure that ignores traditional model formulations in terms of equation structure and in particular the meaning of state variables. Furthermore, no expert knowledge is needed for its application. MAGER is composed of a state variable reduction and a subsequent model learning step generating new differential equations. These dynamic equations are able to reproduce the original model datasets to a large degree. However, while the mapping procedure allows for an interpretation of the new state variables, this may not be the case for the parameters and equations arising from the model learning. In this paper, we use analogies to an oscillatory physical system in order to relate the generated dynamic equations to first principles.

4.2 Models used

4.2.1 Predator-prey model

Rosenzweig and McArthur (1963) introduced a predator-prey model with capacity limited growth of prey and a Holling type II grazing function which is used here in a dimensionless form

$$\begin{aligned}\frac{dx_1}{dt} &= \left(1 - k \cdot x_1 - \frac{x_2}{1 + s \cdot x_1}\right) \cdot x_1 \\ \frac{dx_2}{dt} &= m_p \cdot \left(\frac{x_1}{1 + s \cdot x_1} - 1\right) \cdot x_2,\end{aligned}\tag{4.1}$$

with normalized abundance of prey x_1 and predator x_2 and parameters k , s and m_p explained in Table A.1. The normalized model formulation is chosen to simplify the analysis (see section A.2 for the derivation of (4.1)). Depending on the parameterization, the model is able to simulate coexistence of prey and predator populations either in form of a stable equilibrium (fixed point dynamics) or oscillatory behavior (limit cycle). For the present analysis, only the limit cycle dynamics is considered. For more details on the bifurcation properties of the system see (Rosenzweig and McArthur, 1963; Freedman, 1980; Kot, 2001).

We use model simulations with two different settings of k (Table A.1). Both setups result in time series showing transient oscillatory regimes of different length before a stable limit cycle is reached.

4.2.2 NPZ model

As an extension of the predator-prey system, we use a generic model of food web dynamics on three trophic levels in a chemostat which contains nutrient, phytoplankton and zooplankton biomass densities (Yoshida et al., 2003; Fussmann et al., 2000). The fourth state variable of the original formulation, where the zooplankton biomass was split up into a reproductive and a non-reproductive part, is simplified to reproductive zooplankton with an increased mortality. The system can then be written as

$$\begin{aligned}\frac{dN}{dt} &= \delta \cdot (N_0 - N) - F_P \cdot P \\ \frac{dP}{dt} &= \epsilon_P \cdot F_P \cdot P - F_Z \cdot Z - \delta \cdot P \\ \frac{dZ}{dt} &= \epsilon_Z \cdot F_Z \cdot Z - (\delta + m_z) \cdot Z,\end{aligned}\tag{4.2}$$

with

$$F_P = \mu_{max} \cdot \frac{N}{S_P + N}\tag{4.3}$$

$$F_Z = \xi_{max} \cdot \frac{P}{S_Z + \max(P, P^*)}.\tag{4.4}$$

See Table A.2 for a description of the parameters and their values used in this study. The starting values of N , P and Z were arbitrarily set to again produce limit cycle dynamics with transient oscillations.

4.3 Model reduction

The MAGER approach consists of a sequence of operations which lower the number of state variables of a given model and produce new mathematical descriptions that are able to reproduce the dynamics of the transformed state variables (Figure 4.1). This chapter gives a short overview of the method while a full description can be found in (Bernhardt, 2007).

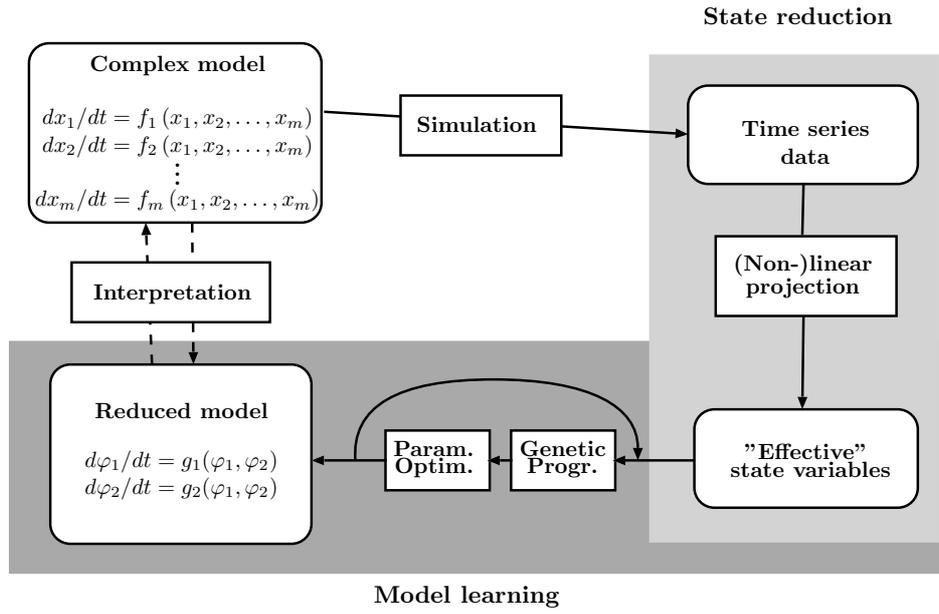


Figure 4.1: Schematic diagram of the model reduction procedure. See text for details.

4.3.1 State transformation and model learning

Starting with the m -dimensional ordinary differential equation (ODE) $dx/dt = \mathbf{f}(\mathbf{x})$, we produced time series data $\mathbf{X} = [\mathbf{x}(t_1), \dots, \mathbf{x}(t_n)]^T$ with n time steps (Appendix A.1). The construction of the new model is solely based on \mathbf{X} and is therefore independent from $\mathbf{f}(\mathbf{x})$. In a second step, Singular Value Decomposition (Golub and Van Loan, 1989) was applied to perform a Principal Component Analysis (PCA; Fukunaga and Koontz, 1970) of \mathbf{X} which provides a mapping to a low-dimensional representation Φ of the data matrix. The columns of Φ contain the component scores φ_i of the corresponding principal components (PC). Typically, only the first l PCs which explain a lower relative amount of data variance (e.g. 95%) are considered in the analysis, i.e. Φ is l -dimensional. Continuing the concept introduced by Wirtz and Eckhardt (1996), these first l PCs will be called "effective" variables.

In a third step, we used a modified Genetic Programming (GP) algorithm (Koza, 1992; Cao et al., 2000; Bernhardt, 2007) to generate new ODEs of the effective variables. In addition, a hybrid parameter optimization scheme consisting of a Genetic Algorithm and a gradient-based optimizer (Zhu et al., 1997; Pham and Karaboga, 2000; Goldberg and Voessner, 1999) was applied to adjust the free parameters of the new model systems. In the GP module, the right-hand sides of the differential equations are encoded as sets of tree structures. These trees are made up of interlinked nodes representing the transformed state variables, numerical constants and arithmetical operators of the new equations. A randomly initialized population of model trees is recombined and transformed by

evolutionary-inspired operators in the course of the algorithm. Figure 4.2 shows an example for a GP-tree representation of a simple ODE. The selection of an individual member for reproduction or transformation in every generation is based on its fitness. The fitness used is a combination of the root-mean-square error (RMSE) as goodness-of-fit measure and the number of GP nodes as a measure of model complexity. Both measures are combined using the weighted average ranking method (Bentley and Wakefield, 1997) with equal weights. The MAGER scheme thus produces new ODE models

$$\frac{d\varphi}{dt} = \mathbf{g}(\varphi), \quad (4.5)$$

with numerical solutions $\varphi_i(t)$ approximating the time series of the l effective variables, i.e. the columns of Φ .

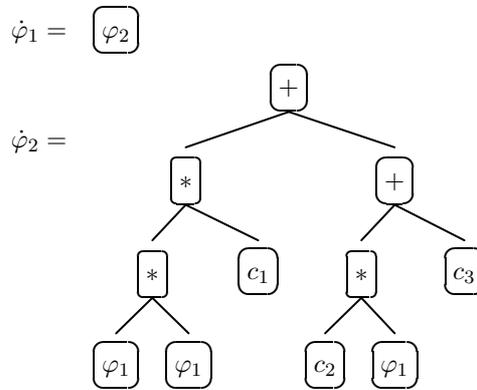


Figure 4.2: Example for a GP-tree representing the ODE $d\varphi_1/dt = \varphi_2$, $d\varphi_2/dt = c_1\varphi_1^2 + c_2\varphi_1 + c_3$.

4.3.2 Reduction of the model systems

Time series datasets of the predator-prey and NPZ models were generated by numerical integration of (4.1) and (4.2) using the parameterizations in Tables A.1 and A.2, respectively. As PCA is sensitive to the scale of variables, the principal components were calculated for the normalized variables \tilde{x}_1 , \tilde{x}_2 and \tilde{N} , \tilde{P} , \tilde{Z} with zero mean and a standard deviation of one. The dimensionality of the datasets was reduced by selecting only the first PC score $\varphi_1(t)$ for the normalized two-dimensional time series of the predator-prey model and the first two PC scores $\varphi_1(t)$, $\varphi_2(t)$ in case of the NPZ dataset. While the first two PCs of the NPZ model explain more than 99% of the data variance, only 50% of the variance can be explained with the predator-prey model's first PC and the reduction must therefore primarily be seen as an introductory example in this case (Bernhardt, 2007). Furthermore, as a first-order ODE with a single state variable is not able to produce oscillations, the predator-prey model

learning scheme was adopted to generate second-order systems

$$\frac{d\varphi_1}{dt} = \varphi_2, \quad \frac{d\varphi_2}{dt} = f(\varphi_1, \varphi_2), \quad (4.6)$$

where φ_2 indicates the time derivative of the effective variable φ_1 .

The GP algorithm was run individually for all datasets using five repetitions with different random initial populations and a maximum tree depth of three. It should be noted that due to the normalization the fitness of learned models becomes insensitive to the values range of the variables and their initial values may be chosen arbitrarily. However, as the choice of initial values of the normalized effective variables $\tilde{\varphi}_i$ may influence the model dynamics, they were introduced as additional model parameters subject to optimization.

Different from the reduction of the predator-prey system, it was further made certain that the effective variables of the NPZ model were positive for all time steps. Without any change of sign of the variables, the new terms of the model equations then have definite meanings as gain or loss terms and their biological interpretation is simplified.

4.4 Results and Discussion

4.4.1 Predator-prey model

Dimensionality reduction

Predator-prey oscillations in the two-dimensional state space of the original model together with the respective first PCs are shown in Figure 4.3. For both settings of the normalized inverse capacity, the first PC provides a projection of the time series onto one dimension explaining the largest percentage of variance. The two simulated oscillatory time series differ in main frequency and amplitude. A lower value of the normalized inverse capacity k permits an increased growth of prey promoting larger amplitudes and lower frequencies of the predator-prey oscillations and shortens the transient phase before the limit cycle is reached (Figure 4.4).

As the PC scores result from linear mappings of the data matrix onto a set of eigenvectors, the effective variables are simple linear combinations of the original variables

$$\varphi = p \cdot (\tilde{x}_2 - \tilde{x}_1), \quad \text{with } p = 0.71. \quad (4.7)$$

φ thus incorporates aspects of both, prey and predator population densities, and describes a negative potential of predation depending on the difference between normalized foragers and food mass. For positive φ , insufficient food availability and a large amount of predators limit trophic interaction while the biomass flux due to feeding grows with a lowering of φ which increases the abundance of predators. We will therefore use the term *feeding limitation* for φ as an abstract characterization of the

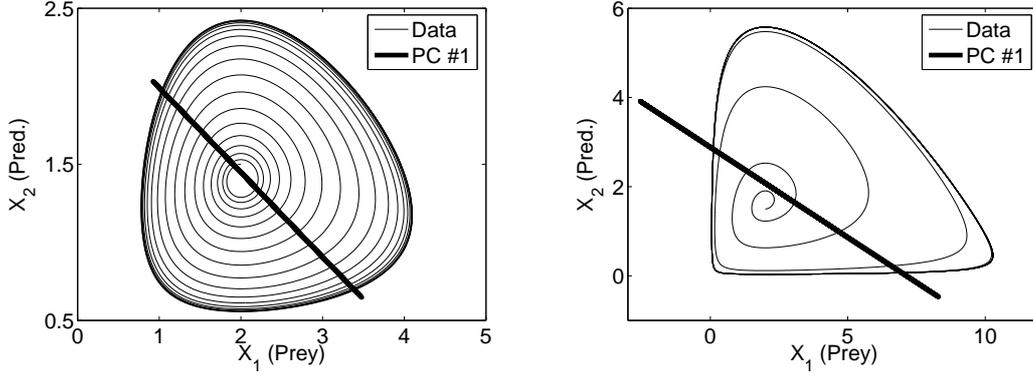


Figure 4.3: Model data of (4.1) using the simulation settings of Table A.1 together with respective first PCs in original state space for (a) $k = 0.15$ and (b) $k = 0.09$.

predator-prey interaction. In a qualitative sense, the feeding limitation corresponds to the predator-prey ratio introduced by Ginzburg (1986). It describes the *interaction* of functional groups, i.e. a structural property of the food chain rather than a biomass.

The predator-prey system as a linear oscillator

The occurrence of a single effective variable incorporating predator and prey densities in (4.1) gives rise to a second-order ODE (see below, section 4.3.2) resembling an oscillator equation. Independent from the black-box MR, an approximation of the latter can be derived analytically from (4.1). The result is used to analyze the models found with the MAGER scheme in section 4.4.1.

Consider a small deviation x'_i from the fixed point x_i^* of prey and predator. A Taylor series expansion of system (4.1) around x_i^* and subsequent introduction of the feeding limitation φ yields an oscillator equation (Appendix A.3)

$$\frac{d^2\varphi'}{dt^2} \approx -\omega^2 \cdot \varphi' - \rho \cdot \frac{d\varphi'}{dt}, \quad (4.8)$$

with the damping/excitation factor $\rho = (2 \cdot k)/(1 - s) - k - s$ and frequency of the undamped oscillator $\omega = \sqrt{m_p \cdot (1 - k - s)}$. Given the apparent analogy to physical systems, for example a damped/driven pendulum or a spring, the terms on the right hand side can be interpreted as mass normalized forces acting on trophic interactions.

Indeed, equation (4.8) is compatible with the "law of balance of biological momentum" (Vadasz and Vadasz, 2002; Ginzburg, 1986) which states that the "rate of change of biological momentum is balanced by the biological forces impressed on the population". The approximation in the present case yields two biological forces, namely the restoring force $-\omega^2 \cdot \varphi'$ and a driving or damping force proportional to the rate of change of feeding limitation. It should be noted that for strictly negative ρ

the oscillator dynamics explodes and the Taylor approximation loses validity. We can nevertheless interpret the original biological parameters in a physical context if we neglect the excitation term by setting $\rho = 0$. The resulting conservative oscillatory system then is completely determined by the frequency or period T , respectively, which follows

$$T = 2\pi \cdot \sqrt{s/(k \cdot m_p)}, \quad (4.9)$$

In terms of the original model parameters, $s/(k \cdot m_p)$ equals the product of capacity and maximum growth rate of prey divided by the product of predator mortality and half-saturation constant (see section A.2). The variability of the feeding regulation is thus given by the ratio of the growth potential of prey and the predators' growth inhibition. In case of a driving force present and large values of the half-saturation, i.e. $\rho < 0$ and $s < 1$, we can further see that excitation is low for small values of the capacity, i.e. when the environmental limits dominate the growth of prey. In conclusion, we find that the biological interactions of different populations produce driving and damping forces in analogy to those observed in physical systems. Working with physical analogies in this context offers a way of understanding the dynamics of the new state variable φ .

However, the validity of the analytical model transformation is limited by the first-order approximation as mentioned above. In particular, the deviation between the calculated and measured period of the oscillations increases for lower values of k , i.e. stronger excitation of the system. In the present case this yields period values of 11.5 vs. 14.8 for $k = 0.15$ and 11.1 vs. 17.1 for $k = 0.09$. The period was calculated according to (4.9) and measured using fast Fourier transform, respectively.

Algorithmic learning of reduced models

Without the shortcomings of the analytical method, the MAGER scheme independently found similar reduced model formulations of the predator-prey system. The best models found, i.e. the least complex sets of equations able to reproduce the limit cycle and transient behavior of (4.1) for $k = 0.15$ and $k = 0.09$, are equal in structure and can be written as

$$\frac{d\varphi_1}{dt} = \varphi_2, \quad \frac{d\varphi_2}{dt} = \frac{-c_1\varphi_1 + c_2\varphi_2}{c_3 - \varphi_1 - c_4\varphi_2} \quad (4.10)$$

with parameter values c_i listed in Table A.3.

While this model is very well able to reproduce the frequency of the respective oscillations and also the major characteristics of the transients, the irregularity of the oscillations for $k = 0.09$ could not be captured in detail (Figure 4.4). Nevertheless, given the overall accuracy of data reproduction for (4.10), we propose an interpretation which is based on the results of the analytical reduction procedure.

It is evident that model (4.10) describes a simple nonlinear oscillator analogous to (4.8). If we set $c_2 = 0$ in (4.10), the ratio $c_1/(c_3 - \varphi_1 - c_4\varphi_2)$ in this model is equivalent to ω^2 in (4.8). The ratio of predator growth inhibition and the growth potential of prey is no longer a constant term but depends

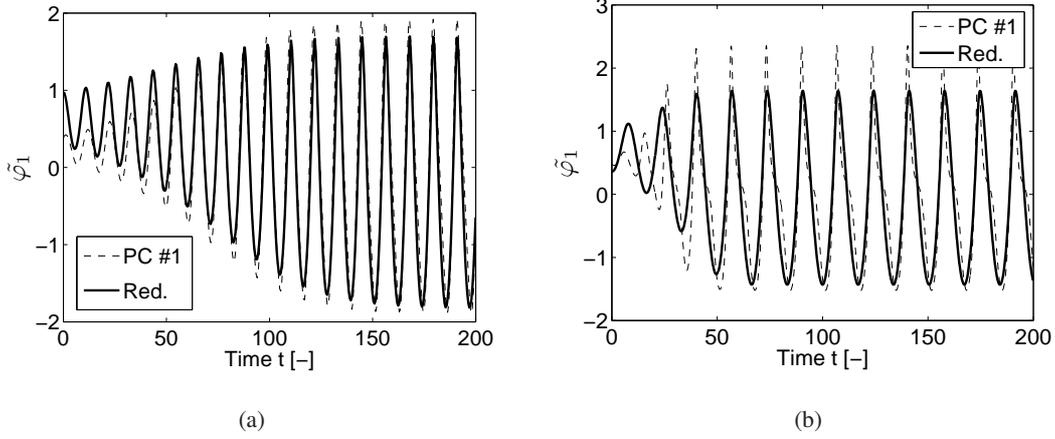


Figure 4.4: Time series produced by the best reduced model (Red.) found for the first principal components (PC #1) of the 2D predator-prey model data with (a) $k = 0.15$ and (b) $k = 0.09$. The RMSE values of the reduced model are 0.32 for (a) and 0.51 for (b).

on φ_1 and φ_2 . More generally, the damping effect exerted by the nonlinear dependency of $d\varphi_2/dt$ on φ_1 and φ_2 plays a central role in balancing the driving force and facilitates the limit cycle dynamics. Thus, the algorithmic reduction is able to reveal a more complex picture of the acting forces compared to the analytical derivation.

4.4.2 NPZ model

Dimensionality reduction and model learning

In case of the chemostat model (4.2) the best reduced two-dimensional system found with the MAGER scheme reads

$$\frac{d\varphi_1}{dt} = -\varphi_2 (c_1 + c_2\varphi_2). \quad (4.11)$$

$$\frac{d\varphi_2}{dt} = \varphi_1/\varphi_2 + c_3\varphi_2 + c_4 \quad (4.12)$$

See Table A.3 for parameter values of (4.11) and (4.12). Apart from small deviations between the transients as well as the amplitudes of the oscillations, the reduced model (4.11)-(4.12) is very well able to reproduce the original model dynamics and in particular the cycle frequency (Figure 4.5).

