Ecological Informatics 24 (2014) 85–89

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Hybrid evolutionary computation quantifies environmental thresholds for recurrent outbreaks of population density

Friedrich Recknagel [a,*], Ilia Ostrovsky [b], Hongqing Cao [a], Qiuwen Chen [c]

[a] *University of Adelaide, School of Biological Sciences, Adelaide 5005, Australia*
[b] *Israel Oceanographic and Limnological Research, Kinneret Limnological Laboratory, Migdal 14950, Israel*
[c] *Center for Eco-Environmental Research, Nanjing Hydraulic Research Institute, Nanjing 210000, China*

ARTICLE INFO

Article history:
Received 30 July 2014
Accepted 30 July 2014
Available online 7 August 2014

Keywords:
Environmental thresholds
Population density outbreaks
Evolutionary computation
IF–THEN–ELSE models
Cyanobacteria blooms

ABSTRACT

Identification of thresholds as indicators for sudden shifts in the community composition has proven difficult. Evolutionary computation learns inferential models from data. We investigated the hypothesis that models built from data of largely different lakes by evolutionary computation reveal similar thresholds for outbreaks of cyanobacteria densities. Here we show that IF–THEN–ELSE models inferred by the hybrid evolutionary algorithm HEA from multidimensional data of the hypertrophic polymictic Lake Taihu (China) and the mesotrophic warm-monomictic Lake Kinneret (Israel) perform fairly good 5-day-ahead forecasting of density outbreaks and indicate thresholds referring to the same environmental factors, such as nitrate and water temperature. The discovered thresholds suggest that hypereutrophic lakes may reach N-limitation at nitrate concentrations that are orders of magnitude higher than in a mesotrophic lake, and that cyanobacteria may grow at much lower water temperatures and within a much wider temperature range at phosphorus sufficiency in a hypereutrophic lake than under phosphorus deficiency in a stratified mesotrophic lake.

Outbreaks in population density of unwanted organisms (e.g. pathogenic bacteria, cyanobacteria, pests) can severely impact on human health and ecological and economic resources. Great efforts are undertaken to qualitatively determine thresholds as indicators for sudden shifts in the community composition (e.g. May 1977; Scheffer et al. 2009). However, multivariate nonlinear methods, such as evolutionary computation, have the capacity to successfully reveal and quantify specific ecological thresholds from the multidimensional ecological data (Sonderegger et al. 2009). We investigated the hypothesis that evolutionary computation can identify and quantify thresholds for seasonal outbreaks of cyanobacteria blooms that are similar for different water bodies.

Cyanobacteria blooms contaminate waters globally by cyanotoxins (Carmichael 1994). Global expansion of cyanobacteria is driven by ongoing eutrophication and climate change (Paerl and Otten 2013). Knowing thresholds of key environmental factors that are associated with sudden outbreaks of cyanobacteria blooms is the main frontier in aquatic ecology, and will inform early warning of these harmful events.

Evolutionary computation learns inferential models from data (Holland 1975; Holland et al. 1986). The hybrid evolutionary algorithm HEA (Cao et al. 2014) has been designed for inductive reasoning of “fittest” IF–THEN–ELSE models from multi-dimensional data patterns in a spiral-like boot-strap scheme (see Fig. 1). Multi-objective optimisation of model structures by genetic programming and of model parameters by differential evolution advance models to perform fairly accurate short-term forecasting and reveal threshold conditions. When HEA is applied to complex ecological parameters, IF-conditions of predictive models reveal thresholds that indicate outbreaks of population densities (Recknagel et al. 2014). Alternative inductive modelling techniques such as artificial neural networks fail to explicitly disclose models and thresholds but allow to interpret underlying relationships by sensitivity analysis (Recknagel et al. 2002).

