# *aring the Circle': Elucidating the Significance of Attri 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

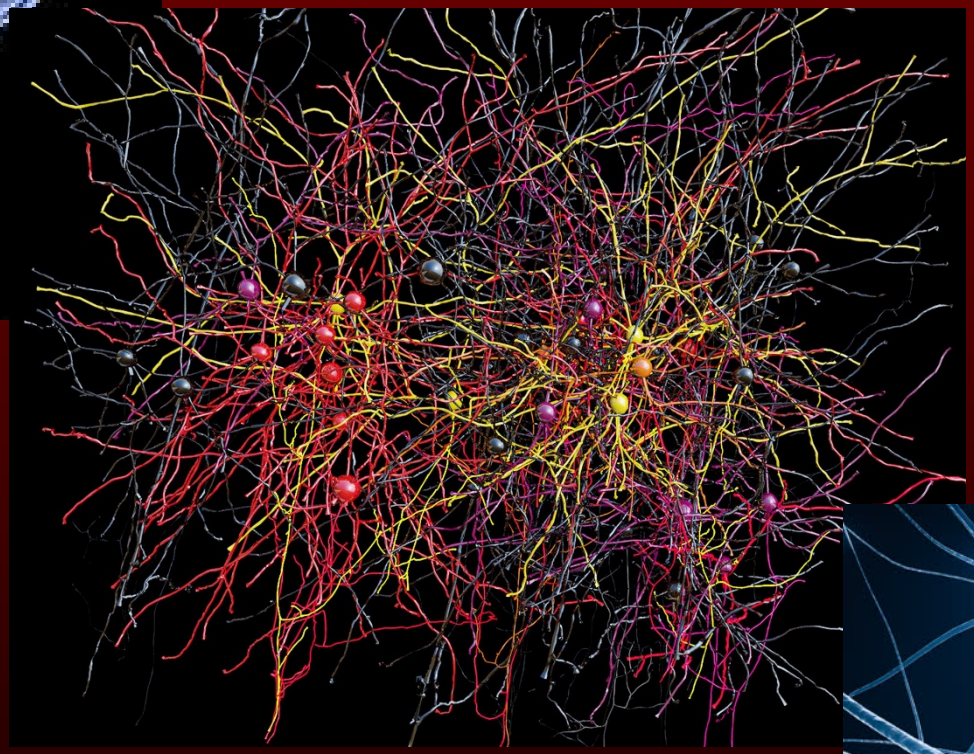
- How to illuminate the various interactions at various states of key attributes

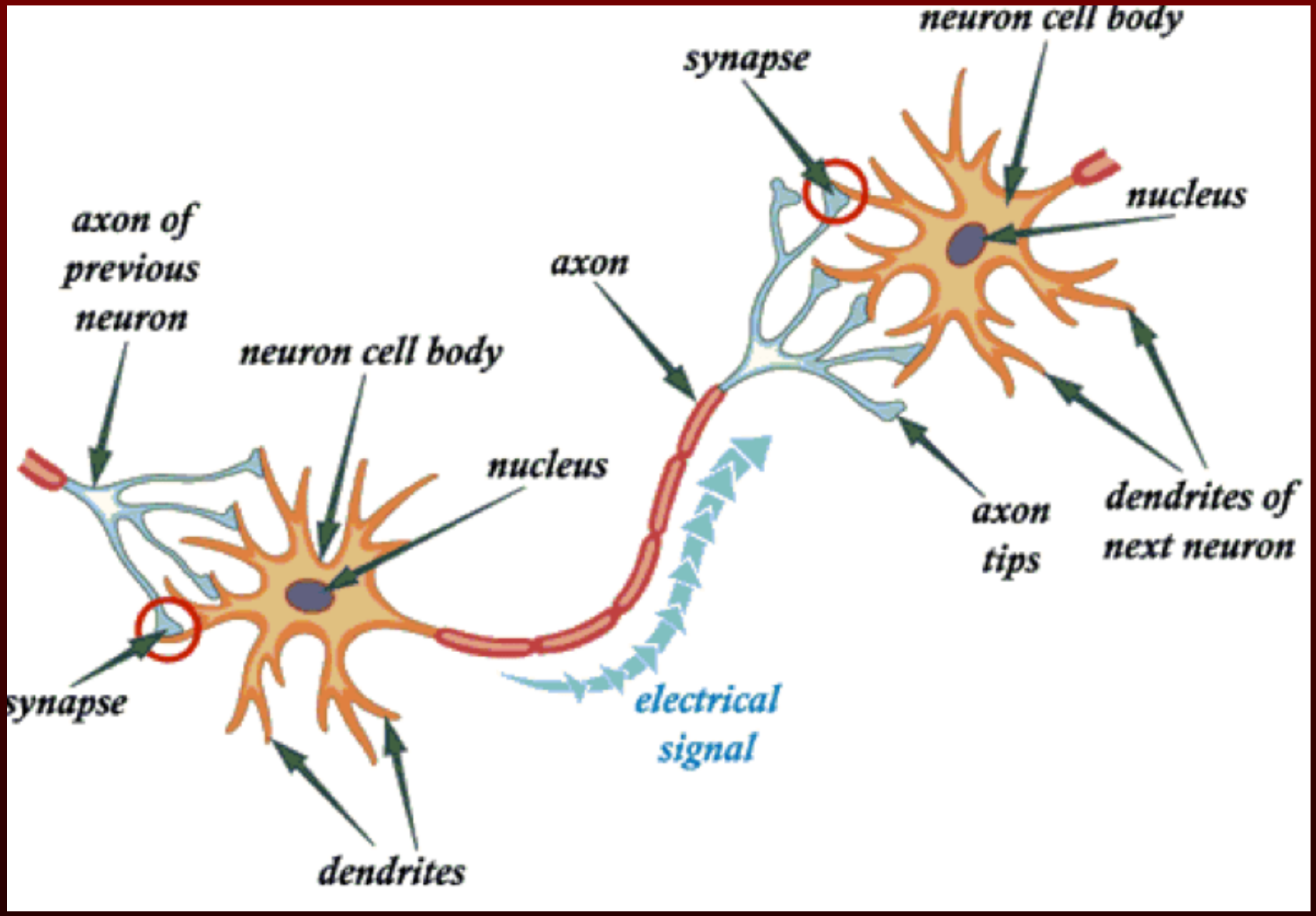
## Artificial Neural Networks

- What are they?

