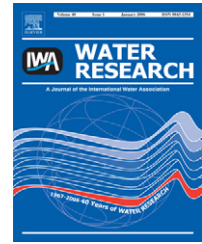


Available at [www.sciencedirect.com](http://www.sciencedirect.com)journal homepage: [www.elsevier.com/locate/watres](http://www.elsevier.com/locate/watres)

# Elucidation and short-term forecasting of microcystin concentrations in Lake Suwa (Japan) by means of artificial neural networks and evolutionary algorithms

Wai Sum Chan<sup>a,b</sup>, Friedrich Recknagel<sup>a,b,\*</sup>, Hongqing Cao<sup>a</sup>, Ho-Dong Park<sup>c</sup>

<sup>a</sup>School of Earth and Environmental Sciences, University of Adelaide, Adelaide 5005, Australia

<sup>b</sup>Cooperative Research Centre for Water Quality and Treatment, Salisbury 5108, Australia

<sup>c</sup>Department of Environmental Sciences, Shinshu University, Matsumoto 390-8621, Japan

## ARTICLE INFO

### Article history:

Received 16 July 2006

Received in revised form

1 December 2006

Accepted 5 February 2007

Available online 3 April 2007

### Keywords:

Lake Suwa

Microcystis

Microcystin

Ordination

Clustering

Forecasting

Explanation

## ABSTRACT

Non-supervised artificial neural networks (ANN) and hybrid evolutionary algorithms (EA) were applied to analyse and model 12 years of limnological time-series data of the shallow hypertrophic Lake Suwa in Japan. The results have improved understanding of relationships between changing microcystin concentrations, *Microcystis* species abundances and annual rainfall intensity. The data analysis by non-supervised ANN revealed that total *Microcystis* abundance and extra-cellular microcystin concentrations in typical dry years are much higher than those in typical wet years. It also showed that high microcystin concentrations in dry years coincided with the dominance of the toxic *Microcystis viridis* whilst in typical wet years non-toxic *Microcystis ichthyoblabe* were dominant. Hybrid EA were used to discover rule sets to explain and forecast the occurrence of high microcystin concentrations in relation to water quality and climate conditions. The results facilitated early warning by 3-days-ahead forecasting of microcystin concentrations based on limnological and meteorological input data, achieving an  $r^2 = 0.74$  for testing.

© 2007 Elsevier Ltd. All rights reserved.

## 1. Introduction

The shallow eutrophic Lake Suwa is being studied for decades with regards to its distinctive patterns of high abundances of *Microcystis* species and high concentrations of microcystin during summer (e.g. Harada et al., 2001; Park et al., 1993, 1998). In situ measurements of limnological variables of Lake Suwa from 1992 to 2003 provided precious information about complex relationships between water quality, meteorology and *Microcystis* population dynamics.

Non-supervised artificial neural networks (ANN) have been demonstrated to be useful tools for ordination, clustering and visualisation of complex data such as of water treatment plants, stream and lake habitats (e.g. Hong et al., 2003; Park

et al., 2003; Recknagel et al., 2006). Supervised ANN and evolutionary algorithms (EA) have successfully been applied to facilitate early warning of sudden outbreaks of toxic blue-green algae in freshwater lakes and rivers by time-series modelling (e.g. Recknagel et al., 1997, 2002; Cao et al., 2006; Jeong et al., 2006). However there is no example yet known for ordination, clustering and forecasting microcystin concentrations in natural water bodies using water quality time-series data.

The present study utilised non-supervised ANN and hybrid EA for unravelling complex ecological relationships in the database of Lake Suwa, and forecasting of microcystin concentrations by means of water quality and meteorological data. The study aimed at following research questions: (i) can

\*Corresponding author. Tel.: +61 8 83033953; Fax: +61 8 83036222.

non-supervised ANN determine seasonal relationships amongst water quality data, *Microcystis* species and microcystin concentrations with regards to rainfall patterns, (ii) can hybrid EA forecast 3-days ahead the timing and magnitudes of extra-cellular microcystin concentrations, and (iii) can hybrid EA discover rule sets to describe conditions for microcystin occurrence in Lake Suwa.

## 2. Materials and methods

### 2.1. Study site and data

The study site of this research was Lake Suwa located in central Japan. Lake Suwa is a shallow, hypertrophic lake with occasional weak thermal stratification in summer between April and October and frequent ice cover in winter between December and February. Lake Suwa is mainly used for recreational activities, but a small proportion of lake water is also used for aquaculture, irrigation and industrial purposes (Park et al., 1998).

*Microcystis* blooms in Lake Suwa have been reported every summer since 1970s (Harada et al., 2001; Yokoyama and Park, 2002), but Watanabe and Oishi (1985) observed that concentrations of toxic microcystin released by *Microcystis* cells varied from month to month and year to year depending on the composition of *Microcystis* species.

Concentrations of both intra- and extra-cellular microcystin were measured complementary to water quality conditions and phytoplankton abundances two to three times per month from June to October between 1992 and 2003 by Park et al. (1998) using standard methods. These data were linearly interpolated to produce daily values required for modelling by means of non-supervised ANN and hybrid EA.

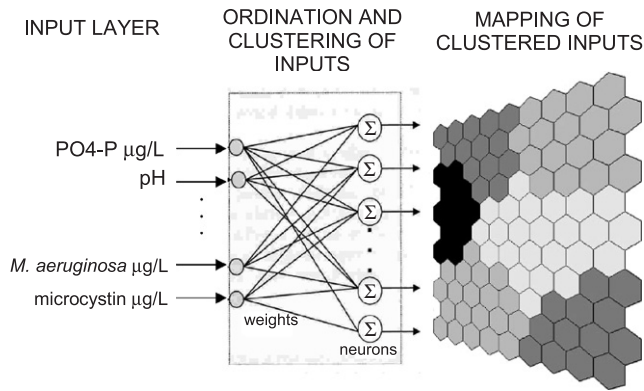
Details of the general characteristics and limnological data of Lake Suwa are summarised in Table 1.