This case study illustrates the capacity of HEA for predictive modelling and threshold discovery of cyanobacteria blooms in two well-studied lakes — the warm-monomictic and mesotrophic Lake Kinneret (Israel) and the shallow-polymictic and hypertrophic Lake Taihu (China). Even though the two lakes differ largely with regards to morphometry, trophic state and climate conditions, both lakes are susceptible to recurrent cyanobacteria blooms. Nine years of limnological data of both lakes summarised in Table 1 were utilized to model and analyse the seasonal and inter-annual dynamics of cyanobacteria biomass by HEA. Key questions of this case study were: (1) can similar driving forces and thresholds for cyanobacteria outgrowth be discovered for different lakes, and (2) how do driving forces and thresholds compare between different lakes.

Fig. 2 illustrates an IF–THEN–ELSE model for cyanobacteria development in Lake Taihu from 2000 to 2008. The Fig. 2A shows good correspondence between measured and 5-day-ahead forecasted cyanobacteria dynamics in Lake Taihu (coefficient of determination, $r^2 = 0.8$). The

---

* Corresponding author.
E-mail address: friedrich.recknagel@adelaide.edu.au (F. Recknagel).

http://dx.doi.org/10.1016/j.ecoinf.2014.07.010
1574-9541/

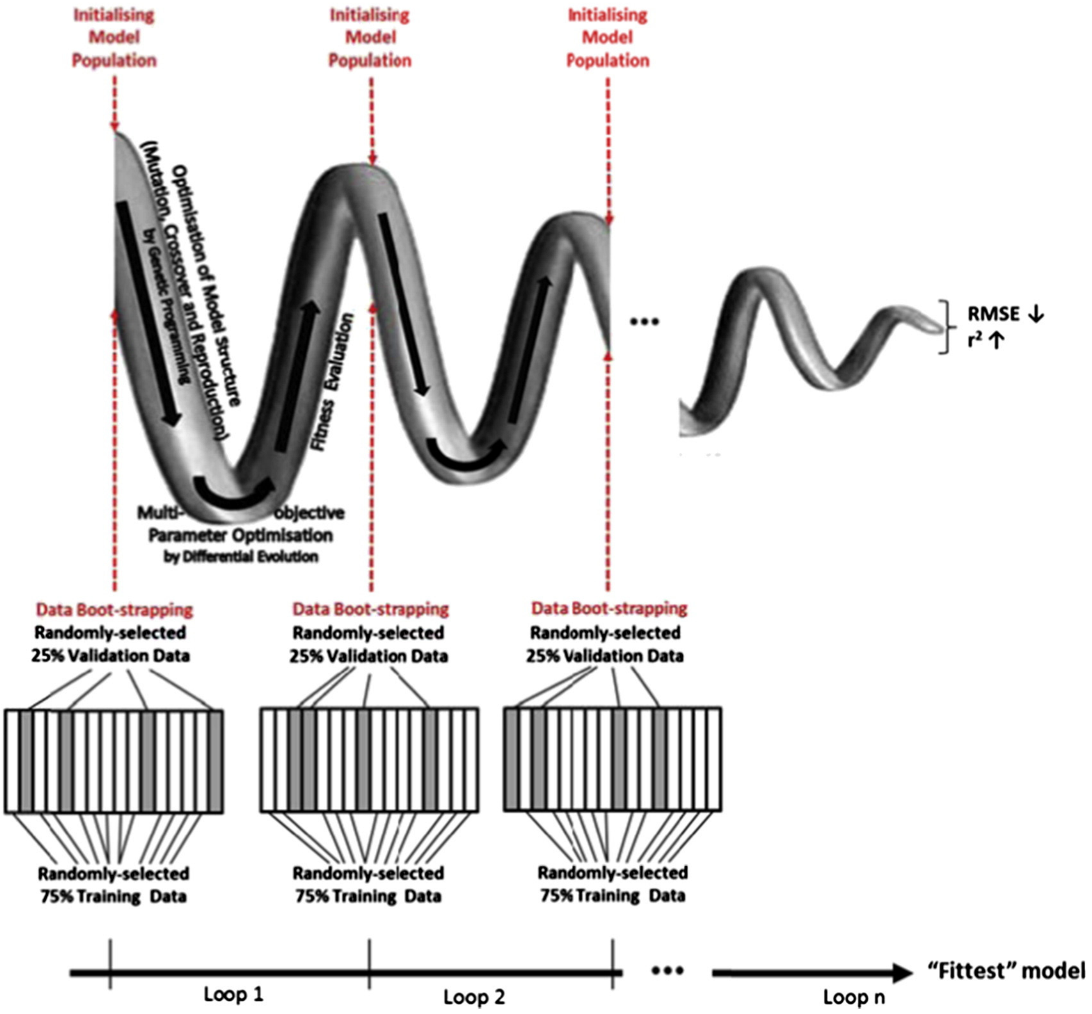

Fig. 1. Conceptual diagram of the hybrid evolutionary algorithm (HEA). HEA evolves 'fittest' (best-matching) models from the multi-dimensional data by cyclically applying genetic programming for the design of IF–THEN–ELSE rules and differential evolution for multi-objective parameter optimisation along thousands of loops. For each loop the model population is initialised and data are boot-strapped. Model fitness is evaluated by lowest root mean square error (RMSE) and highest coefficient of determination (r^2).

underlying model matches well both timing and magnitudes of fast growth events of cyanobacteria observed over nine consecutive years.

Table 1
