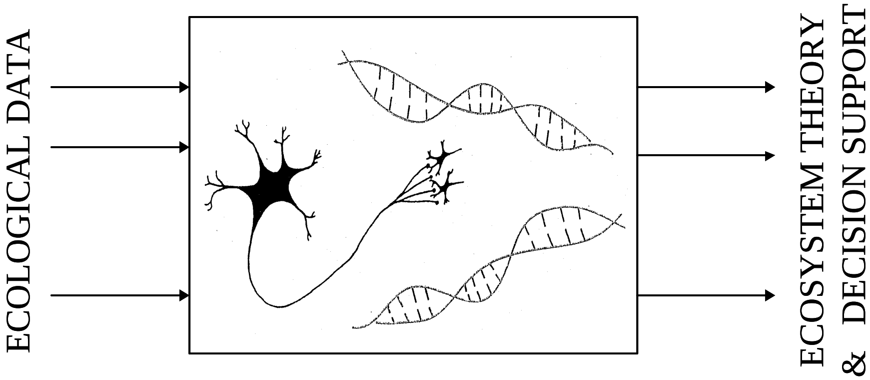
Friedrich Recknagel

Ecological Informatics

Understanding Ecology by Biologically-Inspired Computation

With 152 Figures and 43 Tables



Springer

Preface

In the 50s and 60s cross-sectional data of lake surveys were utilized for steady state assessments of the eutrophication status of lakes by univariate nonlinear regression. This *statistical approach* (see Table 1) became exemplary for river, grassland and forest models and - because of simplicity - widespread for classification of ecosystems.

In the 70s and 80s multivariate time series data were collected from ecosystems such as lakes, rivers, forests and grasslands in order to improve understanding of ecosystem dynamics. Process-based differential equations were used for the computer simulation of food web dynamics and functional group succession. This *differential equation approach* (see Table 1) is still widely used for scenario analysis.

Ecosystems analysis, synthesis and forecasting in the past ten years was very much influenced by inventions in computational technology such as high performance computing and biologically-inspired computation. This *computational approach* (see Table 1) allows to discover knowledge in complex multivariate databases for improving both ecosystem theory and decision support.

Table 1. Concepts for Ecosystems Analysis, Synthesis and Forecasting

	Statistical Regression Approach	**Differential Equations Approach**	**Computational Approach**
Ecosystem Representation	Steady States	Transitional States	Evolving States
Ecosystem Approximation	Univariate Nonlinear / Multivariate Linear	Multivariate Nonlinear	Multivariate Nonlinear
Ecosystem Complexity	Cross-Sectional Nutrient and Abundance Means	Nutrient Cycles and Food Web Dynamics	Species Succession and Ecosystem Evolution
Aquatic Examples	Phosphorus-Chlorophyll Relationship[1,2]; External P-Loading Concept[3]	AQUAMOD[4]; MS-CLEANER[5]; Bierman[6]; Jorgensen[7]; SALMO[8]	Nonlinear Regression[9]; Nonlinear PCA[10]; DELAQUA[11]; ANNA[12]; Evolved Rules[13]; Evolved Equations[14,15]; ECHO[16]; GECKO[17]
Potential Applications	Ecosystem Classification	Scenario Analysis	Ecosystem Forecasting

[1] Sakamoto M (1966) Primary production by phytoplankton community in some Japanese lakes and its dependence on lake depth. Arch. Hydrobiol. 62, 1-28

[2] Dillon P, Rigler F (1974) The phosphorus-chlorophyll relationship in lakes. Limnol.Oceanogr. 19, 135-148

[3] Vollenweider RA (1968) Scientific fundamentals of eutrophication of lakes and flowing waters with special reference to phosphorus and nitrogen. OECD, Paris. OECD/DAS/SCI/68.27

[4] Straskraba M, Gnauck A (1985) Freshwater Ecosystems: Modelling and Simulation. Elsevier, Amsterdam

