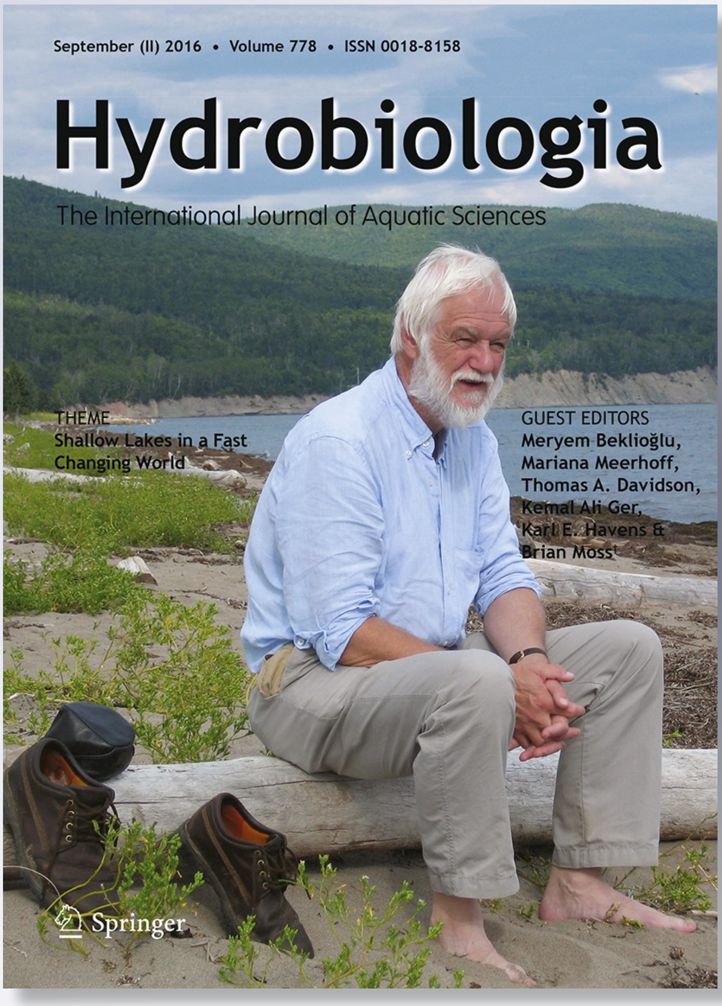
Threshold quantification and short-term forecasting of Anabaena, Aphanizomenon *and* Microcystis *in the polymictic eutrophic Lake Müggelsee (Germany) by inferential modelling using the hybrid evolutionary algorithm HEA*

Friedrich Recknagel, Rita Adrian, Jan Köhler & Hongqing Cao

Hydrobiologia
The International Journal of Aquatic Sciences

ISSN 0018-8158
Volume 778
Number 1

Hydrobiologia (2016) 778:61-74
DOI 10.1007/s10750-015-2442-7



Hydrobiologia (2016) 778:61–74
DOI 10.1007/s10750-015-2442-7

SHALLOW LAKES

Threshold quantification and short-term forecasting of *Anabaena*, *Aphanizomenon* and *Microcystis* in the polymictic eutrophic Lake Müggelsee (Germany) by inferential modelling using the hybrid evolutionary algorithm HEA

Friedrich Recknagel · Rita Adrian ·
Jan Köhler · Hongqing Cao

Received: 29 April 2015 / Revised: 14 July 2015 / Accepted: 10 August 2015 / Published online: 27 October 2015

Abstract Forecasting models for *Anabaena*, *Aphanizomenon* and *Microcystis* have been developed for the hypertrophic phase from 1979 to 1990 and the eutrophic phase from 1997 to 2012 of the polymictic Lake Müggelsee by means of the hybrid evolutionary algorithm HEA. Comparisons of limnological parameters of the two phases revealed not only a distinct seasonal extension of N-limitation but also higher water temperatures that rose earlier and lasted longer between spring and autumn from 1997 to 2012. These differences were reflected by threshold conditions and sensitivity functions of the cyanobacteria-specific models evolved by HEA for the two phases. Seven-day-ahead forecasts matched well timings of peaking biomass observed for the three cyanobacteria but partially failed to predict accurate magnitudes, whereby coefficients of determination r^2 ranged between 0.48 and 0.76 for models in Phase I and between 0.42 and 0.69 in Phase II. The threshold conditions of the models quantified ranges of key predictor variables such as water temperature and transparency, concentrations of NO_3-N and PO_4-P that were symptomatic for sudden outbreaks of high biomass of the three cyanobacteria. Sensitivity functions extracted from 20 best performing models for each of the three cyanobacteria in both phases indicated different abundances between N-fixing *Anabaena* and *Aphanizomenon* compared to non-N-fixing *Microcystis* in response to strengthened N-limitation in Phase II.

Keywords Polymictic eutrophic lake · Cyanobacteria · HEA · Forecasting · Ecological thresholds · Sensitivity analysis · Interrelationships

Introduction

Polymictic eutrophic lakes become increasingly susceptible to nuisance cyanobacteria blooms by more frequent internal nutrient pulses during extended stratified periods in summer under the influence of climate warming (Wilhelm and Adrian, 2008; Wagner & Adrian, 2009). By contrast, extended stratification periods in dimictic lakes may strengthen nutrient limitation in the epilimnion with nutrient pulses occurring during autumn overturn (Adrian et al., 1995; Huisman et al., 2004; Elliott et al., 2006; Jöhnk et al., 2008; Wilhelm & Adrian, 2008; Mooij et al., 2009; Wagner & Adrian, 2009), whereby access to the

Guest editors: M. Beklioğlu, M. Meerhoff, T. A. Davidson, K. A. Ger, K. E. Havens & B. Moss / Shallow Lakes in a Fast Changing World

F. Recknagel (✉) · H. Cao
School of Biological Sciences, University of Adelaide,
Adelaide 5005, Australia
e-mail: friedrich.recknagel@adelaide.edu.au

R. Adrian · J. Köhler
Leibniz-Institute of Freshwater Ecology and Inland
Fisheries, Berlin, Germany

hypolimnetic nutrient pools may be limited to buoyant cyanobacteria species (Reynolds, 1984; Paerl, 1988).

The polymictic eutrophic Lake Müggelsee has been studied extensively with respect to impacts from climate warming (Adrian et al., 2009; Adrian et al., 2012) and the development of cyanobacteria blooms (Wilhelm & Adrian, 2008; Wagner & Adrian, 2009; Huber et al., 2012). The lake has a mean depth of 5 m, a retention time of 1–4 months and a catchment area of ca. 7000 km^2.

Being exposed to both maritime and continental climate, the duration of ice cover and circulation patterns of Lake Müggelsee differ year by year alternating between thermal stratification and wind-driven mixing in summer (Wilhelm & Adrian, 2008; Livingstone & Adrian, 2009; Wagner & Adrian, 2009). Even though the external nutrient load declined by about 50% over the past two decades (Köhler et al., 2005), remobilisation from bottom sediments maintained high phosphorus concentrations in the lake during calm summer periods (Wilhelm & Adrian, 2008). Following factors limiting seasonally phytoplankton development in Lake Müggelsee have been revealed by bioassays (Kolzau et al. 2014) and nutrient ratios (Köhler et al., 2000, 2005): light and temperature in winter and early spring, phosphorus and grazing in late spring, nitrogen in summer.

The spring biomass of cyanobacteria in Lake Müggelsee declined since the 1980s (Köhler et al., 2005) and summer blooms by *Planktothrix agardhii*, *Aphanizomenon flos-aquae*, *Anabaena* and *Microcystis* species became less frequent in recent years. Timing and magnitude of bloom events were largely controlled by total phosphorus and mixing intensity largely driven by meteorological events (Wagner & Adrian, 2009).