# The Brain

is a set of very dense, complex local networks





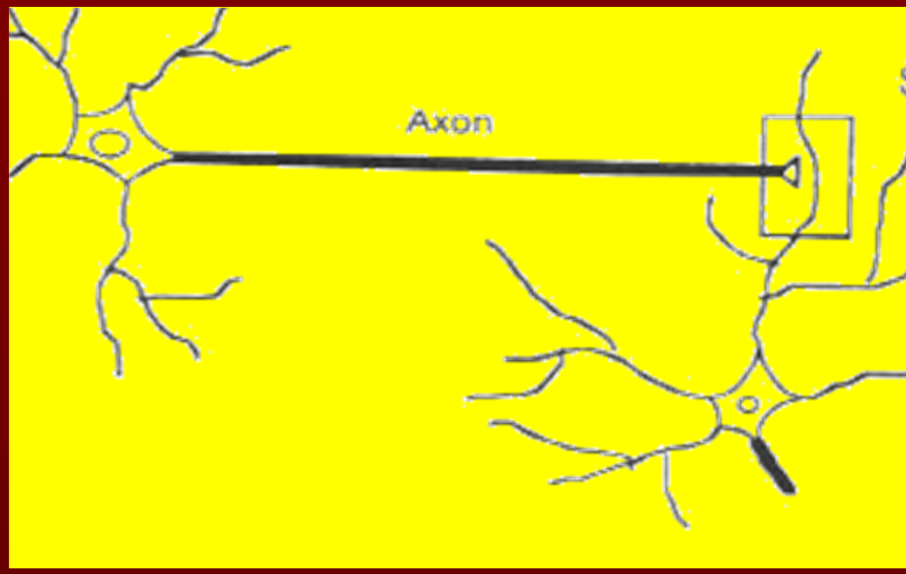

as a cell body, a branching input structure (the dendrite) and a
ranching output structure (the axon)

Axons connect to dendrites via synapses

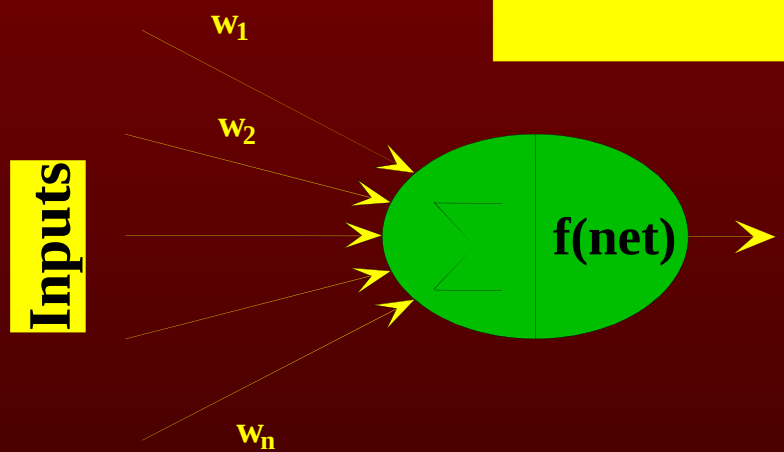
Electro-chemical signals are propagated from the dendritic in

# *Biological Analogy*

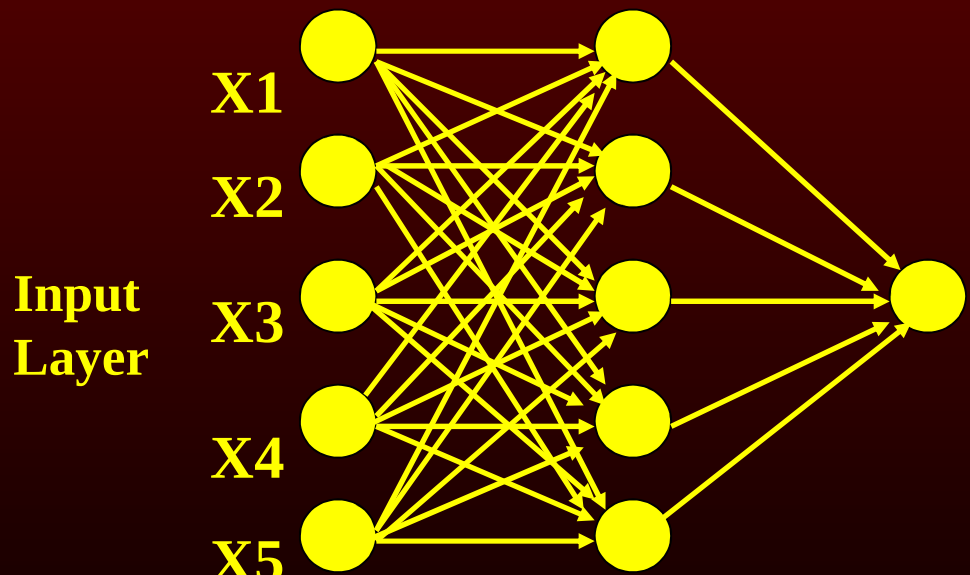
in Neuron




ficial neuron




of processing
ents (PEs) and
nections (weights)
adjustable strengths



# *Artificial Neural Networks*

***ine-learning algorithms that identify data patterns and perf***
***ecision making in a manner imitating cognitive functionalit***

**'*Learning*' (analogous to problem solving) is:**

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

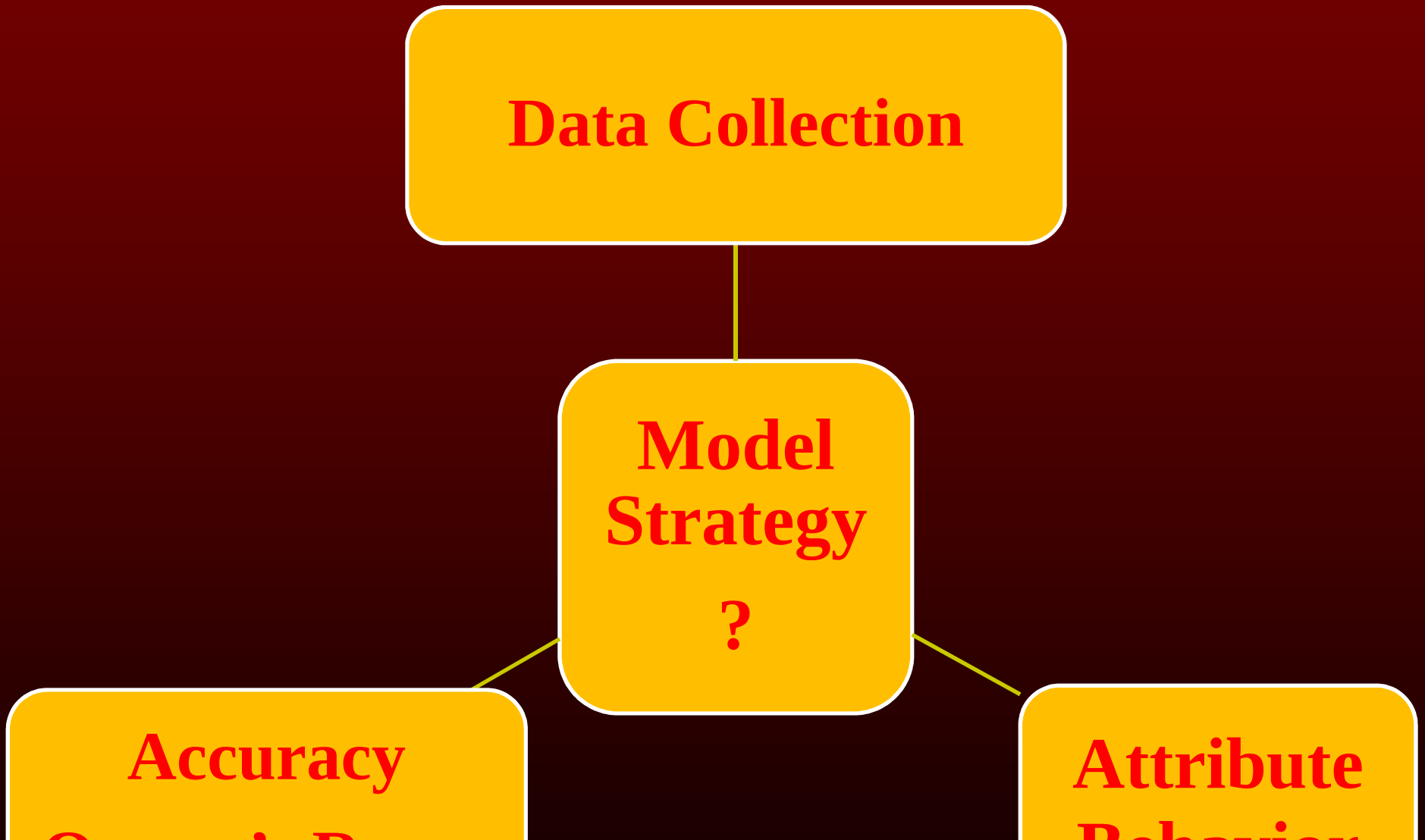
**'*Universal approximators*' – can discover & reproduce any *(linear / non-linear)* trend given enough data & computational (processing) capability**