### 2.2. Non-supervised ANN

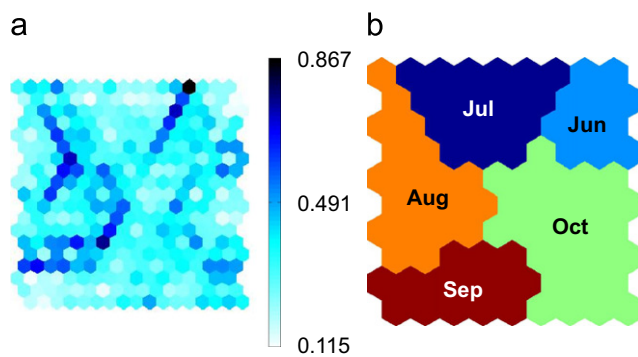
Supervised and non-supervised ANN are designed for different purposes. Supervised ANN aim at finding a transformation from known input to known output patterns. By contrast non-supervised ANN are designed to identify unknown input patterns based on similarities between inputs. The so-called self-organising maps developed by Kohonen (1982) are the most popular non-supervised ANN, which are applied to ordination, clustering and mapping of water quality and phytoplankton data with respect to seasons and rainfall patterns in the context of the present study. The principal approach of non-supervised ANN according to Kohonen (1982) is represented in a simplified manner in Fig. 1. It shows that the neurons of the non-supervised ANN learn to distinguish between similar and dissimilar features of the normalised input data, which are mapped as clustered inputs. The term non-supervised in this context means that the

**Table 1 – General characteristics and measured data of Lake Suwa**

General characteristics of Lake Suwa					
Position	36°3'N, 138°5'E				
Climate	Temperate				
Trophic status	Hypertrophic				
Surface area (km <sup>2</sup> )	13.3				
Volume (ML)	63,000				
Mean depth (m)	5				
Maximum depth (m)	6.8				
Water retention time (days)	50				
Catchment area (km <sup>2</sup> )	515				
Measured data of Lake Suwa June–October of the years 1992–2003					
	Mean	Min/median/max	ANN	EA (1)	EA (2)
Rainfall, RF (mm)	58.2	0/46.9/320.0	✓	✓	✓
Water temperature, WT (°C)	21.3	8.5/21.6/28.0	✓	✓	✓
Secchi depth, SD (cm)	94.5	31.0/99.0/190.0	✓	✓	✓
PH	8.8	6.9/8.9/10.2	✓	✓	✓
Dissolved oxygen, DO (mg/L)	9.01	1.50/9.02/14.10	✓	✓	✓
Nitrate, NO <sub>3</sub> (µg/L)	2.27	0/145.25/1231.0	✓	✓	✓
Phosphate, PO <sub>4</sub> (µg/L)	0.10	0.001/4.63/136.20	✓	✓	✓
NO <sub>3</sub> :PO <sub>4</sub> ratio	212.1	0/33/7930	✓	✓	✓
Chlorophyll a, chl <sub>a</sub> (µg/L)	86.25	7.0/54.0/1661.0	✓	✓	✓
Total <i>Microcystis</i> concentration (× 10 <sup>5</sup> cells/mL)	1.79	0/0.82/22.8	✓	✓	✓
<i>Microcystis aeruginosa</i> , <i>aeru</i> (× 10 <sup>5</sup> cells/mL)	0.24	0/0.05/3.86	✓	✓	✓
<i>Microcystis ichthyoblabe</i> , <i>icht</i> (× 10 <sup>5</sup> cells/mL)	0.63	0/0.16/17.58	✓	✓	✓
<i>Microcystis viridis</i> , <i>viri</i> (× 10 <sup>5</sup> cells/mL)	0.52	0/0/5.77	✓	✓	✓
<i>Microcystis wesenbergii</i> , <i>wese</i> (× 10 <sup>5</sup> cells/mL)	0.26	0/0.03/4.75	✓	✓	✓
Total extra-cellular microcystin (µg/L)	10.3	0/2.4/184.1	✓	✓	✓
Total intra-cellular microcystin (µg/g)	282.3	0/174.4/1905.7	✓	✓	✓



**Fig. 1 – Conceptual diagram of the structure and functioning of non-supervised ANN.**



**Fig. 2 – Ordination and clustering of Lake Suwa in months (June–October) by means of non-supervised ANN and visualised as unified distance matrix map (U-matrix) (a), and as partitioned map (K-means) (b).**

learning algorithm is not guided by known output patterns but learns the patterns from features of the inputs. Those features are expressed by Euclidean distances, which are calculated between the inputs and weights. Similarities between inputs in terms of Euclidean distances can be visualised and partitioned by the unified distance matrix (U-matrix) and the K-means map. Fig. 2 shows the monthly clusters from June to October for Lake Suwa as mapped by the U-matrix and K-means partitioning using the SOM Toolbox of MATLAB 5.3 (Vesanto et al., 2000).

The U-matrix map (Fig. 2(a)) visualises the relative distances between neighbouring data of the input data space as shades of grey. The light areas in the U-matrix visualise neighbouring data with smallest distances belonging to a region or cluster. The black colours represent the biggest distances between neighbouring data and denote borders between clusters. The K-means algorithm partitions the input data space into a specified number of clusters based on the U-matrix. Fig. 2(b) represents the corresponding partitioned map for the 5 months (June–October).