The strong influence of hydrodynamic processes in polymictic lakes complicates predictive modelling of phytoplankton development. As demonstrated for Lake Müggelsee, process-based modelling suits well long-term forecasting by scenario analysis (e.g. Recknagel et al., 1995) and simulation of seasonal population dynamics (e.g. Huber et al., 2008). However, short-term forecasting of timing and magnitudes of species-specific outbreaks of population densities requires alternative modelling techniques that quantify cyanobacteria-specific threshold conditions for population outbreaks (Recknagel et al., 2013; Recknagel et al., 2014c). Wagner & Adrian (2009) applied statistical methods to long-term data of Lake Müggelsee identifying the following threshold conditions for the dominance of cyanobacteria: (1) stratification periods exceeding 3 weeks and exhibiting a Schmidt stability of 0.44 g cm cm^{-2} favoured fast growth of cyanobacteria within a critical TP concentration range (70–215 mg/l), (2) hypolimnetic TP, the intensity of Schmidt stability, epilimnetic TP and epilimnetic TN/TP were most important predictor variables for the occurrence of *Aphanizomenon, Anabaena, Microcystis* and *Planktothrix* and (3) stratification periods exceeding 3 weeks caused a switch in the dominance from non-N-fixing to N-fixing cyanobacteria species affecting the ecosystem structure and functioning. These threshold conditions proved suitable to anticipate cyanobacteria development in Lake Müggelsee during heat waves in 2003 and 2006 (Huber et al., 2010).

Evolutionary computation (Holland, 1975) offers an alternative to rigid, knowledge-constrained process-based modelling. It applies principles of natural selection and evolution to infer multivariate IF–THEN-ELSE models from complex data patterns based on the cognitive principles 'generative creation' and 'choices over open-ended possibilities' (Holland et al., 1986). When applied to population densities, IF conditions of these models reveal explicit thresholds and sensitivity functions between predictor and output variables that quantify ecological interrelationships indicative for outbreaks of population density (Recknagel et al., 2014c). Resulting models are not relying on ad hoc measured data but use routine monitored data, can be easily updated by re-training with annually updated data, are simple and therefor fully transparent, and are easy to implement for early warning and operational control.

Here we apply the hybrid evolutionary algorithm (HEA) (Cao et al., 2013; Recknagel et al., 2014a) for modelling population dynamics of the cyanobacteria *Anabaena*, *Aphanizomenon* and *Microcystis* in two different time periods in Lake Müggelsee. In Phase I (1979–1990), the lake appeared to be hypertrophic, and in Phase II (1997–2012), the lake became eutrophic with decreasing external nutrient loads.

The application of HEA to water quality data patterns of 12 years in Phase I and 16 years in Phase II of Lake Müggelsee aimed at (1) to develop models for 7-day-ahead forecasting of population dynamics of the three cyanobacteria within the two phases, (2) to quantify threshold conditions for sudden outbreaks of population densities within the two phases, and (3) to discover differences in the interrelationships between physical–chemical parameters and the three cyanobacteria between the two phases. The forecasting horizon of 7-day-ahead has been chosen exemplarily since it is typical for operational algal bloom control. Based on outcomes (2) and (3) we aimed at testing the following hypotheses: 1. More frequent seasonal NO_3-N-limitation in Phase II makes the three cyanobacteria more reliant on sediment-sourced NH_4 and N-fixing capability. 2. More frequent summer heatwaves in Phase II expand temperature ranges favouring outbreaks of cyanobacteria population density. 3. Buoyant *Microcystis* is less limited by underwater light in Phase I than *Anabaena* and *Aphanizomenon*.

Materials and methods

Data of Lake Müggelsee

Lake Müggelsee (Fig. 1) has been sampled weekly or biweekly (during winter) since 1979. Until 1987, samples were taken from 0.5, 4 and 7 m depth at the deepest site. Since then, 21 volumetrically weighed subsamples from five stations have been integrated. Secchi disc depth was measured at the deepest site. Concentrations of soluble reactive phosphorus (PO_4-P), nitrate (NO_3-N), ammonium (NH_4-N) and soluble reactive silica (Si) were quantified using standard protocols (APHA, 2005). Water temperature and pH values were measured by a multiprobe sonde (YSI 6600, Yellow Springs Instruments, USA) at 0.5 m. Phytoplankton was counted in sedimentation chambers using an inverted microscope (Utermoehl, 1958). Lengths of counted filaments were always measured. Biomass per cell or filament was calculated assuming simple geometric bodies. Table 1 summarises limnological variables of Lake Mueggelsee utilised in this study.

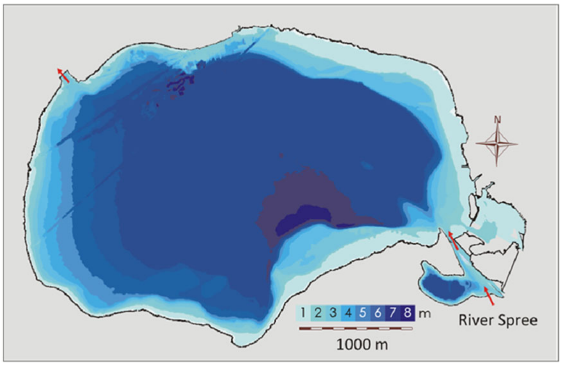


Fig. 1 Bathymetric map of Lake Müggelsee

Hybrid evolutionary algorithm HEA

Evolutionary computation infers models from data (Holland, 1975; Holland et al., 1986). The hybrid evolutionary algorithm HEA (Cao et al., 2013, 2014) has been designed to evolve "fittest IF–THEN-ELSE models" from ecological data by combining genetic programming (GP) for optimising the model structure and differential evolution (DE) for optimising model parameters. GP is applied to search for the optimum model structure by continually recombining arithmetic and logic operators based on principles of crossover, mutation and reproduction as explained in great detail in Recknagel et al. (2014a). The DE is an evolutionary algorithm designed for parameter optimisation (Storn & Price, 1997) that extracts differential information (i.e. distance and direction to global

Table 1 Limnological data of Lake Müggelsee measured from 1979 to 1990 and 1997 to 2012

Limnological variable	1979–1990 Hypertrophic Avg/min/max	1997–2012 Eutrophic
Water temperature (°C)	11.3/0.1/24.6	11.6/0.2/26.6
Secchi depth (m)	1.5/0.4/5.1	2/0.5/6.3
pH	8.2/7/9.4	8.3/7.1/9.5
PO_4-P (μg/l)	58.5/1/474	70.6/2/521
NO_3-N (mg/l)	1.2/0.004/7.4	0.33/0.01/1.91
NH_4-N (mg/l)	0.2/0.005/4.2	0.096/0.01/0.81
Si (mg/l)	3.58/0.01/8.1	4.41/0.05/10.5
Anabaena (mg/l)	0.044/0.001/5.34	0.084/0.001/9
Aphanizomenon (mg/l)	2.33/0.001/73.21	0.55/0.001/29.27
Microcystis (mg/l)	0.36/0.001/12.67	0.082/0.0001/6.23

optimum) from the current population of solutions as guide to search for the global optimum.

The daily-interpolated time series data of Lake Müggelsee from 1979 to 1990 and from 1997 to 2012 provided a wealth of seasonal and interannual patterns of abiotic and biotic limnological variables suitable for modelling by HEA. In order to take full advantage of the information content of the data, a cyclic bootstrap scheme was applied that randomly selected different data subsets for training and testing (see Recknagel et al., 2013) for each of 80 evolved generations of models. After 100 bootstrap runs, the overall "fittest model" of all 80 generations was determined. The fitness of each model was evaluated by the root mean squared error (RMSE) between the measured training data $\hat{y}_i$ and the predicted data y_i defined as follows:

$$\text{Fitness} = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(\hat{y}_i - y_i)^2}.$$

The software HEA automatically carries out sensitivity analyses for the input variables of each discovered model. For this purpose, it calculates output trajectories separately for each input range (mean ± SD) by keeping remaining input variables constant at mean values. Resulting sensitivity curves visualise the output trajectories in percentage terms (0–100%) within their range of each input. In this study, sensitivity curves of 20 models with the highest coefficients of determination r^2 and P values smaller than 0.05 of both Stages I and II were averaged and visualised (Fig. 9).