[5] Park RA, O'Neill RV, Bloomfield JA, Shugart HH, Booth RS, Goldstein RA, Mankin JB, Koonce JF, Scavia D, Adams MS, Clesceri LS, Colon EM, Dettman EH, Hoopes JA, Huff DD, Katz S, Kitchell JF, Koberger RC, La Row EJ, McNaught DC, Petersohn L, Titus JE, Weiler PR, Wilkinson JW, Zahorcak CS (1974) A generalized model for simulating lake ecosystems. Simulation 33-50

[6] Bierman VJ (1976) Mathematical model of the selective enhancement of blue-green algae by nutrient enrichment. In: Canale RP (eds) Modelling Biochemical Processes in Aquatic Ecosystems. Ann Arbour Science Publishers Inc., Ann Arbour, 1-32

[7] Jorgensen SE (1976) A eutrophication model for a lake. Ecol. Modelling 2, 147-162

[8] Recknagel F, Benndorf J (1982) Validation of the ecological simulation model SALMO. Int. Revue Ges.Hydrobiol. 67, 1, 113-125

[9] Lek S, Delacoste M, Baran P, Dimonopoulos I, Lauga J, Aulagnier J (1996) Application of neural networks to modelling nonlinear relationships in ecology. Ecol. Modelling 90, 39-52

[10] Chon TS, Park YS, Moon KH, Cha EY (1996) Patternizing communities by using artificial neural network. Ecol. Modelling 90, 69-78

[11] Recknagel F, Petzoldt T, Jaeke O, Krusche F (1995). Hybrid expert system DELAQUA - a toolkit for water quality control of lakes and reservoirs. Ecol. Modelling 71, 1-3, 17-36

[12] Recknagel F (1997) ANNA - artificial neural network model predicting species abundance and succession of blue-green algae. Hydrobiologia, 349, 47-57

[13] Bobbin J, Recknagel F (2001) Knowledge discovery for prediction and explanation of blue-green algal dynamics in lakes by evolutionary algorithms. Ecol. Modelling 146, 1-3, 253-264

[14] Whigham P, Recknagel F (2001) An inductive approach to ecological time series modelling by evolutionary computation. Ecol. Modelling 146, 1-3, 275-287

[15] Whigham P, Recknagel F (2001) Predicting chlorophyll-a in freshwater lakes by hybridising process-based models and genetic algorithms. Ecol. Modelling 146, 1-3, 243-251

[16] Holland JH (1992) Adaptation in Natural and Artificial Systems. Addison-Wesley, New York

[17] Booth G (1997) Gecko: A continuous 2-D world for ecological modeling. Artif. Life 3, 147-163

The present book focuses on the *computational approach* for ecosystems analysis, synthesis and forecasting called *ecological informatics*. It provides the scope and case studies of ecological informatics exemplary for applications of biologically-inspired computation to a variety of areas in ecology.

Ecological Informatics is defined as interdisciplinary framework promoting the use of advanced computational technology for the elucidation of principles of information processing at and between all levels of complexity of ecosystems - from genes to ecological networks -, and the provision of transparent decisions targeting ecological sustainability, biodiversity and global warming.

Distinct features of ecological informatics are: data integration across ecosystem categories and levels of complexity, inference from data pattern to ecological processes, and adaptive simulation and prediction of ecosystems.

Biologically-inspired computation techniques such as fuzzy logic, artificial neural networks, evolutionary algorithms and adaptive agents are considered as core concepts of ecological informatics.

Fig. 1 represents the current scope of ecological informatics indicating that ecological data is consecutively refined to ecological information, ecosystem theory and ecosystem decision support by two basic computational operations: data archival, retrieval and visualization, and ecosystem analysis, synthesis and forecasting.