Summary of water quality data of Lake Kinneret and Lake Taihu from 2000 to 2008 used in this study.

	Lake Kinneret	Lake Taihu
	Mean ± SD	
Water temperature (WT) °C	22.7 ± 5	17.8 ± 8.36
Secchi depth (SD) m	3 ± 0.6	0.33 ± 0.06
Dissolved oxygen (DO) mg/L	8.6 ± 1.8	9.3 ± 1.6
pH	8.6 ± 0.3	8.1 ± 0.3
Nitrate (NO_s–N) mg/L	0.09 ± 0.12	1.24 ± 0.86
Phosphate ($P0_4$–P) mg/L	0.0002 ± 0.0014	0.04 ± 0.02
N/P (N:P)	54 ± 76.1	41.7 ± 35.9
Bacillariophyta g/m^2	12.6 ± 27.4	7 ± 11.9
Chlorophyta g/m^2	21.3 ± 21	3.9 ± 4.4
Cyanophyta g/m^2	10.7 ± 11.5	25.1 ± 20.6

The model associated the water temperature range 18.8 to 27.3 °C and nitrate concentrations of smaller than 1.85 mg/L with high biomass of cyanobacteria in different years. These conditions agree well with findings that N-limitation of phytoplankton in summer is typical for Lake Taihu (Paerl et al. 2011), and suggest that N_2-fixing cyanobacteria would outcompete non-N_2-fixing cyanobacteria. However, cyanobacteria in Lake Taihu in summer are dominated by non-N_2-fixing *Microcystis* (Chen et al. 2003) that may successfully bypass nitrate limitation by utilizing ammonium (NH_4) released from anaerobic sediments during summer. The Fig. 2B and C illustrates relationships of cyanobacteria with NO_3–N concentrations and water temperature, as discovered by sensitivity analyses of ten best-performing models developed by HEA. Whilst cyanobacteria biomass is inverse-exponentially declining with increased nitrate concentrations, it displays linear growth at increasing water temperatures. Fig. 2D and E eidetic illustrates how the IF-conditions of the model separate 'high' and 'low' cyanobacteria biomasses observed over the various years.

The Fig. 3A illustrates an IF–THEN–ELSE model that achieved good 5-day-ahead forecasting results for seasonal dynamics of cyanobacteria biomass in Lake Kinneret. The model predicts well population dynamics of cyanobacteria between the years 2000 and 2008, but underestimates magnitudes of peak events in 2000 and 2001 reflected by a moderate $r^2 = 0.62$.