Fourteen input variables including RF, WT, SD, pH, NO<sub>3</sub>, PO<sub>4</sub>, chl<sub>a</sub>, total Microcystis, aeru, icht, viri, wese, extra-cellular microcystin, and intra-cellular microcystin (see Table 1) were used for ordination and clustering by means of the SOM

Toolbox for MATLAB 5.3 (Vesanto et al., 2000). Based on the rainfall patterns between June and October of the years 1992–2003 the input variables from the driest years 1992 and 1994 (average monthly rainfall 29.84 and 36.79 mm), and the wettest years 1999 and 2000 (average monthly rainfall 92.61 and 69.39 mm) were linearly interpolated and logistically normalised in order to control the variance of vector components and to remove erroneous values. A total number of 8512 data points was used, and the map size (9 × 10) was selected automatically by the SOM software, which determines the best-fit matrix for the dataset. The map resolution was measured by the mean quantisation error (QE) and map topology preservation was measured by the topographic error (TE) (Vesanto et al., 2000).

### 2.3. Hybrid EA

Hybrid EA are adaptive modelling techniques based on principles of biological evolution, natural selection and genetic variation (Holland, 1975). They search for suitable representations of models by means of genetic operators and the principle of “survival of the fittest”.

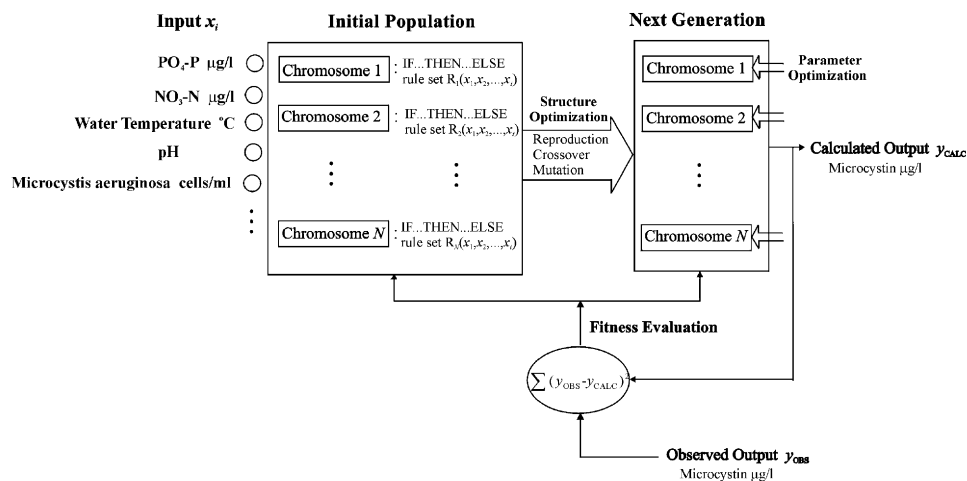
In the context of this research, a hybrid EA designed for rule discovery in water quality time-series (Cao et al. 2006) was applied. The principle framework of the hybrid EA is shown in Fig. 3. The hybrid EA uses genetic programming (GP) (e.g. Koza, 1992) to generate and optimize the structure of rule sets, and a genetic algorithm (GA) (e.g. Mitchell, 1996) to optimise the parameters of a rule set.

The hybrid EA was applied for short-term forecasting of extra-cellular microcystin concentrations in Lake Suwa by means of: (1) physical and chemical lake data (SD, WT, RF, pH, DO, NO<sub>3</sub>:PO<sub>4</sub> ratio, NO<sub>3</sub>, PO<sub>4</sub>, and chl<sub>a</sub>), and (2) data set (1) plus biological data (aeru, icht, viri, and wese) in Table 1. Model 1 was trained by 11 years of data (15,048 data points) and tested by data of an independent year 1998 (1341 data points). Due to the missing biological data of the year 1997, Model 2 was only trained by 10 years of data (19,760 data points) and tested by data of an independent year 1998 (1937 data points).

Three criteria were used to determine the best performing predictive models discovered by hybrid EA: (1) the root mean square error (RMSE), (2) the square of correlation coefficient ( $r^2$ ) of a linear regression between predicted and measured data, and (3) visual comparison between predicted and measured data.

For all applications of the hybrid EA in this study, an initial population of 200, a maximum number of generations of 100, and repetitive runs of 100 were chosen based on expertise gained from trial and error experiments. The applications of the hybrid EA were performed by a Hydra supercomputer (IBM eServer 1350 Linux) with the highest speed of 1.2 Tflops.

Sensitivity analyses were carried out for the best performing predictive rule sets as follows. Firstly, all data were divided according to the THEN and ELSE branches of the IF criteria. Secondly, the values of mean, standard deviation (STDEV), mean–STDEV, and mean+STDEV were calculated for each variable considered in the equations of the THEN or ELSE branch. Finally, the sensitivities of microcystin concentrations to specific variables were calculated by fixing other variables to the mean values in the THEN and ELSE equations.



**Fig. 3 – Conceptual diagram of the hybrid EA for the discovery of predictive rule sets in water quality time-series (Cao et al., 2006).**

### 3. Results

#### 3.1. Ordination and clustering of water quality data by non-supervised ANN

The ordination and clustering of water quality data of Lake Suwa by means of non-supervised ANN has been conducted according to Figs. 1 and 2 distinguishing between typical dry years (1992 and 1994) and wet years (1999 and 2000).