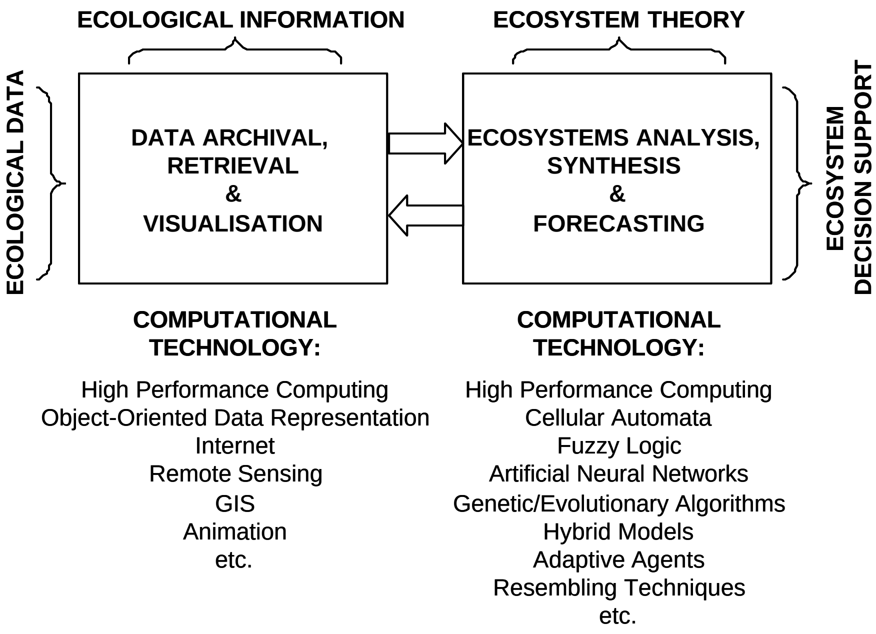


Figure 1. Scope of Ecological Informatics

Computational technologies currently considered being crucial for data archival, retrieval and visualization are:

- High performance computing to provide high-speed data access and processing, and large internal storage (RAM);
- Object-oriented data representation to facilitate data standardization and data integration by the embodiment of metadata and data operations into data structures;
- Internet to facilitate sharing of dynamic, multi-authored data sets, and parallel posting and retrieval of data;
- Remote sensing and GIS to facilitate spatial data visualization and acquisition;
- Animation to facilitate pictorial visualization and simulation.

Following computational technologies are currently considered to be crucial for ecosystems analysis, synthesis and forecasting:

- High performance computing to provide high-speed data access and processing and large internal storage (RAM), and to facilitate high speed simulations;
- Internet and www to facilitate interactive and online simulation as well as software and model sharing;
- Cellular automata to facilitate spatio-temporal and individual-based simulation;
- Fuzzy logic to represent and process uncertain data;
- Artificial neural networks to facilitate multivariate nonlinear regression, ordination and clustering, multivariate time series analysis, image analysis at micro and macro scale;
- Genetic and evolutionary algorithms for the discovery and evolving of multivariate nonlinear rules, functions, differential equations and artificial neural networks; - Hybrid and AI models by the embodiment of evolutionary algorithms in process-based differential equations, the embodiment of fuzzy logic in artificial neural networks or knowledge processing;
- Adaptive agents to facilitate adaptive simulation and prediction of ecosystem composition and evolution.

The present book is an outcome of the *International Conference on Applications of Machine Learning to Ecological Modelling*, 27 November to 1 December 2000, Adelaide, Australia, which concluded with the foundation of the *International Society for Ecological Informatics (ISEI)* (http://www.waite.adelaide.edu.au/ISEI/). The chapters of the present book are based on selected papers of the conference, which are exemplary for current research trends in *ecological informatics*.

Chapters 1 to 5 address principles and ecological application of fuzzy logic, artificial neural networks, genetic algorithms, evolutionary computation and adaptive agents. Salski summarizes concepts of fuzzy logic and discusses applications for knowledge-based modeling, clustering and kriging related to ecotoxicological, geological and population dynamics data. Giraudel and Lek discuss the design and application of unsupervised artificial neural networks for the classification and visualization of multivariate ecological data. They demonstrate the potential of Kohonen-type algorithms by clustering data of forest communities in Wisconsin (USA). Morrall discusses origins and nature of genetic algorithms, and their suitability to induce numerical or rule-based models for ecological applications. Whigham and Fogel provide a scope of evolutionary algorithms and their potential for evolving rules, algebraic and differential equations relevant to ecology. They also address developments on individual and cooperative behaviour, prey-predator algorithms and hierarchical ecosystems based on evolutionary algorithms. Recknagel reflects on Holland’s adaptive agents concept and its potential to more realistically simulate emergent ecosystem structures and behaviours. He distinguishes between individual-based and state variable-based agents, and emphasizes on the embodiment of evolutionary computation in state-variable based agents.

