Inductive reasoning and prediction of population dynamics of *Cylindrospermopsis* in the Wivenhoe Reservoir by means of evolutionary computation

Friedrich Recknagel¹, Philip Orr² and Hongqing Cao¹

¹ School of Earth and Environmental Sciences, University of Adelaide
² South East Queensland Water, Brisbane
Background

The presentation is based on outcomes of the ARC Linkage Project LP0990453 ‘Early warning of cyanobacteria blooms in drinking water reservoirs by means of evolutionary algorithms’ in collaboration with South East Queensland Water and the SA Water Corporation.

The research is focusing on predictive modelling of population dynamics of *Anabaena*, *Microcystis* and *Cylindrospermopsis* in the Myponga Reservoir, South Australia and the Wivenhoe Reservoir, Queensland.

We show only results for *Cylindrospermopsis raciporskii* in the Wivenhoe Reservoir.
Key hypotheses of the project are:

- early warning of cyanobacteria blooms requires operational models driven by *in situ* data from online monitoring
- operational in situ models can be developed by nonlinear inductive evolutionary computation
- nonlinear inductive models reveal predictor variables as well as relationships and thresholds triggering cyanobacteria blooms
Nonlinear inductive modeling by evolutionary computation

Hybrid Evolutionary Algorithm

Set maximal number of generations $t_{\text{max}}$

Initial number of generations $t = 0$

Initial population of models $P(t) = \{\text{model}_1, \text{model}_2, \ldots, \text{model}_n\}$

$t < t_{\text{max}}$

Evaluate fitness of $n$ models in $P(t)$

Recombine $n$ models in $P(t)$ by Vector- and Tree-level Crossover

Recombine $n$ models in $P(t)$ by Tree-level Mutation

Recombine $n$ models in $P(t)$ by Reproduction

Multi-objective parameter optimisation of $n$ models in $P(t)$

Tournament selection of next population of models $P(t+1)$ from $P(t)$

Number of generations $t = t+1$

Fittest model of $t_{\text{max}}$ generations
Nonlinear inductive modeling by evolutionary computation

Boot-Strap Training Mode of the Hybrid Evolutionary Algorithm HEA

Set maximal number of boot-strap runs $r_{\text{max}}$

Initial number of boot-strap runs $r = 0$

If $r < r_{\text{max}}$ then yes, else no

Fittest model after $r_{\text{max}}$ boot-strap runs of HEA

Randomly-selected 25% Validation Data

Randomly-selected 75% Training Data

Hybrid Evolutionary Algorithm HEA

Number of boot-strap runs $r = r + 1$
Wivenhoe Reservoir, Queensland
max vol: 1165 Gl, max depth: 79 m;
thermally stratified
Dam Wall Station:
historical Data 01/1999 – 12/2010
Historical Data

(1/99 – 12/2010)

Cylindrospermopsis cells/ml
(including Chl_a)
7-DAYS-AHEAD

Cylindrospermopsis cells/ml
(excluding Chl_a)
7-DAYS-AHEAD

Measured
Predicted

$\textit{r}^2 = 0.58$

$\textit{r}^2 = 0.57$

Measured
Predicted

$\textit{r}^2 = 0.57$

$\textit{r}^2 = 0.57$
<table>
<thead>
<tr>
<th>Historical Data</th>
<th>Cylindrospermopsis cells/ml (including Chl_a)</th>
<th>Cylindrospermopsis cells/ml (excluding Chl_a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Inputs (1/99 – 12/2010)</td>
<td>IF WT&gt;26.3 THEN Cylindrospermopsis=exp(pH+pH<em>81.1/(Cond+pH)) ELSE Cylindrospermopsis=(Cond+5.6)</em>(Chla-TP)</td>
<td>IF Turb&lt;=57.4 AND Turb&gt;=23.44 THEN Cylindrospermopsis=exp(WT/2.979)+WT ELSE Cylindrospermopsis=exp(WT/(11.4-pH))<em>(exp(Turb/49.5)-TN</em>silica+silica)</td>
</tr>
<tr>
<td></td>
<td>$r^2$ value=0.58</td>
<td>$r^2$ value=0.57</td>
</tr>
<tr>
<td>Elec. Meas. Inputs (1/99 – 12/2010)</td>
<td>IF WT&lt;=25.5 THEN Cylindrospermopsis=(Cond*Turb/192.2-Turb)<em>Chla</em>82.5 ELSE Cylindrospermopsis=exp(pH)<em>exp(pH)/(Cond</em>Cond)*147.9</td>
<td>IF WT&gt;25.6 AND WT&lt;32.7 THEN Cylindrospermopsis=exp(pH)*pH/(Cond-198.8)*100.2 ELSE Cylindrospermopsis=(ln(</td>
</tr>
<tr>
<td></td>
<td>$r^2$ value=0.57</td>
<td>$r^2$ value=0.57</td>
</tr>
</tbody>
</table>
### Historical Data