Results

Average long-term changes in Lake Müggelsee between Phase I and Phase II

Figure 2a–f compare the total annual averages of selected limnological parameters for Phases I and II of the Lake Müggelsee data (Table 1). Figure 2a shows that the average water temperature of Phase II was warming up much faster from mid-March to mid-May

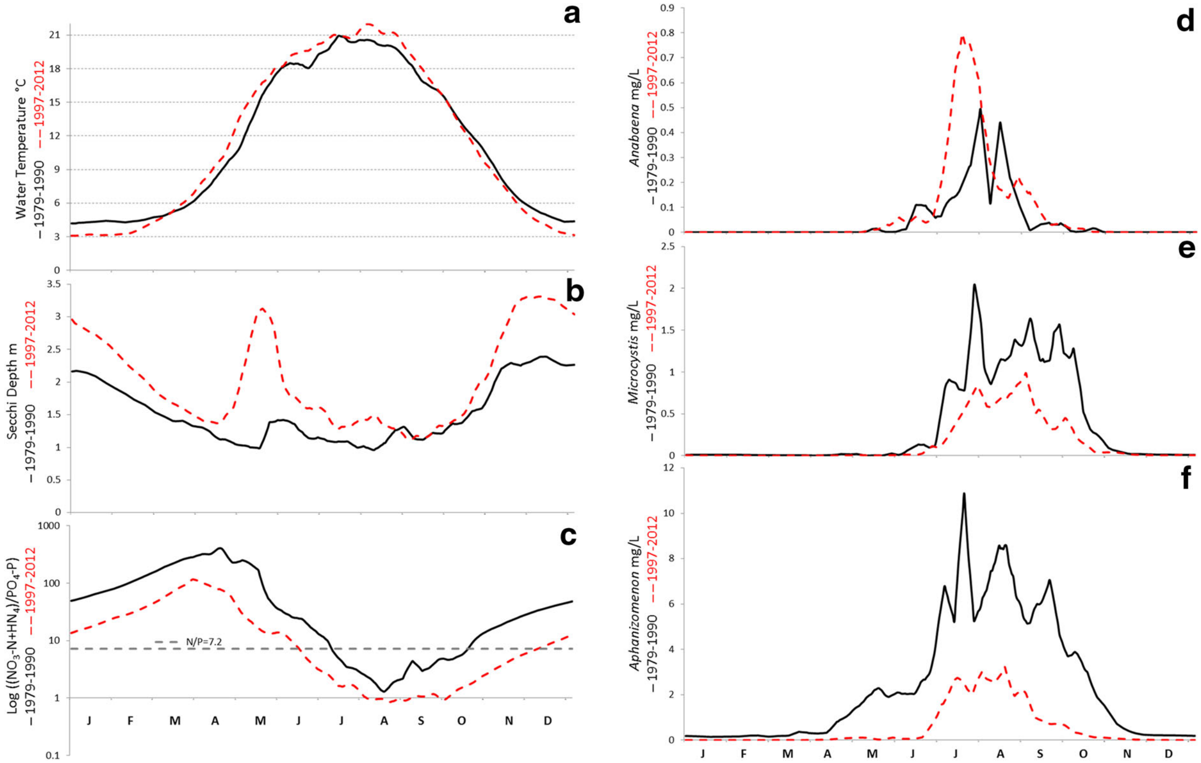


Fig. 2 Annual averages of Phase I (1979–1990) and Phase II (1997–2012) for **a** water temperature °C, **b** Secchi Depth m, **c** the logarithm of the ratio of (NO_3-N + NH_4) to PO_4-P, **d** *Anabaena* mg/l, **e** *Microcystis* mg/l and **f** *Aphanizomenon* mg/l

and reached up to 2°C warmer temperatures from June till August suggesting that stratification periods may have occurred earlier and lasted longer in summer. Interestingly the averaged winter temperatures in Phase II were about 1°C lower in December, January and February.

After averaging the observed Secchi depths (Fig. 2b), it became obvious that the water transparency improved significantly in phase II with a distinct clear water stage in early summer. The averaged ratios of dissolved inorganic nitrogen to dissolved inorganic phosphorus in Fig. 2c clearly reflected that lowered external nutrient loads have successfully extended the period of nitrogen limitation from 3 months in Phase I to 6 months from mid-June to mid-December in Phase II. Figure 2d demonstrated on average much higher biomass of *Anabaena* in July in Phase II compared to Phase I, and may indicate that it benefited from longer lasting thermal stratification (Wagner & Adrian, 2009) and eased competition by chloro- and bacillariophyta subject to zooplankton grazing during the clear water stage observed in

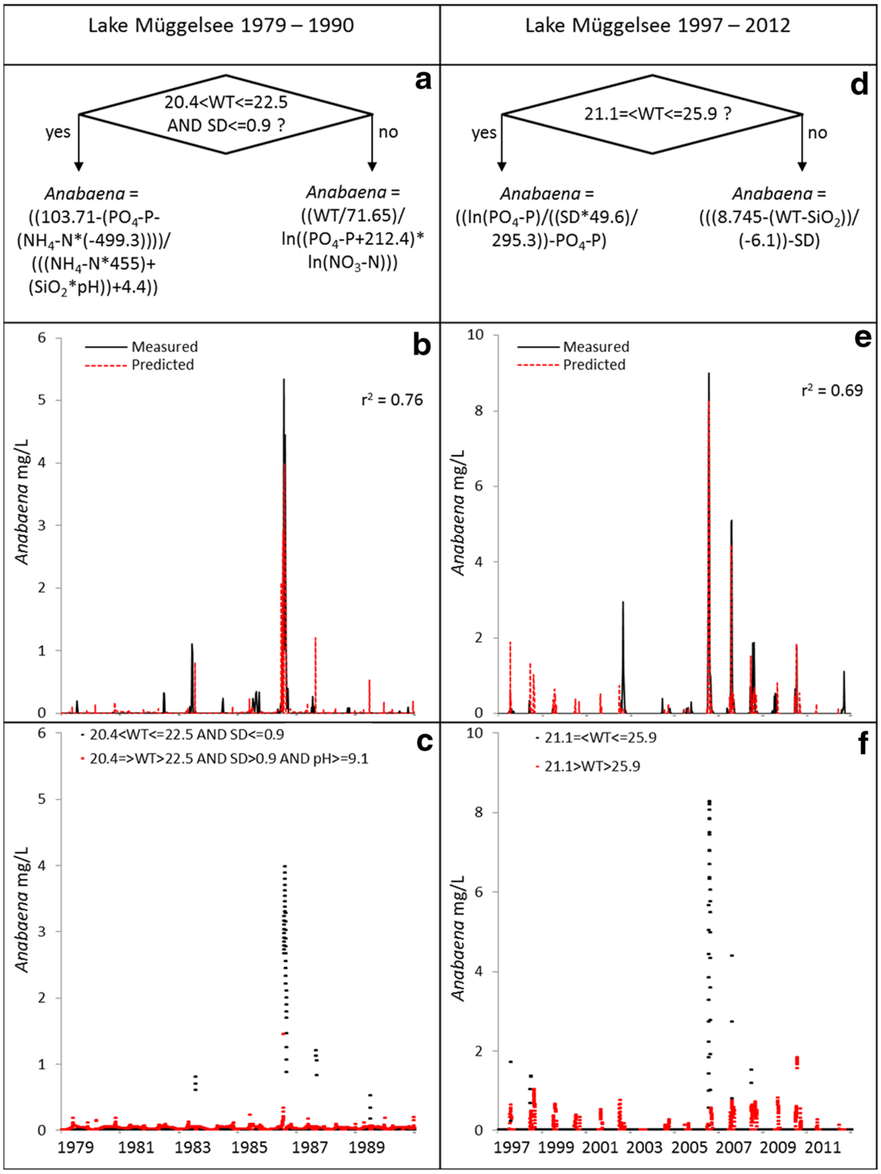


Fig. 3 7-Day-ahead forecasting of *Anabaena* in Lake Müggelsee. Phase I: **a** IF–THEN-ELSE model, **b** validation of the model, **c** illustration of threshold conditions of the model, Phase II: **d** IF–THEN-ELSE model, **e** validation of the model and **f** illustration of threshold conditions of the model

Fig. 2b. This finding does not apply to *Microcystis* and *Aphanizomenon* that show significantly declined average biomass throughout the year in Phase II compared to Phase I (Fig. 2e, f).