Sensitivity analyses of ten best performing models developed by HEA show that cyanobacteria biomass in Lake Kinneret is increasing at

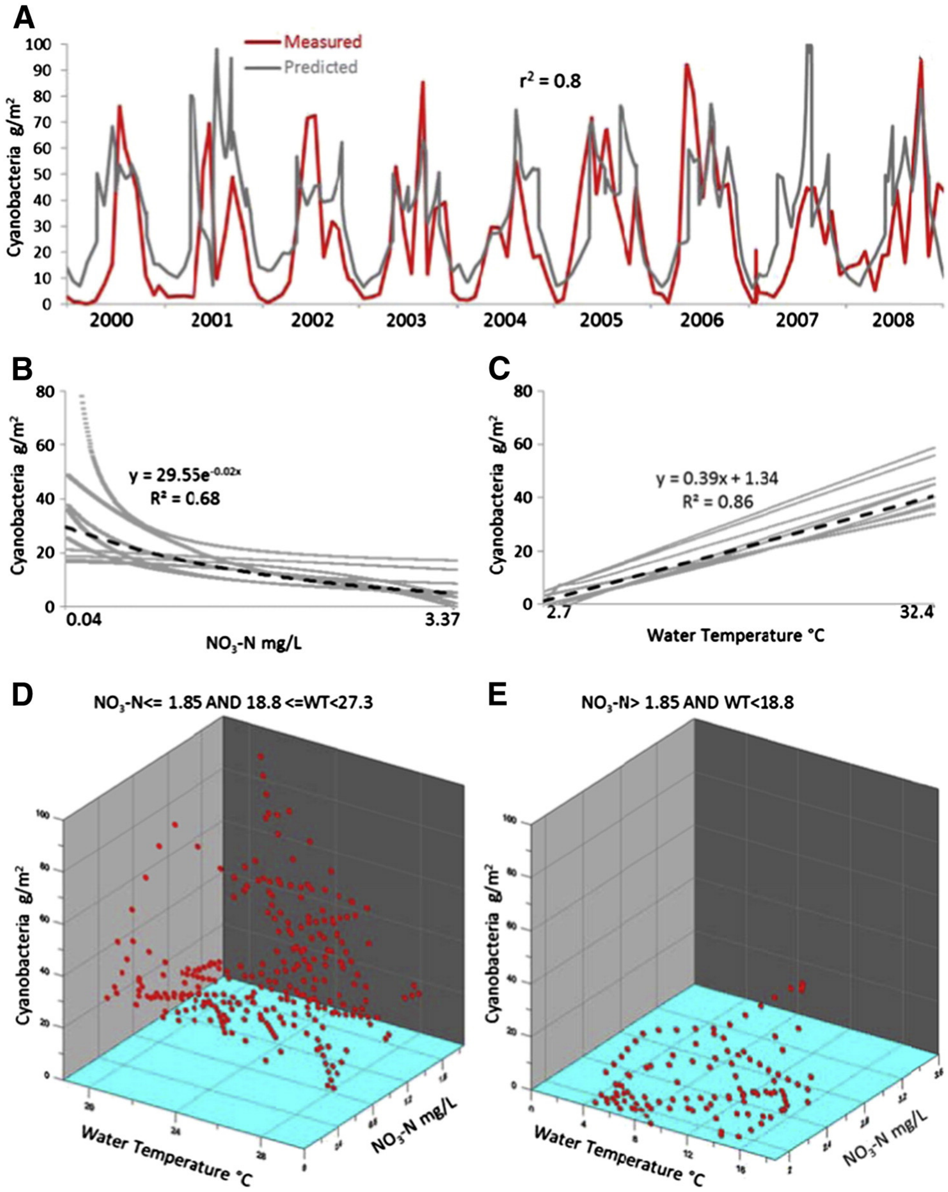

Fig. 2. 5-day-ahead forecasting of cyanobacteria biomass in the upper productive stratum of Lake Taihu from 2000 to 2008 by the model:
IF 18.8 < =WT < 27.3 AND NO_3–N < 1.85
THEN cyanobacteria = (DO ∗ (DO − ln(Secchi ∗ 41.1)))
ELSE cyanobacteria = (ln(15.7/N:P) − ((((−96.95) / DO) / NO_3–N) − WT)).
(A) model validation, (B) nitrate sensitivity of ten best cynobacteria models, (C) water temperature (WT) sensitivity of ten best cynobacteria models, (D) thresholds for high cyanobacteria biomass, (E) thresholds for low cynobacteria biomass. The model was derived based on daily-interpolated monitoring time-series (Taihu Basin Authority) using the HEA. (WT = water temperature, °C; DO = dissolved oxygen, mg/L; Secchi =Secchi depth, m; NO_3–N = nitrate concentration, mg/L; N:P = total nitrogen to total phosphorus ratio.).

declining nitrate concentrations (Fig. 3B), and is increasing along with rising water temperatures (Fig. 3C). Thus, for both lakes the IF-conditions computed by HEA refer to the same environmental factors, such as nitrate and water temperature, but the threshold ranges of IF-conditions differ remarkably. The Lake Kinneret model suggests that nitrate concentrations of smaller than 0.025 mg/L and water temperatures confined between 25.2 and 29.1 °C are associated with high cyanobacteria biomass