Following the discussion in section 4.4.1 we write the effective variables of the NPZ model as linear combinations of the original variables. Without much loss of accuracy we further ignore terms with small coefficients and change the sign of φ_2 to simplify the interpretation which results in

$$\varphi_1 \approx a_3\tilde{Z} - a_2\tilde{P}, \quad (4.13)$$

$$\varphi_2 \approx b_2\tilde{P} - b_1\tilde{N} \quad (4.14)$$

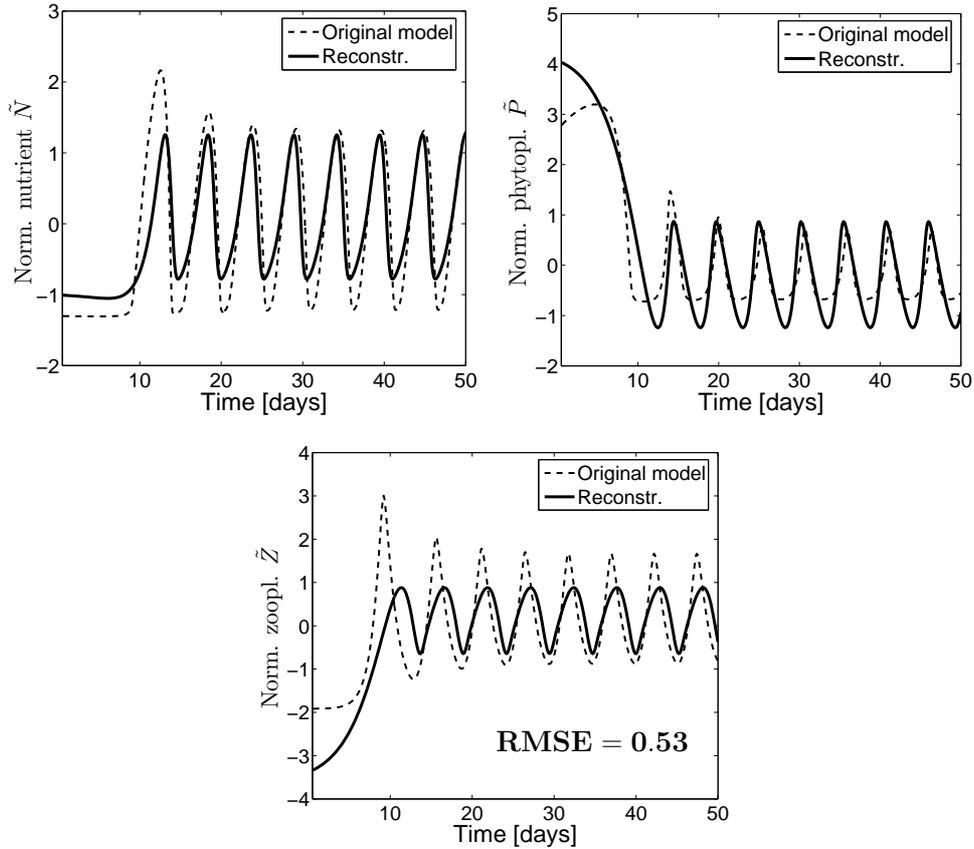


Figure 4.5: Reconstruction of the chemostat time series using model (4.11)-(4.12) in comparison with the normalized simulation data produced by the original NPZ model. The RMSE value of the reduced model is 0.53.

with

$$\begin{aligned} a_3 &= 0.85, & a_2 &= 0.46 \\ b_2 &= 0.66, & b_1 &= 0.74. \end{aligned}$$

We again interpret φ_1 and φ_2 in (4.13) and (4.14) as interaction variables, namely feeding limitations acting as growth inhibitors for predator Z and consumer P in the two oscillatory subsystems, respectively. The formal representation of the dynamics accordingly changes from density-based direct predator-prey and consumer-resource interactions to functions including indirect interactions by the incorporation of structural aspects of the food chain.

Combined interaction- and density-related interpretation

The right hand sides of equations (4.11) and (4.12) directly describe indirect interactions of trophic subsystems. As the learning process (section 4.2.2) was constrained to produce positive effective

variables φ_1 and φ_2 only, the change of sign (section 4.4.2) results in negative φ_2 . Thus, term *I* in (4.11) is a threshold dependent gain or loss term for zooplankton grazing. As long as $c_1 > |c_2\varphi_2|$, a lowering of the phytoplankton density increases the feeding limitation of zooplankton. When φ_2 is too low, however, the reduced supply of food is not sufficient to maintain the zooplankton population and the subsequent death of individuals causes the feeding limitation of the second subsystem to drop. Term *II* in (4.11) is a loss term attributed to the mass distribution between the two trophic subsystems. A strong feeding limitation in the second subsystem, i.e. high values of φ_1 , indicates a biomass transfer from phyto- to zooplankton and an accompanied reduced competition for nutrients on the phytoplankton level. In combination with a high resource availability, i.e. low φ_2 , this leads to good growth conditions for phytoplankton. The negative term *III* then denotes an excitation term induced by nutrient supply. An increasing supply of nutrients in the chemostat system causes favorable growth conditions of phytoplankton and lower values of the feeding limitation but at the same time induces strong density fluctuations eventually leading to extinction, an effect which is also known from the paradox of enrichment. Term *IV* finally antagonizes these loss terms as it describes a feeding-induced constant biomass flow from nutrients to phytoplankton which increases the growth limitation of phytoplankton.

The mechanistic interpretation of the terms substantiates our notion of interaction variables. *Direct* relationships between the interaction variables in terms *I* and *II* can be translated to *indirect* interactions of the density-based original formulations.

In addition to the biological interpretation, the structure of (4.11) and (4.12) again allows for a physical interpretation via a comparison with (4.10) and (4.8). If we understand term *I* as a nonlinear extension of the relation between $d\varphi_1/dt$ and φ_2 in (4.10) we find that the negative term *II* corresponds to the repelling term $-\omega^2\varphi$ in (4.8) and the dependency on φ_2 is similar to the nonlinear damping effect discussed in section 4.4.1. In the same manner, term *III* corresponds to the physical damping/excitation term $-\rho^2 d\varphi/dt$. We can identify another indication for a similarity between this model and a simple linear oscillator by approximately calculating the period of the oscillations (section 4.4.1). In the present case this yields a period of 4.4 which is close to 5.3, the value obtained using fast Fourier transform. It is again obvious, however, that the simple first-order approximation only partially explains the more complex dynamic interactions of the biological system.

4.4.3 Ecological Interaction Models

The reduced models produced by MAGER may be seen as a new model type for simulating species assemblages in biology. For simplicity, the term "species" is used substitutionally for the constituents of food webs, i.e. functional groups as well as nutrients and other resources, in the following discussion. The new approach of describing species dynamics can best be understood by comparison with the traditional approaches of density- and trait-based models (Table 4.1). Density models, e.g. NPZ

or consumer-resource systems, formalize the direct (physical) effects of interactions on the masses or densities of the interacting species/populations. Indirect effects are not described explicitly in these models and emerge from the specific food web structure. Behavioral aspects in the form of species traits, e.g. prey search rates and handling times of predators, are then partially encoded in the model parameters (Real, 1977). The criticism on the treatment of these species traits as constants in the model equations led to model variants where traits variables, such as the diet choice of generalist predators or food qualities, were introduced as dynamic quantities (see e.g. Abrams, 1995; Abrams et al., 1996; Relyea and Yurewicz, 2002; Yamauchi and Yamamura, 2005). In addition to the formulations used in density models, the functional relationships of the state variables in these trait models also account for behavioral adaptations caused by direct species interactions.

The state variables of the reduced models derived with MAGER are combinations of the original densities. Because of their interpretation as feeding limitations, i.e. quantities describing predator-prey and consumer-resource *interactions*, we call this new model type Ecological Interactions Models (EIM) in the following. While density and trait models only incorporate direct interactions, the functional relationships between the interaction variables of EIMs explicitly capture indirect effects as well (see e.g. terms I and II in (4.11) and (4.12)). The state variables of the EIMs found in this study are directly connected to the density-based formulations. However, the observed dynamics of measured data and model simulations may also be caused by a combination of density- and trait-related effects (Křivan and Schmitz, 2004). As the construction of EIM is only based on time series data, we propose that the functionals of the EIM equations may also incorporate the impacts of changing traits. In order to show that it is possible to incorporate trait-density interactions in EIMs, the MAGER approach should be applied to datasets consisting of both, trait and density variables. It should finally be noted that the EIM parameters are functions of the original model variables and parameters which complicates a process-based interpretation. For small reduced systems, however, analogies to known process descriptions suggest a specific meaning of the parameters (see for example the oscillator discussion in section 4.4.1).

4.4.4 Generality of the reduced models

It may be argued that the reduction potential of the MAGER approach is diminished by its adaptation to simulated data. An application of the method to different datasets obtained by parameter variation of the same model in fact produces different reduced models. A collection of simplified models is then needed to reproduce all dynamic regimes of a complex system which may be seen as a drawback of the approach.

However, complex models often serve too many purposes simultaneously (e.g. Lee, 1973). As a result, the complete set of dynamic situations which can be produced with these models, such as limit cycles or chaos, must not necessarily be observable in measured data. Thus, the usefulness of the

Table 4.1: Comparison of density-, trait-based and Ecological Interaction Models.

	Density models	Trait models	EIM
State variables	Mass/density of individual species	Species densities, traits explicitly given	Species interactions (e.g. feeding limitation), combinations of density variables
Parameters	Species- and trait-related factors. E.g. consumption and search rates, handling times	See density models	Functions of density model parameters and variables. Interpretation by analogies
Functional relationships	Direct predator-prey or consumer-resource interactions	Direct interactions and trait adaptations	Direct and indirect species interactions

reproduction of a model's complete bifurcation behavior is questionable. An example is the ongoing discussion of whether or not the paradox of enrichment resulting from studies of simple models is present in experimental data (e.g. Persson et al., 2001; Roy and Chattopadhyay, 2007).

Furthermore, even if the reduced models are typically not able to reproduce the complete spectrum of the original system's dynamic responses, they can nevertheless be fairly general in reproducing different time series within a specific dynamic regime and thus represent prototypes for a broader class of models showing similar dynamics. In order to assess the validity and generalism of the reduced models, we analyze the reproduction capabilities of model (4.10) for different parameterizations of (4.1) with time series showing limit cycle dynamics. This regime contains oscillatory dynamics of the trajectories as well as the existence of an unstable positive fixed-point which is determined by the values of the normalized inverse capacity k and the normalized inverse half saturation s . The corresponding regime boundaries are given by

$$\max(p - \sqrt{p^2 + q}, 0) < k < s \cdot r, \quad (4.15)$$

with

$$p = r(s - 2 \cdot m_p \cdot r), \quad q = r^2(4 \cdot m_p(1 - s) - s^2) \quad \text{and} \quad r = \frac{1 - s}{1 + s}.$$

See (Rosenzweig and McArthur, 1963; Freedman, 1980; Kot, 2001) for details of the stability analysis. While $m_p = 1$ was unchanged, we sampled random parameter sets for s between 0 and 1 and k

between the respective critical values given in (4.15). PCA was then used to produce the reduced datasets from the corresponding time series of (4.1).

Subsequent individual fits of model (4.10) to these datasets using the hybrid parameter optimization scheme of section 4.3.1 show that the single reduced model is able to reproduce most of the new time series to a high degree (Figure 4.6). In most cases, the RMSE values are similar to those obtained for the parameterizations used for model reduction (Table A.1) or lower, i.e. $\text{RMSE} < 0.5$. A value of $\text{RMSE} \approx 1$ thereby corresponds to the approximation of the mean value of the normalized time series, whereas model fits with $\text{RMSE} \approx 0.7$ already reproduce the frequency of the oscillations to a large degree. RMSE values below 0.5 finally indicate close approximations of the transient as well as the oscillatory regime. It should be noted that the larger errors occur for time series approaching the original system's fixed point at the origin. The reduced model, which has only one fixed point, is not able to reproduce the corresponding asymmetric oscillations.

In conclusion, the analysis confirms that the simple reduced formulations found with MAGER are fairly general models for distinct dynamic regimes.

4.5 Conclusion

In this paper we used MR of simple food web models to identify structural explanations for the occurrence of predator-prey and consumer-resource cycles. It should be noted that "clean" oscillations of population densities are rarely observed in measured time series, eventually due to irregularities in the external influences. For some of the more prominent exceptions of regular population cycles in nature see (Turchin, 2003). In controlled systems, the oscillations occur under rather artificial conditions (e.g. Fussmann et al., 2000). However, as the primary interest of our study lay in the search for alternative explanations of model dynamics, the problem of data and measurement uncertainties was ignored in this paper.

Despite the simplicity of the underlying PZ or NPZ formulations, the new scheme was able to find models with an even simpler structure concerning the number of state variables and parameters. The dimensionality of a predator-prey system (section 4.2.1) was reduced from $2(6)$, where the six in brackets denotes the number of parameters and two is the number of state variables, to $1(4)$. However, as a normalized model variant was used with a dimensionality of $2(3)$ the reduction in this case must rather be seen as a model transformation. The NPZ model's complexity (section 4.2.2) could be reduced from $3(11)$ to $2(4)$. Even if the optimized initial values of the state variables are taken into account as further model parameters, this shows the remarkable reduction potential of MAGER even for simple models. The reduction of the models was made possible by the emergence of interaction-related state variables describing the dynamics of the biological systems more efficiently. In all cases the new variables describe the combined consumer-resource interactions of the oscillatory subsystems

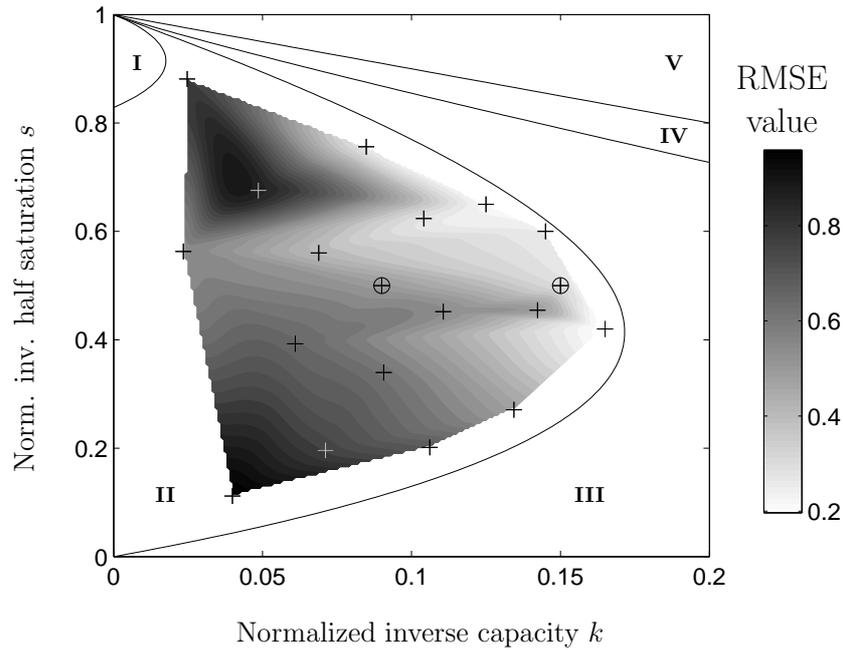


Figure 4.6: RMSE values of (4.10) for the approximation of reduced datasets derived from randomly sampled parameterizations of model (4.1) which all produce limit cycles (+). The parameterizations used for model reduction in section 4.3.2 are marked with circles. Solid lines indicate the boundaries of the system's different stability regions. (I) The fixed point is unstable, trajectories leave its vicinity exponentially, (II) unstable fixed point with oscillatory trajectories reaching a stable limit cycle, (III) stable fixed point and oscillatory trajectories, (IV) stable fixed point which is approached exponentially, (V) no equilibrium exists where predator and prey populations are able to coexist.

in form of feeding limitations of the respective consumer populations. They are explicitly related to the original variables and represented structural food web properties which are determined by the respective trophic interdependencies. The definition of the effective variables further combines the effects of direct and indirect interactions. Terms describing direct interactions between the structural property variables of the reduced chemostat model could be interpreted as indirect interactions of the original density-based variables. In addition, for two and three trophic levels, the comparison with a simple physical oscillator led to the notion of biological driving forces as negative functions of these feeding limitations.

As the model learning was based on pure density variables, there is no formal connection between the effective interaction variables generated by the MAGER scheme and dynamically regulated traits used in population models such as edibility or defense (e.g. Abrams, 1995; Wirtz and Eckhardt, 1996; Yamauchi and Yamamura, 2005). However, the incorporation of such "classical traits" as unobserved quantities in the model learning process is possible and a subject of future work.

It may further be objected in this context that the linear relationships between interaction variables and original state variables resulting from linear PCA restrict the general applicability of the MAGER scheme. For example, traditional traits are often defined as nonlinear functions of density variables which could not be extracted by PCA (e.g. Bruggeman and Kooijman, 2007). However, the use of PCA is not mandatory for the proposed approach. The reduction method has in fact been extended to use nonlinear schemes (Bernhardt, 2007) which may account for apparent limitations of the present approach. Most importantly, we have also shown that even linear combinations of biomasses may result in effective variables which carry a higher information content concerning biological interactions than pure densities.

It was finally demonstrated that the MAGER models represent general formulations of the processes governing a broader dynamic regime even though the learning scheme was trained on single time series. This indicates the potential of the method to search for general principles underlying population dynamics. We propose that the use of structural interaction variables linking density- and trait-related effects can effectively reduce more complicated biological models and may therefore represent a rationale of a new generation of ecosystem models. Applications of the method to data obtained from either measurements or more complex models will be subject to further work.

Chapter 5

Explicit formulation of indirect interactions in a reduced consumer-resource competition model. Part I: Reducing model complexity

Abstract

Theoretical biology has always been concerned with the complexity of biological systems and the appropriate level of detail of their mathematical descriptions. The incorporation of multi-layered interactions and feedbacks between members of a food web, for example, quickly leads to a complicated superposition of direct and indirect effects governing model dynamics. However, it has long been proposed that the main governing processes of a system can be deduced from the most simple model able to reproduce its dynamics.

In this paper, we introduce the first steps to extract, in a condensed form, the constitutive interaction pathways generating the dynamics of a complex array of species competing for essential resources. We use the recently developed Mapping-based Genetic Reduction method (MAGER) to significantly reduce a complex consumer-resource model while preserving the main aspects of its dynamics. A comparison with two other reduction methods, which are based on the omission and aggregation of state variables, respectively, demonstrates the high reduction efficiency of the MAGER approach and the reproduction quality of the simplified models. The biological interpretation of the reduced models given in a companion paper leads to the notion of Ecological Interaction Models which constitute a new class of simply structured models for multi-level and multi-species consumer-resource interactions.

5.1 Introduction

The question of appropriate complexity is essential for model building in biology and environmental sciences. It arises particularly in the research on natural systems as these are characterized by a highly complex interaction network of biological species and environmental forcing. A principal guideline for model building, known as the parsimony principle or Occam's razor, which addresses the problem of complexity may be traced back to sources in the Middle Ages and to ideas of Aristotle (Rodríguez-Fernández, 1999). It states that a simple theory should be used in preference to a more complex one if it is equally well suited to reproduce a given dataset or explain a certain problem. In the present context, we define model complexity as the number of state variables and parameters.

Given a fixed amount of data, lower model complexity leads to lower parameter uncertainty. In addition, it is assumed that the simplest description of a system incorporates only the most important processes and may, in some cases, even point to general underlying principles. However, these principles are often not known or may even not exist so that model development for biological systems typically is a bottom-up approach which starts with simple conceptual models incorporating the processes which are assumed to be most relevant. Further constituents, such as additional functional groups, interactions paths and functional dependencies, are then added in a phenomenological way to increase the model's realism and to provide better data approximations. The resulting models often are very complex with tens or hundreds of state variables and parameters and are, thus, difficult to use and understand. Examples are the European Regional Seas Ecosystem Model (ERSEM, Baretta et al., 1995; Köhler and Wirtz, 2002) and the Gypsy Moth Life System Model (GMLSM, Sharov and Colbert, 1994).

Starting with a complex model, the search for a minimal description able to reproduce the system dynamics now translates to a model *reduction* problem. Different reduction approaches exist which may be classified in *a priori* and *a posteriori* techniques. In *a priori* model reduction, the simplification takes place in the course of model building. In the context of population ecology, this encompasses e.g. the aggregation of individuals into populations, the distinction of functional groups and the selection of relevant interaction pathways (e.g. Fulton et al., 2003). However, the decision on which processes to simplify remains arbitrary and modeling must be performed on a relatively high level of abstraction. Thus, good knowledge of the system and a principal idea of the most relevant processes are needed beforehand. However, as synergies, adaptations and feedbacks in bio-systems obscure the optimal level of model complexity, this knowledge is often lacking. It has, for example, been shown experimentally for a structured predator-prey system, that three predator species with differing hunting strategies and habitat uses may be aggregated into a single functional unit due to their average effect on a shared prey (Sokol-Hessner and Schmitz, 2002). In general, it is easier to increase model complexity than to identify the appropriate level of complexity in advance.

This directly leads to the top-down *a posteriori* reduction methods which are used to simplify complex models while approximately preserving their dynamic behavior. The different approaches in this class will further be categorized in methods based on "educated" simplification, omission of state variables and aggregation.