7-Day-ahead Forecasting

One hundred models have been evolved by HEA for each of the three cyanobacteria in each Phase I (1979–190) and Phase II (1997–2012) based on repeated bootstrap runs. Figures 3, 5 and 7 document the models with the highest coefficients of determination r^2 and P values smaller than 0.05. The selected model for *Anabaena* in Phase I achieved an $r^2 = 0.76$ and forecasted well the timing and magnitudes of two major peaks in 1983 and in 1986 (Fig. 3b). Minor peaks (< 0.5 mg/l) in the remaining years were matched by timing but either over- or underestimated in terms of magnitude. The underlying model (Fig. 3a) identified the water temperature range between 20.4 and 22.5°C and Secchi depths below 0.9 m as indicative for peaking biomass of *Anabaena* in Phase I. Figure 3c clearly illustrates that these threshold conditions prove to be predictive for *Anabaena* biomass greater than 0.5 mg/l. This finding is underpinned by Fig. 4 that shows how the WT and Secchi depth ranges of the model separate high and low values of the measured *Anabaena* biomass. Figure 3e provides the validation result of the *Anabaena* model developed for Phase II as documented in Fig. 3d. This model matched well peaking biomass of *Anabaena* between 2006 and 2010, but failed to predict an *Anabaena* outbreak in 2002. It achieved an $r^2 = 0.69$ and identified a slightly wider water temperature range between 21.1 and 25.9°C as indicative for biomass greater than 1 mg/l in Phase II (Fig. 3f).

The two models for *Aphanizomenon* achieved $r^2 = 0.48$ for Phase I and $r^2 = 0.42$ for Phase II as documented in Fig. 5. The model for Phase I identified WT greater than 22.2 or PO_4-P concentrations smaller than 41.9 µg/l as conditions concurring with biomass of *Aphanizomenon* greater than 10 mg/l (see Fig. 5c). Figure 6a, b illustrate that these conditions separate higher and lower values when applied to the measured biomass of *Aphanizomenon*. The model for Phase II discovered the water temperature range between 20 and 23.4°C and PO_4-P concentrations below 122.9 µg/l as being predictive for biomass of *Aphanizomenon* greater than 3 mg/l (Fig. 5d–f). Whilst this model forecasted reasonably well seasonal and interannual dynamics of *Aphanizomenon* biomass for most of the years between 1997 and 2012, it overestimated biomass magnitudes in 2002 and 2005. Figures 6c, d illustrate how the IF condition of the model divides observed *Aphanizomenon* data between below and above 3 mg/l.

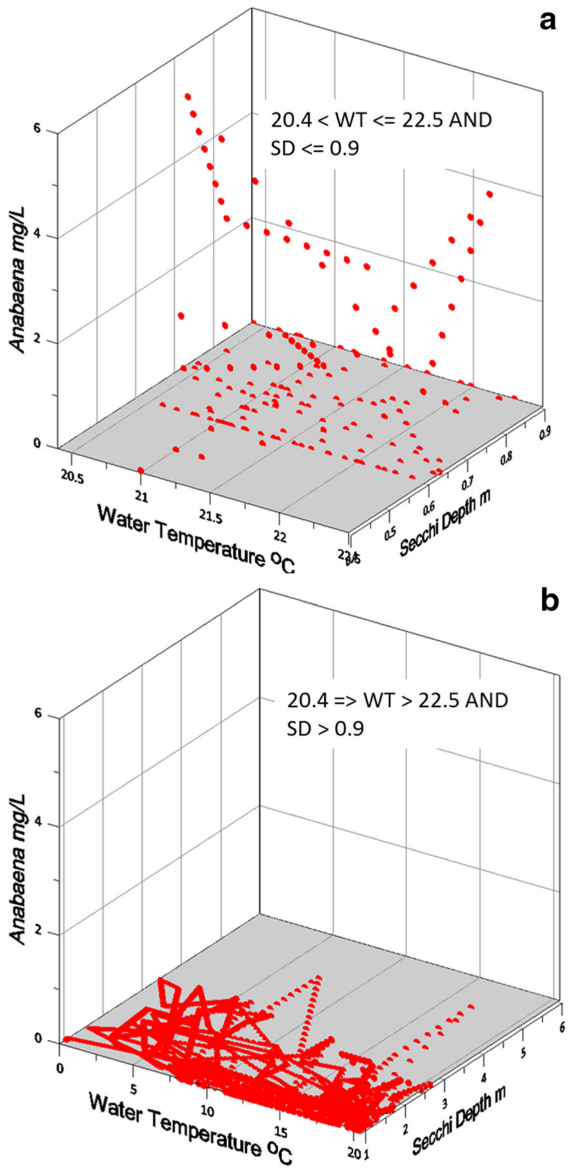


Fig. 4 Functioning of the IF condition as threshold for forecasting of **a** high population densities by the THEN equation and **b** low population densities by the ELSE equation of the model for Anabaena for Phase I

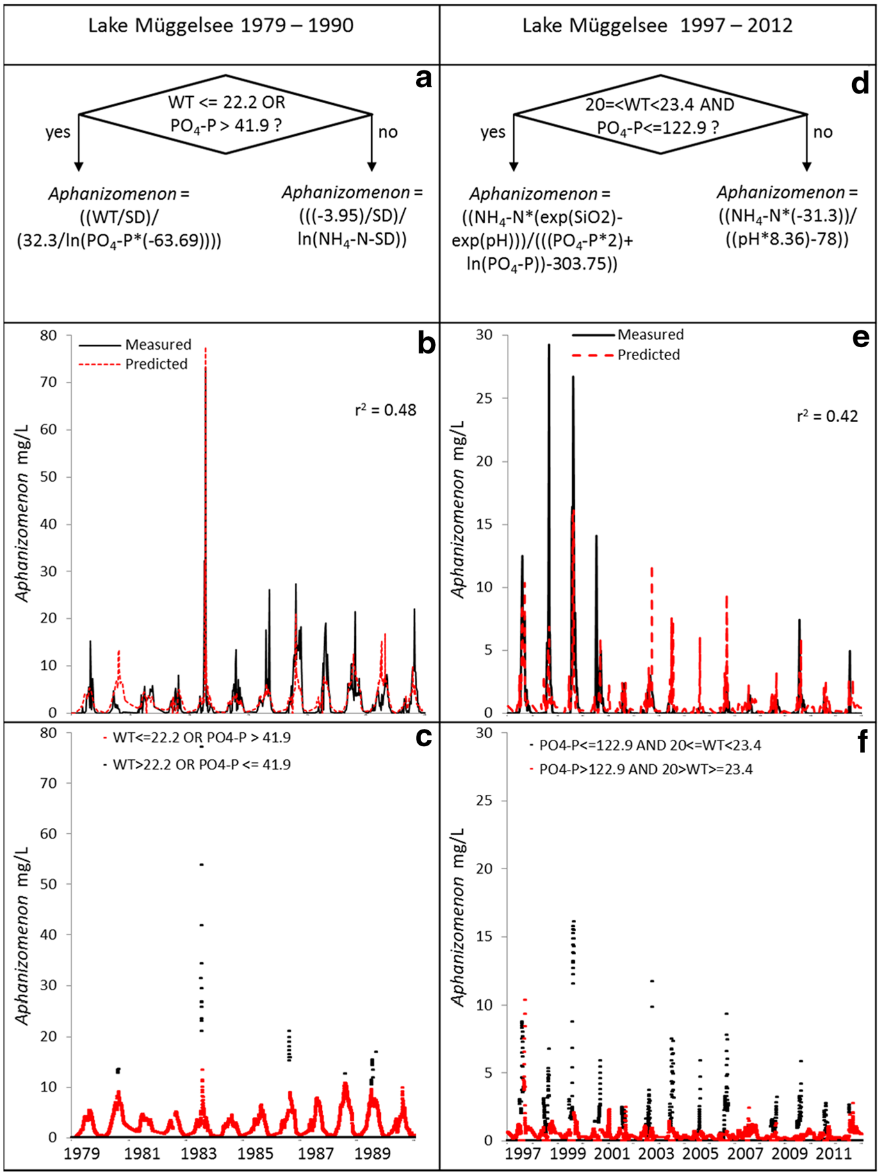


Fig. 5 7-Day-ahead forecasting of *Aphanizomenon* in Lake Müggelsee. Phase I: **a** IF–THEN-ELSE model, **b** validation of the model, **c** illustration of threshold conditions of the model, Phase II: **d** IF–THEN-ELSE model, **e** validation of the model and **f** illustration of threshold conditions of the model

The IF conditions of the *Microcystis* model for Phase I identified water temperatures between 14.5 and 22.2°C and NO_3-N concentrations smaller than 0.041 mg/l as indicative for biomass greater than 1 mg/L (see Fig. 7a, c). The model underestimated the biomass measured for *Microcystis* in 1982 but matched well timing and magnitudes of peaking biomass for the remaining years with an $r^2 = 0.54$ (Fig. 7b). Even though the *Microcystis* model for Phase II selected the same threshold criteria, their ranges were quite different compared to the model for Phase I with water temperatures between 20.9 and 23.7°C and NO_3-N concentrations smaller than 0.2 mg/l (Fig. 7d) that separated well *Microcystis* values below and above 0.5 mg/l (Fig. 7f). It achieved an $r^2 = 0.48$ but failed to forecast high *Microcystis* biomass observed in 2002 and 2006 (Fig. 7e). When the IF conditions of the two models were tested with measured *Microcystis* data of the two phases as illustrated in Fig. 8, higher and lower biomass were clearly selected by these conditions.