over the nine-year period. When applied to the observed data of Lake Kinneret, these threshold conditions separate well 'high' and 'low' cyanobacteria biomass (Figs. 3D and E). The nitrate threshold quantified by HEA suggests that cyanobacteria blooms in Lake Kinneret occur during nitrogen limitation as defined with dissolved inorganic nitrogen concentrations smaller than 0.05 to 0.1 mg/L (Horne and Cummins 1987). It also corresponds with findings that low nitrate concentrations are associated with growth of the invasive N_2-fixing cyanobacterium *Aphanizomenon ovalisporum* (Berman 2001; Gophen et al. 1999). Whilst the N_2-fixing *Cylindrospermopsis raciborskii* formed major summer blooms from 2004 to 2006 and since then co-dominated the summer–fall assemblage with *A. ovalisporum*, the non-N_2-fixing *Microcystis* sp. subdominated during the summers of 2001 and 2003 feeding most likely either on NH_4 supplied from remixed hypolimnetic water or derived nitrogen from amino acids and purines (Berman 2001).

Thus, the HEA has selected the same driving factors, such as nitrate and water temperature, for predicting fast growth events of cyanobacteria in two very different lakes. This finding suggests that late summer nitrogen limitation caused by high nutrient uptake rates may coincide with cyanobacteria blooms no matter if the trophic state is hypereu- or mesotrophic, if the lake is shallow polymictic or warm monomictic, and if the climate is temperate or Mediterranean. In