Educated simplification is similar to *a priori* reduction as far as it is based on the reformulation of existing model processes (see e.g. Van Nes and Scheffer, 2005, for an overview). Brooks et al. (2001), for example, used sensitivity and structure analysis of a wheat yield prediction model to identify the most sensitive parameters and variables. The model was then reduced using, for example, mean values instead of dynamic variables as well as simplified functional relationships and tradeoff terms. Because of the similarity to *a priori* reduction, this method also inherits the former method's

drawbacks. The simplification can only be performed with a thorough knowledge of the system's processes. In addition, it also depends on the knowledge of the complete model formulation which is often not documented in detail.

Another *a posteriori* reduction method is the omission or replacement of state variables. For example, Pahl-Wostl (1997) derived simplified versions of a pelagic food web model by omitting a number of species on each trophic level. The main advantage of this method is that it can be automated. Cox et al. (2006) randomly replaced state variables of a model with their respective mean values. In a subsequent model selection step the most appropriate model was then chosen on the basis of the tradeoff in fitting capabilities and model complexity. For the approximate simulation of steady states, the omission of state variables by replacement with their respective equilibrium constants was already used earlier (McArthur, 1972). However, the use of simple constants instead of variables is typically not possible for dynamic systems which are not in steady state.

The third approach in this context is reduction by aggregation of state variables according to either time scale separation or functional relationships. Time scale separation is also known as singular perturbation (Nayfeh, 1973; May, 1977). This method relies on the existence of different time scales and a stable manifold, e.g. a fixed point or a limit cycle, which is quickly reached by a subset of the variables. Aggregated macro variables, which are constants of motion of these so-called fast variables, are then introduced and the system is rewritten in terms of the aggregated states. The method has been used in spatially structured population models (Auger and Poggiale, 1996b, 1998) and simple food chain models (Kooi et al., 1998; Rinaldi and Scheffer, 2000). Macro variables used in (Auger and Poggiale, 1996b), for example, are the total population densities of different spatial patches which are made up of sets of subpopulations.

Aggregation according to functional relationships is the basis of the Effective Variable Approximation method (EVA) (Wirtz and Eckhardt, 1996). This approach introduces mean functional traits as aggregated variables with corresponding differential equations derived from the original model formulation. The reduction potential of EVA is based on the quantification of tradeoffs in parameter space which are used to reduce the number of the new functional traits.

The major drawback of the aggregation approaches discussed so far is their dependency on specific properties of the complex models. In the case of singular perturbation, this relates to the existence of separate temporal scales of the model dynamics which may not be present. Similarly, the applicability of the EVA approach is limited if the tradeoffs in parameter space which uniquely relate the functional traits to each other are missing. Furthermore, the ecological models reduced by aggregation may, in principle, deviate dynamically from the originals (Schaffer, 1981). In the singular perturbation based approach, for example, a limit cycle of the fast variables is approximated with the respective temporal mean value. Additionally, all reduction methods discussed have the common disadvantage of a high analytical effort and, apart from the random replacement of model parts with constant val-

ues, an automation of the methods is impossible as the details of reduction are specific for each model.

In this paper, we demonstrate the application of the Mapping-based Genetic Reduction method (MAGER, Bernhardt, 2007), a recently developed model reduction technique which eliminates some of the shortcomings of the aforementioned approaches. It proceeds automatically, directly aims at reproducing the original dynamics, has a high reduction performance and produces process-based models which may be interpretable in system-specific terms. MAGER is a black-box procedure which can be applied to ordinary differential equation (ODE) models. The method's ability to reproduce irregular data results from its data-adaptive nature. It only depends on model-generated or measured data and no knowledge or analysis of the original equations is needed. The reduction performance and general applicability of the method derives from its independence from the former model structure. The reduced systems are not based on traditional formulations and can typically not be derived analytically (Bernhardt and Wirtz, 2007). The new structures may thus give new insights into alternative process formulations.

We use the MAGER scheme to reduce a prominent consumer-resource (CR) model with a number of phytoplankton species competing for essential resources. CR models play a central role in environmental modeling as important building blocks of complex ecosystem models. In the context of the parsimony principle, CR models are good candidates for testing data-driven model reduction as their dynamics is the result of a complex mixture of direct and indirect interactions between the species. This first part of the paper introduces the technical aspects and demonstrates the performance of the reduction procedure. The process-oriented biological interpretation of the reduced models and the associated identification of the system's key processes will be given in the second part of the paper.

5.2 Model system and data

Starting point of the study is a well-established consumer-resource competition model (Léon and Tumpson, 1975; Tilman, 1982; Huisman and Weissing, 2001b) with n_P phytoplankton species and n_N nutrients. Let P_i and N_j be the abundance of species i and the availability of resource j , respectively. Net growth of species i is given by

$$\frac{dP_i}{dt} = P_i \cdot (\mu_i - \omega_i) \quad i = 1, \dots, n_P \quad (5.1)$$

with mortality ω_i and gross production μ_i which is determined by the most limiting resource according to

$$\mu_i = \min_j \left(\frac{g_i N_j}{k_{ji} + N_j} \right) \quad j = 1, \dots, n_N \quad (5.2)$$

with k_{ji} denoting the half-saturation constant for resource j of species i and g_i the maximal growth rate. The time evolution of the abiotic resource j is controlled by a supply term and resource con-

sumption,

$$\frac{dN_j}{dt} = D \cdot (S_j - N_j) - \sum_i c_{ij} \cdot \mu_i \cdot P_i \quad j = 1, \dots, n_N \quad (5.3)$$

where D describes the nutrient turnover rate, S_j is the supply concentration and c_{ij} quantifies the content of nutrient j in species i .

With the parameterizations used in (Huisman and Weissing, 2001b) for $n_P = 5$ and $n_N = 3$ the model is able to generate oscillatory dynamics of the coexisting species with transient regimes of different length. Depending on the initial conditions, two different sets of dominant species emerge characterizing two distinct oscillatory. We generate two eight-dimensional time series ($n = n_P + n_N$), each corresponding to one of the oscillatory states (Figure 5.1) using a common set of parameters but different initial values (Appendix B.1). The matrix \mathbf{X}^a , with time series of the n variables $x_j(t_h)$ ($j = 1, \dots, n; x_1(t_h) = P_1(t_h), \dots, x_8(t_h) = N_3(t_h)$) in columns and m time steps ($t_h = 0, 0.1, \dots, 200$), denotes the resulting dataset showing a dominance of P_1, P_4 and P_5 and \mathbf{X}^b is characterized by the dominance of P_1, P_2 and P_3 . The data matrices serve as individual inputs for the model reduction procedure. The time course of the correlation dimension D_C (8.13) shows that the dimensionality of the time series changes in time and two different regimes can be found (Figure 5.1(b)). For \mathbf{X}^a , the dimensionality of the time series drops from a value near 3 for the mixture of transient and limit cycle regime to $D_C \approx 2$ for the limit cycle alone starting at $t \approx 80$. The same separation of transient and limit cycle regime according to dimensionality is found for \mathbf{X}^b (not shown). In order to study how the complexity of the dynamics affects the model reduction results, the subsequent analysis is performed for the complete as well as the transient-free datasets. The complete datasets will be denoted by the corresponding lower case letters ($\mathbf{X}^a, \mathbf{X}^b$) whereas upper case letters will be used for the transient-free datasets ($\mathbf{X}^A, \mathbf{X}^B$). Figure 5.2 shows the individual steps of the reduction scheme.

5.3 Data preparation

We employ Principal Component Analysis (PCA) for dimensionality reduction of the competition model. As the values of the different state variables strongly vary in scale, we normalize the transient-free datasets to zero mean and unit variance before PCA is applied (case S in Figure 5.2). Without normalization, the dominant species have an exceeding impact on the PCA results in comparison with the inferior species and resources. In case of the complete datasets, the variables drop to values close to zero after a short transient regime (Figure 5.1), so that normalization of the variables artificially enhances the importance of the transient compared to the limit cycle dynamics (case s in Figure 5.2). Therefore, the variables are only centered by subtracting the mean values in this case.

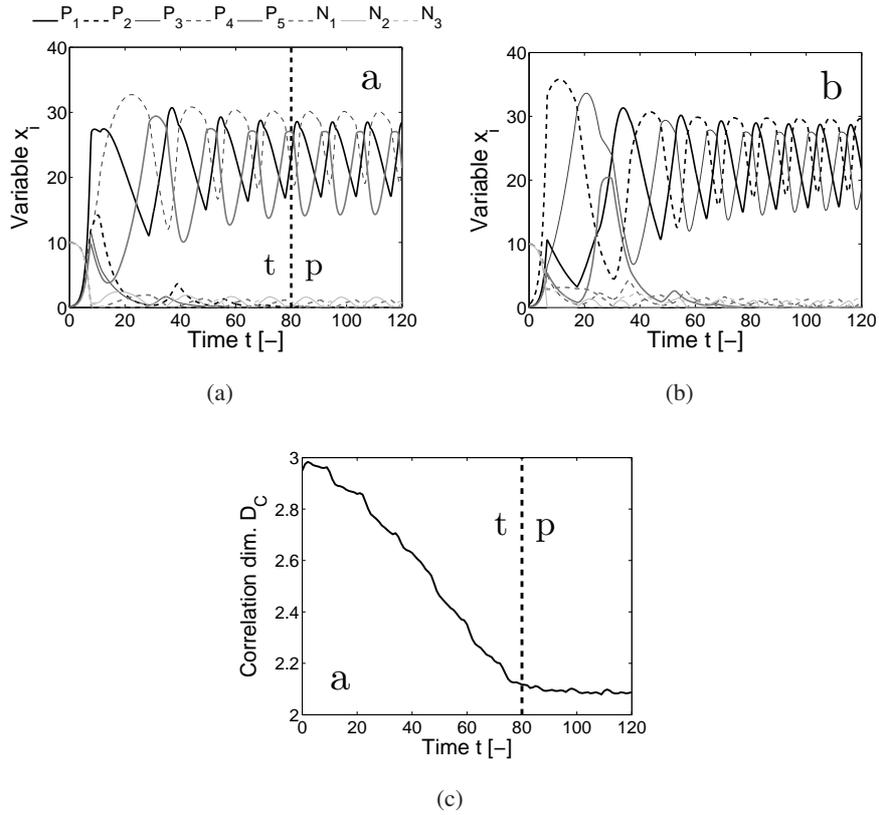


Figure 5.1: Datasets (a) \mathbf{X}^a and (b) \mathbf{X}^b showing the CR model's two different oscillatory regimes. The dominant species in (a) are P_1, P_4 and P_5 , whereas P_1, P_2 and P_3 dominate in (b). The time course of the correlation dimension (Appendix B.3) for \mathbf{X}^a is given in (c). The transient regime is marked with "t" whereas "p" indicates the periodic oscillatory regime.

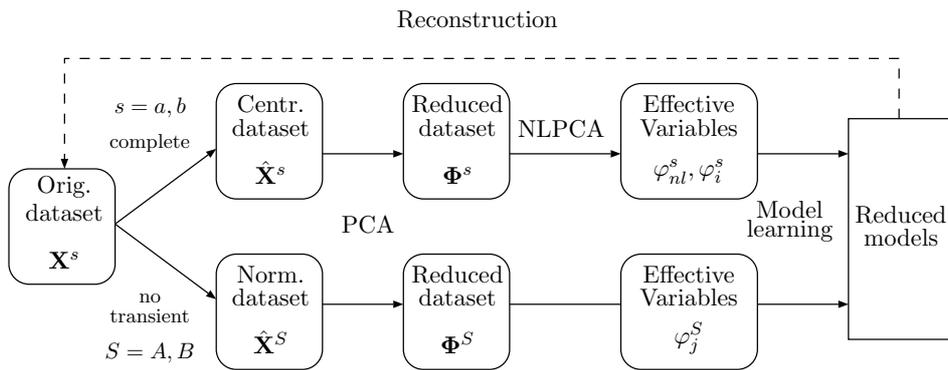


Figure 5.2: Overview of the model reduction steps and notation used. See text for details.

5.4 Linear dimensionality reduction

A PCA extracts principal components (PC) which are used to reconstruct a part of the variability inherent to the data (Pearson, 1901; Golub and Van Loan, 1989). The number of retained PCs determines the quality of approximation. We use a critical value of the fraction of explained variance ($FEV_c \approx 90\%$) that is accounted for by a subset of all principal components to determine this number (Appendix B.3).

The separation of transient and limit cycle regimes according to the D_C values (section 5.2) is supported by the PCA results: the FEV_c criterion is met by the first three PCs in cases a and b . A further linear reduction to two dimensions is not possible as it causes overlaps of the trajectories in phase space. However, if only the long-term dynamics of the limit cycles is considered, i.e. by dropping the transient regimes, the first two PCs are sufficient to comply with FEV_c in both cases. In conclusion, the limit cycle dynamics of the system is essentially two-dimensional and the analysis is performed for the PCA reduced data matrices of the complete ($\Phi^s \in \mathbb{R}^{m \times 3}$ with $s = a, b$) as well as the transient-free ($\Phi^S \in \mathbb{R}^{m \times 2}$ with $S = A, B$) datasets. The individual PCs, with time series given in the columns of the reduced datasets, are denoted by the respective lowercase letters, i.e. φ_i^s with $i = 1, 2, 3$ as well as φ_j^S with $j = 1, 2$. Following Bernhardt (2007), the principal components will be called "effective variables".

5.5 Nonlinear dimensionality reduction

The higher the dimensionality of the dataset, the more difficult is the learning of an appropriate model structure. In fact, no simple models incorporating the three effective state variables extracted with PCA for the two datasets could be found with the MAGER scheme. However, if the PCA reduction results are plotted against the corresponding two-dimensional projections of the time series (Figure 5.3) we can see that the two-dimensional data of PCs #1 and #2 of Φ^a and PCs #2 and #3 of Φ^b can be approximated with a one-dimensional curve. Thus, the dimensionality of the PCA transformed datasets can further be reduced by one with a nonlinear extension of PCA. We apply a nonlinear Principal Component Analysis (NLPCA) scheme based on an auto-associative neural network (Hsieh, 2001; Kramer, 1991) which facilitates the reconstruction of the dataset by inverse mapping from the NLPCs. The combination of the reconstructed linear and nonlinear PCs is then compared with the original time series data. See Appendix B.3 for the definition of FEV in the nonlinear case.

Thus, the first two PCs of Φ^a are replaced by the resulting nonlinear principal component, φ_{nl}^a . The combination of this nonlinear effective variable and the third linear PC still complies with the FEV_c criterion. For Φ^b , the NLPC is constructed from the second and third PC.

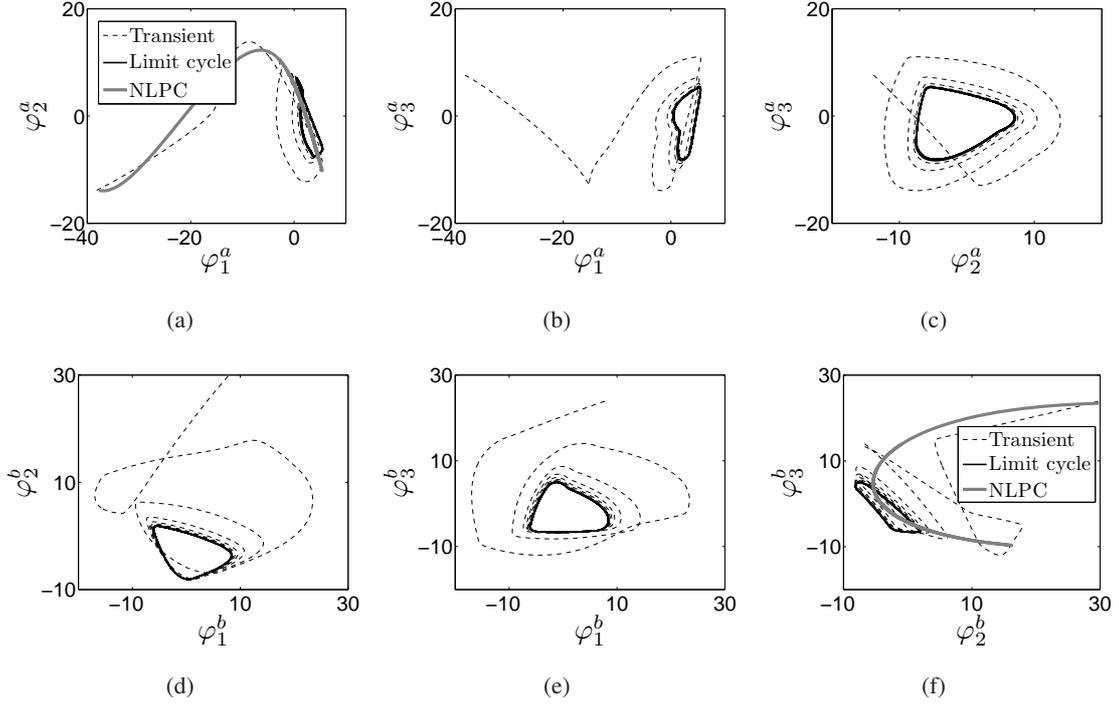


Figure 5.3: (a)-(c) The three different projections of the lasso-shaped reduced time series data Φ^a onto pairs of the principal components φ_i . (d)-(f) The same depictions for Φ^b . The transient (denoted by "t" in Figure 5.1) is drawn with dashed lines whereas the oscillatory regime ("p" in Figure 5.1) is indicated by solid lines. Figures (a) and (f) also show the learned nonlinear principal components as bold gray lines (see text for details).

5.6 Model learning and reconstruction

In the following, we give a short summary of the genetic model learning steps of MAGER while further details can be found in (Bernhardt, 2007). After the identification of effective variables (sections 5.4 and 5.5), a modified Genetic Programming (GP) algorithm in combination with a hybrid parameter optimization scheme consisting of a Genetic Algorithm (GA) and a gradient-based optimizer is used to generate new model structures. The right-hand sides of the new differential equations are thereby encoded as sets of tree structures in the GP module. These trees are made up of interlinked nodes representing the effective state variables, numerical constants and arithmetical operators of the new equations. A random initial population of the model trees is recombined and transformed by evolutionary-inspired operators in the course of the algorithm. The GA module is used to optimize the parameters of the new models and the initial values of the effective state variables. The selection and transformation of the best members is then based on their fitness, i.e. their ability to reproduce the reduced time series. Here, we use a combined measure which is sensitive to the course of the

transient as well as the amplitude of the oscillations of the limit cycle regime (Appendix B.3). The root-mean-square error (RMSE) is used for the transient-free datasets Φ^A and Φ^B . As the scale information of the original state variables is lost by normalization and (non-)linear mapping, the scale of the effective variables is irrelevant for the interpretation of the new models and the comparison between the variables of the reduced models and the effective states is based on normalized values $\tilde{\varphi}_i$ with zero mean and unit standard deviation. As the model learning step is based on random initializations, the GP module is run ten times and the normalized time series data of the resulting best models, i.e. the smallest models able to reproduce the dynamics, are finally mapped back into the original eight-dimensional state space in order to compare between reduced and original model formulations.

5.7 Model results and discussion

5.7.1 Results for the transient-free time series

The simplest models of Φ^A and Φ^B produced by the reduction procedure follow

$$\frac{d\varphi_1^S}{dt} = p_1 \cdot \varphi_2^S + p_2, \quad \frac{d\varphi_2^S}{dt} = -p_3 \cdot \varphi_1^S + p_4, \quad (5.4)$$

with $S = A, B$ (Table B.1). Equation (5.4) can be rewritten as $\ddot{\varphi}_1^S = -p_1 p_3 \varphi_1^S + p_1 p_4$ which is the equation of a linear oscillator with constant external force $p_1 p_4$. Detailed characteristics of the time series, such as asymmetries of the oscillations, cannot be captured by (5.4) as it produces only harmonic oscillations. Instead, the slightly more complicated models follow

$$\frac{d\varphi_1^B}{dt} = \varphi_2^B - p_1, \quad \frac{d\varphi_2^B}{dt} = -\frac{\varphi_1^B - p_2}{\varphi_2^B} \quad (5.5)$$

for Φ^B and

$$\frac{d\varphi_1^A}{dt} = -\frac{p_1}{\varphi_2^A} - p_2, \quad \frac{d\varphi_2^A}{dt} = -\varphi_1^A \cdot (\varphi_1^A - p_3) \quad (5.6)$$

for Φ^A (Figure 5.4). These models are capable of reproducing the original datasets to a very high degree ($FEV = 0.98$ for Φ^A and Φ^B). System (5.5) is a simple nonlinear extension of model (5.4) with a modified repelling term for $d\varphi_2^A/dt$. Similar reduced models have already been found and discussed for simple (i.e. mono-species) food web models (Bernhardt and Wirtz, 2007).