Results in Fig. 4 indicate distinctive differences both in magnitudes as well as seasonal occurrences of *Microcystis* abundances between dry and wet years. Whilst maximum cell numbers peak in the dry years at 515,000 cells/mL in autumn (September and October), they peak in the wet years at 97,000 cells/mL in late summer (August). Fig. 4 reveals that in the dry years *Microcystis viridis* contributes up to 85% (441,000 cells/mL) and *Microcystis wesenbergii* up to 15% (80,000 cells/mL) to the maximum cell numbers in autumn. By contrast, Fig. 4 shows that in the wet years *Microcystis ichthyoblabe* solely forms the maximum cell number of 97,000 cells/mL from July to October.

Fig. 5 illustrates the corresponding patterns for extra-cellular microcystin and nutrient concentrations as well as rainfall for the dry and wet years. In concurrence with highest *Microcystis* abundances in Fig. 4, the extra-cellular microcystin concentrations in Fig. 5 for the dry years are highest in October peaking at 102 μg/L and for the wet years are highest in September peaking at 3.6 μg/L.

Fig. 5 also indicates significant differences in the nutrient patterns between the wet and the dry years. Interestingly, the highest PO<sub>4</sub> concentrations occur in the dry years peaking at 109 μg/L in September whilst PO<sub>4</sub> concentrations in the wet years peak at 9.3 μg/L in August. By contrast, NO<sub>3</sub> concentrations in the dry years only peak at 290 μg/L in October but reach the highest concentrations at 590 μg/L in both August and October in the wet years. Fig. 5 also shows the distinctively different rainfall levels between the selected dry years with high rainfall events up to 70 mm in late

summer and October, and the selected wet years with high rainfall events up to 187 mm in October.

#### 3.2. Predictive modelling of microcystin concentrations by hybrid EA

Two models for 3-days-ahead forecasting of extra-cellular microcystin concentrations in Lake Suwa were developed by means of hybrid EA based on: (1) physical and chemical lake data only, and (2) physical, chemical and biological data (see Section 2.3).

The training and testing results are shown in Figs. 6(a) and (b) for model (1) whereas in Figs. 7(a) and (b) for model (2). For the training results, both models (Figs. 6(a) and 7(a)) indicate good correlations between measured and predicted data, with high  $r^2$  values of 0.64 and 0.82, respectively. In particular, model (2) successfully captures the extreme variations of microcystin concentrations in Lake Suwa.

For the independent testing of the models (Figs. 6(b) and 7(b)), the hybrid EA effectively forecast the right timing of three peaks of microcystin concentrations in mid June, late July and September of 1998. However, under-estimation of microcystin concentration appears to be a problem in model (1) (Fig. 6(b)) whereas over-estimation occurs in model (2) (Fig. 7(b)).

#### 3.3. Rule set discovery and sensitivity analysis by hybrid EA

The hybrid EA discovered the predictive rule set in Fig. 8(a) for the forecasting of microcystin concentrations by means of physical and chemical inputs as documented in Fig. 6. The IF criterion PO<sub>4</sub> > 70.022 μg/L of the rule set either activates the THEN or the ELSE equation in order to determine the daily microcystin concentrations in Lake Suwa. The input sensitivities of microcystin either for the THEN equation (Fig. 8(b)) or the ELSE equation (Fig. 8(c)) illustrate clearly the two different cases. Microcystin concentrations of up to 100 μg/L are

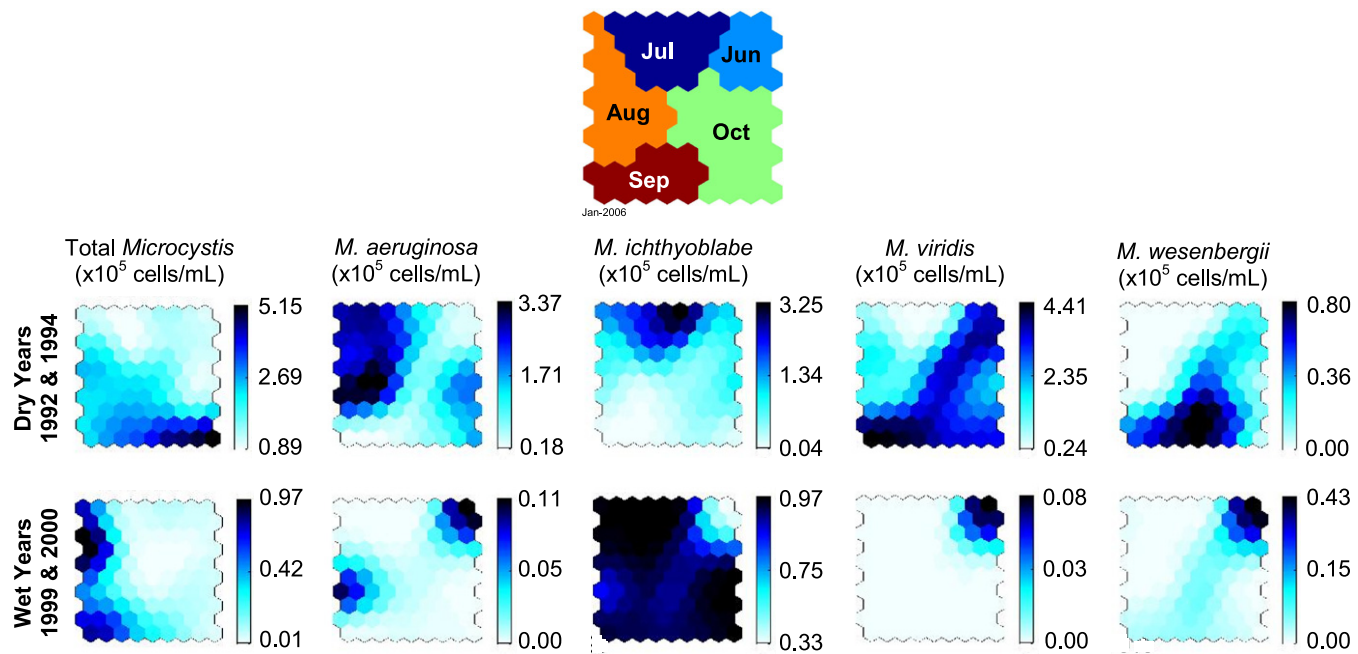


Fig. 4 – Component planes of four *Microcystis* species compositions in Lake Suwa with regards to seasonality (months) and rainfall patterns (typical dry and wet years).