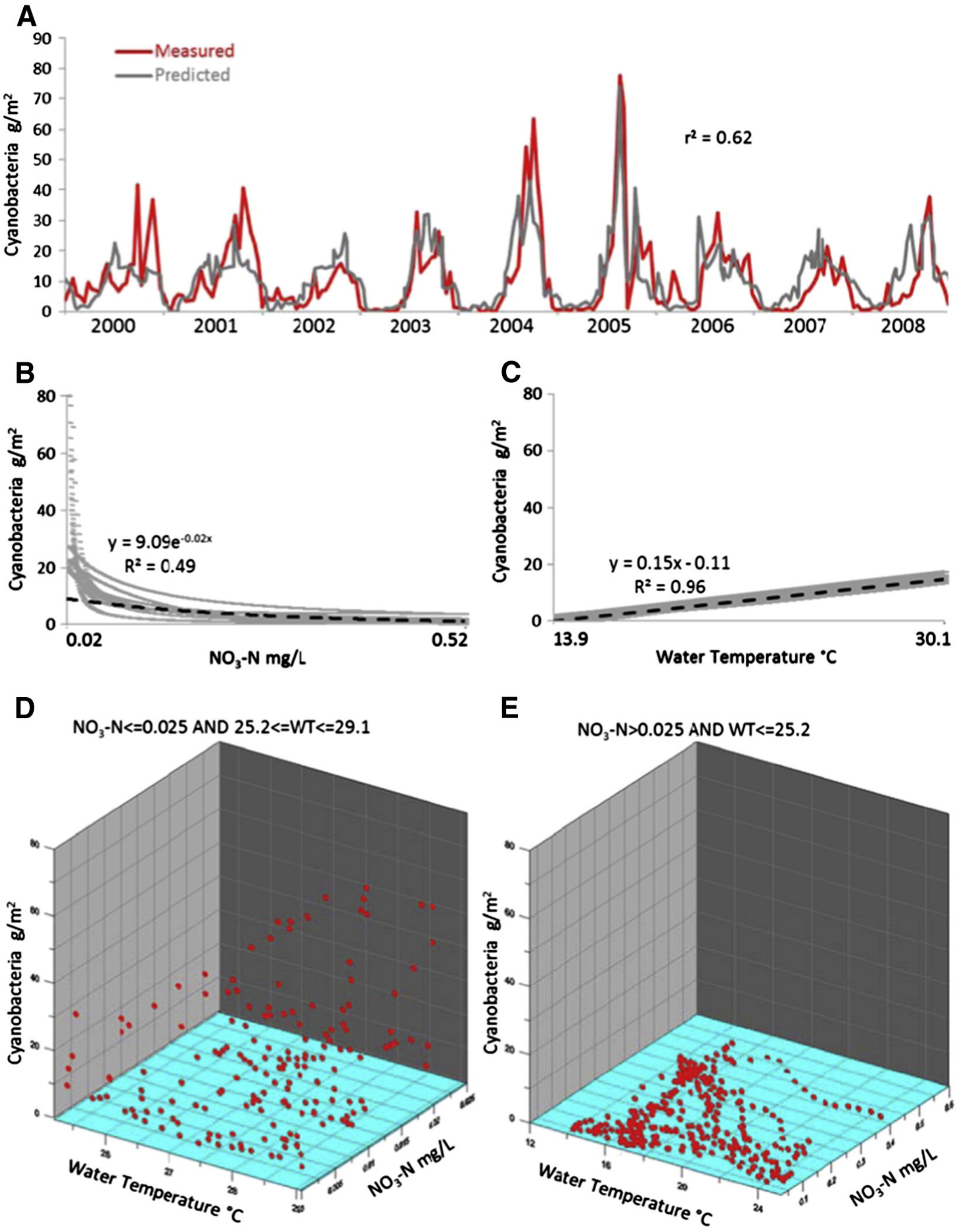

Fig. 3. 5-day-ahead forecasting of cyanobacteria biomass in the upper productive stratum of Lake Kinneret from 2000 to 2008 by the model:
IF $NO_3\text{–}N <= 0.025$ AND $25.2 <= WT <= 29.1$
THEN cyanobacteria $= ((((-191) / \text{Secchi}^2) * (-328.5)) / (\text{zooplankton} + 2 * \text{N:P} + 143.2))$
ELSE cyanobacteria $= (399.8/((\text{N:P} + 26.3) + (\text{bacillariophyta}^2 * NO_3)))$.
(A) model validation, (B) nitrate sensitivity of ten best cyanobacteria models, (C) water temperature sensitivity of ten best cyanobacteria models, (D) thresholds for high cyanobacteria biomass, (E) thresholds for low cyanobacteria biomass. The model was derived based on daily-interpolated monitoring time-series (the Lake Kinneret Database) using the HEA. (bacillariophyta =bacillariophyta biomass, g/m^2; other abbreviations as on Fig. 2).

accordance with nutrient stoichiometry it also suggests that a hypereutrophic lake may reach N-limitation at nitrate concentrations that are two orders of magnitude higher than in a mesotrophic lake. The quantified temperature ranges indicate that cyanobacteria may grow at much lower water temperatures (e.g. 19 °C) and within a much wider temperature range (e.g. 19 to 29 °C) at phosphorus sufficiency in a hypertrophic lake than under phosphorus deficiency in a stratified mesotrophic lake.