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

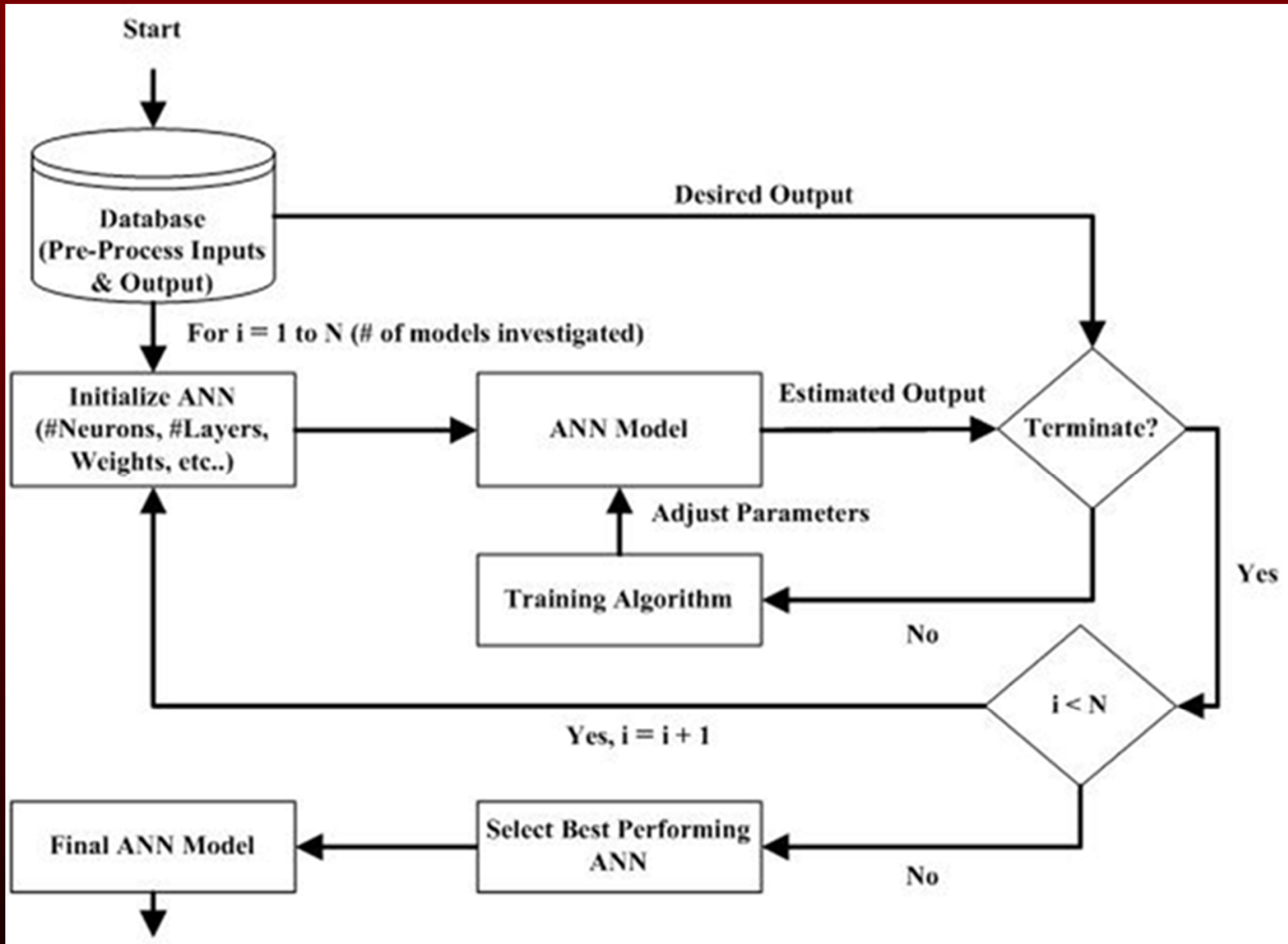
**Disadvantage - *(superficially ? ?)* lack a declarative knowledge structure**