#### All Inputs

<table>
<thead>
<tr>
<th>Cylindrospermopsis cells/ml (including Chl_a)</th>
<th>Cylindrospermopsis cells/ml (excluding Chl_a)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>THRESHOLDS</strong></td>
<td><strong>THRESHOLDS</strong></td>
</tr>
<tr>
<td>- WT&gt;26.3</td>
<td>- 23.3&lt;=TURB&lt;=57.4</td>
</tr>
<tr>
<td>- WT&lt;=26.3</td>
<td>- 23.3&gt;TURB&gt;57.4</td>
</tr>
</tbody>
</table>

#### Historical Data

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WT&lt;=25.5</strong></td>
<td>- 25.6&lt;=WT&lt;32.7</td>
</tr>
<tr>
<td><strong>WT&gt;25.5</strong></td>
<td>- 25.6&lt;=WT&gt;32.7</td>
</tr>
</tbody>
</table>

#### Example Graphs

- For Cylindrospermopsis cells/ml, data is shown with thresholds for WT and TURB.
- For Chl_a, data is shown with thresholds for WT.

#### Analysis

- The data suggests a significant variation in Cylindrospermopsis cell counts across different WT and TURB thresholds.
- Chl_a values also show variability, particularly in the WT range.

---

*Note: The image contains graphs that are not fully visible due to the limitation in text representation.*
Cylindrospermopsis cells/ml (including Chl_α)  

THEN-Branch

ELSE-Branch

Cylindrospermopsis cells/ml (excluding Chl_α)  

THEN-Branch

ELSE-Branch

Historical Data  

All Inputs (1/99 –12/2010)

Historical Data  

Wivenhoe Reservoir, Queensland
max vol: 1165 GL, max depth: 79 m;
thermally stratified
Dam Wall Station:
Online Data 09/2007 – 12/2010
<table>
<thead>
<tr>
<th>Historical Data</th>
<th>Cylindrospermopsis cells/ml (including Chl_a) 7-DAYS-AHEAD</th>
<th>Cylindrospermopsis cells/ml (excluding Chl_a) 7-DAYS-AHEAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Inputs (1/99 –12/2010)</td>
<td><img src="image" alt="Measured vs Predicted" /> $r^2 = 0.58$</td>
<td><img src="image" alt="Measured vs Predicted" /> $r^2 = 0.57$</td>
</tr>
<tr>
<td>Electron. Meas. Inputs (1/99 –12/2010)</td>
<td><img src="image" alt="Measured vs Predicted" /> $r^2 = 0.57$</td>
<td><img src="image" alt="Measured vs Predicted" /> $r^2 = 0.57$</td>
</tr>
<tr>
<td>Online Data (9/2007 to 12/2010)</td>
<td><img src="image" alt="Measured vs Predicted" /> $r^2 = 0.65$</td>
<td><img src="image" alt="Measured vs Predicted" /> $r^2 = 0.68$</td>
</tr>
<tr>
<td>Historical Data</td>
<td>Cylindrospermopsis cells/ml (including Chl_a)</td>
<td>Cylindrospermopsis cells/ml (excluding Chl_a)</td>
</tr>
<tr>
<td>-----------------</td>
<td>---------------------------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>All Inputs</td>
<td>IF WT&gt;26.3 THEN Cylindrospermopsis=exp(pH+pH*81.1/(Cond+pH)) ELSE Cylindrospermopsis=(Cond+5.6)*(Chla-TP)</td>
<td>IF Turb&lt;=57.4 AND Turb&gt;=23.44 THEN Cylindrospermopsis=exp(WT/2.979)+WT ELSE Cylindrospermopsis=exp(WT/(11.4-pH))<em>(exp(Turb/49.5)-TN</em>silica+silica)</td>
</tr>
<tr>
<td></td>
<td>(r^2) value=0.58</td>
<td>(r^2) value=0.57</td>
</tr>
<tr>
<td>Elect. Meas. Inputs</td>
<td>IF WT&lt;=25.5 THEN Cylindrospermopsis=(Cond*Turb/192.2-Turb)<em>Chla</em>82.5 ELSE Cylindrospermopsis=exp(pH)<em>exp(pH)/(Cond</em>Cond)*147.9</td>
<td>IF WT&gt;25.6 AND WT&lt;32.7 THEN Cylindrospermopsis=exp(pH)*pH/(Cond-198.8)*100.2 ELSE Cylindrospermopsis=(ln(</td>
</tr>
<tr>
<td></td>
<td>(r^2) value=0.57</td>
<td>(r^2) value=0.57</td>
</tr>
</tbody>
</table>