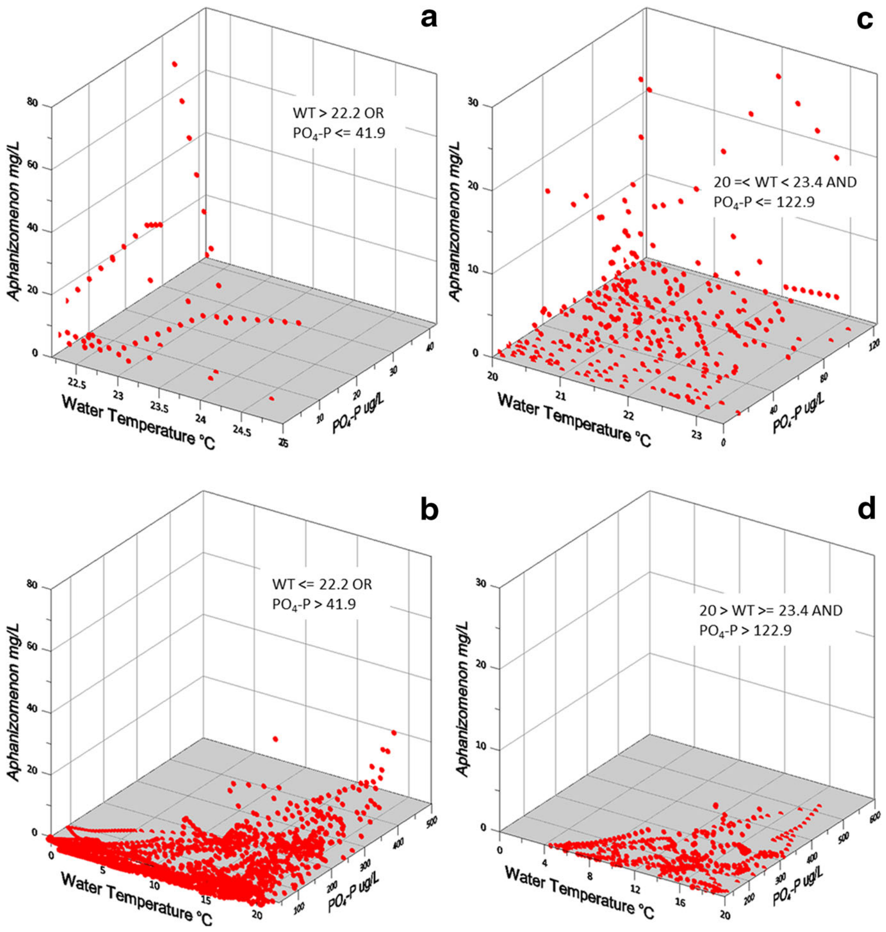


Fig. 6 Functioning of the IF condition as threshold for forecasting of: **a** high population densities by the THEN equation and **b** low population densities by the ELSE equation of the model for *Aphanizomenon* for Phase I, **c** high population densities by the THEN equation and **d** low population densities by the ELSE equation of the model for *Aphanizomenon* for Phase II

Interrelationships

Figure 9 shows averaged sensitivity functions of the three cyanobacteria in relation to Secchi depth and concentrations of NO_3-N and NH_4 for Phase I (1979–1990) and Phase II (1997–2012) extracted from 20 cyanobacteria-specific models for each phase with the highest coefficients of determination r^2 and p-values smaller than 0.05. As Fig. 9a, b indicate, both *Anabaena* and *Aphanizomenon* were more limited by underwater light in Phase I than in Phase II. Whilst the highest biomass of both cyanobacteria corresponded with the lowest Secchi depths in both phases most likely as result of self-shading, biomass remained still high at 4–5 mg/l at highest Secchi depths in Phase II but tended to be close to zero in Phase I. This may reflect the fact that the Phase I was less nutrient limited (see Fig. 2c) but more light limited (see Fig. 2b) compared to Phase II.

Results for *Microcystis* displayed a similar trend in Phase II with becoming extinct at Secchi depths greater than 2 m but responding slightly positive to increasing Secchi depths in Phase I (Fig. 9c). The finding for Phase I may refer to the fact that *Microcystis* withstands underwater light limitation at less nutrient limiting conditions by buoyancy enabled by its cell-intern gas vacuoles also reflected by almost neutral sensitivity to Secchi depth.

Anabaena displayed positive relationships with increasing NO_3-N concentrations in both phases (Fig. 9d). Figure 9e suggests a slightly negative relationship of *Aphanizomenon* with NO_3-N that is more distinct in Phase I under less nitrogen-limiting conditions than in Phase II. The fact that biomass of *Aphanizomenon* remained still high with 8 mg/l at NO_3-N concentrations of 1.2 mg/l in Phase II demonstrates that it may take advantage of its N-fixing capacity.