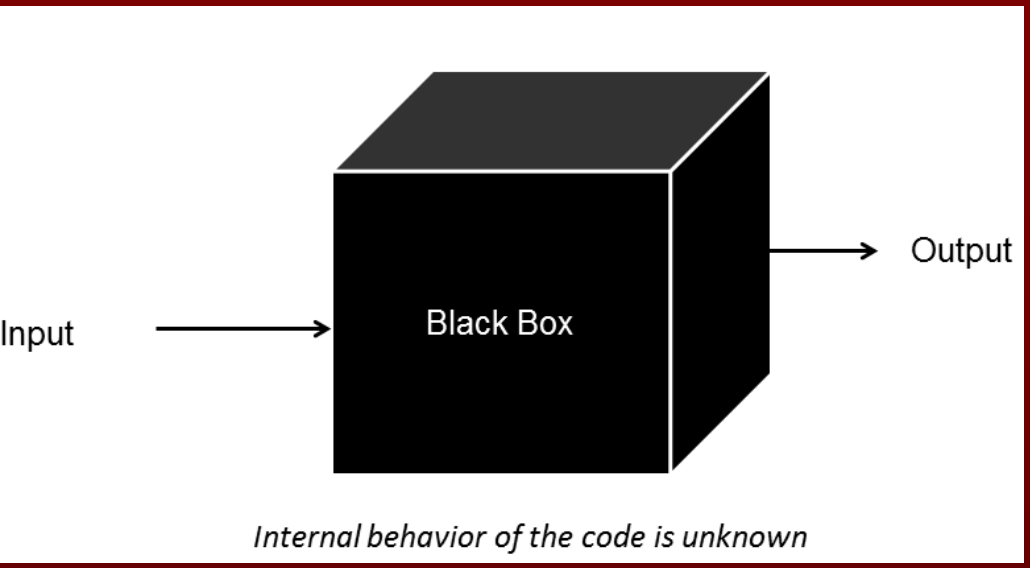
- ✓ **a *'Black Box'* (i.e. no global equation)**

# Early Days: Interested in “Model Accuracy”



# Modeling Approach

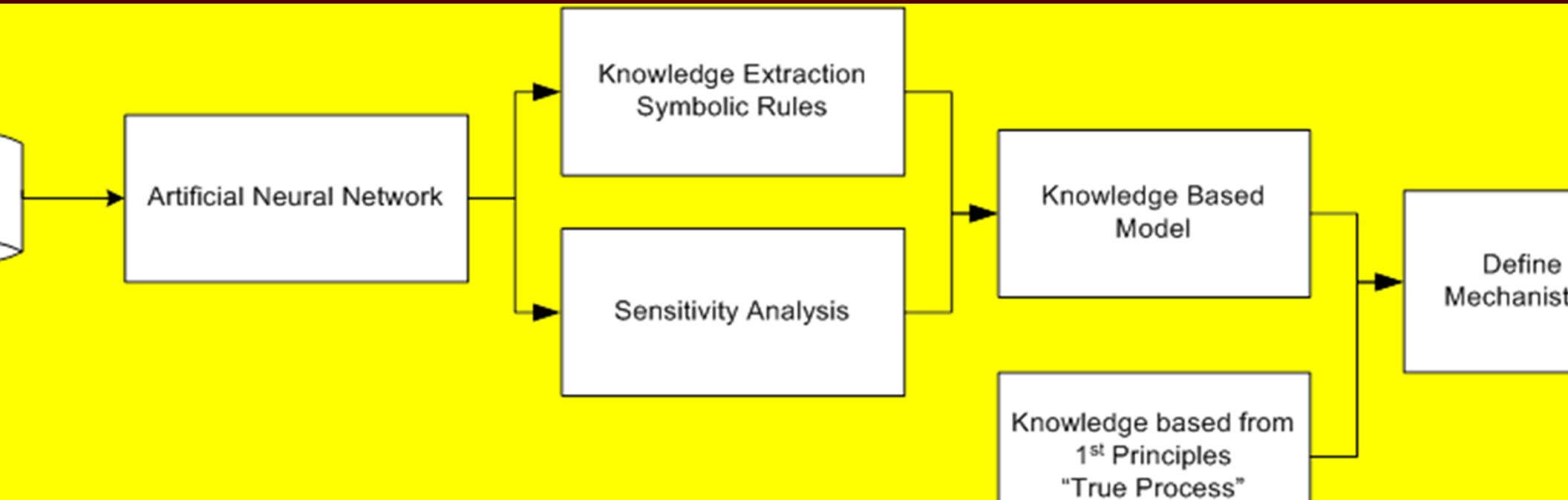





**KNOWLEDGE EXTRACTION defined:**

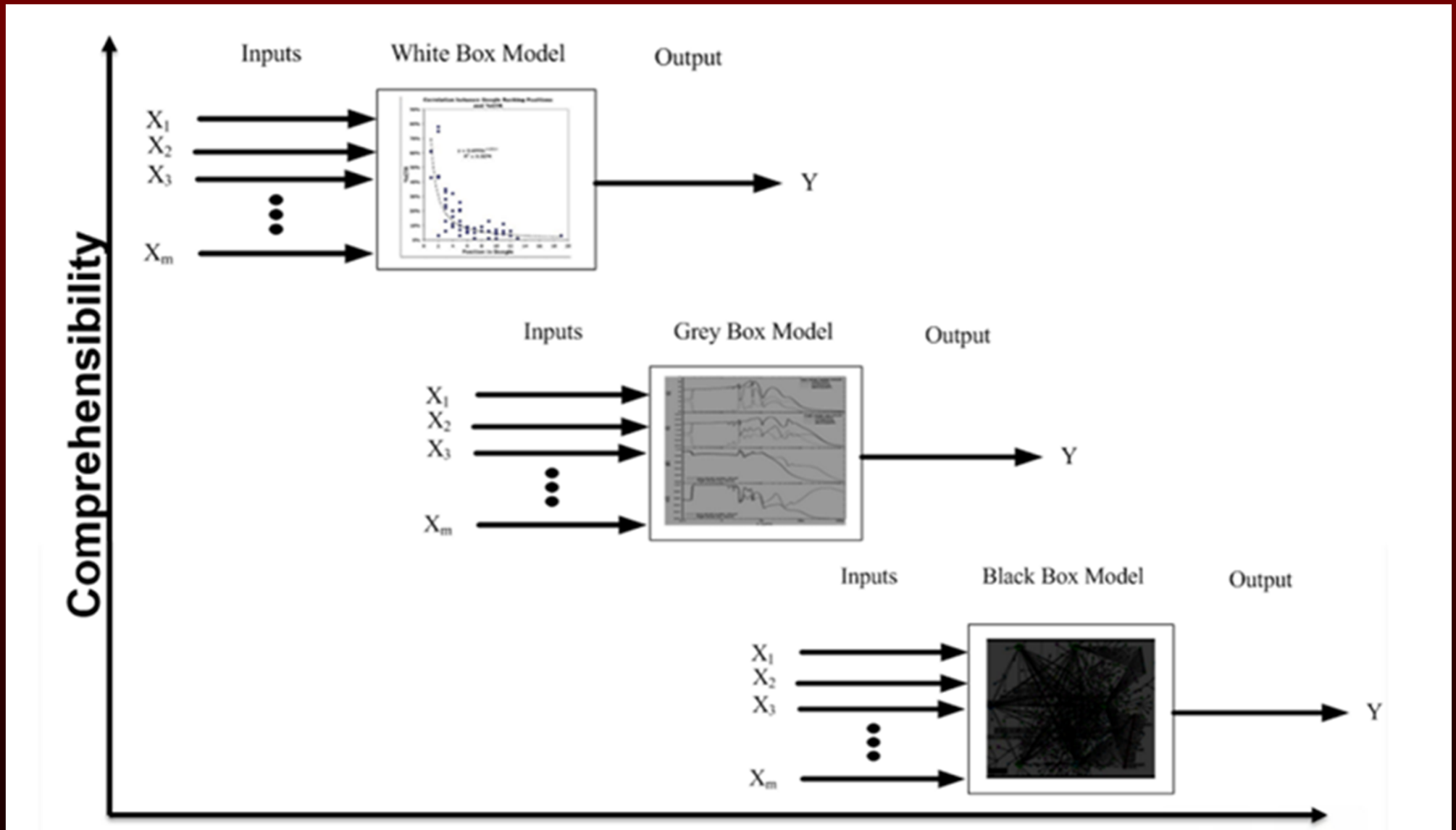
**is the creation of knowledge from structured (relational databases, XML) and unstructured (text, documents, images) sources [https://en.wikipedia.org/wiki/]**

# Is there a way illuminate the black box?



Multiple Variable Interactions while looking at various states!