### Historical Data

**All Inputs (1/99 –12/2010)**

- IF WT>26.3 THEN Cylindrospermopsis=exp(pH+pH*81.1/(Cond+pH)) ELSE Cylindrospermopsis=(Cond+5.6)*(Chla-TP)
- **\(r^2\) value=0.58**


- IF WT<=25.5 THEN Cylindrospermopsis=(Cond*Turb/192.2-Turb)*Chla*82.5 ELSE Cylindrospermopsis=exp(pH)*exp(pH)/(Cond*Cond)*147.9
- **\(r^2\) value=0.57**

### Online Data

**All Inputs (9/2007 to 12/2010)**

- IF WT\_online<25.7 THEN Cylindrospermopsis=(WT\_online-25.2)* (pH\_online+Chla\_online)*98.5+3868.6 ELSE Cylindrospermiosis=exp(pH\_online)+562.1)* (ln(|Chla\_online-14|)-(Conductivity\_online-pH\_online)* 44.5-68.3)*115.3
- **\(r^2\) value=0.65**

**Conductivity\_online\_<=296.5**

- THEN Cylindrospermosis=((Conductivity\_online+ Turbidity\_online)+Turbidity\_online)*(WT\_online-16.1)) ELSE Cylindrospermiosis=(((WT\_online-(-10.9))*82.5)- (Conductivity\_online+(Conductivity\_online+135.1))^(WT\_online+ (exp(pH\_online)/(Conductivity\_online+pH\_online))))
- **\(r^2 = 0.68\)**
### Cylindrospermopsis cells/ml (including Chl-a) THRESHOLDS

<table>
<thead>
<tr>
<th>Historical Data</th>
<th>Cylindrospermopsis cells/ml (excluding Chl-a) THRESHOLDS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Elec. Meas. Inputs</strong></td>
<td><strong>Historical Data All Inputs</strong></td>
</tr>
<tr>
<td>100000</td>
<td>80000</td>
</tr>
<tr>
<td>20000</td>
<td>70000</td>
</tr>
<tr>
<td>40000</td>
<td>60000</td>
</tr>
<tr>
<td>80000</td>
<td>50000</td>
</tr>
<tr>
<td>30000</td>
<td>40000</td>
</tr>
<tr>
<td>60000</td>
<td>30000</td>
</tr>
<tr>
<td>70000</td>
<td>20000</td>
</tr>
<tr>
<td>80000</td>
<td>10000</td>
</tr>
</tbody>
</table>

- WT>26.3
- WT<=26.3

### Cylindrospermopsis cells/ml (excluding Chl-a) THRESHOLDS

<table>
<thead>
<tr>
<th><strong>Historical Data Elec. Meas. Inputs</strong></th>
<th><strong>Online Data</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>100000</td>
<td>60000</td>
</tr>
<tr>
<td>20000</td>
<td>50000</td>
</tr>
<tr>
<td>40000</td>
<td>40000</td>
</tr>
<tr>
<td>80000</td>
<td>30000</td>
</tr>
<tr>
<td>30000</td>
<td>20000</td>
</tr>
<tr>
<td>60000</td>
<td>10000</td>
</tr>
</tbody>
</table>

- WT<25.5
- WT>25.5

- WT<25.7
- WT>=25.7

<table>
<thead>
<tr>
<th><strong>Online Data</strong></th>
<th><strong>Online Data</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>9/2007 to 12/2010</strong></td>
<td><strong>EC &gt; 296.5</strong></td>
</tr>
<tr>
<td>10000</td>
<td>50000</td>
</tr>
<tr>
<td>20000</td>
<td>40000</td>
</tr>
<tr>
<td>30000</td>
<td>30000</td>
</tr>
<tr>
<td>40000</td>
<td>20000</td>
</tr>
<tr>
<td>50000</td>
<td>10000</td>
</tr>
</tbody>
</table>