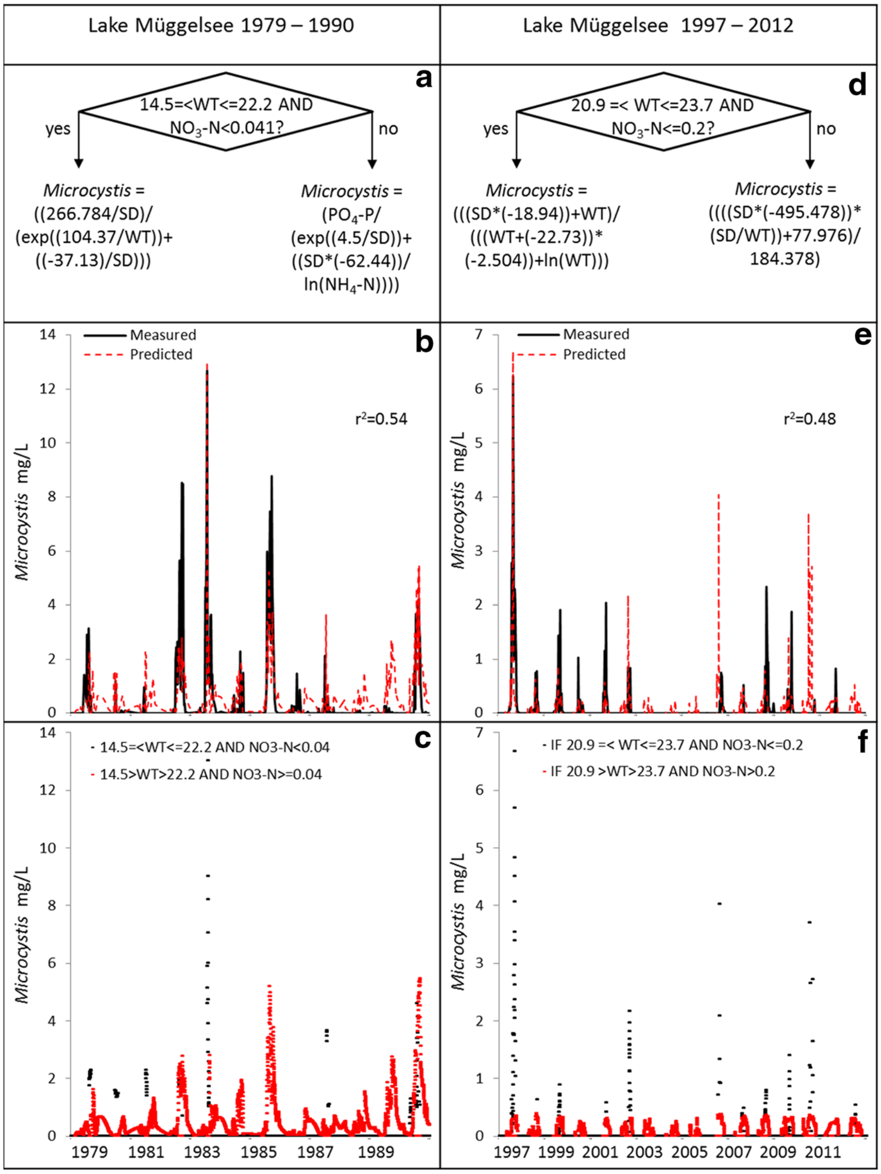


Fig. 7 7-Day-ahead forecasting of *Microcystis* in Lake Müggelsee. Phase I: **a** IF–THEN-ELSE model, **b** validation of the model, **c** illustration of threshold conditions of the model, Phase II: **d** IF–THEN-ELSE model, **e** validation of the model and **f** illustration of threshold conditions of the model

Even though NO_3-N concentrations were up to 4 times higher in Phase I compared to Phase II, *Microcystis* displayed highest biomass at lowest NO_3-N concentrations and vice versa in Phase I (Fig. 9f) indicating seasonally high N-consumption by all phytoplankton phyla. In Phase II, *Microcystis* biomass grew hyperbolically with increasing NO_3-N concentrations towards a plateau at about 3.5 mg/l. Figure 9f clearly shows that the growth of *Microcystis* is inhibited by the extended nitrogen limitation in Phase II (see also Fig. 2c).

Negative logarithmic relationships with ammonium appeared to be typical for the three cyanobacteria in Phase I reflecting their seasonal competition for N (Fig. 9g–i). However, in Phase II, *Anabaena* and *Aphanizomenon* displayed not only positive linear relationships with increasing NH_4 concentrations indicating their preference but also increasing reliance on ammonium under extended nitrate limitation in that phase. *Microcystis* appeared to be neutral to changes in ammonium concentrations in Phase II.

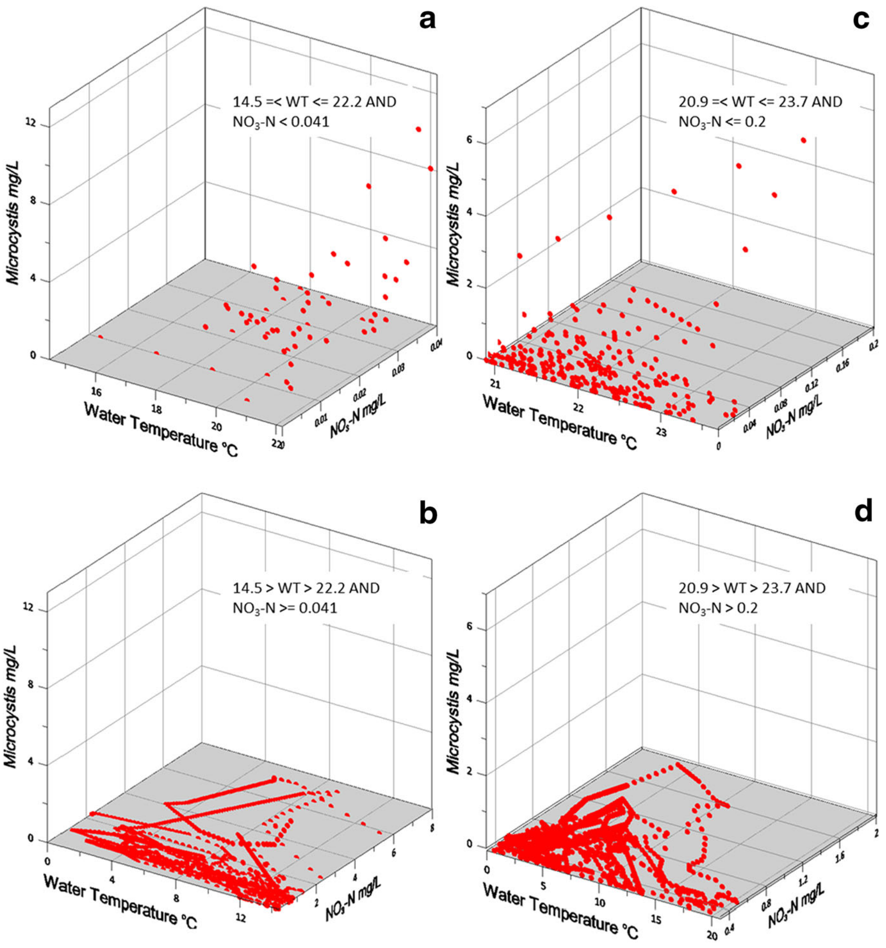


Fig. 8 Functioning of the IF condition as threshold for forecasting of: **a** high population densities by the THEN equation and **b** low population densities by the ELSE equation of the model for *Microcystis* for Phase I, **c** high population densities by the THEN equation and **d** low population densities by the ELSE equation of the model for *Microcystis* for Phase II

Discussion

(1) Differences in annual averages of physical parameters of Lake Müggelsee revealed that water temperatures were rising earlier and lasting up to 2°C higher during summer in Phase II, and Secchi depth was up to 1 m higher throughout the year including a distinct clear water stage in early summer in Phase II compared to Phase I. These findings were reflected by higher and broader ranged temperature thresholds of the models supposed to be favouring the development of *Anabaena*, *Aphanizomenon* and *Microcystis*. The sensitivity analyses displayed high biomass of *Anabaena* and *Aphanizomenon* at high water transparency in Phase II.

(2) The strengthening of nitrogen limitation throughout summer and autumn in Lake Müggelsee in Phase II was also reflected by threshold conditions of the cyanobacteria models. Typical concentrations of PO_4-P concurring with high *Aphanizomenon* biomass in Phase I were below 41.9 μg/l and in Phase II below 122.9 μg/l. It indicated that even though three times more PO_4-P was available in Phase II, it had not been utilised by phytoplankton most likely because of N-limiting stoichiometry. In addition, sensitivity analyses revealed lesser reliance of *Anabaena* and *Aphanizomenon* on NO_3-N concentrations but increasing utilisation of ammonium in Phase II possibly combined with N-fixation from gaseous sources. By contrast, the non-N-fixing *Microcystis* demonstrated increasing reliance on NO_3-N in Phase II.

(3) The models for the three cyanobacteria forecasted well the timing of peaking biomass in both phases. However, magnitudes of high peak events were partially underestimated and magnitudes of low peak events overestimated.