Our drive to Mechanistic Model: Grey Box => WHITE BOX



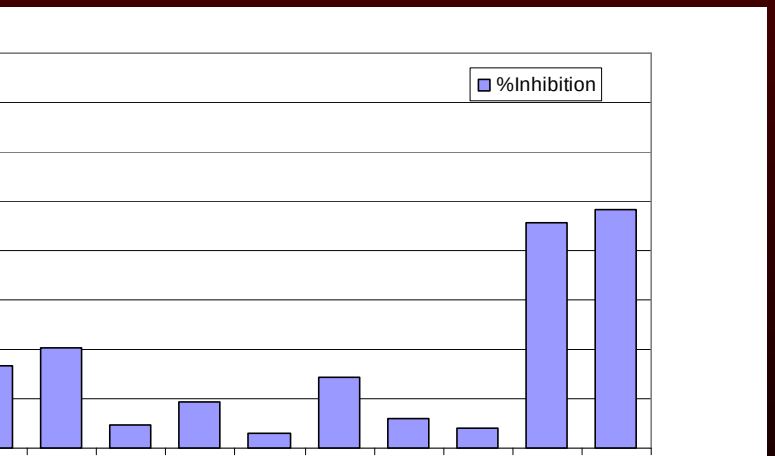

## 1st ATTEMPT:

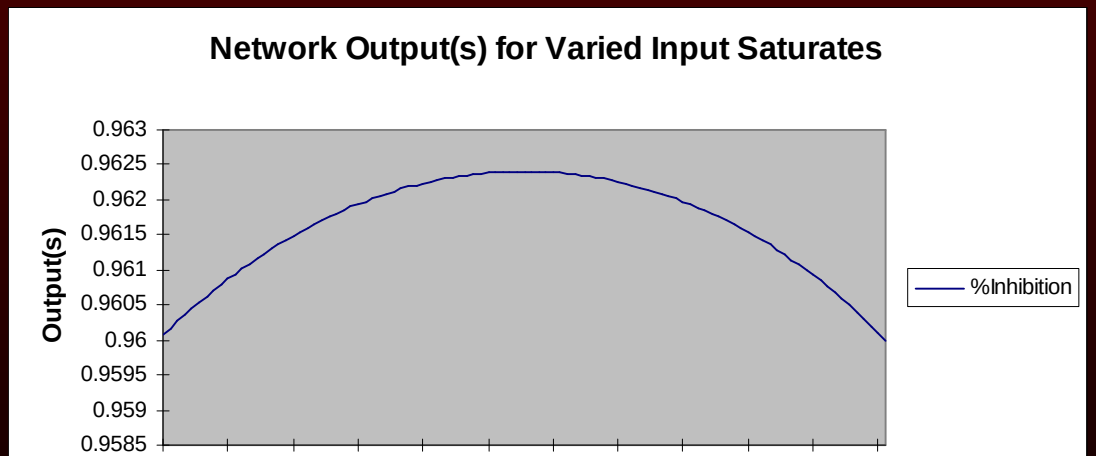
- Included all attributes collected
- Sensitivity about the means