Chapters 6 to 9 provide case studies for the prediction and elucidation of stream ecosystems by means of machine learning techniques. Goethals, Dedecker, Gabriels and de Pauw demonstrate applications of classification trees and artificial neural networks for the bioassessment of the Zwalm river system in Belgium.

Schleiter, Obach, Wagner, Werner, Schmidt and Borchardt carried out a comprehensive study of the Breitenbach stream (Germany) based on a variety of unsupervised and supervised learning algorithms for artificial neural networks. They draw interesting conclusions regarding suitability of different algorithms for bioindication of stream habitats and input sensitivity of streams. Chon, Park, Kwak and Cha provide a summary of achievements in the structural classification and dynamic prediction of macroinvertebrate communities in Korean streams by artificial neural networks. They also discuss patterning of organizational aspects of macroinvertebrate communities. Huong, Recknagel, Marshall and Choy study relationships between environmental factors, stream habitat characteristics and the occurrence of macroinvertebrate taxa in the Queensland stream system (Australia) by means of a neural network based sensitivity analysis.

Chapters 10 to 12 contain examples of time series analysis of river water quality by artificial neural networks. Jeong, Recknagel and Joo apply recurrent neural networks to explain and predict the seasonal abundance and succession of different algae species in the River Nakdong (Korea). Validation results reveal a reasonable correspondence between seven days ahead forecasts and observations of algal abundance. Information on favouring conditions and processes for certain algal species discovered by a comprehensive sensitivity analysis comply well with domain knowledge. Bowden, Maier and Dandy combine super- and unsupervised artificial neural networks as well as genetic algorithms for automated input determination of neural networks in order to forecast the abundance of an algae species in the River Murray (Australia). Gevrey, Lek and Oberdorff apply two approaches of sensitivity analysis for the study of riverine fish species by means of artificial neural networks.

Chapters 14 to 17 provide case studies for the application of fuzzy logic, artificial neural networks and evolutionary algorithms to freshwater lakes and marine fishery systems. Karul and Soyupak compare results for the chlorophyll-a estimation in three Turkish lakes achieved by multiple regression and artificial neural networks. Wilson and Recknagel design a generic neural network model for forecasting algal blooms that is validated by means of six lake databases. It considers bootstrapping, bagging and time-lagged training as crucial techniques for minimising prediction errors. Bobbin and Recknagel apply evolutionary algorithms to discover rules for the abundance and succession of blue green algae species in the hypereutrophic Lake Kasumigaura (Japan). Resulting rules correspond with literature findings, reveal hypothetical relationships and are able to predict timing and magnitudes of algal dynamics.

Reick, Gruenewald and Page address the issue of data quality in the context of ecological time-series analysis and prediction. They describe cross-validation and automated training termination of neural networks applied for multivariate time-series predictions of marine zooplankton in the German Northern Sea. Chen combines fuzzy logic and artificial neural networks in order to classify fish stock-recruitment relationships in different environmental regimes near the West Coast Vancouver Island (Canada) and southeast Alaska (USA).

Chapters 18 to 20 provide examples for the classification of ecological images at micro and macro scale by artificial neural networks. Wilkins, Boddy and

Dubelaar demonstrate possibilities for the identification of marine microalgae by the analysis of flow cytometric pulse shapes with the help of neural networks. Robertson and Morison applied a probabilistic neural network for the automation of age estimation in three fish species. Thin-sections of sagittal otoliths viewed with transmitted light were used for all species, and the number of opaque increments used to estimate the age. The neural network correctly classified a larger range of age classes. Foody gives a representative summary of neural network algorithms currently used for the pattern recognition and classification of remotely sensed landscape images.

At this point I want to thank all of the authors who responded with great enthusiasm to my request for chapters to the theme of the book and delivered on time. I am also grateful to 24 colleagues and friends in Australia and overseas who significantly improved the quality of chapters by their critical reviews.