- EC > 296.5
- EC <= 296.5
### Cylindrospermopsis cells/ml

**THEN-Branch**
- (including Chl_a)

**ELSE-Branch**
- (excluding Chl_a)

#### Historical Data

<table>
<thead>
<tr>
<th>All Inputs (1/99 – 12/2010)</th>
<th>THEN-Branch</th>
<th>ELSE-Branch</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Historical Data</strong></td>
<td><strong>Cond:</strong> 220.7~477.3</td>
<td><strong>WT:</strong> 14.2~27.9</td>
</tr>
<tr>
<td><strong>pH:</strong> 7.7~9</td>
<td><strong>Conduc.vity_online:</strong> 241~498</td>
<td><strong>pH:</strong> 7.4~8.7</td>
</tr>
<tr>
<td><strong>Chla:</strong> 0.1~15.6</td>
<td><strong>pH_online:</strong> 7.3~8.9</td>
<td><strong>Turb:</strong> 0~39.4</td>
</tr>
<tr>
<td><strong>Turb:</strong> 0~38.7</td>
<td><strong>Conduc.vity_online:</strong> 240~311</td>
<td><strong>WT:</strong> 16.6~39.4</td>
</tr>
<tr>
<td></td>
<td><strong>WT:</strong> 16.6~30.2</td>
<td><strong>silica:</strong> 0~8.5</td>
</tr>
<tr>
<td></td>
<td><strong>WT:</strong> 15.2~27</td>
<td><strong>TN:</strong> 0.3~0.7</td>
</tr>
</tbody>
</table>

#### Electric. Meas. Inputs

<table>
<thead>
<tr>
<th>(1/99 – 12/2010)</th>
<th>THEN-Branch</th>
<th>ELSE-Branch</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Historical Data</strong></td>
<td><strong>Cond:</strong> 195.2<del>501.1 &amp; Chla: 0.1</del>15.6</td>
<td><strong>pH:</strong> 7.4~8.7</td>
</tr>
<tr>
<td><strong>Turb:</strong> 0~38.7</td>
<td><strong>pH:</strong> 7.8~8.8</td>
<td><strong>Turb:</strong> 0~39.4</td>
</tr>
<tr>
<td><strong>pH:</strong> 7.8~8.8</td>
<td><strong>pH:</strong> 7.7~8.8</td>
<td><strong>WT:</strong> 16.6~30.2</td>
</tr>
</tbody>
</table>

#### Online Data

<table>
<thead>
<tr>
<th>(9/2007 to 12/2010)</th>
<th>THEN-Branch</th>
<th>ELSE-Branch</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Online Data</strong></td>
<td><strong>WT:</strong> 15.1~26.4</td>
<td><strong>WT:</strong> 16.1~30.7</td>
</tr>
<tr>
<td><strong>pH:</strong> 7~9</td>
<td><strong>Conduc.vity_online:</strong> 240~311</td>
<td><strong>pH:</strong> 7.3~9.1</td>
</tr>
<tr>
<td><strong>Chla:</strong> 0.8~11.7</td>
<td><strong>pH_online:</strong> 7~9</td>
<td></td>
</tr>
</tbody>
</table>
Towards *in situ* operational models for early warning of cyanobacteria blooms
Towards *in situ* operational models for early warning of cyanobacteria blooms

**Real-time *in situ* Water Quality Measurements by Hydrolab DataSonde 5X**
- Water Temperature °C
- Diss. Oxygen mg/l
- Turbidity NTU
- pH
- Ammonium NH₄ mg/l
- Conductivity mS/cm
- Total Chlorophyll µg/l
- Phycocyanin µg/l

**Data Archiving in the Ecological Data Warehouse**
- On-line Data
- Historical Data

**Data Acquisition by Hydrolab Process Monitor**
- On-line Data

**Data Merger and Validation**
- Time Series

**Data Preprocessing Module**

**Real-Time Forecasting**
- Model

**Forecasting Module**
- Early Warning for Operational Raw Water Control if Algal Bloom is Imminent
- Chlorophyll-a µg/l
- *Anabaena circinalis* cells/ml
- *Cylindrospermopsis* cells/ml
- *Microcystis* cells/ml
Towards operational models for early warning of cyanobacteria blooms

