Squaring the Circle: Elucidating the Significance of Attribute State Variation in Artificial Neural Networks

Gary R. Weckman
Ohio University
Palm Island Enviro-Informatics & Business Solutions, LLC

David F. Millie
Palm Island Enviro-Informatics & Business Solutions, LLC
Florida Institute of Human & Machine Cognition, Inc.
What is ‘Squaring the Circle’
– Try to do something very difficult or impossible

Elucidating the Significance of Attribute State Variation
– How to illuminate the various interactions at various states of key attributes

Artificial Neural Networks
– What are they?
The Brain is a set of very dense, complex local networks.
as a cell body, a branching input structure (the dendrite) and a branching output structure (the axon)

Axons connect to dendrites via synapses

Electro-chemical signals are propagated from the dendritic input, through the synapses, and along the axon to the next neuron.
Biological Analogy

in Neuron

Artificial neuron

Set of processing elements (PEs) and connections (weights) with adjustable strengths

\[ f(\text{net}) \]

\[ \sum \]

\[ w_1, w_2, \ldots, w_n \]

of processing elements (PEs) and connections (weights) adjustable strengths
Artificial Neural Networks

ine-learning algorithms that identify data patterns and perform decision making in a manner imitating cognitive functionality

‘Learning’ (analogous to problem solving) is:

✓ adaptive - knowledge is altered, updated, & stored (via weights)
✓ iterative - examples to generalizations

‘Universal approximators’ – can discover & reproduce any (linear / non-linear) trend given enough data & computational (processing) capability

✓ No expert knowledge required
✓ Few (if any) ‘formal’ assumptions - i.e. Gaussian requirements, etc.

Disadvantage - (superficially ? ?) lack a declarative knowledge structure
✓ a ‘Black Box’ (i.e. no global equation)
Early Days: Interested in “Model Accuracy”
Modeling Approach

Start

Database
(Pre-Process Inputs & Output)

Desired Output

For i = 1 to N (# of models investigated)

Initialize ANN
(#Neurons, #Layers, Weights, etc..)

ANN Model

Estimated Output

Terminate?

Yes

Adjust Parameters

Training Algorithm

No

Yes, i = i + 1

Final ANN Model

Select Best Performing ANN

No
KNOWLEDGE EXTRACTION defined:

is the creation of knowledge from structured (relational databases, XML) and unstructured (text, documents, images) sources


Is there a way illuminate the black box?
Multiple Variable Interactions while looking at various states!

Our drive to Mechanistic Model: Grey Box  $\Rightarrow$  WHITE BOX
1st ATTEMPT:

- Included all attributes collected
- Sensitivity about the means
- Found many limitations to current method

How are we to explain a more complex situation?
### Bathymetric Contours

- **5 m intervals**

<table>
<thead>
<tr>
<th></th>
<th>Inner</th>
<th>Outer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depth</td>
<td>14.0 / 5.09</td>
<td>40.5 / 13.66</td>
</tr>
<tr>
<td>Volume</td>
<td>1,554</td>
<td>1,217</td>
</tr>
<tr>
<td>3)</td>
<td>7.91</td>
<td>16.63</td>
</tr>
</tbody>
</table>

- **58 - River mouth to Lake proper**

### Ocean’s & Human Health Initiative (2003-2005)

**Sampling Sites**

- **2003-2005**

### Multiple Stressor Program (2008-2010)

**Sampling Sites**

- **2008-2010**

---

*ALGAL BLOOM ADVISORY*

A harmful algal bloom has been detected at this location. Users are encouraged to avoid ingesting water and avoid surface scum.

---

**Dreissena polymorpha**

(Oactual size is 20 mm)

**Dreissena bugensis**

(Oactual size is 15 mm)

---

**CAUTION**

BEACH MUCK ADVISORY
WHEN BEACH MUCK IS PRESENT
FOR YOUR SAFETY

Avoid the water. Do not enter the water. Wash thoroughly with soap and water if contact with beach muck occurs.