- Found many limitations to current method

**How are we to explain a more complex situation?**

## Variable Behavior

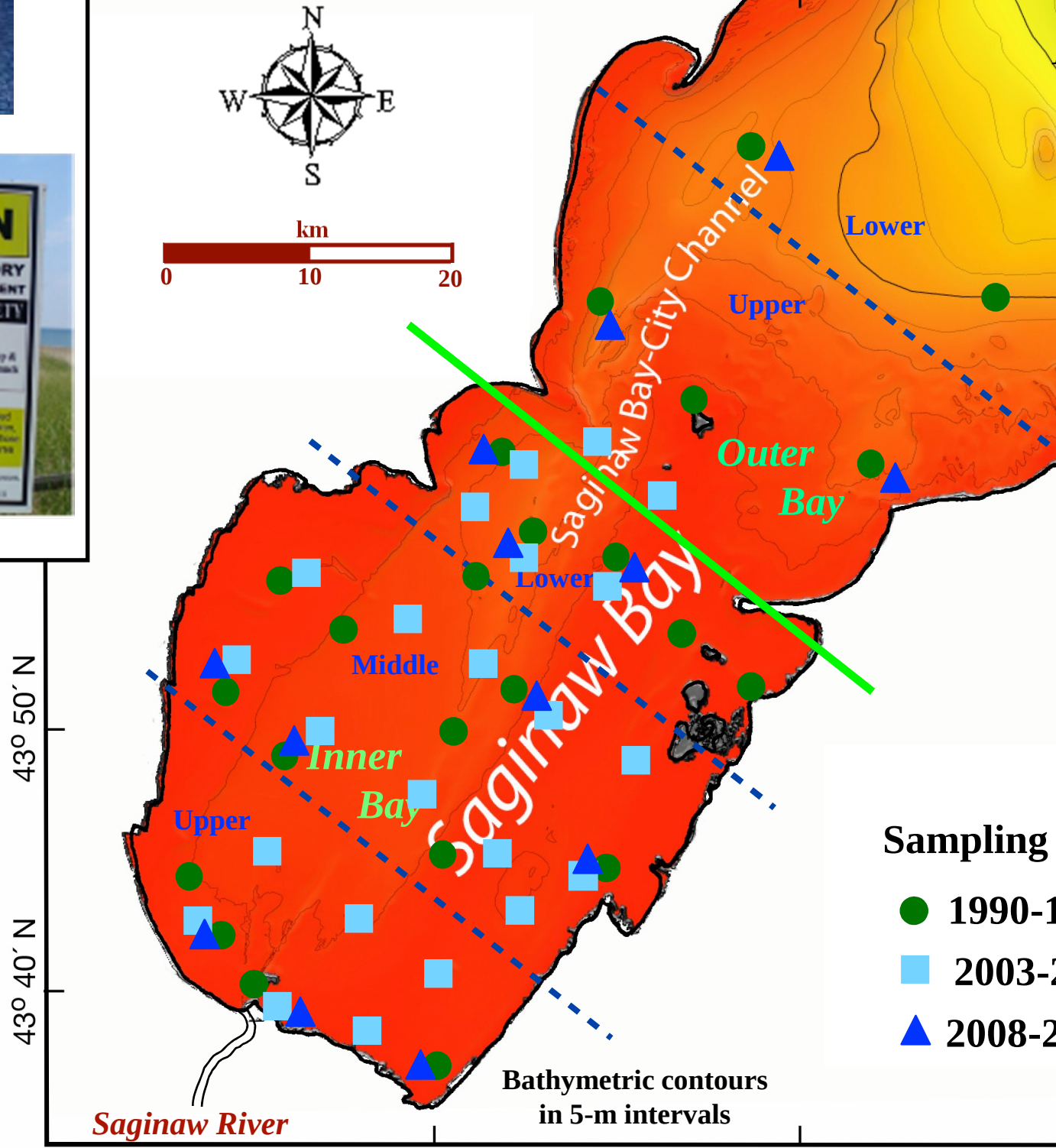





| | *Inner* | *Outer* |
|---|---|---|
| | 14.0 / 5.09 | 40.5 / 13.66 |
| | 1,554 | 1,217 |
| $^{3}$) | 7.91 | 16.63 |