The analogy of (5.5) with a linear oscillator supports the notion of biological forces in population models (Ginzburg, 1986). This analogy also leads to a fundamental similarity between biological and physical oscillatory systems (Bernhardt and Wirtz, 2007). The linear oscillator, as a general description of an undamped oscillating system, is the simplest model for this kind of dynamics. Thus, due to the structural flexibility of the MAGER scheme in combination with its data-adaptivity, the method is indeed able to find the minimal model in this case. MAGER is further able to reproduce the

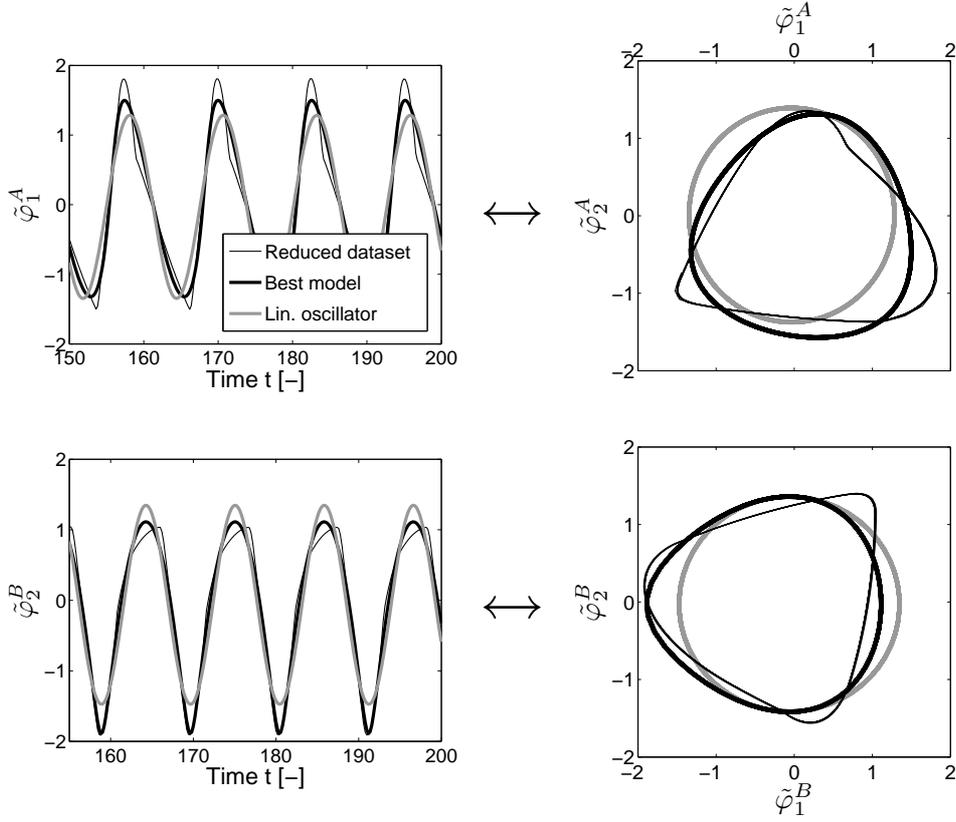


Figure 5.4: Time series and state space plots produced by the best reduced models (5.5) and (5.6) for Φ^A and Φ^B in comparison with the simple linear oscillator (5.4).

irregularities of the oscillations. However, this requires additional complexity of the resulting models which includes the incorporation of nonlinear terms in equations (5.5) and (5.6). Further biological interpretations of the learned models are given in the second part of this paper.

5.7.2 Results for the complete time series

The best reduced model found for the reconstruction of the transient and oscillatory regimes of Φ^a requires five parameters:

$$\begin{aligned} \frac{d\varphi_3^a}{dt} &= -\frac{p_1}{\varphi_{nl}^a} - \varphi_{nl}^a - p_2 \\ \frac{d\varphi_{nl}^a}{dt} &= -\frac{1}{\varphi_3^a} \cdot p_3 \varphi_{nl}^a (1 + p_4 \varphi_{nl}^a) - p_5 \varphi_3^a. \end{aligned} \quad (5.7)$$

For Φ^b , the best model found has a more complex structure,

$$\begin{aligned} \frac{d\varphi_1^b}{dt} &= p_1\varphi_{nl}^b - p_2 \\ \frac{d\varphi_{nl}^b}{dt} &= \left(\frac{p_3\varphi_{nl}^b}{p_4 + \varphi_{nl}^b} - \varphi_{nl}^b - \varphi_1^b \right) \cdot \left(\frac{p_5\varphi_{nl}^b}{p_6 + \varphi_{nl}^b} - \varphi_1^b \right). \end{aligned} \quad (5.8)$$

See Table B.1 for parameter values of both models. Figure 5.5 shows the reconstructions of \mathbf{X}^a and \mathbf{X}^b using (5.7) and (5.8). We can see that the reduced models are very well able to reproduce not only the transient regime of the consumer-resource time series but also the onset and amplitude of the limit cycle dynamics as well as the asymmetries of the oscillations. The original and approximated time series only differ with respect to the frequencies of the oscillations. However, as the combined error measure does not penalize deviations from the original frequency, this kind of approximation error was expected. Additionally, the inferior species of the original model are only reproduced to a lesser degree which can be attributed to the centering of the datasets (see section 5.3). Note that, in comparison with the results for $\widehat{\mathbf{X}}^a$, the nonlinear state reduction of $\widehat{\mathbf{X}}^b$ was affected by a larger approximation error (section 5.5). Nevertheless, the main course of the transient dynamics and the amplitude of the oscillatory regime of the dominant species as well as the nutrients can again be reproduced very well. This application of the MAGER scheme shows that the approach is also appropriate to reduce models showing more irregular dynamics than limit cycle oscillations. The increased complexity of the dynamics thereby requires a higher number of nonlinear terms. Nevertheless, general concepts like the notion of biological forces may still be found in these systems. Model (5.8), for example, is a nonlinear extension of the linear oscillator. Finally, although the learning scheme is based on single model parameterizations and is in general not able to reproduce the complete dynamic behavior of a system, the reduced models learned from more irregular dynamics may nevertheless be fairly general in reproducing different time series in a specific dynamic regime (Bernhardt and Wirtz, 2007). They may thus represent prototypes for a broader class of models showing similar dynamics.

5.7.3 Performance of the reduction procedure

The efficiency of the MAGER reduction can be assessed by a comparison with two analytical MR approaches, namely the EVA approach (Wirtz and Eckhardt, 1996) and the simple omission of state variables (see e.g. Pahl-Wostl, 1997). The complete centered dataset $\widehat{\mathbf{X}}^a$ serves as the data basis for all methods.

The application of the EVA method to the consumer-resource model (5.1)-(5.3) results in a reduced formulation incorporating functional descriptions of the tradeoffs in resource requirements and nutrient quotas (Wirtz and Bernhardt, 2008). The simplified model (EVA1) has a dimensionality of 3(13), where the 13 in brackets denotes the number of parameters and three is the number of state variables, whereas the original model's dimensionality is 8(44). The other reduced formulations are produced

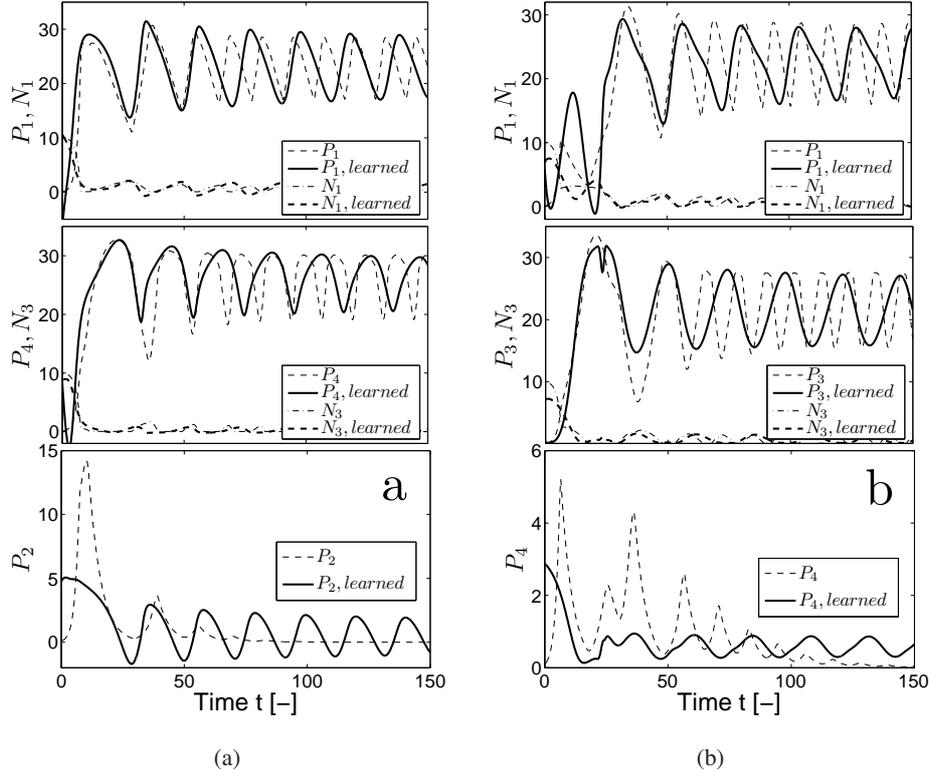


Figure 5.5: Reconstruction of the original datasets using (a) the learned reduced model (5.7) for \mathbf{X}^a and (b) model (5.8) for \mathbf{X}^b .

by a stepwise omission of (O1) the inferior species P_2 and P_3 , (O2) the inferior species as well as P_5 and (O3) all species except for P_1 . The nutrients are left unaltered. The dimensionalities of these models are 6(28) for O1, 5(20) for O2 and 4(12) for O3. The parameters of EVA1 and O1-3 are optimized using the hybrid GA optimization scheme and the combined error measure consisting of RMSE and E_p (8.15).

Figure 5.6 shows the negative correlation between the number of parameters and the deviation from the original model $E_C = (RMSE + E_p)/2$. The comparison includes the best 20 systems found with the MAGER approach. The best MAGER model (5.7) is of complexity 2(5). The accuracy/complexity tradeoffs of the data-adaptive and analytical reduction approaches are indicated with gray lines in Figure 5.6. While the reduction efficiencies, i.e. the slopes of these lines which represent the error reduction per number of added parameters, are comparable for the methods, there is a significant displacement of the tradeoff. Models reduced with the MAGER scheme have much less degrees of freedom compared to the simplified systems obtained with the other two reduction techniques while being comparable or better in terms of error values. In fact, in addition to the number of free parameters, the state dimensionality of the GP reduced systems is also lower than that of O1-3 and EVA1.

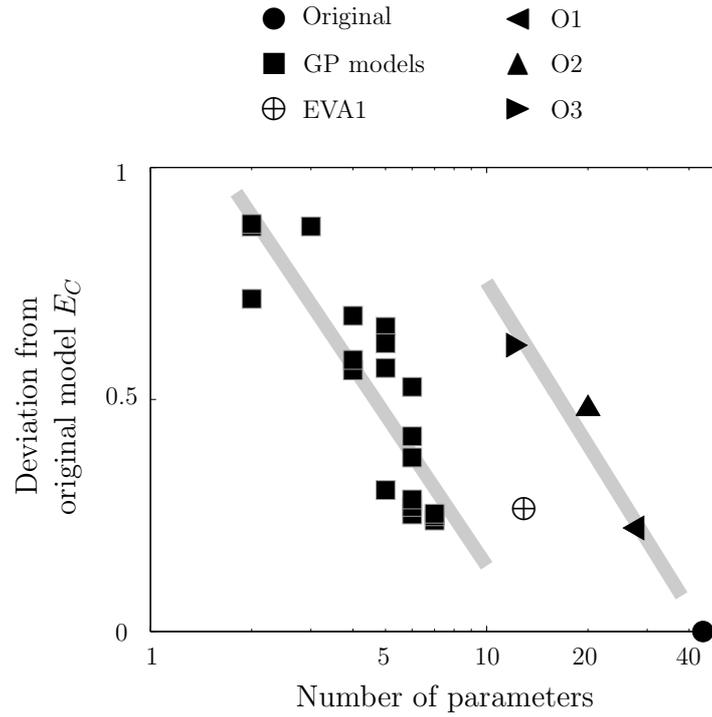


Figure 5.6: Number of parameters versus reduction error E_C for the best 20 reduced models found with the GP algorithm (■) and the original model (●) without P_2 and P_3 (O1, ◀), without P_2 , P_3 and P_4 (O2, ▲) and with P_1 only (O3, ▶). Also shown is the EVA model (EVA1, ⊕). The gray bold lines indicate the displacement of the accuracy/complexity tradeoff for MAGER in comparison with the analytical MR approaches (see text for details).

The smallest models found with MAGER have only two or three parameters and coincide with linear oscillators similar to (5.4) (not shown). However, it should be noted that models with E_C values of around 0.5 or worse can in general not reproduce all details of the time series, i.e. the transient and/or limit cycle. In comparison with O1-O3, the EVA model gives a better approximation of the transient regime which explains the corresponding low error value. However, only the relatively complex O1 model is able to concurrently reproduce the limit cycle and transient similar to the best reduced systems of the MAGER approach.

The better performance of the MAGER scheme is a consequence of its independence from the former model structure and its adaptivity to data. The data-driven fitness assignment of the learning scheme thereby facilitates the generation of good models with high fitting capabilities whereas the structural reformulation leads to simpler system descriptions. We suggest that the limitations of reduction efficiency for the analytical approaches arise from the lack of one or both of these properties. The method of state variable omission is not data-adaptive and does not alter the model structures either which re-

sults in larger models and a higher probability that the dynamics of the learned time series cannot be reproduced. The same applies for the EVA approach as well as the other aggregation methods discussed in section 5.1. These approaches are not based on model results or depend on the existence of specific dynamic regimes such as steady states. Furthermore, the model structure is only changed to a small degree by aggregation or, in the case of the EVA approach, only as far as the incorporation of tradeoff functions is concerned. In addition, the concurrent reproduction of transient and oscillatory regimes in most cases could not be accomplished by these methods.

5.8 Conclusion

In this paper, we used the new Mapping-based Genetic Reduction method to simplify a complex consumer-resource model. The method was applied to reconstruct the transient and limit cycle dynamics for two parameter sets. We demonstrated that the combination of state variable aggregation and evolutionary learning of new model structures reduced the parameter and state dimensionality very effectively. More specifically, the number of state variables could be reduced by 75% in comparison with the original model (5.1)-(5.3) whereas the number of parameters was reduced by over 90% for the complete dataset and by over 95% for the transient-free time series. A comparison of the MAGER results with reduced models obtained with the Effective Variable Aggregation method (EVA) as well as by omission of state variables further showed that the genetic reduction was able to produce much smaller models while being comparable or better in terms of reduction error. If the number of parameters is considered as the sole measure of complexity, the MAGER approach was able to reduce the original model by over 50% more than the omission of state variables. Traditional aggregation methods typically are not able to reproduce any detail of irregular dynamics. The associated dynamic mismatch has been described as "dynamic emergence" (Auger and Poggiale, 1998) but it is understood as a kind of reduction error in the present context. The improved reduction performance of MAGER can thereby be attributed to the combination of data-adaptive learning and reformulation of model structure. Finally, different from other aggregation procedures, no restrictions concerning the complexity of the original models, such as linearity of the differential equations, apply for the new scheme.

So far, the model systems have only been interpreted in structural or mathematical terms, i.e. by a qualitative comparison with (non-)linear oscillators. This result, however, already is a fundamental finding which supports the picture of biological oscillators and the notion of biological forces in these systems (Ginzburg, 1986; Vadasz and Vadasz, 2002; Vandermeer, 2004; Bernhardt, 2007; Bernhardt and Wirtz, 2007). Similar to physical oscillatory systems, the reduced models incorporate repelling forces which are proportional to the deviation of the effective states from their mean. The new states can thereby be related to indirect interactions between the species leading to the notion of Ecologi-

cal Interaction Models, as shown in the second part of the paper. Consequently, the reduced models indicate that the interaction intensities between coexisting species tend to be self-stabilizing in environmental systems. In order to substantiate this proposition, the MAGER approach has to be utilized for the reduction and transformation of more complex models as well as measured environmental datasets in future applications.

Chapter 6

Explicit formulation of indirect interactions in a reduced consumer-resource competition model. Part II: Biological implications

Abstract

The identification of the principles underlying the competitive coexistence of species is one of the major challenges in theoretical biology. A number of model studies analyzing the influences of different aspects of species networks on coexistence, such as the role of direct and indirect interactions, have been performed in this context. The analyses in these studies are mainly based on simple formulations, for example Lotka-Volterra competition models, and the resulting proposed prerequisites for coexistence differ depending on the specific model used.

In the present paper, the question of general biological principles is investigated using a data-adaptive point of view. We use simulation data of a complex consumer-resource (CR) model and apply the new Mapping-based Genetic Reduction method (MAGER) to derive simplified system descriptions. Following the parsimony principle, the simplest possible model able to explain the data is proposed to be the best abstraction of coexistence in CR networks. The method is based on time-series data and does not use any knowledge about the system-specific interactions of the biological state variables. The resulting Ecological Interaction Models (EIM) incorporate the indirect density-mediated effects of the dominant species and the interactions between the new effective consumer variables can be traced back to tradeoffs in the community level descriptions of resource requirements and nutrient quotas. Thus, this work presents a new simple representation of competitive coexistence in the CR network. This meta level description explicitly formulates the original model dynamics in terms of oscillations between different states of community resource usage characteristics. Furthermore, the analysis of the reduced models shows that the transient regime of the time series can be explained by competition-free logistic growth.

6.1 Introduction

The effect of competition among species on coexistence in biological networks has repeatedly been investigated since the formulation of the famous competitive exclusion principle (Volterra, 1928; Gause, 1934). In particular, the violation of this principle and the mechanisms governing the coexistence of species assemblages based on a small number of essential resources have been addressed in many

studies (e.g. Hutchinson, 1961). Apart from spatially explicit or behavior-based approaches where species competition is weakened by the existence of ecological niches, recent model studies also emphasize the role of asynchronous oscillatory dynamics for species coexistence (Abrams, 2006; Vandermeer, 2006). While competition for living resources is usually described by competition models of the Lotka-Volterra type (Volterra, 1928; Lotka, 1932; Chesson, 2000; Vandermeer and Pascual, 2006), competition for abiotic resources is formulated using consumer-resource (CR) models (Tilman, 1982; Huisman and Weissing, 2001a; Abrams, 2006). Both groups of models incorporate implicit formulations of indirect species interactions by means of shared resources or prey pools. Depending on the specific model used, different prerequisites for oscillatory coexistence have been pointed out, such as the nonlinearity of the functional response of consumers to resource density (Abrams, 2006), trade-offs in resource requirements and the accompanied relation between requirements and consumption (Huisman and Weissing, 2001a) as well as the degree of diet specialization (Vandermeer and Pascual, 2006). The model study of Vandermeer and Pascual (2006) on specialization revealed a crucial dependency of coexistence on a dominance of indirect (mutualism) over direct trophic interactions (feeding).

The importance of trait- (TMII) and density-mediated indirect interactions (DMII) for coexistence in food webs has also been demonstrated in many studies. Quite a number of different indirect interactions with both positive and negative effects on species abundances and traits have been documented (see e.g. Wootton, 1994; Strauss, 1991, for an overview). Among others, these interactions include trophic cascades, apparent competition and indirect mutualism. By measuring the net or effective interaction of species, Lawlor (1979) as well as Roberts and Stone (2004) showed that the multitude of DMII may in fact reverse the negative direct competitive effects in LV models with varying competition parameters. In a combined model and experimental study, van Veen et al. (2005) analyzed the contribution of indirect effects to the stability of insect communities and found that the concurrence of DMII and TMII stabilized the investigated three-species system. Further experimental demonstrations of the outstanding role of indirect effects for coexistence were accomplished e.g. for parasite-host communities (Krasnov et al., 2005) and marine intertidal food webs (Menge, 1995).