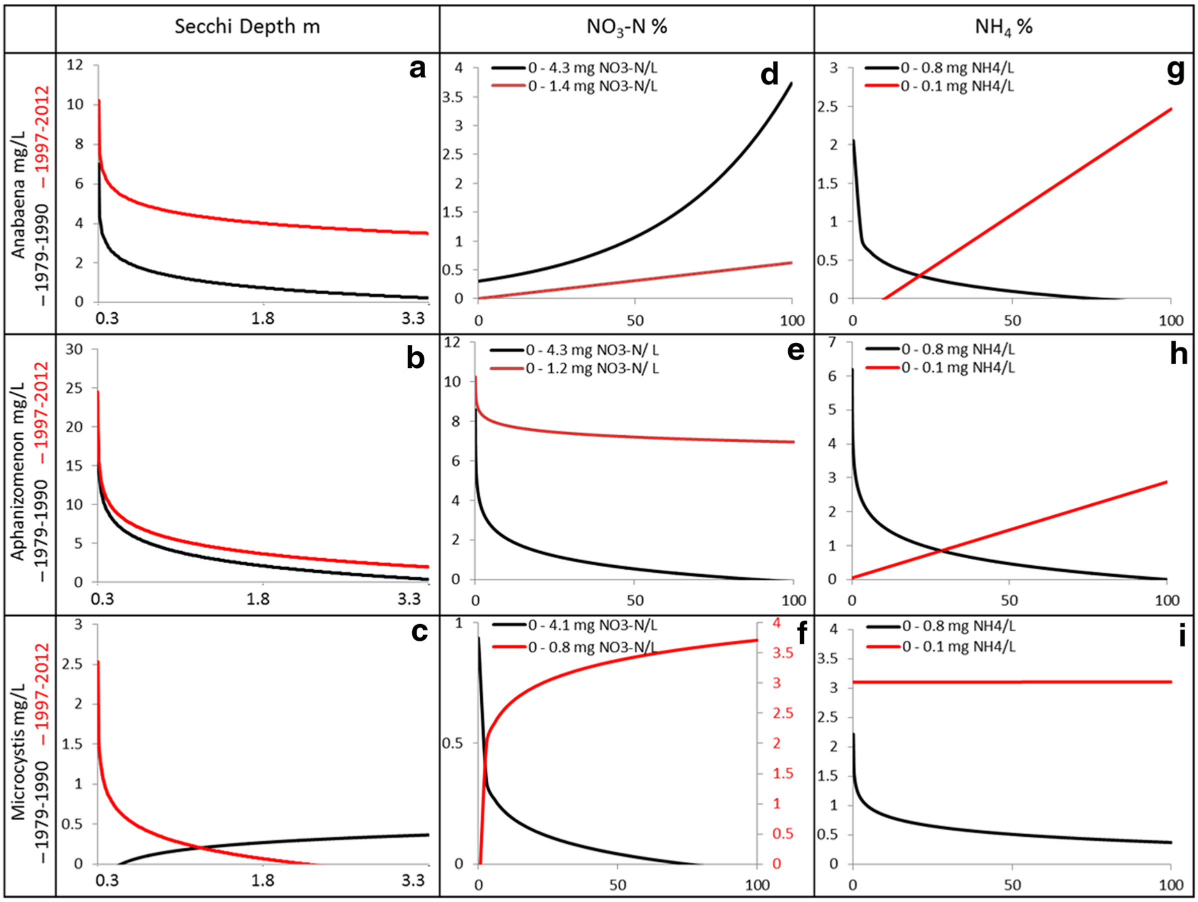


Fig. 9 Average sensitivity functions of 20 cyanobacterium-specific models for phase I and phase II revealing interrelationships between: Secchi depth and: **a** *Anabaena*, **b** *Aphanizomenon* and **c** *Microcystis*; NO_3-N and **d** *Anabaena*, **e** *Aphanizomenon* and **f** *Microcystis*; NH_4 and **g** *Anabaena*, **h** *Aphanizomenon* and **i** *Microcystis*

(4) Threshold conditions of the *Anabaena* models indicating biomass greater than 0.5 mg/l in Phase I and biomass greater than 1 mg/l in Phase II included the water temperature range between 20.4 and 22.5°C, and Secchi depths smaller than 0.9 m in Phase 1 and the water temperature range between 21.1 and 25.9°C in Phase II. The temperature ranges corresponded well with optimum temperatures for maximum growth of *Anabaena* (Reynolds, 1984), whereby the wider range to a much higher upper temperature conformed with the higher summer temperatures in Phase II (see Fig. 2a) and distinct heatwaves in 2003 and 2006 (Huber et al., 2010). The selection of diminished water transparency as threshold in Phase I could be attributed to shading by high algal biomass during the peak growing season.

(5) Threshold conditions of the *Aphanizomenon* models indicating biomass greater than 10 mg/l in Phase I and greater than 3 mg/l in Phase II included water temperature and PO_4-P concentrations in both phases. Whilst water temperatures greater than 22.2°C were selected for Phase I, the range between 20 and 23.4°C was revealed for Phase II corresponding well with optimum temperatures for maximum growth of *Aphanizomenon* (Reynolds, 1984). However, there were distinct differences in the PO_4-P concentrations suggested to be smaller than 41.9 µg/l in Phase I compared to be smaller than 122.9 µg/l in Phase II. Since on average *Aphanizomenon* was much less abundant and grew mainly in late summer in Phase II compared to Phase I (see Fig. 2f), the N-limitation was most recognisable at that time as well (Fig. 2c) leaving

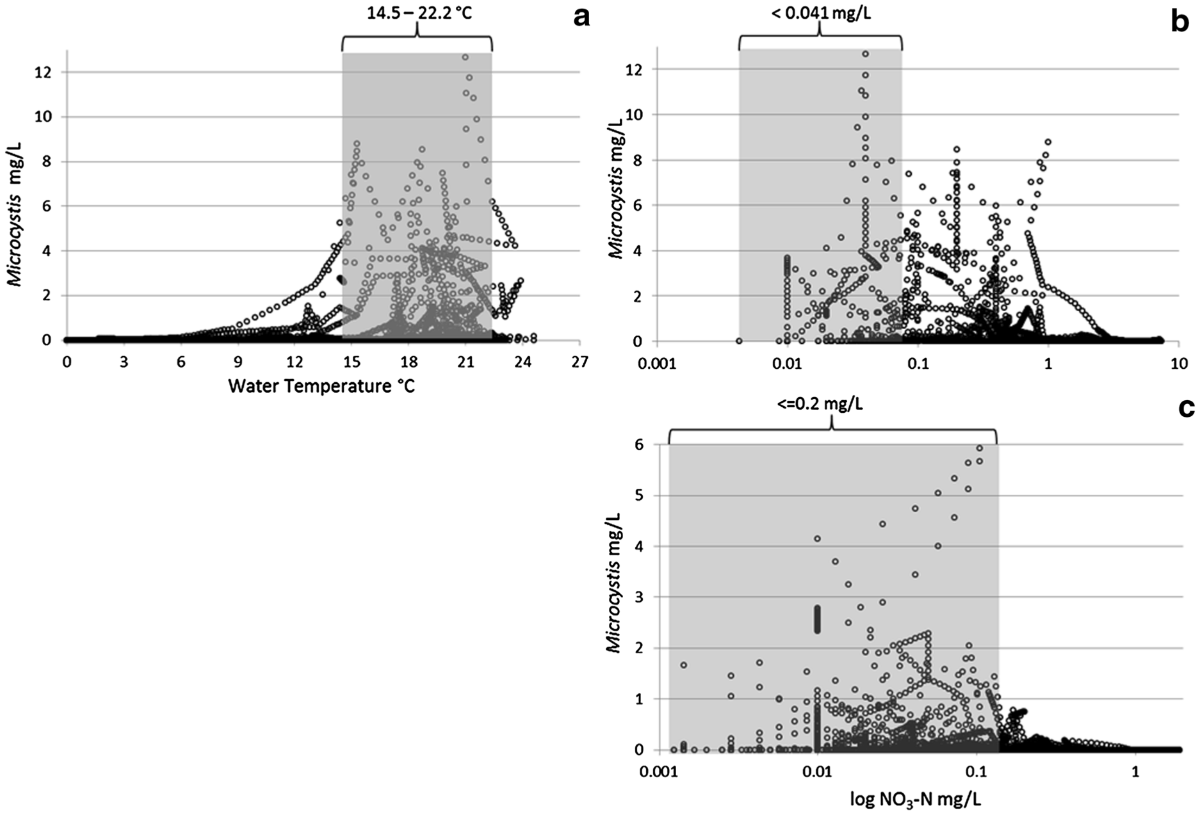


Fig. 10 Relationship between: **a** observed water temperatures and biovolumes of *Microcystis* in Lake Müggelsee in Phase I, **b** observed NO_3-N concentrations and biovolumes of *Microcystis* in Lake Müggelsee in Phase I, and **c** observed NO_3-N concentrations and biovolumes of *Microcystis* in Lake Müggelsee in Phase II

more unutilised PO_4-P in the water that perhaps occasionally has been enriched by PO_4-P remobilised from sediments by wind-induced turbulences.