Finally I express my thanks to Dr. Christian Witschel and Agata Oelschlaeger of the Geosciences Editorial Team of the Springer Verlag for their close collaboration in producing the book.

Friedrich Recknagel
Adelaide, 15 April 2002

Contents

Contributors

Jason Bobbin
Department of Soil and Water, Adelaide University
Glen Osmond 5064, Australia
E-mail: jason@jbobbin.waite.adelaide.edu.au

Lynne Boddy
Cardiff School of Biosciences, Cardiff University, Cardiff CF10 3TL, U.K.
E-mail: BoddyL@Cardiff.ac.uk

Dietrich Borchardt
University of Kassel,
Institute of Aquatic Resources Research and Management,
Kurt-Wolters-Str. 3, D-34125 Kassel
Germany

Gavin Bowden
Department of Civil and Environmental Engineering,
Adelaide University, Adelaide 5005
Australia
E-mail: gbowden@civeng.adelaide.edu.au

Eui Young Cha
Division of Electronics and Computer Sciences,
Pusan National University, Pusan
Korea

Din Chen
International Pacific Halibut Commission
University of Washington, Seattle WA 98195
USA
E-mail: din@iphc.washington.edu

Tae-Soo Chon
Division of Biological Sciences,
Pusan National University, Pusan
Korea
E-mail: tschon@hyowon.cc.pusan.ac.kr

Satish Choy
Queensland Department of Natural Resources and Mining,
1345 Ipswich Rd, Rocklea 4106
Australia
E-mail: Satish.Choy@dnr.qld.gov.au

Graeme Dandy
Department of Civil and Environmental Engineering
Adelaide University, Adelaide 5005
Australia

A Dedecker
Laboratory of Environmental Toxicology and Aquatic Ecology
Ghent University
J. Plateaustraat 22, B-9000 Ghent
Belgium

N. de Pauw
Laboratory of Environmental Toxicology and Aquatic Ecology
Ghent University
J. Plateaustraat 22, B-9000 Ghent
Belgium

George B.J. Dubelaar
Dubelaar Research Instruments Engineering (DRIE)
Zeelt 2, 2411 DE Bodegraven
The Netherlands

Gary B. Fogel
Natural Selection, Inc., La Jolla, CA 92037
U.S.A.
E-mail: gfogel@natural-selection.com

Giles M. Foody
Department of Geography, University of Southampton
Highfield, Southampton, SO17 1BJ
U.K.
E-mail: G.M.Foody@soton.ac.uk

Wim Gabriels
Laboratory of Environmental Toxicology and Aquatic Ecology
Ghent University
J. Plateaustraat 22, B-9000 Ghent
Belgium
E-mail: wim.gabriels@rug.ac.be

Peter Goethals
Laboratory of Environmental Toxicology and Aquatic Ecology
Ghent University
J. Plateaustraat 22, B-9000 Ghent
Belgium
E-mail: peter.goethals@rug.ac.be

Muriel Gevrey
CESAC (Center d'Ecologie des Systemes Aquatiques et Continentaux)
University of Paul Sabatier, Toulouse
FRANCE
E-mail: gevrey@cict.fr

Jean-Luc Giraudel
CESAC (Center d'Ecologie des Systemes Aquatiques et Continentaux)
University of Paul Sabatier, Toulouse
FRANCE
E-mail: giraudel@montesquieu.u-bordeaux.fr

Antje Gruenewald
University Hamburg, Department of Computer Science
Vogt-Koelln-Str. 30
22527 Hamburg
Germany

Huong Hoang
Department of Soil and Water
Adelaide University, Glen Osmond 5064
Australia

Kwang-Seuk Jeong
Department of Biology
Pusan National University
Jang-Jeon Dong, Gum-Jeong Gu, Busan, 609-735
Korea
E-mail: pow5150@hotmail.com, pow0606@hananet.net

Gea-Jae Joo
Department of Biology
Pusan National University
Jang-Jeon Dong, Gum-Jeong Gu, Busan, 609-735
Korea
E-mail: gjjoo@pusan.ac.kr

Cueneyt Karul
Department of Environmental Engineering
Middle East Technical University, 06531, Ankara
Turkey