In former model-based studies, the intensity of competition and the degree of coexistence were exclusively determined by the setting of specific growth and consumption parameters. In particular, coexistence emerged from implicit descriptions of indirect interactions. While the degree of coexistence may be quantified using, for example, a common index of species diversity, no explicit variables describing the competitive interactions have been proposed so far. The missing direct representation of interactions makes it difficult to qualitatively assess the dependency of DMII- or TMII-based coexistence on internal or external factors. Furthermore, without an explicit description of competitive dynamics, the understanding of the mechanisms facilitating coexistence remains phenomenological. In the present study, we follow a model building strategy based on functional relationships which leads

to an explicit formulation of DMIs. We analyze alternative descriptions of a classical CR model incorporating competition among different consumer species (Huisman and Weissing, 2001a). These new representations, which were obtained with the Mapping-based Genetic Reduction (MAGER) procedure in the first part of this paper, approximate the original model dynamics on the basis of newly constructed minimized model structures. The use of MAGER as a means to explicitly formulate the CR model's inherent indirect effects was motivated by earlier findings on interactions of predator-prey systems using this approach (Bernhardt and Wirtz, 2007). We further suggest that the simplest model able to reproduce the model-generated datasets provides insights into the system's most relevant processes. Thus, the application of the MAGER scheme to the CR model helps to extract the key variables and their mutual dynamic inter-dependencies which effectively represent important indirect interaction pathways. In addition, it helps to assess the conditions for the occurrence of competitive coexistence. In the context of CR models, this coexistence is given in the form of sustained species cycles.

The analysis will be related to the tradeoffs in resource requirements and the accompanied relation between requirements and consumptions formulated in the original CR model. We investigate how these tradeoffs, which are implicitly formulated by the parameterizations of CR models, re-emerge in the reduced model formulations.

6.2 Method overview

Details of the simplification procedure as well as the original model are given in the first part of the paper. The eight-dimensional CR model was run with two different initial conditions which resulted in time series with different sets of dominant species. The simulated and centered datasets $\widehat{\mathbf{X}}^a$ and $\widehat{\mathbf{X}}^b$ as well as their transient-free subsets $\widehat{\mathbf{X}}^A$ and $\widehat{\mathbf{X}}^B$ were used as inputs for a (non-)linear Principal Component Analysis (PCA/NLPCA) to reduce their dimensionality. The choice of the number of retained components in the PCA procedure was based on the fraction of explained variance (FEV) for the subsets of components. All datasets could be reduced to two (non-)linear principal components (PC/NLPC) which served as new state variables for the reduced model systems. These *effective consumer variables* (ECV) will be denoted by φ_i^s and φ_i^S with $s = a, b$ for the complete datasets and $S = A, B$ for the transient-free cases, respectively. A numeric subscript indicates the i th linear PC whereas nonlinear components are denoted by φ_{nl}^s . A Genetic Programming (GP)/Genetic Algorithm (GA) hybrid was then used to construct new models as dynamical systems of φ_i^s and φ_i^S . During model learning, the initial values of the state variables were subject to optimization and the method was restricted to only produce time series with positive variables in order to simplify the process-based interpretation.

The simulated time series of the best models fit the corresponding datasets very well (see Figure 6.1

which depicts the simulation results for the best models approximating the complete datasets (section 6.4.2)). The slight deviations in the frequency of the oscillations were not captured by the GP/GA model optimization as the applied fitness measure ignores the frequency and phase information of the limit cycle regime.

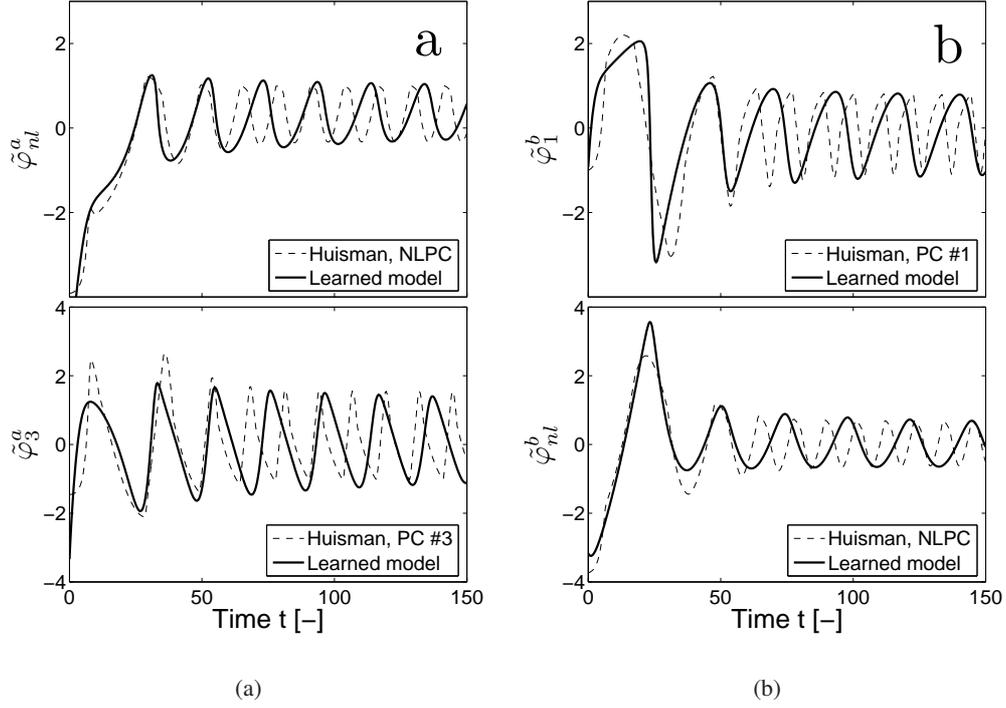


Figure 6.1: Simulated time series of the best reduced models found with the MAGER approach in comparison with the original CR model data projected onto the corresponding principal components. The figure shows the two setups *a* and *b* which are characterized by different assemblages of coexisting consumer species.

6.3 Limit cycle dynamics

6.3.1 Effective consumer variables

The ECVs obtained with PCA are linear combinations of the centered phytoplankton abundances and nutrient concentrations. For the transient-free datasets, the first two PCs of $\hat{\mathbf{X}}^A$ can be written as

$$\begin{aligned}\varphi_1^A &= u_1 \hat{P}_1 - u_2 \hat{P}_4, \\ \varphi_2^A &= v_1 \hat{P}_5 - (v_2 \hat{P}_4 + v_3 \hat{P}_1)\end{aligned}\quad (6.1)$$

and the first two PCs of $\widehat{\mathbf{X}}^B$ follow

$$\begin{aligned}\varphi_1^B &= u_1\widehat{P}_3 - (u_2\widehat{P}_1 + u_3\widehat{P}_2), \\ \varphi_2^B &= v_1\widehat{P}_2 - (v_2\widehat{P}_1 + v_3\widehat{P}_3).\end{aligned}\quad (6.2)$$

See Table C.1 for parameter values. Note that the sign of φ_2^A has been changed to simplify the process-based interpretation. Without much loss of accuracy, terms with small parameters have been ignored in (6.1) and (6.2) as the remaining terms explain the main part of the data variance, i.e. $FEV > 0.95$. It is obvious that the PCs primarily incorporate the dominant and most abundant species of the multivariate time series whereas the influences of the other species and nutrients are ignored. This effect can be attributed to the data preparation where no scaling of the variables was used (Bernhardt and Wirtz, 2008). However, the calculation of the correlation coefficients between the individual species of $\widehat{\mathbf{X}}^A$ as well as φ_1^A and φ_2^A shows that this approximation error affects the inferior phytoplankton species only as each of the nutrients negatively correlates with one of the dominant species (Figure 6.2). In a similar manner, the nutrients of $\widehat{\mathbf{X}}^B$ are negatively correlated to P_1 , P_2 and P_3 (not shown).

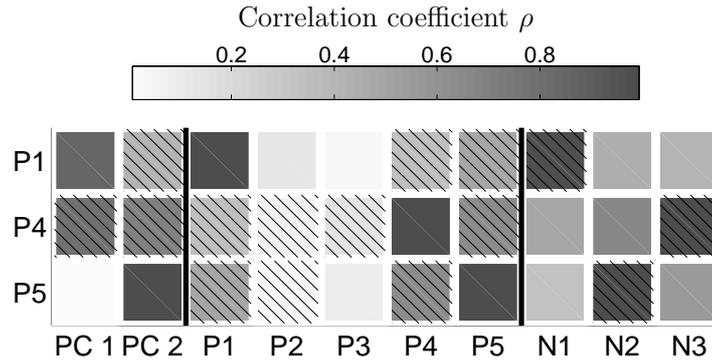


Figure 6.2: Correlation coefficients for the dominant species P_1 , P_4 and P_5 in $\widehat{\mathbf{X}}^A$ with all species and resources as well as the first two principal components (PC1 and PC2). Hatched patterns indicate negative values.

The negative correlations further indicate that each nutrient acts as a main limiting resource for one of the dominant phytoplankton species.

These growth limitations are determined by the nutrient quotas of the individual species. As outlined by Huisman and Weissing (2001a), the competitive strength of each individual species can be calculated using the respective resource requirements $R_j^*(i)$ according to

$$R_j^*(i) = \frac{\omega_i \cdot k_{ji}}{g_i - \omega_i}, \quad (6.3)$$

where ω_i denotes the mortality, g_i is the maximal growth rate of species P_i and k_{ji} denotes the half saturation constant for nutrient N_j of this species. Figure 6.3 shows the occupation of different niches

in the three-dimensional resource requirement space by the individual consumer species. The species

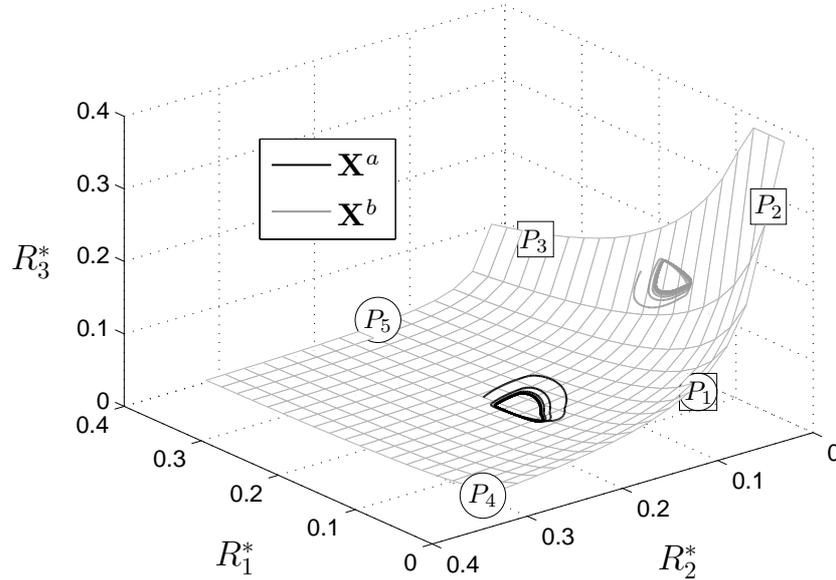


Figure 6.3: Tradeoffs in resource requirements $R_j^*(i)$ of species P_i for resource N_j . The fitted surface illustrates the form of the two-dimensional tradeoff in resource requirement. The community resource requirements $\langle R_j^* \rangle$ (see equation (6.4)) for the oscillatory part of the datasets are shown as gray and black solid lines. The system oscillates between the species resource requirements as extremal points with clearly separated limit cycles for the two different initial conditions.

with the lowest requirement for resource j has the highest competitive ability for this resource compared to the other consumers. Species are clearly separated from each other in this picture which indicates different diet preferences. For example, P_1 is a typical generalist which can be displaced by P_4 or P_2 as it gets limited by N_1 (Table C.2). The same applies for the other species and the observed limit cycle dynamics critically depends on these tradeoffs of species-specific values of resource requirements $R_j^*(i)$ and resource consumptions/quotas c_{ji} (Huisman and Weissing, 2001a). Figure 6.3 further shows that the existence of the two oscillatory states, i.e. the coexistence of either P_1, P_4 and P_5 or P_1, P_2 and P_3 , may be attributed to differences in the resource competition setup.

Given these well-known findings, the definition of the ECVs in equations (6.1) and (6.2) indicates that the single species description of competition in the CR model has been changed to a new formulation which incorporates community level information. This link between the effective consumer variables and the community level of the biological system can be illustrated by a comparison of the ECVs with the community resource requirements $\langle R_j^* \rangle$ and community resource quotas $\langle C_j^* \rangle$ as analytical means

of up-scaling by aggregation,

$$\langle R_j^* \rangle = \frac{\sum_i R_j(i)^* \cdot P_i}{\sum_i P_i}, \quad \langle C_j^* \rangle = \frac{\sum_i c_{ji} \cdot P_i}{\sum_i P_i}. \quad (6.4)$$

These community properties are in part strongly correlated with the ECVs (Figure 6.4) which again supports the finding that the dynamics of the reduced system is governed by tradeoffs in resource requirements and quotas similar to the ones present in the original system. According to the informa-

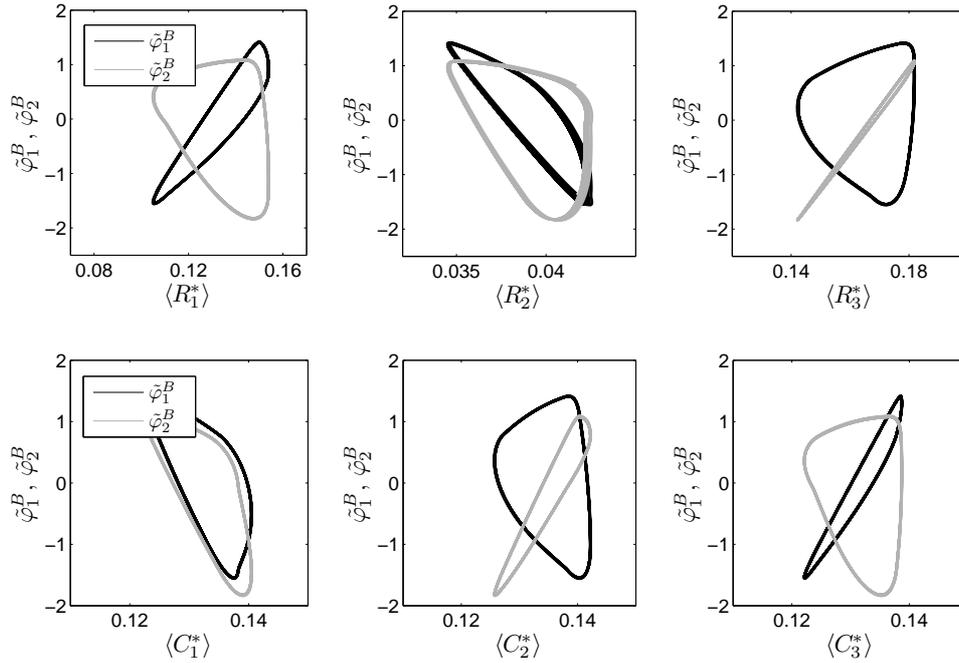


Figure 6.4: Community resource requirements $\langle R_j^* \rangle$ and community resource quotas $\langle C_j^* \rangle$ versus normalized ECVs $\tilde{\varphi}_1^B$ and $\tilde{\varphi}_2^B$ for the limit cycle regime in $\hat{\mathbf{X}}^B$. Similar correlations are found for $\tilde{\varphi}_1^A$ and $\tilde{\varphi}_2^A$ of $\hat{\mathbf{X}}^A$ (not shown).

tion inherent to Figures 6.3 and 6.4, the new ECVs describe weighted community compositions which serve as indicators for the system's adaptation to resource-specific internal tradeoffs in resource requirement and quota configurations. In addition, the effective consumer variables also capture the configuration *changes* which are mediated by the indirect inter-dependencies between the dominant species in the CR network (Figures 6.5 and 6.6). We will simplify the following discussion by concentrating on the resource requirement tradeoffs but it should be noted that the situation is similar for the tradeoffs in resource quotas.

According to Figure 6.5(a), φ_1^A describes a shift in community composition ($P_4 \rightarrow P_1$) resulting from an adaptation to a shortage in the second resource and the accompanied limitations. In addition, this ECV captures the indirect interaction between the two species, namely the competitive growth

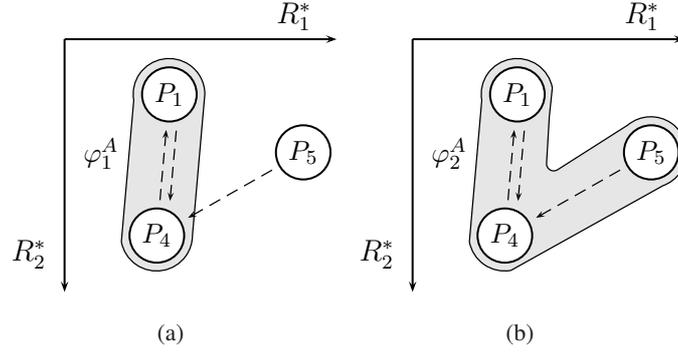


Figure 6.5: Resource requirement tradeoffs effecting the competition between the dominant species in $\widehat{\mathbf{X}}^A$. (a) φ_1^A captures relations between P_1 and P_4 and (b) φ_2^A describes the difference between the high requirement for resource 1 of P_5 on the one hand and the low R_1^* values of P_4 as well as P_1 on the other hand (see equation (6.1)). Positions of the dominant species in the two-dimensional requirement space follow Figure 6.3. Dashed arrows indicate negative indirect interactions of one species on another due to utilization of shared resources which results from the parameterizations of half-saturation constants k_{ji} and nutrient quotas c_{ji} for species P_i and nutrients N_j (Table C.2). Gray areas indicate the connection of the individual ECVs to abundance changes of the original species which, in turn, reflect the existence of DMIs.

limitation of P_4 induced by P_1 which is a strong competitor for N_3 , the limiting resource of P_4 (Table C.2). The analogous community shift resulting from an adaptation to the affinity for resource 1 is described by the second ECV in form of a biomass distribution between P_5 as well as the first and fourth species (Figure 6.5(b)). As these species are all linked together via their indirect interactions, φ_2^A also incorporates the impact of P_5 on the CR food web. This impact is more complicated as, in addition to the competitive growth limitation of P_4 due to consumption of N_3 by P_5 , a change in P_5 abundance also indirectly affects P_1 via the nutrient utilization of P_4 (Table C.2). Along this interaction chain, the affinity for resource 1 changes from a low value for P_5 to high values for P_4 and P_1 .

The basic meaning of the ECVs is similar for $\widehat{\mathbf{X}}^B$ even though the competitive situation is different. Here, the mutual limitations between the dominant species are arranged in a cyclic fashion (Figure 6.6). Similar to φ_2^A of $\widehat{\mathbf{X}}^A$, the new variables reveal the requirement tradeoffs of resources 3 and 1 which are again mediated by active food web interactions. P_3 is a strong competitor for N_2 , the limiting nutrient for P_2 , and P_2 itself affects P_1 via the consumption of N_1 . In a similar manner, from the perspective of P_2 , we find a DMII of P_2 on P_3 via the growth limitation of P_1 .

Summing up the preceding discussion, the effective consumer variables of both setups, $\widehat{\mathbf{X}}^A$ and $\widehat{\mathbf{X}}^B$, represent biomass distributions between the dominant species in relation to community resource requirements and quotas. The ECVs can further be understood as condensed descriptions of density-

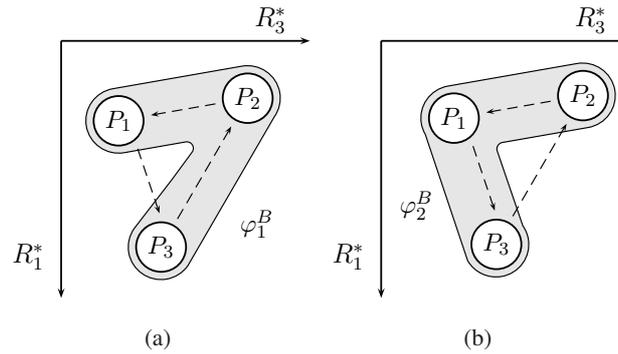


Figure 6.6: Same as Figure 6.5 for the competition of the dominant species in $\widehat{\mathbf{X}}^b$.

mediated indirect interactions within the CR network. Based on their integrative role, the ECVs will be called "community resource usage characteristics" in the following discussion. The term "usage" thereby relates to the combination of requirement and consumption of specific nutrients. By explicitly incorporating these resource usage characteristics in new ODE models (sections 6.3.2 and 6.4.2), the aggregation scheme of MAGER further translates the *indirect* species interactions mediating the internal tradeoffs in resource usage to *direct* interactions of ECVs as new competition factors. Because of this transformation, we will use the term *Ecological Interaction Models* (EIM) as a description of the new model class (Bernhardt and Wirtz, 2007).

6.3.2 Ecological Interaction Models

The best reduced models of the transient-free datasets follow the equations of simple oscillators (not shown). The slightly more complicated learned models able to reproduce the asymmetries of the oscillations follow

$$\frac{d\varphi_1^B}{dt} = \varphi_2^B - p_1, \quad \frac{d\varphi_2^B}{dt} = -\frac{\varphi_1^B - p_2}{\varphi_2^B} \quad (6.5)$$

for $\widehat{\mathbf{X}}^B$ and

$$\frac{d\varphi_1^A}{dt} = -\frac{p_1}{\varphi_2^A} - p_2, \quad \frac{d\varphi_2^A}{dt} = -\varphi_1^A \cdot (\varphi_1^A - p_3) \quad (6.6)$$

for $\widehat{\mathbf{X}}^A$. Note again that the sign of φ_2^A has been changed. See Table C.3 for parameter values.