---

**Bathymetric contours**

in 5-m intervals

---

**Saginaw Bay**

**Outer Bay**

**Upper Bay**

**Middle Bay**

**Inner Bay**

**Lower**

**Saginaw Bay-City Channel**

**Saginaw River**
Predicting Saginaw Bay Chl a (1991-1996)
MLP - 1 Hidden Layer of 4 Processing Elements

**Hydrological Predictors:** $^\circ$C, Secchi, $K_d$, Cl, NO$_3$, NH$_4$, SRP, TP, SiO$_2$, PSiO$_2$, DOC, POC

<table>
<thead>
<tr>
<th>Data Set</th>
<th>$r$</th>
<th>$p$-Value</th>
<th>MSE / NMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Training</strong> (n = 586)</td>
<td>0.79</td>
<td>$&lt; 0.0001$</td>
<td>10.69 / 0.39</td>
</tr>
<tr>
<td><strong>Cross Validation</strong> (n = 146)</td>
<td>0.85</td>
<td>$&lt; 0.0001$</td>
<td>4.89 / 0.27</td>
</tr>
<tr>
<td><strong>Test</strong> (n = 244)</td>
<td>0.89</td>
<td>$&lt; 0.0001$</td>
<td>8.08 / 0.25</td>
</tr>
</tbody>
</table>

Measured Chlorophyll a ($\mu$g L$^{-1}$)
Existing Knowledge Extraction Tools

Neural Interpretation Diagram

• Decomposition method to visualize
  – Determine significance of input variables
  – Based on the magnitude of interconnecting weights

Connected Weights

• Decomposition method that uses weights of an ANN to determine:
  – Input Significance to model
  – Nodes Significance to ANN

• Procedure
  – Calculate “connected weights” for all possible paths of the network
positive (excitatory) weights

Garson’s Algorithm
Relative Share of Prediction

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>POC</td>
<td>11%</td>
</tr>
<tr>
<td>DOC</td>
<td>9%</td>
</tr>
<tr>
<td>°C</td>
<td>9%</td>
</tr>
<tr>
<td>Secchi</td>
<td>6%</td>
</tr>
<tr>
<td>CL</td>
<td>6%</td>
</tr>
<tr>
<td>PSiO₂</td>
<td></td>
</tr>
<tr>
<td>Chl a</td>
<td></td>
</tr>
</tbody>
</table>

Analysis (±1 SD)

Target (µg L⁻¹)

$\Delta$ Chl a (µg L⁻¹)

0.0 0.7 1.4 2.1 2.8 3.5
Multi-Variable Sensitivity Analysis (circa 2006!)

**Inner Bay**

- Chlorophyll $a$ $\mu$g L$^{-1}$
  - Black: $< 1.50$
  - Red: $1.50 - 2.50$
  - Green: $2.51 - 4.00$
  - Yellow: $4.01 - 9.00$
  - Blue: $9.01 - 15.00$
  - Purple: $> 15$

**Outer Bay**

- Total Phosphorus (mg L$^{-1}$)
  - Black: $< 1.50$
  - Red: $1.50 - 2.50$
  - Green: $2.51 - 4.00$
  - Yellow: $4.01 - 9.00$
  - Blue: $9.01 - 15.00$
  - Purple: $> 15$

**Secchi Depth (m)**

- **Temperature (°C)**
  - 30
  - 25
  - 20
  - 15
  - 10
  - 5
  - 0
Decision Trees

- Symbolic Knowledge Extraction Technique
- Most commonly used decision tree induction algorithm – (Quinlan)
- Recursive partitioning of the data
- Drawback: Amount of data reaching each node decreases with the depth of the tree
- Alternative: TREPAN
Different Project: Crude Oil Impact

New Set of Tools:

Limitations to Sensitivity:

ANNs were created for “high” and “low” % Crude Oil. Sensitive results were very different

Crude Oil <= 20%  

Crude Oil >= 50%
Training; n = 151
r = 0.95, NMSE = 0.10

Cross-Validation; n = 50
r = 0.88, NMSE = 0.22

Test; n = 49
r = 0.98, NMSE = 0.05

Local Sensitivity

- + 1 SD, Common Variation
- + 2 SD, Disturbance Variation

Range of Sensitivity

SD CHL a

Distinct differences based on range of Sensitivity
CHL as a function of TP & TEMP

Modeled Chlorophyll a (g L⁻¹)

1.98 + (0.03*TP)
    it SE = 0.41, Fstat = 29857.36

2.23 + (0.002*TEMP²)
    it SE = 1.03, Fstat = 6323.88

Half-Maximal Abundance Concentrations / Conditions:

TP₅₀ = 51.8 g L⁻¹
TEMP₅₀ = 22.6 °C
WndSpd₃₋₅₀ = 18.0 km hr⁻¹
Taking Into Account the Interactions and/or Synergisms of Co-Limiting Nutrients
Generalized Equation for 2 variable interaction with output (CHL a)
**Iterations: ANNs Models**

\[
\text{CHL}_\text{Grey-Box} = [\text{CHL } a]_{1\text{st iteration}} + [\text{CHL } a]_{2\text{nd iteration}} + \ldots + [\text{CHL } a]_{n\text{th iteration}} + r_n
\]

**Multiple ANN models utilizing 2 variables at a time to predict Output**

**Iterations: Additive Models**

**Finalized Combined Model**

\[
[\text{CHL } a]_{\text{Grey-Box}} = -13.3 + 2.25
\]

\[
+ \left(0.41 \times 1.16 + 5.0E17 \times (\text{TEMP}^{6.592} \times \text{SA})\right)
\]

\[
+ \left(0.35 \times \left(\frac{0.72 - 0.01 \times \text{pH} + 0.115 \times \text{log}_10 \text{PH} + \text{e}^{\log_10 \text{pH}}}{1 - 1.08 \times \text{log}_10 \text{TEMP}}\right)\right)
\]

\[
+ \left(\frac{0.12 \times (70.89 + 12.69 \times \text{log}_10 \text{CURSPD} - 12)}{1 - 1.343 \times \text{CURSPD} + 0.01}\right)
\]

\[
+ \left(\frac{0.06 \times (8.01 - 1.98 \times \text{log}_10 \text{CURDIR} - 2.69 \times \text{log}_10 \text{CURDIR} + 0.03 \times \text{W})}{1 - 0.01 \times \text{CURDIR} + 0.03 \times \text{W}}\right)
\]
Global Sensitivity

Sensitivity about Means

Local Sensitivity
Does not consider variable interactions as states change

Developed Global Sensitivity
Looks at how variables interact as their states change!
## Culprit: Correlation

|            | Tot.Par | Ave_WndSpd | Ave_WndDir | Ave_BP | Ave_TotPrecip | Ave_AmbC | Ave_RH% | Ave_CurrentSpd | Ave_CurrentDir | Ave_Urea | Ave_WatC | Ave_PSU | Ave_pH | Ave_Turb | Ave_DO |
|------------|---------|------------|-----------|--------|---------------|----------|---------|---------------|---------------|-----------|----------|---------|--------|---------|--------|---------|--------|
| Tot.Par    | 1.00    |            |           |        |               |          |         |               |               |           |          |         |        |         |        |         |
| Ave_WndSpd | -0.30   | 1.00       |           |        |               |          |         |               |               |           |          |         |        |         |        |         |
| Ave_WndDir | 0.33    | -0.19      | 1.00      |        |               |          |         |               |               |           |          |         |        |         |        |         |
| Ave_BP     | 0.18    | -0.21      | -0.23     | 1.00   |               |          |         |               |               |           |          |         |        |         |        |         |
| Ave_TotPrecip | -0.38 | 0.15       | 0.06      | -0.13  | 1.00          |          |         |               |               |           |          |         |        |         |        |         |
| Ave_AmbC   | 0.43    | -0.41      | 0.64      | -0.09  | -0.04         | 1.00     |         |               |               |           |          |         |        |         |        |         |
| Ave_RH%    | -0.37   | 0.12       | 0.12      | -0.29  | 0.28          | 0.11     | 1.00    |               |               |           |          |         |        |         |        |         |
| Ave_CurrentSpd | -0.04 | 0.20       | -0.10     | -0.25  | -0.15         | -0.17    | -0.26   | 1.00          |               |           |          |         |        |         |        |         |
| Ave_CurrentDir | 0.00  | -0.03      | -0.03     | 0.04   | -0.12         | -0.09    | -0.34   | 0.68          | 1.00          |           |          |         |        |         |        |         |
| Ave_Urea   | 0.20    | 0.20       | 0.40      | -0.15  | 0.14          | 0.15     | -0.03   | -0.12         | -0.06         | 1.00      |          |         |        |         |        |         |
| Ave_WatC   | 0.35    | -0.26      | 0.43      | 0.03   | 0.08          | 0.83     | -0.04   | -0.26         | -0.16         | 0.21       | 1.00     |         |        |         |        |         |
| Ave_PSU    | 0.30    | 0.15       | 0.55      | -0.02  | 0.12          | 0.38     | 0.03    | -0.43         | -0.42         | 0.50       | 0.50     | 1.00     |        |         |        |         |
| Ave_pH     | -0.19   | -0.02      | -0.31     | -0.26  | -0.11         | -0.32    | 0.01    | 0.49          | 0.40          | -0.40      | -0.47    | -0.73    | 1.00    |        |         |        |
| Ave_Turb   | -0.18   | 0.03       | -0.12     | -0.09  | -0.05         | -0.26    | -0.17   | 0.48          | 0.23          | -0.05      | -0.28    | -0.34    | 0.34    | 1.00    |        |         |
| Ave_DO     | -0.05   | -0.04      | -0.22     | -0.31  | -0.24         | -0.45    | -0.17   | 0.43          | 0.20          | -0.26      | -0.64    | 0.47     | 0.39    | 0.24    | 1.00    |        |
**Global Sensitivity**