Inn-Sil Kwak
Division of Biological Sciences
Pusan National University, Pusan
Korea

Sovan Lek
CESAC (Center d'Ecologie des Systemes Aquatiques et Continentaux)
University of Paul Sabatier, Toulouse
FRANCE
E-mail: lek@cict.fr

Holger Maier
Department of Civil and Environmental Engineering
Adelaide University, Adelaide 5005
Australia
E-mail: hmaier@civeng.adelaide.edu.au

John Marshall
Queensland Department of Natural Resources and Mining
1345 Ipswich Rd., Rocklea 4106
Australia
E-mail: Jonathan.Marshall@dnr.qld.gov.au

Donna Morrall
The Procter & Gamble Co.
Environmental Science Department, Miami Valley Laboratories
P.O. Box 538707, Cincinnati, OH 45253-8707, U.S.A.
E-mail: morrall.dd@pg.com

Alexander K. Morison
Marine and Freshwater Resources Institute
PO Box 114, Queenscliff, VIC 3225
AUSTRALIA
E-mail: Sandy.morison@nre.vic.gov.au

Michael Obach
Limnologische Fluss-Station Schlitz der Max-Planck-Gesellschaft
P.O.Box 260, D-36105 Schlitz

Germany
E-mail: michaelo@neuro.informatik.uni-kassel.de

Thierry Oberdorff
Institut dEcologie et de Gestion de la Biodiversite, MNHN,
Lab. Ichtyologie appliquee, 43 rue Cuvier, 75005 Paris
France

Bernd Page
University Hamburg, Department of Computer Science
Vogt-Koelln-Str. 30
22527 Hamburg
Germany

Young Seuk Park
CESAC (Center d'Ecologie des Systemes Aquatiques et Continentaux)
University of Paul Sabatier, Toulouse
FRANCE

Friedrich Recknagel
Department of Soil and Water, University of Adelaide
Glen Osmond 5064
Australia
E-mail: Friedrich.Recknagel@adelaide.edu.au

Christian H. Reick
Alfred-Wegener-Institute
PMB 120161, 27515 Bremerhaven
Germany
E-mail: creick@AWI-Bremerhaven.DE

Simon G. Robertson
Marine and Freshwater Resources Institute
PO Box 114, Queenscliff, VIC 3225
AUSTRALIA
E-mail: simon.robertson@nre.vic.gov.au

Arkadiusz Salski
Institute of Computer Science and Applied Mathematics
University of Kiel
Olshausenstr. 40, 24118 Kiel
Germany
E-mail: asa@email.uni-kiel.de

Ingrid M. Schleiter

University of Kassel
Institute of Aquatic Resources Research and Management
Kurt-Wolters-Str. 3, D-34125 Kassel
Germany
E-mail: schleiter@hrz.uni-kassel.de
Hans-Heinrich Schmidt
Limnologische Fluss-Station Schlitz der Max-Planck-Gesellschaft
P.O.Box 260, D-36105 Schlitz
Germany

Selcuk Soyupak
Department of Environmental Engineering
Middle East Technical University
06531 Ankara
Turkey
E-mail: soyupak@metu.edu.tr

Ruediger Wagner
Limnologische Fluss-Station Schlitz der Max-Planck-Gesellschaft
P.O.Box 260, D-36105 Schlitz
Germany
E-mail: RWAGNER@MPIL-SCHLITZ.MPG.DE

Heinrich Werner
University of Kassel,
Dept. of Mathematics and Computer Sciences
Research Group Neural Networks
Heinrich-Plett-Str. 40, D-34132 Kassel
Germany
E-mail: werner@neuro.informatik.uni-kassel.de

Malcolm F. Wilkins
Cardiff School of Biosciences, Cardiff University
Cardiff CF10 3TL
U.K.
E-mail: WilkinsMF@Cardiff.ac.uk

Peter Whigham
University of Otago, Department of Information Science
Dunedin
New Zealand
E-mail: pwhigham@infoscience.otago.ac.nz

Hugh Wilson
Department of Soil and Water, Adelaide University
Glen Osmond 5064

Australia
E-mail: hwilson@waite.adelaide.edu.au