Model (6.5) is a simple nonlinear extension of the linear oscillator model with a modified repelling term for $d\varphi_2^B/dt$. Similar reduced models have already been discussed for other food web models by Bernhardt and Wirtz (2007). In equation (6.5), the direct interactions of the resource usage characteristics differ for the two variables which may be qualitatively attributed to the cyclic dependencies of the dominant species in the food web. φ_2^B is related to the community resource requirement for nutrient 3 (Figure 6.6). Thus, a decreasing competitive pressure caused by an abundance of N_3 permits a community structure with higher requirement for this resource and concurrent higher values of

φ_2^B . As the resource usage characteristic φ_1^B is correlated to the community quotas/consumption of N_3 , the subsequent higher usage of this nutrient is described by the positive effect of φ_2^B on $d\varphi_1^B/dt$ in equation (6.5). On the species level, this corresponds to a biomass shift from the second to the third consumer species along the network's indirect interaction pathway (Figure 6.6(b)). The resulting community composition is characterized by a higher requirement for resource 1. Subsequently, the indirect species interactions support the invasion of the first consumer which strongly consumes this nutrient (Figure 6.6(a)). The accompanied change in resource usage characteristics is expressed by the negative influence of φ_1^B on the temporal change of φ_2^B as the resulting community is in a state of low requirement for the third resource.

The reduced model (6.6) of $\widehat{\mathbf{X}}^A$ incorporates the main changes in resource usage characteristics of the competitive system in an analogous way. In this case, the characteristics are related to the first and second nutrient. At first, in a situation of abundant N_1 , the community is characterized by high requirements for this resource. It then shifts to a state of strong N_1 consumption ($P_5 \rightarrow P_1$) along the indirect interactions. This shift is expressed by the reciprocal influence of φ_2^A on $d\varphi_1^A/dt$ as the first resource usage characteristic is correlated with $\langle C_1^* \rangle$. The negative impact of φ_1^A on $d\varphi_2^A/dt$ may further be explained by the shared usage of N_3 by P_1 and P_5 (Table C.2).

However, both reduced models lack important direct interactions of the resource usage characteristics which are able to restore the starting points of the oscillations. In model (6.5) on the one hand, this relates to a community shift from a state of high N_1 quotas back to high requirements for the third resource and consumption of N_2 . Alternatively, in terms of species interactions, we may also say that no state variable capturing the indirect positive effect of P_1 on P_2 in $\widehat{\mathbf{X}}^B$ is present (Figure 6.6). On the other hand, in model (6.6), no direct ECV interaction leading to a decrease of the first resource usage characteristic can be found.

Instead, this "rewinding" is achieved by the constant terms in both equations which act as external driving factors. p_1 in (6.5) and p_2 in (6.6) support the indirect interactions leading to a decrease of φ_1^B and φ_1^A , respectively. Furthermore, p_2 in (6.5) and p_3 in (6.6) are responsible for the concurrent increase of the second ECVs which reverses the interactions between P_1 and P_2 as well as P_4 and P_5 , respectively. Strictly speaking, these positive driving factors are different from the two other external driving factors as they are connected to the resource usage characteristics φ_2^B and φ_1^A . However, these constants have a special meaning as they are necessary for the occurrence of positive oscillations. Finally, these "interaction support" and "interaction reversal" factors must be in balance in order to allow for sustained positive (or negative in case of φ_2^A) oscillations between the respective states of community resource requirements and quotas (Figure 6.7). These terms thus extend the set of direct ECV inter-dependencies and replace additional state variables or more complicated terms which could alternatively close the gap of the missing interactions.

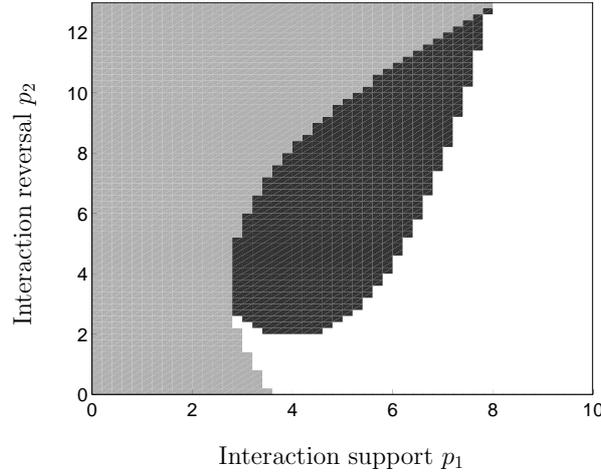


Figure 6.7: Results of a parameter variation of the interaction support p_1 and the interaction reversal factor p_2 showing the effects of these factors on the dynamics of model (6.6). The resource usage characteristics coexist in the region shaded in dark gray. Light gray indicates numerical truncations of the simulations and unrealistic zero crossings of the variables occurred in the white region. The initial values of φ_1^B and φ_2^B were left unaltered (Table C.3).

6.4 Modification by transients

6.4.1 Effective consumer variables

As outlined in section 6.2, NLPCA was used in addition to linear Principal Component Analysis to further reduce the dimensionality of the complete datasets including the transients to the limit cycle regimes. Unlike PCA, however, the neural network-based NLPCA does not provide simple algebraic interpretations of the principal components. Therefore, we use a modification of the MAGER scheme to find functional relationships between the linear and the nonlinear principal components. As shown in the first part of the paper, single nonlinear components could be used to replace the first and second PC of $\widehat{\mathbf{X}}^a$ and φ_2^b and φ_3^b of $\widehat{\mathbf{X}}^b$, respectively. Thus, the PCs served as input variables for the MAGER scheme and the GP part of MAGER was changed to produce nonlinear functions, instead of ordinary differential equations, which approximate the nonlinear components.

For $\widehat{\mathbf{X}}^a$ the linear principal components follow

$$\varphi_1^a = u_1 \widehat{P}_5 + u_2 \widehat{P}_4 + u_3 \widehat{P}_1, \quad (6.7)$$

$$\varphi_2^a = v_1 \widehat{P}_5 - (v_2 \widehat{P}_4 + v_3 \widehat{P}_1), \quad (6.8)$$

$$\varphi_3^a = w_1 \widehat{P}_1 - w_2 \widehat{P}_4. \quad (6.9)$$

For a better interpretability, the signs in (6.8) and (6.9) have been changed. A comparison with the ECVs of the limit cycle regime (equation (6.1)) shows the similarity between the resource usage characteristics φ_2^a and φ_2^A as well as φ_3^a and φ_1^A . Finally, the first component may be seen as a weighted sum of the dominant species' biomass. Based on these findings, we can easily interpret the simplest functional form of the NLPC found with MAGER which reads

$$\varphi_{nl}^a = a_1\varphi_2^a - a_2\varphi_1^a. \quad (6.10)$$

Note that φ_{nl}^a has also been changed to negative values. Thus, the nonlinear component is similar to the second linear ECV in the transient-free case, the only difference being the additional weighted biomass sum φ_1^a in (6.10). This additional term is responsible for the reproduction of the transient phase whereas φ_2^a mainly reproduces the limit cycle dynamics. In fact, FEV drops from 97.7% to 3.3% if we neglect φ_1^a in the transient phase. The dynamics of the limit cycle regime, however, can approximately be explained by φ_2^a alone (FEV=81%). The analogy of the resource usage characteristics for the transient-free and complete dataset is not changed by this additional weighted biomass sum. Similar to φ_2^A , φ_{nl}^a is related to the community requirements for the first resource and community quotas of N_2 (not shown).

In case b with the complete dataset $\widehat{\mathbf{X}}^b$, the linear relationship for φ_1^b follows

$$\varphi_1^b = u_1\widehat{P}_2 - u_2\widehat{P}_1. \quad (6.11)$$

The competitive situation is more complicated than in case a and no simple functional form relating the second and third linear PCs to the nonlinear component could be produced with MAGER. However, when the learning of the functional description was based on the centered species instead of the PCs, a simple algebraic interpretation of the NLPC could be found,

$$\varphi_{nl}^b = v_1\widehat{P}_3 - \widehat{P}_1 \cdot (v_2\widehat{P}_3 + v_3). \quad (6.12)$$

The signs of φ_1^b and φ_{nl}^b have also been changed so that both variables are negative. Similar to φ_2^B , the linear PC φ_1^b is related to $\langle R_3^* \rangle$ and the community quotas of the second resource $\langle C_2^* \rangle$. For φ_{nl}^b , the interpretation in terms of a simple biomass distribution between the dominant species is partially lost because of the more complex structure of equation (6.12). However, φ_{nl}^b is similar to φ_1^B in terms of tradeoffs in resource usage (not shown) and may still be interpreted as a nonlinear resource usage characteristic. Thus, the new variables in case b again capture the community tradeoff between R_1^* and R_3^* as well as the tradeoff between the quotas of the second and third nutrient.

In conclusion, the ECVs of the full datasets describe community resource usage characteristics similar to those in the transient-free case. However, details of the definitions of the ECVs have changed in order to account for the increased complexity introduced by the system's transient dynamics.

6.4.2 Ecological Interaction Models for the complete dynamics

The best reduced models reproducing the dynamics of the complete datasets follow

$$\frac{d\varphi_3^a}{dt} = \overbrace{-\frac{p_1}{\varphi_{nl}^a}}^I - \overbrace{\varphi_{nl}^a}_{II} - p_2 \quad (6.13)$$

$$\frac{d\varphi_{nl}^a}{dt} = \underbrace{-1/\varphi_3^a}_{III} \cdot \underbrace{p_3\varphi_{nl}^a \cdot (1 + p_4\varphi_{nl}^a)}_{IV} - \underbrace{p_5\varphi_3^a}_{V}. \quad (6.14)$$

for $\widehat{\mathbf{X}}^a$ and

$$\frac{d\varphi_1^b}{dt} = p_1\varphi_{nl}^b - p_2 \quad (6.15)$$

$$\frac{d\varphi_{nl}^b}{dt} = \left(\frac{p_3\varphi_{nl}^b}{p_4 + \varphi_{nl}^b} - \varphi_{nl}^b - \varphi_1^b \right) \cdot \left(\frac{p_5\varphi_{nl}^b}{p_6 + \varphi_{nl}^b} - \varphi_1^b \right) \quad (6.16)$$

for $\widehat{\mathbf{X}}^b$. See Table C.3 for parameter values.

The interactions of the ECVs in equations (6.13) and (6.14) can again be interpreted using the dominant competitive relationships within the CR network outlined in section 6.3.1. As discussed in the last section, the interpretation of the ECVs for the transient-free and complete datasets is qualitatively similar. The same applies for the dynamic impact of the second on the first competition factor in (6.13). We again find positive effects of (negative) φ_{nl}^a on φ_3^a which corresponds to a shift in community composition to a state of high N_1 quotas in case of N_1 abundance. These interactions are expressed by terms *I* and *II* in (6.13) which represent two different ways of positive influences between the characteristics. Here, term *I* describes the shift in community composition based on the DMIs of the network (see section 6.3.2). As the endpoints of this interaction chain, i.e. P_5 and P_1 , also compete for the shared resource N_1 and concurrently limit P_4 (Table C.2), we can also understand the increase of the positive effect for smaller values of φ_{nl}^a which reduces the competitive pressure for P_1 (term *II*). Furthermore, p_2 can again be interpreted as an "interaction support" factor responsible for the reconstruction of the starting point of the oscillations in analogy to p_1 in (6.6).

Similarities with the transient-free case can also be found for the dynamic equation of the nonlinear resource usage characteristic $d\varphi_{nl}^a/dt$. First, term *V* again describes the positive influence between the usage characteristics which was similarly captured by $p_3\varphi_1^A$ in equation (6.6). Thus, p_5 in term *V* may also be seen as an "interaction reversal" factor in analogy to p_3 in (6.6). Second, term *IV* formally corresponds to a logistic growth term for negative state variables with growth rate p_3 and capacity $1/p_4$. An analysis of the time evolution of this term shows that it has a high positive value during the transient phase in comparison with term *III*. The logistic growth factor then drops to a small negative constant value in the oscillatory regime. Because of this constant, equation (6.14) reduces to a formulation which is very similar to the transient-free form in the limit cycle regime. This

dynamic change can easily be explained by the initial abundance of nutrients. At the beginning of the simulation, the community biomass is low and the resources supply exceeds the demand which leads to a reduced competitive pressure and an initial competition-free growth phase. Subsequently, the increase in community biomass and the concurrent short supply of nutrients causes the competitive dynamics and the cyclic shifts in community resource usages.

It was not possible to find a simple biological interpretation of the best model for Φ^b in equations (6.15) and (6.16). In physical terms, however, this model corresponds to the ordinary differential equation of a nonlinear oscillator describing the cyclic shifts in community resource usage similar to (6.6) in section 6.3. Thus, in comparison with the transient-free case, we find a stronger dependency of the detailed model structures on the initial conditions for the complete datasets. While the different initial conditions of the transient-free cases *A* and *B* lead to very similar models, the increased complexity of the complete dynamics reduces the similarity between the EIMs.

6.5 Conclusion

In this paper, we discussed the simplified formulations of an eight-dimensional consumer-resource model obtained with the new MAGER technique. The model reduction and analysis was performed for two time series with different initial values. An additional distinction was made between the complete datasets including the transient dynamics and the periodic oscillating regimes alone. The dynamics of the original system could be reproduced to a high degree using two state variables only. Because of the method's use of low-dimensional mappings, it was further possible to reconstruct the original biological variables from the simplified descriptions.

The new effective consumer variables of the reduced models incorporated community-level descriptions of the competitive DMIs. Instead of single species, they described community resource usage characteristics which were related to tradeoffs in resource requirements and resource quotas. Accordingly, the oscillatory species dynamics of the CR model was translated to cyclic shifts in community composition representing different states of resource usage in the new Ecological Interaction Models (EIM).

The best EIMs could be interpreted as nonlinear oscillators. This formal correspondence between physical and biological systems has already been discussed earlier (Ginzburg, 1986; Vandermeer, 1993, 2004; Gertsev et al., 2008) as well as in our recent studies (Bernhardt, 2007; Bernhardt and Wirtz, 2007) and in the first part of the paper. In one of the investigated cases, the separate analysis of oscillatory and transient dynamics further revealed different underlying principles governing the two dynamic regimes. According to the reduced EIM, the CR dynamics could be described by an initial competition-free growth phase and the subsequent dynamic switching between different states of resource usage due to mutual limitations. Furthermore, some of the reduced models incorporated ad-

ditional external factors which were responsible for the reconstruction of the initial oscillatory states. Thus, these factors bypassed the use of additional state variables capturing the missing usage characteristic of the third resource.

The relative importance of direct and indirect interactions for the occurrence of species coexistence cannot be deduced from this model study as no direct interactions, for example predator-prey relations, are present in the CR model. However, the application of MAGER to a simple predator-prey model showed that the dynamics could similarly be described by interaction-related variables. We therefore propose that the EIM formulation is a very general and effective description of biological coexistence.

The discussion of the reduction results of the complete dataset $\widehat{\mathbf{X}}^b$ also indicated some limitations of the MAGER approach. In particular, it may be expected that more complicated dynamics can only be captured with more complex learned models which may be difficult to interpret in simple biological terms. The evaluation of the maximum model complexity limiting the suitable application of MAGER has still to be done. In further studies, we will also investigate the relative importance of direct and indirect interactions as well as the interplay of trait- and density-mediated effects. This will include the reduction of models incorporating species traits as additional state variables as well as direct species interactions. Aiming at the derivation of suitable data-generating processes, the MAGER approach will further be applied to measured datasets.

Summing up the discussions of the last chapters, the MAGER scheme developed in this thesis has been shown to be a promising automatic model reduction method which also offers new possibilities to examine and understand complex (biological) systems in process-based terms. Regarding the model reduction performance of MAGER, I just highlight its good reproduction capabilities (e.g. of transient dynamics) and high reduction potential in comparison with other methods. The data-based change to a higher-level descriptive "language" and new model structures is a central property of MAGER in this context. Apart from these technical benefits, the application to biological networks further indicated that new information on important system properties can be gained with MAGER. For example, the explicit incorporation of species inter-relations in the reduced systems of chapter 6 can be seen as an emergence of dominant processes as these interactions were not given explicitly in the original model.

However, the method has its own drawbacks, some of which have been addressed in the last chapters, and a number of extensions and improvements of MAGER can be pointed out. The first point of criticism addresses the general applicability of MAGER. As the method is based on low-dimensional model dynamics/attractors embedded in high-dimensional variable spaces, its reduction potential is confined to the existence and dimensionality of such structures. It should also be noticed that an increase of the number of state variables drastically increases the size of the search space for model learning. Thus, the reducibility of model datasets using the proposed combination of linear and non-linear mappings must be re-evaluated for any new application. However, as outlined in section 3.10, there is some evidence that many models produce low-dimensional dynamics despite their own complexity. As shown in chapter 5, the check for reducibility should thereby also be extended to parts of the dynamics and submodels within the complex over-all structure. It is, in fact, improbable that reduced models found with MAGER will be able to replace large ecosystem models like ERSEM. Instead, the method should be seen as an analysis tool for large-scale models which can help to give alternative explanations for occurring (sub-)dynamics.

It may also be criticized that the method produces no single "universal" model which completely reproduces the bifurcation behavior of the original system (see section 4.4.4). However, remembering the equifinality discussion in chapter 1, we could state that the aim of finding a single "optimal" model is ill-posed and that the comparative use of multiple "minimum-realistic" models should be recom-

mended (Fulton et al., 2003). Nevertheless, the argument could further be weakened by refining the structure optimization scheme of the MAGER approach. In order to account for different dynamic situations, a multi-objective extension of the GA optimization could be applied to find optimized parameter sets for multiple time series. However, the successful use of such a technique requires solutions for a number of new problems which are, for example, related to the normalization and (non-)linear state reduction of multiple time series.

It has further been discussed in chapter 3 that parameter tuning of the hybrid GP/GA module is very difficult to perform because of the long operation times of MAGER and the diversity of the GP models. Nevertheless, the optimization procedures can eventually be improved by investigating the effects of these control parameters on the over-all reduction performance in more detail. In addition, the use of more sophisticated versions of the genetic procedures could enhance the model learning capabilities of MAGER. This includes, for example, Pareto-optimal GA (Coello Coello, 2000) or modified global/local search hybrids (e.g. Goldberg and Voessner, 1999).

Extensions of MAGER should also include other quality measures to account for a variety of dynamic situations. As shown in the preceding chapters, the appropriate type of fitness measure depends on the desired focus on dynamic details and ranges from unspecific, e.g. the RMSE, to very specific for a given situation, e.g. the combination of transient and limit cycle dynamics (chapter 5). Further examples for unspecific fitness measures derive from information theory, such as the Akaike or Bayesian Information Criterion (AIC,BIC) which have been used in model selection applications (Cox et al., 2006). The selection of appropriate fitness measures is certainly a major crucial aspect of successful applications of the new reduction method.

Apart from more complex model reduction studies, MAGER can also be used for data-based model building. Future applications will, for example, be based on measured data from chemostat experiments (e.g. Kooijman, 2000; Fussmann et al., 2000; Becks et al., 2005). In this respect, the model learning capabilities of MAGER can be improved by incorporating new state variable components of the GP trees. Two extensions of this kind, which are connected to unobserved variables and external forcing, have already been implemented in MAGER but have not been used in applications yet. I suppose that the addition of unobserved state variables to the new ODE formulations allows for a treatment of classic traits in biological systems. The incorporation of external forcing, i.e. variables which are only used as input for the learned dynamic systems, is another important aspect for model building.

It is obvious that the mapping-based reduction step of MAGER is of minor importance for model building applications. Even without this step, however, some of these new features, such as the incorporation of unobserved (trait-)variables in combination with automatic functional interpretations, still offer new fields of application in comparison with related GP-based model building studies (e.g. Cao et al., 2000).

A Appendix of Chapter 4

A.1 Tables

Table A.1: Simulation settings for the predator-prey model (3.15).

$[x_1, x_2]_{t=0}$	[2, 1.5]
Time t	[0, 0.5, ..., 200]
Normalized inverse capacity of prey growth k	0.15, 0.09
Normalized inverse half saturation for predation s	0.5
Normalized predator mortality m_p	1

A.2 Normalization of the Rosenzweig-McArthur model

Rosenzweig and McArthur (1963) introduced a predator-prey model with capacity limited growth and Holling type II grazing term:

$$\begin{aligned} \frac{dP}{dt} &= r \cdot \left(1 - \frac{P}{K}\right) \cdot P - \mu \cdot \frac{P}{P+S} \cdot Z \\ \frac{dZ}{dt} &= \epsilon \mu \cdot \frac{P}{P+S} \cdot Z - M \cdot Z, \end{aligned} \quad (8.1)$$

where N and P denote prey and predator abundance, respectively (see Table A.4 for parameter descriptions).