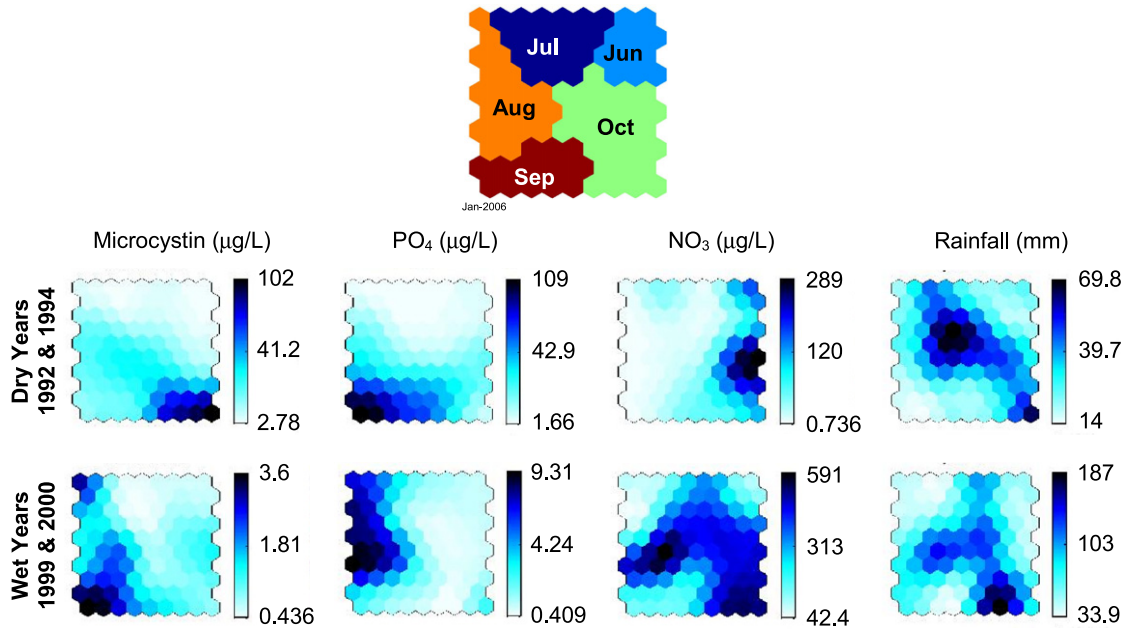


Fig. 5 – Component planes of extra-cellular microcystin concentration, nutrient contents and rainfall in Lake Suwa with regards to seasonality (months) and rainfall patterns (typical dry and wet years).

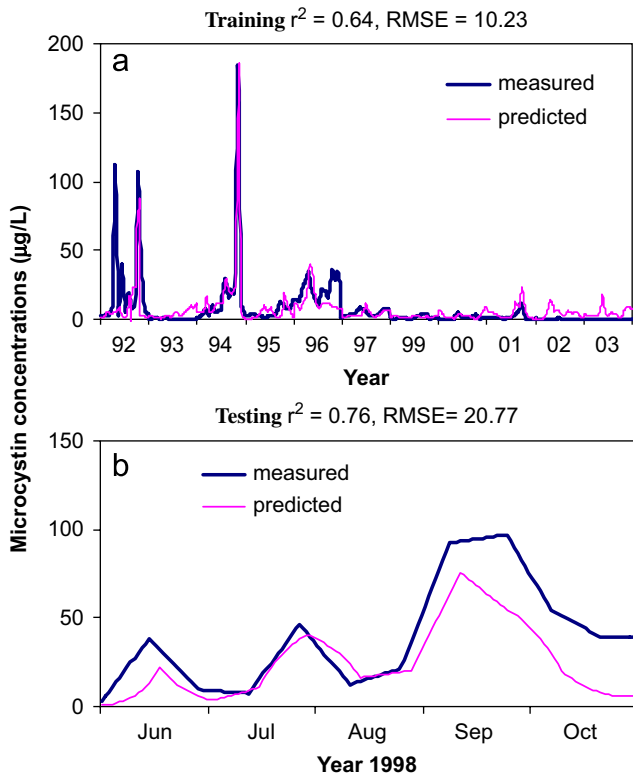


Fig. 6 – Three-days-ahead forecasting of extra-cellular microcystin concentrations in Lake Suwa by physical and chemical inputs: (a) training results and (b) testing results.