What's next?
Spatially-explicit forecasting of cyanobacteria assemblages in drinking water reservoirs by multi-objective evolutionary computation
What's next?
Spatially-explicit forecasting of cyanobacteria assemblages in drinking water reservoirs by multi-objective evolutionary computation

GENETIC PROGRAMMING

**OPTIMISATION OF RULE STRUCTURES BY REPRODUCTION, MUTATION AND CROSSOVER**

IF \( g(X_1, ..., X_p, P_m, ..., P_n) \) THEN
- \( Y_1 = f_1(x_1, ..., X_p, P_m, ..., P_n) \)
- \( Y_2 = f_2(x_1, ..., X_p, P_m, ..., P_n) \)
- \( \vdots \)
- \( Y_i = f_i(x_1, ..., X_p, P_m, ..., P_n) \)

ELSE
- \( Y_1 = f_1(x_1, ..., X_p, P_{m}, ..., P_{n}) \)
- \( Y_2 = f_2(x_1, ..., X_p, P_{m}, ..., P_{n}) \)
- \( \vdots \)
- \( Y_i = f_i(x_1, ..., X_p, P_{m}, ..., P_{n}) \)

**DIFFERENTIAL EVOLUTION**

**OPTIMISATION OF RULE PARAMETERS BY CROSSOVER**

IF \( g(x_1, ..., X_p, P_{m\text{opt}}, ..., P_{n\text{opt}}) \) THEN
- \( Y_1 = f_1(x_1, ..., X_p, P_{m\text{opt}}, ..., P_{n\text{opt}}) \)
- \( Y_2 = f_2(x_1, ..., X_p, P_{m\text{opt}}, ..., P_{n\text{opt}}) \)
- \( \vdots \)
- \( Y_i = f_i(x_1, ..., X_p, P_{m\text{opt}}, ..., P_{n\text{opt}}) \)

ELSE
- \( Y_1 = f_1(x_1, ..., X_p, P_1, ..., P_n) \)
- \( Y_2 = f_2(x_1, ..., X_p, P_1, ..., P_n) \)
- \( \vdots \)
- \( Y_i = f_i(x_1, ..., X_p, P_1, ..., P_n) \)

**MULTI-OBJECTIVE OPTIMISATION OF FITNESS FUNCTION**

\[ \sum_{i=1}^{k} (Y_i \text{ (obs)} - Y_i) \rightarrow \text{MIN} \]

"Fittest" spatio-temporal forecasting model for Microcystis
Conclusions:

- Nonlinear inductive models by evolutionary computation inform about ecological relationships and thresholds determining cyanobacteria population dynamics.

- Nonlinear inductive models by evolutionary computation provide short-term forecasting of cyanobacteria population dynamics.

- Nonlinear inductive models by evolutionary computation suit as *in situ* operational models for early warning of outbreaks of cyanobacteria blooms.

- Multi-objective evolutionary computation allows to develop models for spatially-explicit forecasting of cyanobacteria assemblages across sampling stations in lakes.
Acknowledgements:

We thank the ARC, SEQWater and SA Water for providing funding, and thank SEQWater and SA Water for their constructive support.
Early Warning of Cyanobacteria Blooms

Cyanobacteria Forecast on 23/12/2008

- **Cylindrospermopsis at Wivenhoe-3001 (70.3%)**
  - Mode 1: Safe
  - Mode 2: Normal
  - Mode 3: Warning

- **Cylindrospermopsis at North Pine Dam-10001 (39.5%)**
  - Mode 1: Safe
  - Mode 2: Normal
  - Mode 3: Alarm

- **Cylindrospermopsis at Somerset-2001 (10.3%)**
  - Mode 1: Safe
  - Mode 2: Normal
  - Mode 3: Warning
Early Warning of Cyanobacteria Blooms

Cyanobacteria Forecast on 23/02/2009

- **Anabaena at Wivenhoe-30001 (0.8%)**
  - Normal
  - Warning
  - Alarm

- **Cylindropermopsis at Wivenhoe-30001 (100%)**
  - Normal
  - Warning
  - Alarm

- **Microcystis at Wivenhoe-30001 (60%)**
  - Normal
  - Warning
  - Alarm
Early Warning of Cyanobacteria Blooms

Cyanobacteria Forecast on 23/02/2009

Anabaena at North Pine Dam: 10001 (0.0%)

Cylindrospermopsis at North Pine Dam: 10001 (0.0%)

Microcystis at North Pine Dam: 10001 (0.0%)