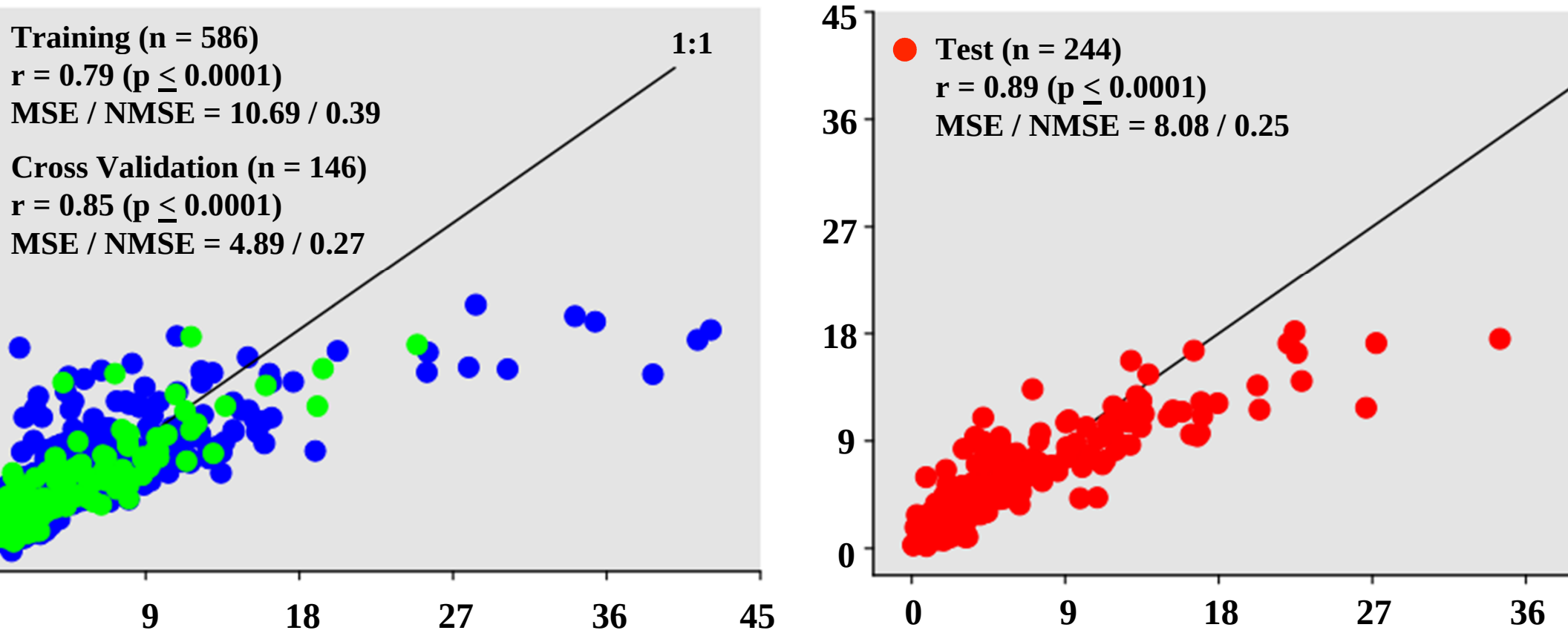
**58 - River mouth to Lake proper**



# Predicting Saginaw Bay Chl a (1991-1996)

## MLP - 1 Hidden Layer of 4 Processing Elements

**gical Predictors: °C, Sechhi, $K_d$, Cl, $NO_3$, $NH_4$, SRP, TP, $SiO_2$, $PSiO_2$, DOC, I**



# isting Knowledge Extraction Too

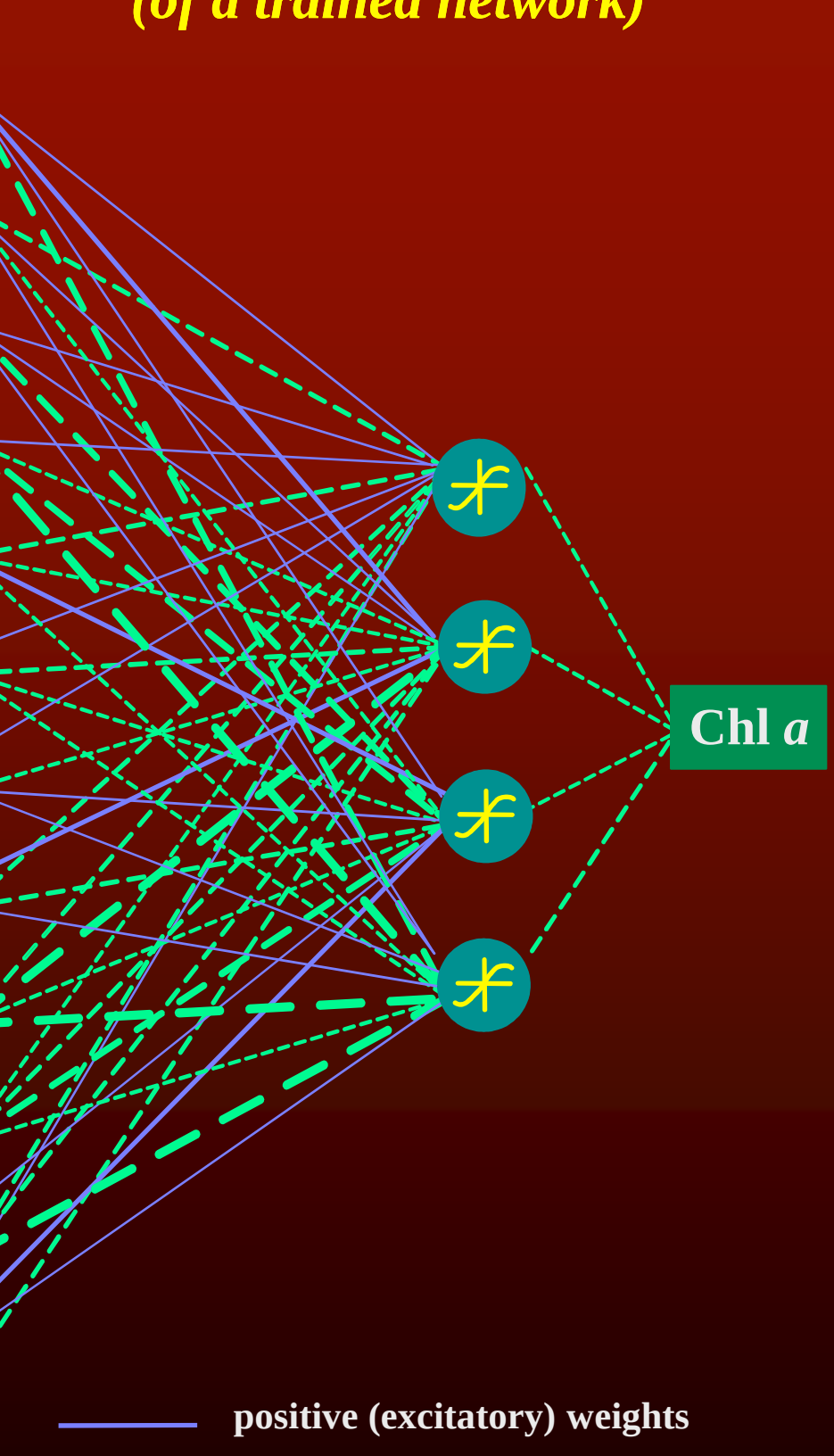
## Neural Interpretation Diagram

- Decomposition method to visual
  - Determine significance of input variables
  - Based on the magnitude of interconnecting weights