Three of the six model parameters of (8.1) can be eliminated by normalization of N , P and time, resulting in an equivalent dimensionless system. Using inverse values of the capacity $\beta = 1/K$ and half saturation constant $\kappa = 1/S$ and introducing normalization constants

$$t_n = \frac{1}{r}, \quad N_n = \frac{M}{\epsilon \mu \kappa}, \quad P_n = \frac{r}{\mu \kappa}, \quad (8.2)$$

Table A.2: Simulation settings of model (4.2)-(4.3). The parameter values slightly differ from the ones used by Yoshida et al. (2003) to account for the change in model formulation (section 4.2.2). Note that the dimensions of phytoplankton and zooplankton for the NPZ model are individual cells counted and phytoplankton values relate to 10^9 cells.

Parameter	Description	Value
$[N, P, Z]_{t=0}$	Initial values of state variables	[0.45, 3.2, 3]
t	Time steps of the simulation [days]	[0, 0.1, ..., 200]
δ	Dilution rate [1/d]	0.25
N_0	Nutrient concentration of the inflow [$\mu\text{mol}/l$]	242.4
ϵ_P	Conversion efficiency of primary producers [$10^9\text{cells}/\mu\text{mol}$]	0.05
μ_{max}	Max. nutrient growth rate of algae [$\mu\text{mol}/(d \cdot 10^9\text{cells})$]	66.0
S_P	Half saturation constant of nutrient uptake [$\mu\text{mol}/l$]	13
ξ_{max}	Maximum clearance rate of Z [1/d]	$3.3 \cdot 10^{-4}$
S_Z	Half saturation constant of grazing [$10^9\text{cells}/l$]	0.88
ϵ_Z	Conversion efficiency of grazers	5400
P^*	Critical phytoplankton concentration [$10^9\text{cells}/l$]	1.324
m_z	Mortality of zooplankton [1/d]	0.455

Table A.3: Parameter values and RMSE of the reduced models learned. $\varphi_i(0)$ denote optimized initial values of the effective variables.

Model	c_1	c_2	c_3	c_4	$\varphi_1(0)$	$\varphi_2(0)$	RMSE
(3.19) with $k = 0.15$	2.627	0.288	7.779	0.192	-1.885	—	0.32
(3.19) with $k = 0.09$	1.266	0.662	8.401	1	1.853	—	0.51
(4.11)(4.12)	9.274	2	3.296	29.15	0.355	3.85	0.53

system (8.1) can be rewritten in the form of (3.15) for normalized abundances of prey $x_1 = N/N_n$ and predator $x_2 = P/P_n$ using the normalized parameters

$$k = \beta N_n = \frac{MS}{\epsilon\mu K}, \quad (8.3)$$

$$s = \kappa N_n = \frac{M}{\epsilon\mu} \quad \text{and} \quad (8.4)$$

$$m_p = Mt_n = \frac{M}{r}. \quad (8.5)$$

Table A.4: Description of parameters of the original Rosenzweig-McArthur model (8.1).

r	Maximal growth rate of prey P
K	Carrying capacity
μ	Maximal ingestion rate of predator Z
S	Half saturation constant for predation
ϵ	Assimilation efficiency of predation
M	Predator mortality

A.3 Derivation of the pendulum equation for (3.15)

In the vicinity of the non-trivial fixed point $\{x_1^*, x_2^*\}$ of (3.15), an approximate transformation into a pendulum equation can be reached analytically. For

$$\frac{dx_1}{dt} = F_1(x_1, x_2), \quad \frac{dx_2}{dt} = F_2(x_1, x_2), \quad (8.6)$$

with F_1 and F_2 given in (3.15) and the steady-state values x_1^* and x_2^* of predator and prey densities,

$$x_1^* = \frac{1}{1-s}, \quad x_2^* = \frac{1-k-s}{(1-s)^2}, \quad (8.7)$$

we follow the temporal evolution of a small perturbation around the equilibrium, i.e. $x_i = x_i^* + x'_i$, $i = 1, 2$. A first-order Taylor approximation of (3.15) leads to

$$\begin{aligned} \frac{d}{dt}x'_1 &\approx \left[\frac{\partial}{\partial x_1} F_1 \right]_{x_1^*, x_2^*} \cdot x'_1 + \left[\frac{\partial}{\partial x_2} F_1 \right]_{x_1^*, x_2^*} \cdot x'_2 \\ &= f_1 \cdot x'_1 - x'_2, \end{aligned} \quad (8.8)$$

with $f_1 = k + s - (2 \cdot k)/(1-s)$ and

$$\begin{aligned} \frac{d}{dt}x'_2 &\approx \left[\frac{\partial}{\partial x_1} F_2 \right]_{x_1^*, x_2^*} \cdot x'_1 + \left[\frac{\partial}{\partial x_2} F_2 \right]_{x_1^*, x_2^*} \cdot x'_2 \\ &= f_2 \cdot x'_1, \end{aligned} \quad (8.9)$$

with $f_2 = m_p \cdot (1-k-s)$.

Note that f_1 and f_2 are not altered by normalization of x_1 and x_2 . We obtain the oscillator equation for a small deviation φ' of the effective variable $\varphi = p_1 \tilde{x}_1 + p_2 \tilde{x}_2$ by differentiation of φ' using (8.8) and (8.9):

$$\frac{d^2 \varphi'}{dt^2} \approx -f_2 \cdot \varphi' + f_1 \cdot \frac{d\varphi'}{dt}. \quad (8.10)$$

B Appendix of Chapter 5

B.1 Parameter settings of the competition model

The parameter setting of the competition model follows the setup for five species and three resources of Huisman and Weissing (2001b). The matrix of half-saturation constants follows

$$\mathbf{K} = \begin{pmatrix} 0.2 & 0.05 & 1.0 & 0.05 & 1.2 \\ 0.2 & 0.1 & 0.05 & 1.0 & 0.4 \\ 0.15 & 0.95 & 0.35 & 0.1 & 0.05 \end{pmatrix}. \quad (8.11)$$

The entries of \mathbf{K} are individual half saturation constants k_{ji} of species i (in columns) and nutrients j (in rows). The initial values of the species and nutrients used are $N_j(0) = \{10, 10, 10\}$ and $P_i(0) = \{0.1, 0.1, 0.1\}$ resulting in a dominance of species 1,4 and 5 and $P_i(0) = \{0.1, 0.5, 0.1\}$ for a dominance of species 1,2 and 3. All other parameter values are left unaltered, i.e. growth rate $\mu_i = 1$ and mortality $\omega_i = 0.25$ of species i as well as nutrient turnover rate $D = 0.25$ and supply concentration of nutrient j , $S_j = 10$. The matrix of nutrient quotas \mathbf{C} with entries c_{ji} denoting the content of nutrient j in species i is given by

$$\mathbf{C} = \begin{pmatrix} 0.2 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.1 & 0.2 & 0.1 & 0.1 & 0.2 \\ 0.1 & 0.1 & 0.2 & 0.2 & 0.1 \end{pmatrix}. \quad (8.12)$$

B.2 Tables

Table B.1: Parameter values of the reduced models learned. $\varphi_x(0)$ and $\varphi_y(0)$ denote the optimized initial values of the first and second or linear and nonlinear effective variables, respectively.

Model	p_1	p_2	p_3	p_4	p_5	p_6	$\varphi_x(0)$	$\varphi_y(0)$
(5.4), Φ^A	1	-8.01	0.25	1.6	-	-	5.23	6.99
(5.4), Φ^B	0.34	-1.42	1	6.03	-	-	3.85	6.04
(5.5)	10	1	3.1	-	-	-	1.58	-9.86
(5.6)	3.12	5.9	-	-	-	-	3.6	3.97
(5.7)	9.19	6.54	0.46	0.24	0.04	-	-9.62	-8.25
(5.8)	0.05	0.11	9.76	4.88	5.03	2.35	-1.5	-2.66

B.3 Quality measures

Dimensionality reduction

We use two measures to quantify the appropriate dimensionality of the reduced datasets. First, the dimensionality of different temporal regimes can be determined by calculating the time course of the correlation dimension D_C (Baker and Gollub, 1990; Theiler, 1986). $D_C(\mathbf{X}_s(t))$ is calculated for a section \mathbf{X}_s with length n_s starting at t according to

$$D_C(\mathbf{X}_s(t)) = \lim_{\epsilon \rightarrow 0} \frac{\log(C(\epsilon))}{\log(\epsilon)}, \quad \text{with} \quad C(\epsilon) = \frac{n_\epsilon}{n_s^2}. \quad (8.13)$$

Here, $C(\epsilon)$ denotes the correlation integral, $n_s = 1000$ and n_ϵ is the number of points with distance smaller than ϵ from each other. D_C is calculated for subsequent values of t by fitting a straight line to the slope of the double logarithmic plot of $C(\epsilon)$ versus ϵ obtained for varied small values of ϵ .

As a second measure, the fraction of explained variance (FEV) is used to select the number of retained principal components which can be linear or non-linear for dimensionality reduction. The FEV for the linear PCA is given by the ratio of individual eigenvalues of the data covariance matrix to the sum of all eigenvalues or the ratio of the squared singular values if singular value decomposition is used to calculate the PCs, respectively (Cattell, 1966). For NLPCA, the FEV (FEV^{nl}) is calculated according to

$$FEV^{nl} = 1 - \frac{\sum \|\hat{\mathbf{Y}}^s - \hat{\mathbf{X}}^s\|^2}{\sum \|\hat{\mathbf{X}}^s\|^2}, \quad (8.14)$$

where $s = a, b$ and $\hat{\mathbf{Y}}^s$ denotes the centered reconstructed datasets (Hsieh, 2004). A critical value of $FEV_c \approx 90\%$ is used to select the appropriate number of principal components. As outlined in sections 3.7.2 and 5.4, the FEV values of the linear PCA confirm the results on data dimensionality obtained by calculating the correlation dimension.

Model learning

In (Bernhardt, 2007), the reproduction error of the individual learned models was measured using a combination of the mean root-mean-square error (RMSE) of the normalized reduced time series as goodness-of-fit measure and the number of GP nodes as a measure of model complexity. However, for time series showing oscillatory regimes with different frequencies, see for example the transient and limit cycle dynamics of \mathbf{X}^a and \mathbf{X}^b (Figure 5.1), the RMSE may not be the most appropriate choice. Due to the relatively large errors which may arise even from small phase shifts between the data and simulated time series we hence propose a combined measure which is sensitive to the course of the transient and the amplitude of the oscillations but is insensitive to the frequency and phase of the limit

cycle regime. In this measure, the RMSE is used to quantify the approximation error of the transient phase, while the error of the oscillatory regime (E_p) is calculated as

$$E_p = \frac{1}{l \cdot n_o} \sum_{i=1}^l \sum_{j=1}^{n_o} \left| \phi_{ij}^{opt} - \overline{\varphi_i^{opt}} \right|. \quad (8.15)$$

Here, l is the number of reduced state variables, n_o denotes the number of local optima of the simulated oscillations, ϕ_{ij}^{opt} is the j th maximum or minimum of the effective variable approximation for φ_i , respectively, and $\overline{\varphi_i^{opt}}$ is the mean amplitude of all local maxima or minima of φ_i . For fitness calculations, the average of E_p for the maxima and minima of the time series is used. The measure thus captures the mean deviation from the lower and upper bounds of the data oscillations. If the number of maxima and minima is less than two, E_p is set to 9999 as a model which shows no oscillations is considered as a bad approximation of the original system.

The separation of transient and limit cycle phase is thereby set at $t_s = 80$ for both datasets, i.e. Φ^a and Φ^b . The RMSE is calculated up to this threshold and E_p is used beyond. The results of the two regimes are then combined using the weighted average ranking method (Bentley and Wakefield, 1997) with equal weights which is also used in the GP module (Bernhardt, 2007). For the transient-free datasets Φ^A and Φ^B , the RMSE is used for the whole time series as no frequency and phase changes occur in these cases.

C Appendix of Chapter 6

C.1 Tables

Table C.1: Parameter values for the algebraic interpretations of the effective model variables.

Equation	u_1	u_2	u_3	v_1	v_2	v_3	w_1	w_2
(6.1)	0.77	0.62	-	0.84	0.46	0.27	-	-
(6.2)	0.8	0.54	0.22	0.76	0.61	0.21	-	-
(6.7)-(6.9)	0.7	0.55	0.3	0.66	0.59	0.42	0.8	0.55
(6.10)	0.05	0.12	-	-	-	-	-	-
(6.11)-(6.12)	0.8	0.54	-	0.11	0.004	0.03	-	-

Table C.2: Species-specific nutrient usages resulting from the settings of half saturation constants and nutrient quotas in the original CR model (Bernhardt and Wirtz, 2008). For P_1 , the brackets indicate that this species is a weak competitor for N_3 compared to P_5 but also to P_4 .

Species	P_1	P_2	P_3	P_4	P_5
competes for	(N_3)	N_1	N_2	N_1	N_3
gets limited by	N_1	N_2	N_3	N_3	N_2

Table C.3: Parameter values of the reduced models learned. $\varphi_x(0)$ and $\varphi_y(0)$ denote the optimized initial values of the effective consumer variables. For equations (6.13)-(6.16), $\varphi_x(0)$ refers to the linear and $\varphi_y(0)$ to the nonlinear ECV.

Model	p_1	p_2	p_3	p_4	p_5	p_6	$\varphi_x(0)$	$\varphi_y(0)$
(6.6)	3.12	5.9	-	-	-	-	3.6	3.97
(6.5)	10	1	3.1	-	-	-	1.58	-9.86
(6.13)(6.14)	9.19	6.54	0.46	0.24	0.04	-	-9.62	-8.25
(6.15)(6.16)	0.05	0.11	9.76	4.88	5.03	2.35	-1.5	-2.66

Bibliography

- Abrams, P. A. (1995). Implications of dynamically variable traits for identifying, classifying and measuring direct and indirect effects in ecological communities. *American Naturalist*, 146:112–134.
- Abrams, P. A. (2006). The prerequisites for and likelihood of generalist-specialist coexistence. *American Naturalist*, 167:329–342.
- Abrams, P. A., Menge, B. A., Mittelbach, G. G., Spiller, D., and Yodzis, P. (1996). The role of indirect effects in food webs. In Polis, G. and Winemiller, K., editors, *Food Webs: Integration of Patterns and Dynamics*, pages 371–395. Chapman & Hall, New York.
- Aguierre, L. A., Freitas, U. S., Letellier, C., and Maquet, J. (2001). Structure-selection techniques applied to continuous-time nonlinear models. *Physica D*, 158:1–18.
- Anderson, T. R. (2005). Plankton functional type modelling: running before we can walk? *Journal of Plankton Research*, 27:1073–1081.
- Ando, S., Sakamoto, E., and Iba, H. (2002). Evolutionary modeling and inference of gene network. *Information Sciences*, 145:237–259.
- Antoulas, A. C., Sorensen, D. C., and Gugercin, S. (2001). A survey of model reduction methods for large-scale systems. *Contemporary Mathematics*, 280:193–219.
- Auger, P. and Poggiale, J.-C. (1996a). Aggregation and emergence in hierarchically organized systems: population dynamics. *Acta Biotheoretica*, 44:301–316.
- Auger, P. and Poggiale, J.-C. (1996b). Emergence of population growth models: fast migration and slow growth. *Journal of Theoretical Biology*, 182:99–108.

- Auger, P. and Poggiale, J.-C. (1998). Aggregation and Emergence in systems of ordinary differential equations. *Mathematical and Computer Modelling*, 27:1–21.
- Baker, G. L. and Gollub, J. P. (1990). *Chaotic Dynamics*. Cambridge University Press, Cambridge.
- Baraldi, A. and Alpaydin, E. (2002). Constructive feedforward ART clustering networks - Part II. *IEEE Transactions on Neural Networks*, 13:662–677.
- Baretta, J., Ebenhöh, W., and Ruardij, P. (1995). The European Regional Seas Ecosystem Model, a complex marine ecosystem model. *Netherlands Journal of Sea Research*, 33:233–246.
- Bauer, H.-U. and Pawelzik, K. R. (1992). Quantifying the neighborhood preservation of Self-Organizing Feature Maps. *IEEE Transactions on Neural Networks*, 3:570–579.
- Beckerman, A. P., Uriarte, M., and Schmitz, O. J. (1997). Experimental evidence for a behavior-mediated trophic cascade in a terrestrial food chain. *Proceedings of the National Academy of Sciences USA*, 94:10735–10738.
- Becks, L., Hilker, F. M., Malchow, H., Jürgens, K., and Arndt, H. (2005). Experimental demonstration of chaos in a microbial food web. *Nature*, 435:1226–1229.
- Bentley, P. J. and Wakefield, J. P. (1997). Finding acceptable solutions in the pareto-optimal range using multiobjective genetic algorithms. In Chawdhry, P. K., Roy, R., and Pant, R. K., editors, *Soft Computing in Engineering Design and Manufacturing*, volume 5, pages 231–240, London. Springer.
- Bernhardt, K. (2007). Finding alternatives and reduced formulations for process-based models. *Evolutionary Computation*, accepted.
- Bernhardt, K. and Wirtz, K. W. (2007). Oscillation of structural properties in foodweb models. *Theoretical Biology*, submitted.
- Bernhardt, K. and Wirtz, K. W. (2008). Explicit formulation of indirect interactions in a reduced consumer-resource competition model. Part I: Reducing model complexity. *Theoretical Population Biology*, in prep.
- Beven, K. (2001). On modelling as collective intelligence. *Hydrological Processes*, 15:2205–2207.
- Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, 320:18–36.
- Brooks, R. J., Semenov, M. A., and Jamieson, P. D. (2001). Simplifying Sirius: sensitivity analysis and development of a meta-model for wheat yield prediction. *European Journal of Agronomy*, 14:43–60.

- Bruggeman, J. and Kooijman, S. A. L. M. (2007). A biodiversity-inspired approach to aquatic ecosystem modeling. *Journal of Limnology and Oceanography*, 52:1533–1544.
- Cao, H., Kang, L., Chen, Y., and Yu, J. (2000). Evolutionary Modeling of systems of Ordinary Differential Equations with Genetic Programming. *Genetic Programming and Evolvable Machines*, 1:309–337.
- Carreira-Perpiñán, M. Á. (1997). A review of dimension reduction techniques. Technical Report CS-96-09, University of Sheffield.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1:245–276.
- Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, 31:343–366.
- Coello Coello, C. A. (2000). An updated survey of GA-based multiobjective optimization techniques. *ACM Computing Surveys*, 32:109–143.
- Costanza, R. and Sklar, F. (1985). Articulation, accuracy and effectiveness of mathematical models: a review of freshwater wetland applications. *Ecological Modelling*, 27:45–68.
- Cox, G. M., Gibbons, J. M., Wood, A. T. A., Craigon, J., Ramsden, S. J., and Crout, N. M. J. (2006). Towards the systematic simplification of mechanistic models. *Ecological Modelling*, 198:240–246.
- de Villemagne, C. and Skelton, R. E. (1987). Model reduction using a projection formulation. *International Journal of Control*, 46:2141–2169.
- DeAngelis, D. L. and Rose, K. A. (1992). Which individual-based approach is most appropriate for a given problem? In DeAngelis, D. L. and Gross, L. J., editors, *Individual-based models and approaches in ecology: populations, communities and ecosystems*, pages 67–87. Chapman & Hall, New York.
- Dieckmann, U. and Law, R. (1996). The dynamical theory of coevolution: a derivation from stochastic ecological processes. *Journal of Mathematical Biology*, 34:579–612.
- Ebenhöh, W. (1996). Stability in models versus stability in real ecosystems. *Senckenbergiana Maritima*, 27:251–254.
- Eshelman, L. J. and Schaffer, J. D. (1992). Real-coded Genetic Algorithms and Interval-Schemata. In Whitley, L. D., editor, *Foundations of Genetic Algorithms*, volume 2, pages 187–202. Morgan Kaufmann, San Mateo, CA.