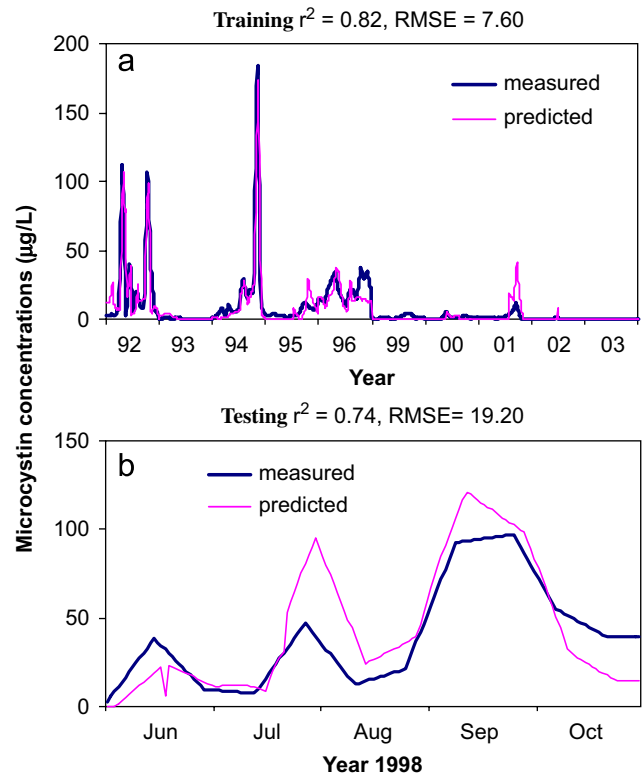
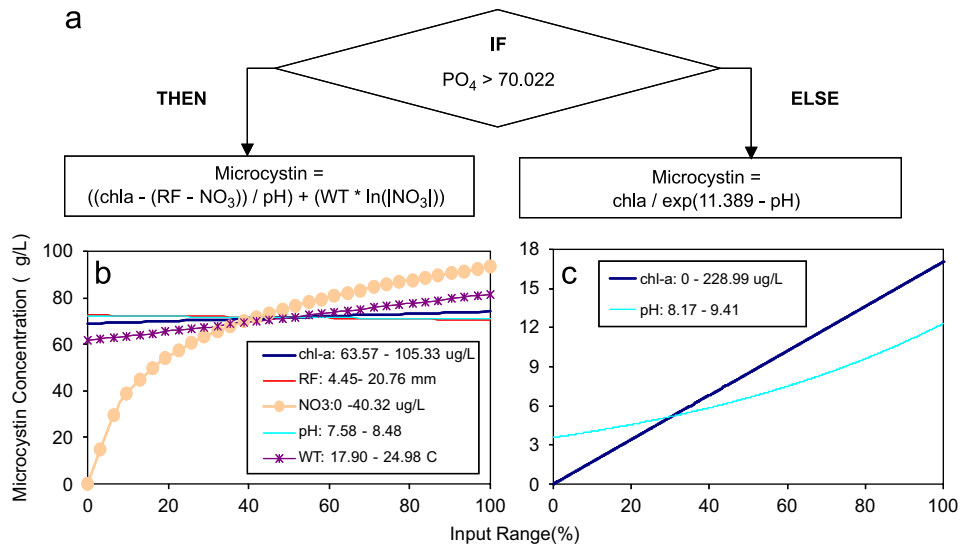


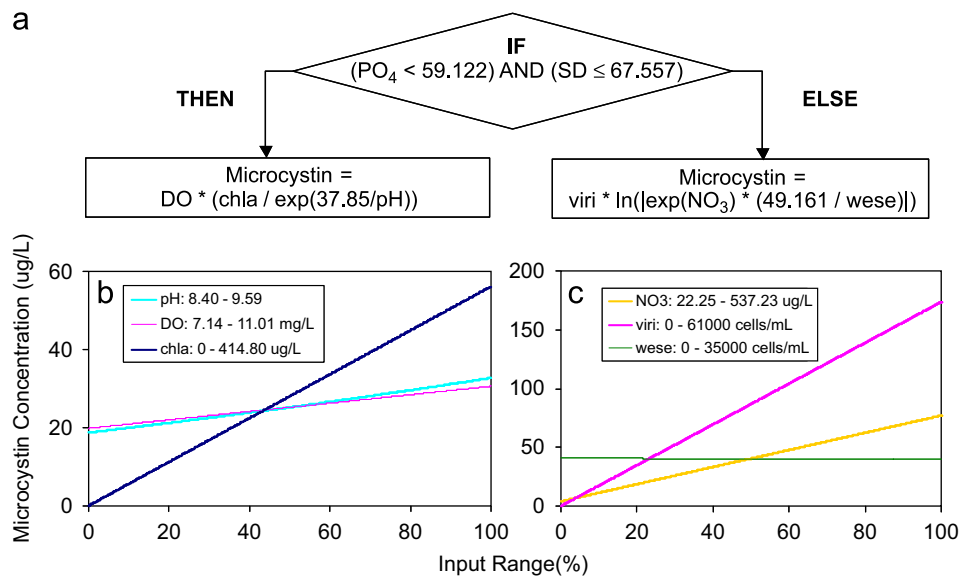
Fig. 7 – Three-days-ahead forecasting of extra-cellular microcystin concentrations in Lake Suwa by physical, chemical and biological inputs: (a) training results and (b) testing results.

primarily determined by both high nitrate concentration of up to 40 µg/L and water temperatures between 18 and 25 °C. Even though chlorophyll *a*, rainfall and pH are although considered as inputs for the *THEN* equation, their sensitivity

curves show little influence on the calculated microcystin concentrations. On the other hand, the *ELSE* equation reflects conditions where the phosphate concentration is below 70.022 µg/L and the resulting microcystin



**Fig. 8 – Rule set and sensitivity analysis for extra-cellular microcystin concentration in Lake Suwa forecasted by physical and chemical inputs: (a) rule set structure, (b) sensitivity analysis of THEN equation, and (c) sensitivity analysis of ELSE equation with variations ( $\pm 1$  STDEV).**



**Fig. 9 – Rule set and sensitivity analysis for extra-cellular microcystin concentration in Lake Suwa forecasted by physical, chemical and biological inputs: (a) rule set structure, (b) sensitivity analysis of THEN equation, and (c) sensitivity analysis of ELSE equation with variations ( $\pm 1$  STDEV).**

ranges between 0 and 18  $\mu\text{g/L}$ . The corresponding sensitivity analysis in Fig. 8(c) indicates that the microcystin concentrations increase with increasing values of chlorophyll *a* and pH.