(6) Threshold conditions of *Microcystis* models indicating biomass greater than 4 mg/l in Phase I and greater than 0.5 mg/l in Phase II included water temperature and concentrations of NO_3-N. The temperature range between 14.5 and 22.2°C in Phase I was somewhat surprising since optimum temperatures for *Microcystis* growth are known to be at around 23°C (Reynolds, 1984) which is well matched by the temperature range from 20.9 to 23.7°C in Phase II. However, Fig. 10a confirms that biomass of 6–8 mg/l has been observed at temperatures at around 15°C in Phase I most likely related to the transition from late summer to autumn and thus been recognised as 'high biomass' by the model. The NO_3-N concentrations that were identified by the models to coincide with high biomass of *Microcystis* in summer were smaller than 0.041 mg/l for Phase I and smaller than 0.2 mg/l for Phase II. These findings by the models correspond with the observed relationships between *Microcystis* and NO_3-N as shown in Fig. 10b, c. In Phase I between 1979 and 1990, there is a pattern of biomass greater than 8 mg/L at NO_3-N concentrations below 0.05 mg/l reflecting the fastest N-consumption at fastest growth of *Microcystis,* whilst highest biomass of greater than 1 mg/l occur at NO3-N concentrations below 0.2 mg/l in Phase II.

Overall, this study has demonstrated that the hybrid evolutionary algorithm HEA is a powerful tool for unravelling and synthesising complex relationships in ecological data by developing models that allow short-term forecasting of sudden outbreaks of population densities as prerequisite for early warning of cyanobacteria blooms. The resulting models also provide access to precious information contained in complex data by quantifying thresholds and interrelationships that explain environmental conditions which are symptomatic for outbreaks of cyanobacteria population density.

The functionality of HEA is steadily upgraded, and forecasting accuracy of models is expected to improve by continuingly testing and integrating alternative algorithms for multi-objective evolutionary computation. Future research on modelling Lake Müggelsee will focus on the succession and phenology of the phytoplankton phyla chlorophyta, bacillariophyta, cyanophyta and dinophyta taking into account not only physical–chemical predictor variables but also interrelationships with the zooplankton groups copepoda, cladocera and rotifer as well as with *Dreissena polymorpha*.

Inferential modelling by HEA suits very well for meta-analysis of lakes classified by trophic state and circulation type (Recknagel et al., 2008, 2014b, c), and is in preparation for a wide range of lakes worldwide.

Acknowledgements We thank two anonymous reviewers for their instructive comments that have significantly improved the manuscript. This research was partially funded by the EU Project LIMNOTIP under the FP7 ERA-Net Scheme (Biodiversa 01LC1207A).

References

Adrian, R., R. Deneke, U. Mischke, R. Stellmacher & P. Lederer, 1995. A long-term study of Heiligensee (1975–1992). Evidence for effects of climate change on the dynamics of eutrophied lake ecosystems. Archive for Hydrobiology 133: 315–337.

Adrian, R., C. M. O'Reilly, H. Zagarese, S. B. Baines, D. O. Hessen, W. Keller, D. M. Livingstone, R. Sommaruga, D. Straile, E. van Donk, G. A. Weyhenmeyer & M. Winder, 2009. Lakes as sentinels of climate change. Limnology and Oceanography 54: 2283–2297.

Adrian, R., D. Gerten, V. Huber, C. Wagner & S. R. Schmidt, 2012. Windows of change: temporal scale of analysis is decisive to detect ecosystem responses to climate change. Marine Biology 159: 2533–2542.

APHA, 2005. Standard Methods for the Examination of Water and Wastewater, 21st ed. American Public Health Association, Washington, DC.

Cao, H., F. Recknagel & P. Orr, 2013. Enhanced functionality of the redesigned hybrid evolutionary algorithm HEA demonstrated by predictive modelling of algal growth in the Wivenhoe Reservoir, Queensland (Australia). Ecological Modelling 252: 32–43.

Cao, H., F. Recknagel & P. T. Orr, 2014. Parameter optimisation algorithms for evolving rule models applied to freshwater ecosystem. IEEE Transactions on Evolutionary Computation 18(6): 793–806.

Huber, V., R. Adrian & D. Gerten, 2008. Phytoplankton response to climate warming modified by trophic state. Limnology and Oceanography 53(1): 1–13.

Elliott, J. A., I. D. Jones & S. J. Thackeray, 2006. Testing the sensitivity of phytoplankton communities to changes in water temperature and nutrient load, in a temperate lake. Hydrobiologia 559: 401–411.

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

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

Huber, V., R. Adrian & D. Gerten, 2010. A matter of timing: heat wave impact on crustacean zooplankton. Freshwater Biology 55: 1769–1779.

Huber, V., C. Wagner, D. Gerten & R. Adrian, 2012. To bloom or not to bloom: contrasting responses of cyanobacteria to different heat waves explained by critical thresholds of abiotic drivers. Oecologia 169: 245–256.

Huisman, J., J. Sharples, J. M. Stroom, P. M. Visser, W. E. A. Kardinaal, J. M. H. Verspagen & B. Sommeijer, 2004. Changes in turbulent mixing shift competition for light between phytoplankton species. Ecology 85: 2960–2970.

Jöhnk, K. D., J. Huisman, J. Sharples, B. Sommeijer, P. M. Visser & J. M. Stroom, 2008. Summer heatwaves promote blooms of harmful cyanobacteria. Global Change Biology 14: 495–512.

Köhler, J., H. Behrendt & S. Hoeg, 2000. Long-term response of phytoplankton to reduced nutrient load in the flushed Lake Müggelsee (Spree system, Germany). Archiv für Hydrobiologie 148: 209–229.

Köhler, J., S. Hilt, R. Adrian, A. Nicklisch, H. P. Kozerski & N. Walz, 2005. Long-term response of a shallow, moderately flushed lake to reduced external phosphorus and nitrogen loading. Freshwater Biology 50: 1639–1650.

Livingstone, D. M. & R. Adrian, 2009. Modeling the duration of intermittent ice cover on a lake for climate-change studies. Limnology and Oceanography 54(5): 1709–1722.

Mooij, W. M., L. N. De Senerpont Domis & J. H. Janse, 2009. Linking species- and ecosystem-level impacts of climate change in lakes with a complex and a minimal model. Ecological Modelling 220: 3011–3020.

Paerl, H. W., 1988. Nuisance phytoplankton blooms in coastal, estuarine and inland waters. Limnology and Oceanography 33: 823–847.

Recknagel, F., M. Hosomi, T. Fukushima & D.-S. Kong, 1995. Short- and long-term control of external and internal phosphorus loads in lakes—a scenario analysis. Water Research 29(7): 1767–1779.

Recknagel, F., H. Cao, C. van Ginkel, D. van der Molen, H. Park & N. Takamura, 2008. Adaptive agents for forecasting seasonal outbreaks of blue-green algal populations in lakes categorised by circulation type and trophic state. Verhandlungen Internationale Verein Limnologie 30(2): 191–197.

Recknagel, F., I. Ostrovsky, H. Cao, T. Zohary & X. Zhang, 2013. Ecological relationships, thresholds and time-lags determining phytoplankton community dynamics of Lake Kinneret, Israel elucidated by evolutionary computation and wavelets. Ecological Modelling 255: 70–86.

Recknagel, F., I. Ostrovsky & H. Cao, 2014a. Model ensemble for the simulation of plankton community dynamics of lake Kinneret (Israel) induced from in situ predictor variables by evolutionary computation. Environmental Modelling & Software 61: 380–392.

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

Recknagel, F., I. Ostrovsky, H. Cao & Q. Chen, 2014c. Hybrid evolutionary computation quantifies environmental thresholds for recurrent outbreaks of population density. Ecological Informatics 24: 85–89.

Reynolds, C., 1984. The Ecology of Freshwater Phytoplankton. Cambridge University Press, Cambridge.

Storn, R. & K. Price, 1997. Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization 11: 341–359.

Utermoehl, H., 1958. Zur Vervollkommnung der quantitativen Phytoplankton Methodik. Mitt. Internationale Ver. Theoretische und Angewandte Limnologie 9: 1–38.

Wilhelm, S. & R. Adrian, 2008. Impact of summer warming on the thermal characteristics of a polymictic lake: consequences for oxygen, nutrients and phytoplankton. Freshwater Biology 53: 226–237.

Wagner, C. & R. Adrian, 2009. Cyanobacteria blooms: quantifying the effects of climate change. Limnology and Oceanography 54(6): 2460–2468.