Overall this study has demonstrated that evolutionary computation can build predictive models of such complex and fast-evolving events such as cyanobacterial blooms in waters. Thresholds identified and quantified by evolutionary computation enable aquatic and terrestrial ecologists to disclose key limiting factors affecting population dynamics and study underlying biological and genetic mechanisms. The case study has also shown that evolutionary computation suits as an excellent tool for meta-analysis of samecategory systems such as lakes and reservoirs with largely different environmental conditions.

Acknowledgments

We thank the Kinneret Limnological Laboratory for making data from its Lake Kinneret database available. Specifically, we thank T. Zohary for algae and cyanobacteria biomasses, A. Rimmer for temperature data, A. Nishri and the Mekorot Water company for conducting the for chemical measurements, K.D. Hambright and G. Gal for zooplankton data and field crew for collecting the samples and Secchi depth data. We also thank the Monitoring Bureau of the Taihu Basin Authority for providing limnological data of Lake Taihu from 2000 to 2008. We acknowledge Rita Adrian for her valuable comments on the manuscript.

References

Berman, T., 2001. The role of DON and the effect of N:P ratios on occurrence of cyanobacterial blooms: implications from the outgrowth of *Aphanizomenon* in Lake Kinneret. Limnol. Oceanogr. 46 (2), 443–447.

Cao, H., Recknagel, F., Orr, P., 2014. Parameter optimization algorithms for evolving rule models applied to freshwater ecosystems. IEEE Trans. Evol. Comput.

Carmichael, W.W., 1994. The toxins of cyanobacteria. Scientific American 78–86 (January 1994).

Chen, Y., et al., 2003. Long-term dynamics of phytoplankton assemblages: *Microcystis*-domination in Lake Taihu, a large shallow lake in China. J. Plankton Res. 25 (1), 445–453.

Gophen, M., et al., 1999. Nitrogen deficiency, phosphorus sufficiency, and the invasion of Lake Kinneret, Israel, by the N_2-fixing cyanobacterium *Aphanizomenon ovalisporum*. Aquat. Sci. 61, 293–306.

Holland, J.H., 1975. Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbour, Mich.

Holland, J.H., et al., 1986. Induction. Process of Inference, Learning, and Discovery. MIT Press.

Horne, A.J., Commins, M.L., 1987. Macronutrient controls on nitrogen fixation in planktonic cyanobacterial population. Mar.Freshw.Res. 21, 413–423.

May, R.M., 1977. Thresholds and breakpoints in ecosystems with a multiplicity of stable states. Nature 269, 471–477.

Paerl, H.W., Otten, T.G., 2013. Blooms bite the hand that feeds them. Science 342, 433–434.

Paerl, H.W., et al., 2011. Controlling harmful cyanobacterial blooms in hyper-eutrophic lake (Lake Taihu, China): the need for a dual nutrient (N&P) management strategy. Water Res. 45, 1973–1983.

Recknagel, F., Bobbin, J., Wilson, H., Whigham, P., 2002. Comparative application of artificial neural networks and genetic algorithms for multivariate time-series modelling of algal blooms in freshwater lakes. J. Hydroinf. 4 (2), 125–133.

Recknagel, F., Orr, P., Cao, H., 2014. Inductive reasoning and forecasting of population dynamics of *Cylindrospermopsis raciborski* in three sub-tropical reservoirs by evolutionary computation. Harmful Algae 31, 26–34.

Scheffer, M., et al., 2009. Early-warning signals for critical transitions. Nature 461, 53–59.

Sonderegger, D.L., Wang, H., Clements, W.H., Noon, B.R., 2009. Using SiZer to detect thresholds in ecological data. Front. Ecol. Environ. 7 (4), 190–195.