Fig. 9(a) represents the rule set discovered for forecasting microcystin by means of physical, chemical and biological inputs as documented in Fig. 7. When the IF criteria  $\text{PO}_4 < 59.122$  AND  $\text{SD} \leq 67.557$  are fulfilled, the THEN equation forecasts a lower range of extra-cellular microcystin concentrations. The sensitivity analysis in Fig. 9(b) illustrates that microcystin concentrations of up to 60  $\mu\text{g/L}$  are primarily

determined by high chlorophyll *a* concentrations but also driven by increasing pH and DO values.

The sensitivity curves for the ELSE equation in Fig. 9(c) indicate that high microcystin concentrations of up to 170  $\mu\text{g/L}$  are caused by extensive growth of *M. viridis* peaking at 610,000 cells/mL. It also shows that the microcystin concentrations are sensitive to increasing  $\text{NO}_3$  concentration of up to 70  $\mu\text{g/L}$ . Even though the sensitivity analysis suggests that the abundance of *M. wesenbergii* does not significantly affect changes in microcystin concentrations, it certainly contributes to the microcystin production as reflected by the rule set.

## 4. Discussion

### 4.1. Ordination and clustering

Non-supervised ANN have unravelled complex relationships between *Microcystis* species and microcystin concentrations for different seasons and rainfall intensities. The big differences in the abundance of *Microcystis* and therefore microcystin concentration between typical dry and typical wet years seem to reflect the impact of the Asian monsoon in summer. As reported by Ahn and Jones (2000) extreme cyanobacterial blooms are unlikely to develop in Korean freshwaters during summers with intense monsoonal rain that cause unfavourable conditions for *Microcystis* growth such as strong turbulence and rapid flushing. The microcystin concentration in water may also be diluted by heavy rainfall. However as suggested by the results in Fig. 4 the low microcystin content in wet years may also be largely caused by a significant change in the *Microcystis* species composition in the lake. Whilst high abundances of toxic *Microcystis aeruginosa* and *M. viridis* were typical in late summer and autumn of dry years the non-toxic *M. ichthyoblabe* dominated in wet years.

The component planes in Fig. 5 indicate significant difference in nutrient concentrations between dry and wet years. The substantially lower concentration of phosphate in wet years is likely the result of aerobic sediments that trap oxidised phosphate and prevent internal  $\text{PO}_4$  loads in Lake Suwa that otherwise are typical for shallow eutrophic lakes. In addition the phosphate flushing and dilution effects may outweigh external  $\text{PO}_4$  loadings during intensive monsoonal rains. By contrast the high nitrate content in wet years may be caused by increased external loading as rainwater dissolves atmospheric nitrogen such as  $\text{NO}_3^-$  and  $\text{NH}_4^+$ , and dissolves organic N (Whitall et al., 2003). The component planes in Fig. 5 support such correlation between high  $\text{NO}_3^-$  concentration and high rainfall in August and October of wet years.

On the contrary, the high  $\text{PO}_4$  and low  $\text{NO}_3^-$  concentrations in dry years indicated by Fig. 5 might be explained by excessive internal  $\text{PO}_4$  loads on the one hand and denitrification losses of nitrogen on the other hand from anaerobic sediments. This may coincide with the fact that cyanobacteria uptake nitrate rapidly but phosphate slowly in summer due to the low storage of nitrogen but high storage of phosphorus in *Microcystis* cells during spring (Ahn et al., 2002). Hence, *Microcystis* cells primarily utilized internal phosphorus accumulated in spring before taking in external phosphorus from water.

### 4.2. Forecasting

Hybrid EA assembled successfully two rule-based models from limnological time-series data. They proved to be valid for 3-days-ahead forecasting of microcystin concentrations in Lake Suwa as indicated by high  $r^2$  values as well as correct timing and approximate magnitudes of microcystin concentrations in Figs. 6 and 7. The difference between the two models is the nature of input variables. The physical and

chemical inputs of model (1) in Fig. 6 can be measured in situ electronically and manually. Thus, model (1) is more suitable for operational forecasting and early warning. Model (2) includes additional biological data such as *Microcystis* species abundances, which require time-consuming species identification and cell counting. Hence the use of biological input data makes the model (2) impractical for short-term forecasting but provides better insights into the causality of *Microcystis* blooms.

### 4.3. Explanation by rule sets and sensitivity analysis

The predictive rule sets for microcystin discovered by hybrid EA in time-series data of Lake Suwa reveal the importance of certain inputs and their relationships with particular environmental conditions by means of sensitivity analyses.

In both models, chlorophyll *a* concentration appeared to be an important driving variable for the prediction of microcystin. As *Microcystis* species dominate the phytoplankton community of Lake Suwa they are the major component of its chlorophyll *a* during summer. Thus, microcystin released from toxic *Microcystis* species is directly correlated with the chlorophyll *a* concentration.

Another variable considered in both rule sets is pH. Many studies have reported high competitiveness of cyanobacteria in warm eutrophic waters with high pH or low dissolved  $\text{CO}_2$  conditions (Reynolds, 1984; Shapiro, 1990). Hence, microcystin levels are likely to increase with increasing pH, which implies growing populations of *Microcystis* consisting of both toxic and non-toxic species.

The most interesting result discovered from the second rule set is that *M. viridis* and *M. wesenbergii* were selected by the hybrid EA for the prediction of high microcystin events of up to  $170 \mu\text{g/L}$  (Fig. 9(c)). The sensitivity curves clearly indicate that a blooming *M. viridis* population causes high microcystin concentrations. On the other hand, *M. wesenbergii* has little effect on microcystin concentration when compared to *M. viridis*. These are explained by the low toxicity of *M. wesenbergii*, but the high toxicity of *M. viridis* in Lake Suwa. Both the structures of the rule sets and the sensitivity analyses correspond well with findings from field and laboratory experiments.