- Fahse, L., Wissel, C., and Grimm, V. (1998). Reconciling classical and individual-based approaches in theoretical population ecology: a protocol for extracting population parameters from individual-based models. *American Naturalist*, 152:838–852.
- Flynn, K. J. (2003). Modelling multi-nutrient interactions in phytoplankton; balancing simplicity and realism. *Progress in Oceanography*, 56:249–279.
- Fodor, I. K. (2002). A survey of dimension reduction techniques. Technical Report UCRL-ID-148494, Lawrence Livermore National Laboratory.
- Franks, P. J. S. (2002). NPZ models of plankton dynamics: their construction, coupling to physics and application. *Journal of Oceanography*, 58:379–387.
- Freedman, H. I. (1980). *Deterministic Mathematical Models in Population Ecology*. Marcel Dekker, Inc., New York.
- Friedrichs, M. A. M., Hood, R. R., and Wiggert, J. D. (2006). Ecosystem model complexity versus physical forcing: Quantification of their relative impact with assimilated Arabian Sea data. *Deep-Sea Research II*, 53:576–600.
- Fukunaga, K. and Koontz, W. (1970). Application of Karhunen-Loeve Expansion to Feature Selection and Ordering. *IEEE Transactions on Computers*, 311.
- Fulton, E. A., Smith, A. D. M., and Johnson, C. R. (2003). Effect of complexity on marine ecosystem models. *Marine Ecology Progress Series*, 253:1–16.
- Fussmann, G. F., Ellner, S. P., Hairston Jr., N. G., Jones, L. E., Shertzer, K. W., and Yoshida, T. (2005). Ecological and evolutionary dynamics of experimental plankton communities. *Advances in Ecological Research*, 37:221–243.
- Fussmann, G. F., Ellner, S. P., Shertzer, K. W., and Nelson Jr., G. H. (2000). Crossing the Hopf Bifurcation in a Live Predator-Prey System. *Science*, 290:1358–1360.
- Gause, C. F. (1934). *The struggle for existence*. Williams & Wilkins, Baltimore.
- Gertsev, V. I., Gertseva, V. V., and Matson, S. E. (2008). Population as an oscillating system. *Ecological Modelling*, 210:242–246.
- Ginzburg, L. R. (1986). The theory of population dynamics: I. Back to first principles. *Journal of theoretical Biology*, 122:385–399.

- Goldberg, D. E. and Voessner, S. (1999). Optimizing global-local search hybrids. In Banzhaf, W., Daida, J., Eiben, A. E., Garzon, M. H., Honavar, V., Jakiela, M., and Smith, R. E., editors, *GECCO-99: Proceedings of the Genetic and Evolutionary Computation Conference*, volume 1, pages 220–228, Orlando, Florida. Morgan Kaufmann.
- Golub, G. H. and Van Loan, C. F. (1989). *Matrix Computations*. Johns Hopkins University Press, Baltimore, 2nd edition.
- Goodman, N. (1983). *Fact, Fiction, and Forecast*. Harvard University Press, Cambridge, MA, 4th edition.
- Grimme, E. J. (1997). *Krylov projection methods for model reduction*. PhD thesis, Univ. Illinois, Urbana-Champaign, IL.
- Hastie, T. and Stuetzle, W. (1989). Principal curves. *Journal of the American Statistical Association*, 84:502–516.
- Hindmarsh, A. C., Brown, P. N., Grant, K. E., Lee, S. L., Serban, R., Shumaker, D. E., and Woodward, C. S. (2005). SUNDIALS: Suite of Nonlinear and Differential/Algebraic Equation Solvers. *ACM Transactions on Mathematical Software*, 31:363–396.
- Holling, C. S. (1959). The components of predation as revealed by a study of small-mammal predation of the European pine sawfly. *Canadian Entomology*, 91:293–320.
- Håkanson, L. (1995). Optimal size of predictive models. *Ecological Modelling*, 78:195–204.
- Hsieh, W. W. (2001). Nonlinear Principal Component Analysis by Neural Networks. *Tellus*, 53A:599–615.
- Hsieh, W. W. (2004). Nonlinear multivariate and time series analysis by neural network methods. *Review of Geophysics*, 42. RG1003.
- Huisman, J. and Weissing, F. J. (1999). Biodiversity of plankton by species oscillations and chaos. *Nature*, 402:407–410.
- Huisman, J. and Weissing, F. J. (2001a). Biological conditions for oscillations and chaos generated by multispecies competition. *Ecology*, 82:2682–2695.
- Huisman, J. and Weissing, F. J. (2001b). Fundamental unpredictability in multispecies competition. *American Naturalist*, 157:488–493.
- Hutchinson, G. E. (1961). The paradox of the plankton. *American Naturalist*, 45:137–145.

- Iwasa, Y., Levin, S. A., and Andreasen, V. (1989). Aggregation in model ecosystems II. Approximate aggregation. *IMA Journal of Mathematics Applied in Medicine & Biology*, 6:1–23.
- Kégl, B., Krzyzak, A., Linder, T., and Zeger, K. (2000). Learning and design of Principal Curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:281–297.
- Kennel, M. B., Brown, R., and Abarbanel, H. D. I. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A*, 45:3403–3411.
- Kessler, M. and Haynes, T. (1999). Depth-Fair Crossover in Genetic Programming. In *Proceedings of the ACM Symposium on Applied Computing*, pages 319–323.
- Kiviluoto, K. (1996). Topology preservation in Self-Organizing Maps. In *Proceedings of IEEE International Conference on Neural Networks*, volume 1, pages 294–299.
- Köhler, P. and Wirtz, K. W. (2002). Linear understanding of a huge aquatic ecosystem model using a group-collecting sensitivity study. *Environmental Modelling & Software*, 17:613–625.
- Kohlmeier, C. and Ebenhöf, W. (1995). The stabilizing role of cannibalism in a predator-prey system. *Bulletin of Mathematical Biology*, 57:401–411.
- Kohonen, T. (1997). *Self-organizing maps*. Springer, Berlin, 2nd edition.
- Kooi, B. W., Poggiale, J. C., and Auger, P. (1998). Aggregation methods in food chains. *Mathematical and Computer Modelling*, 27:109–120.
- Kooijman, S. A. L. M. (2000). *Dynamic energy and mass budgets in biological systems*. Cambridge University Press, Cambridge.
- Kot, M. (2001). *Elements of Mathematical Ecology*. Cambridge University Press, Cambridge.
- Koza, J. R. (1992). *Genetic Programming: on the programming of computers by means of natural selection*. MIT Press, Cambridge, MA.
- Kramer, M. A. (1991). Nonlinear principal component analysis using autoassociative neural networks. *AIChE Journal*, 37:233–243.
- Krasnov, B. R., Mouillot, D., Shenbrot, G. I., Khokhlova, I. S., and Poulin, R. (2005). Abundance patterns and coexistence processes in communities of fleas parasitic on small mammals. *Ecography*, 28:453–464.
- Křivan, V. and Schmitz, O. J. (2004). Trait and density mediated indirect interactions in simple food webs. *Oikos*, 107:239–250.

- Lall, S., Krysl, P., and Marsden, J. E. (2003). Structure-preserving model reduction for mechanical systems. *Physica D*, 184:304–318.
- Lawlor, L. R. (1979). Direct and indirect effects of n-species competition. *Oecologia*, 43:355–364.
- Le Quéré, C., Harrison, S. P., Prentice, I. C., Buitenhuis, E. T., Aumont, O., Bopp, L., Claustre, H., Da Cunha, L. C., Geider, R., Giraud, X., Klaas, C., Kohfeld, K. E., Legendre, L., Manizza, M., Platt, T., Rivkin, R. B., Sathyendranath, S., Uitz, J., Watson, A. J., and Wolf-Gladrow, D. (2005). Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biology*, 11:2016–2040.
- Lee, D. B. (1973). Requiem for large-scale models. *Journal of The American Institute of Planners*, 39:163–178.
- Léon, J. A. and Tumpson, D. B. (1975). Competition between two species for two complementary or substitutable resources. *Journal of Theoretical Biology*, 50:185–201.
- Logan, J. (1994). In defense of big ugly models. *American Entomologist*, 40:202–207.
- Lotka, A. J. (1932). The growth of mixed populations: two species competing for a common food supply. *Journal of the Washington Academy of Sciences*, 22:461–469.
- Malaeb, Z. A., Summers, J. K., and Pugesek, B. H. (2000). Using structural equation modeling to investigate relationships among ecological variables. *Environmental and Ecological Statistics*, 7:93–111.
- Malthouse, E. C. (1998). Limitations of nonlinear PCA as performed with generic neural networks. *IEEE Transactions on Neural Networks*, 9:165–173.
- Martinetz, T. M. and Schulten, K. J. (1991). A "neural-gas" network learns topologies. In *Artificial Networks*, pages 397–402. North-Holland, Amsterdam.
- Matear, R. J. (1995). Parameter optimization and analysis of ecosystem models using simulated annealing: a case study at station p. *Journal of Marine Research*, 53:571–607.
- Matthies, H. G. and Meyer, M. (2003). Nonlinear Galerkin Methods for the Model Reduction of Nonlinear Dynamical Systems. *Computers and Structures*, 81:1277–1286.
- May, R. M. (1977). Thresholds and breakpoints in ecosystems with a multiplicity of state variables. *Nature*, 269:471–477.
- McArthur, R. H. (1972). *Geographical ecology*. Harper and Row, New York, NY.

- Menge, B. A. (1995). Indirect effects in marine rocky intertidal interaction webs: patterns and importance. *Ecological Monographs*, 65:21–74.
- Mühlenbein, H. and Schlierkamp-Voosen, D. (1993). Predictive models for the Breeder Genetic Algorithm, I: continuous parameter optimization. *Evolutionary Computation*, 1:25–49.
- Nayfeh, A. H. (1973). *Perturbation Methods*. John Wiley, New York.
- Pahl-Wostl, C. (1997). Dynamic structure of a food web model: comparison with a food chain model. *Ecological Modelling*, 100:103–123.
- Palmer, J. R. and Totterdell, I. J. (2001). Production and export in a global ecosystem model. *Deep-Sea Research I*, 48:1169–1198.
- Pearson, K. (1901). On lines and planes of closest fit to a system of points in space. *Philosophical Magazine*, 2:557–572.
- Persson, A., Hansson, L.-A., Brönmark, C., Lundberg, P., Pettersson, L. B., Greenberg, L., Nilsson, P. A., Nyström, P., Romare, P., and Tranvik, L. (2001). Effects of enrichment on simple aquatic food webs. *American Naturalist*, 157:654–669.
- Pham, D. T. and Karaboga, D. (2000). *Intelligent optimization techniques*. Springer, Berlin, Heidelberg, New York.
- Phillips, J. (2000). Projection frameworks for model reduction of weakly nonlinear systems. In *Proceedings of the 37th Design Automation Conference*, pages 184–189, Los Angeles, CA.
- Preisser, E. L., Bolnick, D. I., and Benard, M. F. (2005). Scared to death? the effects of intimidation and consumption in predator-prey interactions. *Ecology*, 86:501–509.
- Raick, C., Soetaert, K., and Grégoire, M. (2006). Model complexity and performance: How far can we simplify. *Progress in Oceanography*, 70:27–57.
- Real, L. (1977). The kinetics of functional response. *American Naturalist*, 111:289–300.
- Relyea, R. A. and Yurewicz, K. L. (2002). Predicting community outcomes from pairwise interactions: integrating density- and trait-mediated effects. *Oecologia*, 131:569–579.
- Rewieński, M. J. (2003). *A Trajectory Piecewise-Linear Approach to Model Order Reduction of Nonlinear Dynamical Systems*. PhD thesis, Massachusetts Institute of Technology.
- Rinaldi, S. and Scheffer, M. (2000). Geometric analysis of ecological models with slow and fast processes. *Ecosystems*, 3:507–521.

- Roberts, A. and Stone, L. (2004). Advantageous indirect interactions in systems of competition. *Journal of Theoretical Biology*, 228:367–375.
- Rodríguez-Fernández, J. L. (1999). Ockham's razor. *Endeavour*, 23:121–125.
- Rosenzweig, M. L. and McArthur, R. H. (1963). Graphical representation and stability conditions of predator-prey interactions. *American Naturalist*, 97:209–223.
- Rowley, C. W., Colonius, T., and Murray, R. M. (2004). Model reduction for compressible flows using POD and Galerkin projection. *Physica D*, 189:115–129.
- Roy, S. and Chattopadhyay, J. (2007). The stability of ecosystems: a brief overview of the paradox of enrichment. *Journal of Biosciences*, 32:421–428.
- Schaffer, W. M. (1981). Ecological abstraction: the consequences of reduced dimensionality in ecological models. *Ecological Monographs*, 51:383–401.
- Schmitz, O. J. (1993). Trophic exploitation in grassland food chains: simple models and a field experiment. *Oecologia*, 93:327–335.
- Schölkopf, B., Smola, A., and Müller, K.-R. (1998). Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 10:1299–1319.
- Sharov, A. and Colbert, J. (1994). Gypsy Moth Life System Model. In *Integration of Knowledge and User's Guide*, page 232. Virginia Polytechnic Institute and State University, Blacksburg, VA.
- Simon, H. A. (2002). Science seeks parsimony, not simplicity: searching for pattern in phenomena. In Zellner, A., Keuzenkamp, H. A., and McAleer, M., editors, *Simplicity, Inference and Modelling: Keeping it sophisticatedly simple*. Cambridge University Press, Cambridge.
- Sirovich, L. (1987). Turbulence and the dynamics of coherent structures, I-III. *Quarterly Applied Mathematics*, 45:561–590.
- Sivakumar, B. (2007). Dominant processes concept, model simplification and classification framework in catchment hydrology. *Stochastic Environmental Research and Risk Assessment*, DOI: 10.1007/s00477-007-0183-5.
- Smirnov, D. A., Bezruchko, B. P., and Seleznev, Y. P. (2002). Choice of dynamical variables for global reconstruction of model equations from time series. *Physical Review E*, 65:026205.
- Sober, E. (2002). What is the problem of simplicity? In Zellner, A., Keuzenkamp, H. A., and McAleer, M., editors, *Simplicity, Inference and Modelling: Keeping it sophisticatedly simple*. Cambridge University Press, Cambridge.

- Sokol-Hessner, L. and Schmitz, O. J. (2002). Aggregate effects of multiple predator species on a shared prey. *Ecology*, 83:2367–2372.
- Stenseth, N. C., Falck, W., Bjørnstad, O. N., and Krebs, C. J. (1997). Population regulation in snowshoe hare and Canadian lynx: asymmetric food web configurations between hare and lynx. *Proceedings of the National Academy of Science USA*, 94:5147–5152.
- Strauss, S. Y. (1991). Indirect effects in community ecology: their definition, study and importance. *Trends in Ecology and Evolution*, 6:206–210.
- Tenenbaum, J. B., de Silva, V., and Jangford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, 290:2319–2323.
- Theiler, J. (1986). Spurious dimension from correlation algorithms applied to limited time-series data. *Physical Review A*, 34:2427–2432.
- Tilman, D. (1982). *Resource competition and community structure*. Princeton University Press, Princeton, New Jersey.
- Turchin, P. (2003). *Complex population dynamics*. Princeton University Press, New Jersey.
- Vadasz, P. and Vadasz, A. S. (2002). The neoclassical theory of population dynamics in spatially homogeneous environments. (I) Derivation of universal laws and monotonic growth. *Physica A*, 309:329–359.
- Van Nes, E. H. and Scheffer, M. (2005). A strategy to improve the contribution of complex simulation models to ecological theory. *Ecological Modelling*, 185:153–164.
- van Veen, F. J. F., van Holland, P. D., and Godfray, H. C. J. (2005). Stable coexistence in insect communities due to density- and trait-mediated indirect effects. *Ecology*, 86:1382–1389.
- Vandermeer, J. (1993). Loose coupling of predator-prey cycles: entrainment, chaos, and intermittency in the classic Mac-Arthur consumer-resource equations. *American Naturalist*, 141:687–716.
- Vandermeer, J. (2004). Coupled oscillations in food webs: Balancing competition and mutualism in simple ecological models. *American Naturalist*, 163:857–867.
- Vandermeer, J. (2006). Oscillating populations and biodiversity maintenance. *BioScience*, 56:967–975.
- Vandermeer, J. and Pascual, M. (2006). Competitive coexistence through intermediate polyphagy. *Ecological Complexity*, 3:37–43.

- Vesanto, J., Himberg, J., Alhoniemi, E., and Parhankangas, J. (1999). Self-organizing map in Matlab: the SOM toolbox. In *Proceedings of the Matlab DSP Conference*, pages 35–40, Espoo, Finland.
- Villmann, T., Der, R., Herrmann, M., and Martinetz, T. (1997). Topology preservation in Self-Organizing Feature Maps: exact definition and measurement. *IEEE Transactions on Neural Networks*, 8:1291–1303.
- Volterra, V. (1928). Variations and fluctuations of the number of individuals in animal species living together. *Journal du Conseil International pour l'Exploration de la mer*, 3:3–51.
- Wirtz, K. W. (2002). A generic model for changes in microbial kinetic coefficients. *Journal of Biotechnology*, 97:147–162.
- Wirtz, K. W. and Bernhardt, K. (2008). Effective variables in a consumer-resource model. in prep.
- Wirtz, K. W. and Eckhardt, B. (1996). Effective variables in ecosystem models with an application to phytoplankton succession. *Ecological Modelling*, 92:33–53.
- Wirtz, K. W. and Wiltshire, K. (2005). Long-term shifts in marine ecosystem functioning detected by inverse modeling of the Helgoland Roads time-series. *Journal of Marine Systems*, 56:262–282.
- Wootton, J. T. (1994). The nature of indirect effects in ecological communities. *Annual Reviews of Ecological Systems*, 25:443–466.
- Yamauchi, A. and Yamamura, N. (2005). Effects of defense evolution and diet choice on population dynamics in a one-predator-two-prey system. *Ecology*, 86:2513–2524.
- Yoshida, T., Jones, L. E., Ellner, S. P., Fussmann, G. F., and Hairston Jr., N. G. (2003). Rapid evolution drives ecological dynamics in a predator-prey system. *Nature*, 424:303–306.
- Young, P. C. and Garnier, H. (2006). Identification and estimation of continuous-time, data-based mechanistic (DBM) models for environmental systems. *Environmental Modelling & Software*, 21:1055–1072.
- Zhu, C., Byrd, R. H., and Nocedal, J. (1997). L-BFGS-B: Algorithm 778: L-BFGS-B, FORTRAN routines for large-scale bound constrained optimization. *ACM Transactions on Mathematical Software*, 23:550–560.

Acknowledgments

First of all, I would like to thank my supervisor, Kai W. Wirtz, for his scientific enthusiasm as well as the motivation and cordial support he gave me. Continuous support was also provided by my colleagues Jan Holstein, Gunnar Brandt, Cora Kohlmeier and Stefan Kotzur. We often couldn't assist each other with the details of our works but talking about them always helped to stay on top of things and eventually led to new ways to proceed. I am further grateful to Jorn Bruggeman, Bob Kooi and Bernd Blasius for helpful comments on the manuscripts. The legibility of the text was substantially improved by the contributions of Conny, Suniti, Gunnar and Adelheid who offered corrections even on short notice.

Many people supported me during the (sometimes difficult) years I worked on this thesis. I am probably unable to express how happy I feel about having good friends like that and I can only hope that they all know. I am also lucky to have wonderful parents who always give me strength and confidence even in hard times when these feelings are scarce and difficult to pass on. Support and Encouragement were also offered by Maren who was burdened with being at my side during the last months of the work and always kept smiling.

Finally, I wish to thank my brothers, Jan and Ralph, for the wonderful years we spent together and apart. You took a part of me when you left. You will never be forgotten.

Curriculum Vitae

Knut Bernhardt

geboren am 03.01.1978 in Kassel

Ledig, deutsch

- 05/2007 - 11/2007 Wissenschaftliche Hilfskraft am Institut für Küstenforschung, GKSS Forschungszentrum, Geesthacht
- 03/2003 - heute Promotionstudent am Institut für Chemie und Biologie des Meeres (ICBM), AG Mathematische Modellierung, Carl von Ossietzky Universität Oldenburg
- 09/2002 - 03/2003 Wissenschaftliche Hilfskraft am Lehrstuhl Ökologische Modellbildung, Universität Bayreuth
- 11/1997 - 09/2002 Geoökologiestudium, Universität Bayreuth
Titel der Diplomarbeit: *Zeitlokale Analysen und multifraktale Modelle zur Landschafts- und Flussnetzwerkentstehung*
- 16.06.1997 Allgemeine Hochschulreife am Oberstufengymnasium Herderschule, Kassel

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Dissertation selbstständig verfasst habe und nur die angegebenen Hilfsmittel verwendet habe. Teile der Dissertation wurden bereits veröffentlicht bzw. sind zur Veröffentlichung eingereicht, wie an den entsprechenden Stellen angegeben. Die Dissertation hat weder in Teilen noch in ihrer Gesamtheit einer anderen wissenschaftlichen Hochschule zur Begutachtung in einem Promotionsverfahren vorgelegen.

Oldenburg, den 7. März 2008

Knut Bernhardt