Each Variable has its own distribution of values ($

Impact of Correlation on State Behavior

<table>
<thead>
<tr>
<th>N</th>
<th>Secchi</th>
<th>TSS</th>
<th>TP</th>
<th>TDP</th>
<th>SRP</th>
<th>NH4</th>
<th>NO3</th>
<th>CL</th>
<th>Sol_Si</th>
<th>POC</th>
<th>DOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>5σ</td>
<td>1.57</td>
<td>-0.70</td>
<td>-0.98</td>
<td>-0.48</td>
<td>-0.25</td>
<td>0.02</td>
<td>-0.02</td>
<td>-0.57</td>
<td>-0.16</td>
<td>-1.16</td>
<td>-0.80</td>
</tr>
<tr>
<td>5σ</td>
<td>0.53</td>
<td>-0.67</td>
<td>-0.59</td>
<td>-0.04</td>
<td>-0.14</td>
<td>0.09</td>
<td>0.41</td>
<td>-0.02</td>
<td>-0.40</td>
<td>-0.79</td>
<td>0.04</td>
</tr>
<tr>
<td>5σ</td>
<td>-0.17</td>
<td>-0.08</td>
<td>-0.16</td>
<td>-0.11</td>
<td>-0.09</td>
<td>-0.09</td>
<td>-0.04</td>
<td>-0.14</td>
<td>-0.09</td>
<td>-0.26</td>
<td>-0.14</td>
</tr>
<tr>
<td>5σ</td>
<td>-0.40</td>
<td>-0.02</td>
<td>0.14</td>
<td>0.13</td>
<td>0.04</td>
<td>-0.26</td>
<td>-0.24</td>
<td>-0.16</td>
<td>0.39</td>
<td>0.35</td>
<td>-0.06</td>
</tr>
<tr>
<td>5σ</td>
<td>-0.68</td>
<td>0.50</td>
<td>0.31</td>
<td>-0.37</td>
<td>-0.06</td>
<td>-0.49</td>
<td>-0.35</td>
<td>-0.06</td>
<td>0.20</td>
<td>0.87</td>
<td>0.15</td>
</tr>
</tbody>
</table>
Global Variation Across States

**Network Output(s) for Varied Input TP**

\[ y = 0.1604x^{1.35} \]
\[ R^2 = 0.9766 \]

**Network Output(s) for Varied Input Secchi**

\[ y = -7.938 \ln(x) + 12.678 \]
\[ R^2 = 0.927 \]

Significant difference
Global versus Local Sensitivity
Global (State Based) versus Local (Means) Sensitivity

Sensitivity: State Based versus Means

Input Name

- WndSpd
- Air_TMP
- NH4
- Air_TMP
- WndSpd_3D
- WndDir_3D
- WndDir
- Gust
- TDP
- NO3
- CL
- GUST_3D
- GUST_3D
- Srf_Tmp
- Sol_Si
- SRP
- DOC
- Bay
- Secchi
- PON
- TSS
- POC
- TP

Chl_a 1 std dev
SB Chl_a 1 std dev
New Grey Box

ANN Model Output

$\text{Iteration}^t$
impact 1st Iteration
Repeat Iterations
Equation Model: summed after each Iteration for remaining attributes
Grey Box versus ANN

Grey Box: Deviations still high versus ANN

Possible Improvements – more detail breakdown in deviations
State Based Sensitivity

Machine-learning algorithms capable of autonomously unearthing and reproducing complex patterns within sizeable data quantities afford great potential for fueling ecological hypothesis creation and ‘intelligent’ knowledge derivation.