## 5. Conclusions

Non-supervised ANN successfully revealed the seasonal succession of *Microcystis* species and microcystin dynamics in relation to different rainfall and environmental conditions of Lake Suwa. Hybrid EA assembled rule sets that proved to be valid for both prediction and explanation of microcystin dynamics. Outcomes of this research can facilitate early warning of high microcystin concentrations in Lake Suwa.

## REFERENCES

- Ahn, K.-G., Jones, J.R., 2000. Factors regulating bluegreen dominance in a reservoir directly influenced by the Asian monsoon. *Hydrobiologia* 432, 37–48.

- Ahn, C.-Y., Chung, A.-S., Oh, H.-M., 2002. Rainfall, phycocyanin, and N:P ratios related to cyanobacterial blooms in a Korean large reservoir. *Hydrobiologia* 474, 117–124.
- Cao, H., Recknagel, F., Kim, B., Takamura, N., 2006. Hybrid evolutionary algorithm for rule set discovery in time-series data to forecast and explain algal population dynamics in two lakes different in morphometry and eutrophication. In: Recknagel, F. (Ed.), *Ecological Informatics*, second ed. Springer, New York, pp. 347–367.
- Harada, K., Mayumi, T., Shimada, T., Fujii, K., Kondo, F., Park, H.-D., Watanabe, M.F., 2001. Co-production of microcystins and aeruginopeptins by natural cyanobacterial bloom. *Environ. Toxicol.* 16, 298–305.
- Holland, J.H., 1975. *Adaptation in Natural and Artificial Systems*. Addison-Wesley, New York.
- Hong, Y.-S.T., Rosen, M.R., Bhamidimarri, R., 2003. Analysis of a municipal wastewater treatment plant using a neural network-based pattern analysis. *Water Res.* 37, 1608–1618.
- Jeong, K.-S., Recknagel, F., Joo, G.-J., 2006. Prediction and elucidation of population dynamics of the blue-green algae *Microcystis aeruginosa* and the diatom *Stephanodiscus hantzschii* in the Nakdong river-reservoir system (South Korea) by a recurrent artificial neural network. In: Recknagel, F. (Ed.), *Ecological Informatics*, second ed. Springer, New York, pp. 255–273.
- Kohonen, T., 1982. *Self-Organization and Associative Memory*. Springer, Berlin.
- Koza, J.R., 1992. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, Cambridge.
- Mitchell, M., 1996. *An introduction to genetic algorithms*. MIT Press, Cambridge, MA.
- Park, H.-D., Watanabe, M.F., Harada, K.-I., Suzuki, M., Hayashi, H., Okino, T., 1993. Seasonal variations of *Microcystis* species and toxic heptapeptide microcystins in Lake Suwa. *Environ. Toxicol. Water Qual.* 8, 425–435.
- Park, H.-D., Iwami, C., Watanabe, M.F., Harada, K., Okino, T., Hayashi, H., 1998. Temporal variabilities of the concentrations of intra- and extra-cellular microcystin and toxic *Microcystis* species in a hypertrophic lake, Lake Suwa, Japan (1991–1994). *Environ. Toxicol. Water Qual.* 13, 61–72.
- Park, Y.-S., Verdonschot, P.F.M., Chon, T.-S., Lek, S., 2003. Patterning and predicting aquatic macroinvertebrate diversities using artificial neural network. *Water Res.* 37 (8), 1749–1758.
- Recknagel, F., French, M., Harkonen, P., Yabunaka, K., 1997. Artificial neural network approach for modelling and prediction of algal blooms. *Ecol. Modelling* 96 (1–3), 11–28.
- Recknagel, F., Bobbin, J., Whigham, P., Wilson, H., 2002. Comparative application of artificial neural networks and genetic algorithms for multivariate time-series modelling of algal blooms in freshwater lakes. *J. Hydroinformatics* 4 (2), 125–133.
- Recknagel, F., Cao, H., Kim, B., Takamura, N., Welk, A., 2006. Unravelling and forecasting algal population dynamics in two lakes different in morphometry and eutrophication by neural and evolutionary computation. *Ecol. Informatics* 1 (2), 133–151.
- Reynolds, C.S., 1984. *The Ecology of Freshwater Phytoplankton*. Cambridge University, Cambridge, MA.
- Shapiro, J., 1990. Current beliefs regarding dominance by blue-greens: the case for the importance of CO<sub>2</sub> and pH. *Verh. Int. Ver. Theor. Ang. Limnol.* 24, 38–54.
- Vesanto, J., Himberg, J., Alhoniemi, E., Parhankangas, J., 2000. SOM Toolbox for Matlab 5. Helsinki University of Technology, Finland.
- Watanabe, M.F., Oishi, S., 1985. Effects of environmental factors on toxicity of a cyanobacterium (*Microcystis aeruginosa*) under culture conditions. *Appl. Environ. Microbiol.* 49, 1342–1344.
- Whitall, D., Hendrickson, B., Paerl, H., 2003. Importance of atmospherically deposited nitrogen to the annual nitrogen budget of the Neuse River estuary, North Carolina. *Environ. Int.* 29, 393–399.
- Yokoyama, A., Park, H.-D., 2002. Mechanism and prediction for contamination of freshwater bivalves (unionidae) with the cyanobacterial toxin microcystin in hypereutrophic Lake Suwa, Japan. *Environ. Toxicol.* 17, 424–433.