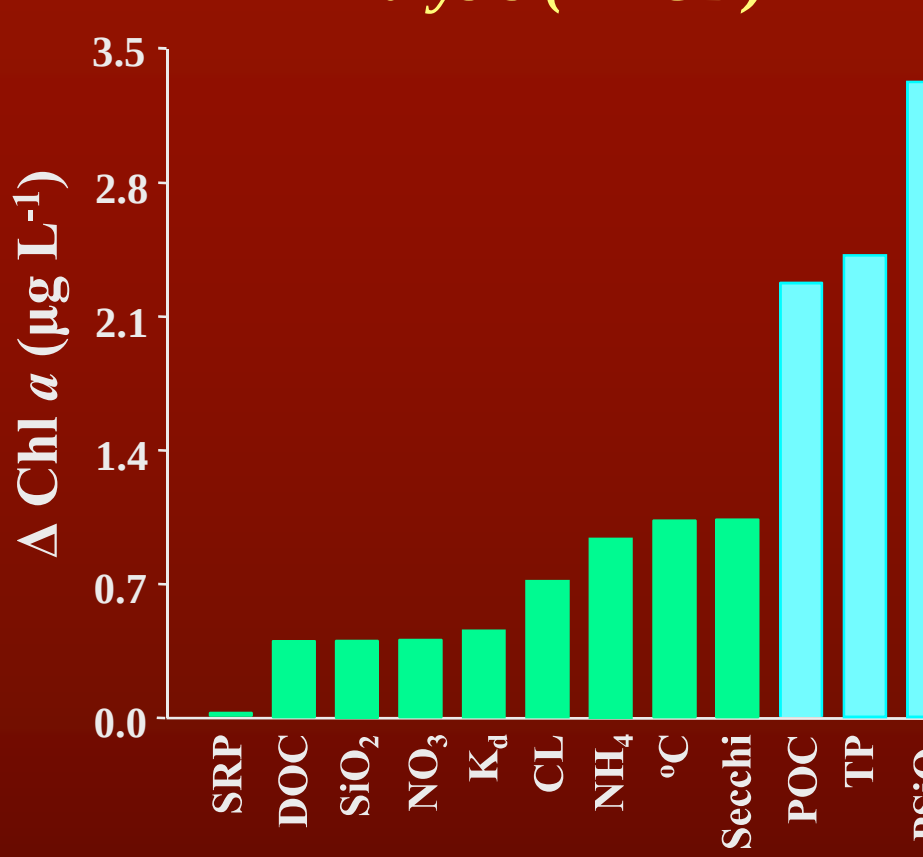
## Connected Weights

- 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

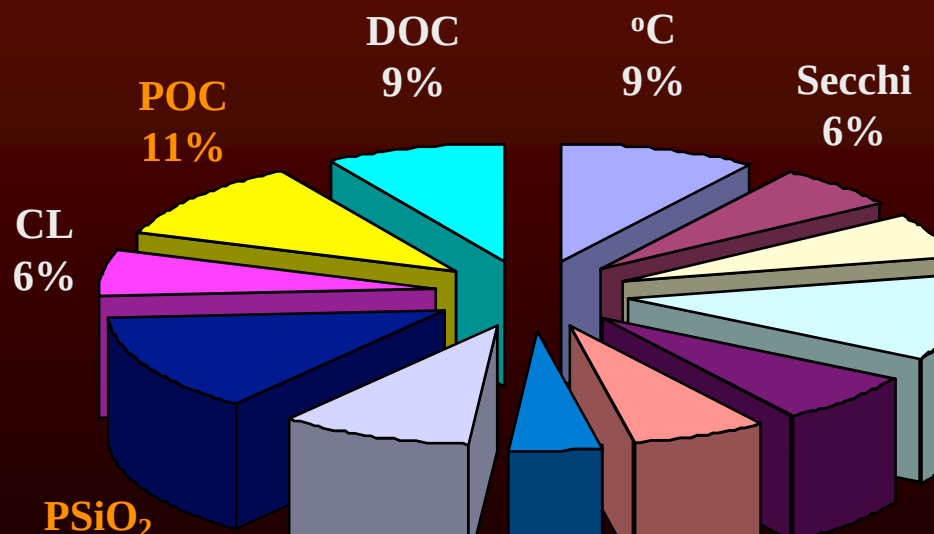
*(of a trained network)*



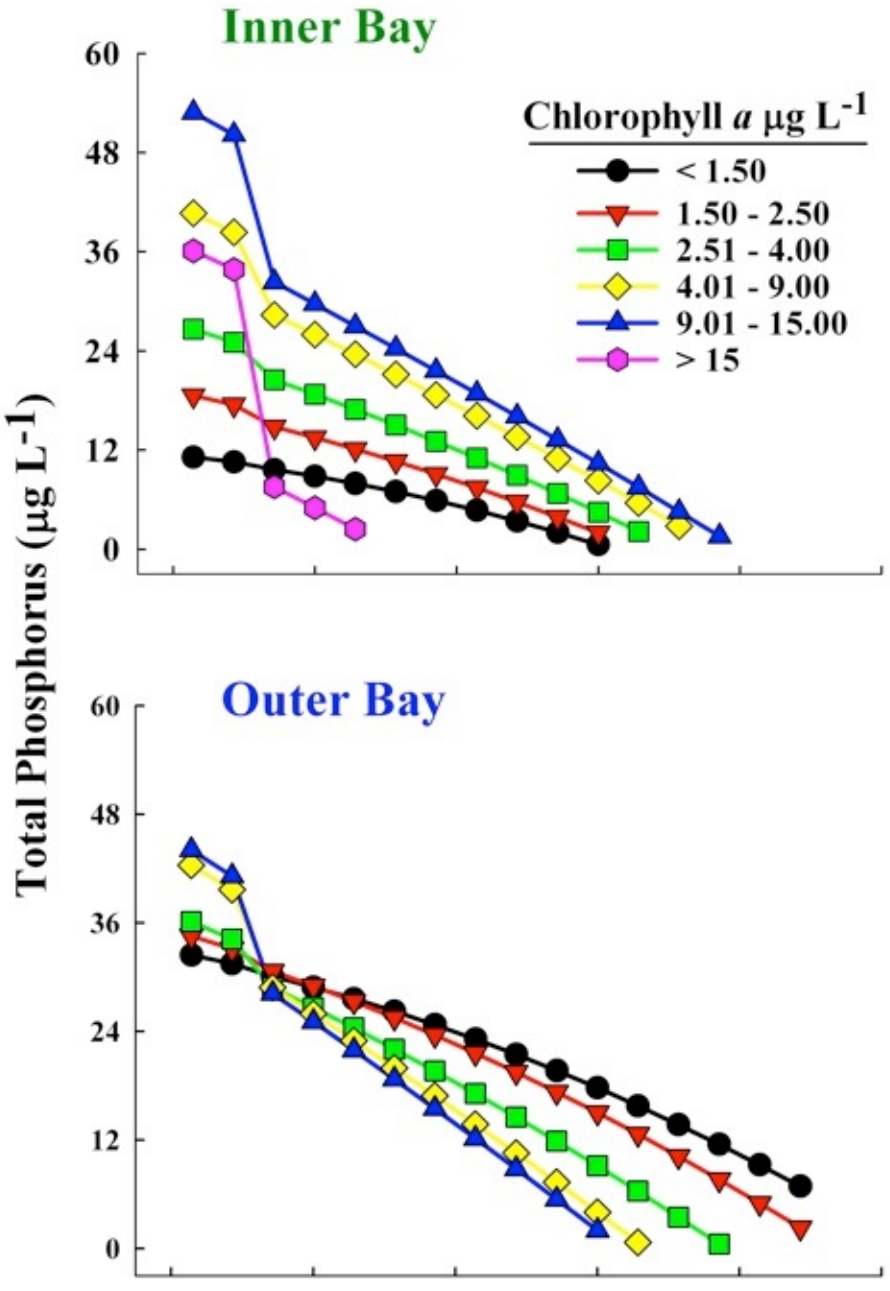

positive (excitatory) weights

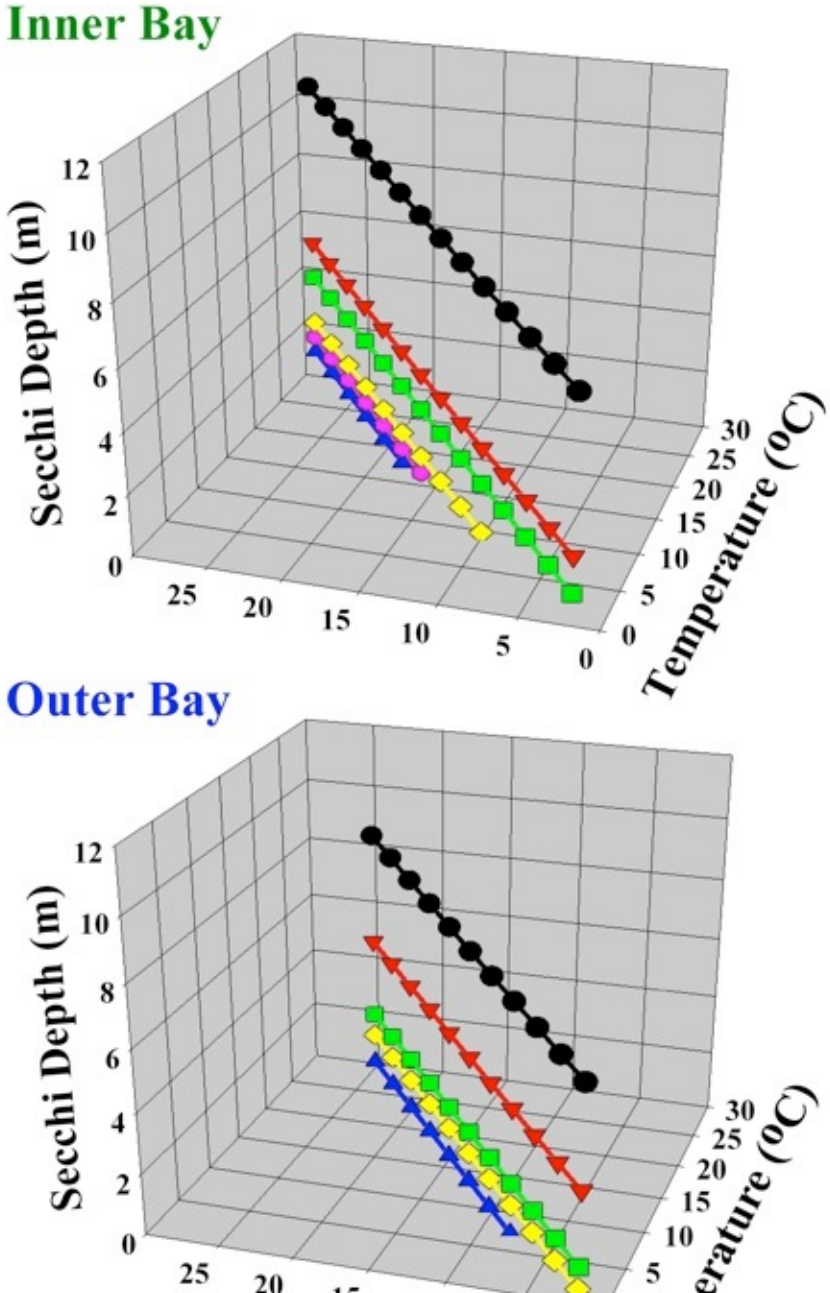



## *Garson's Algoritthm*
## Relative Share of Prediction

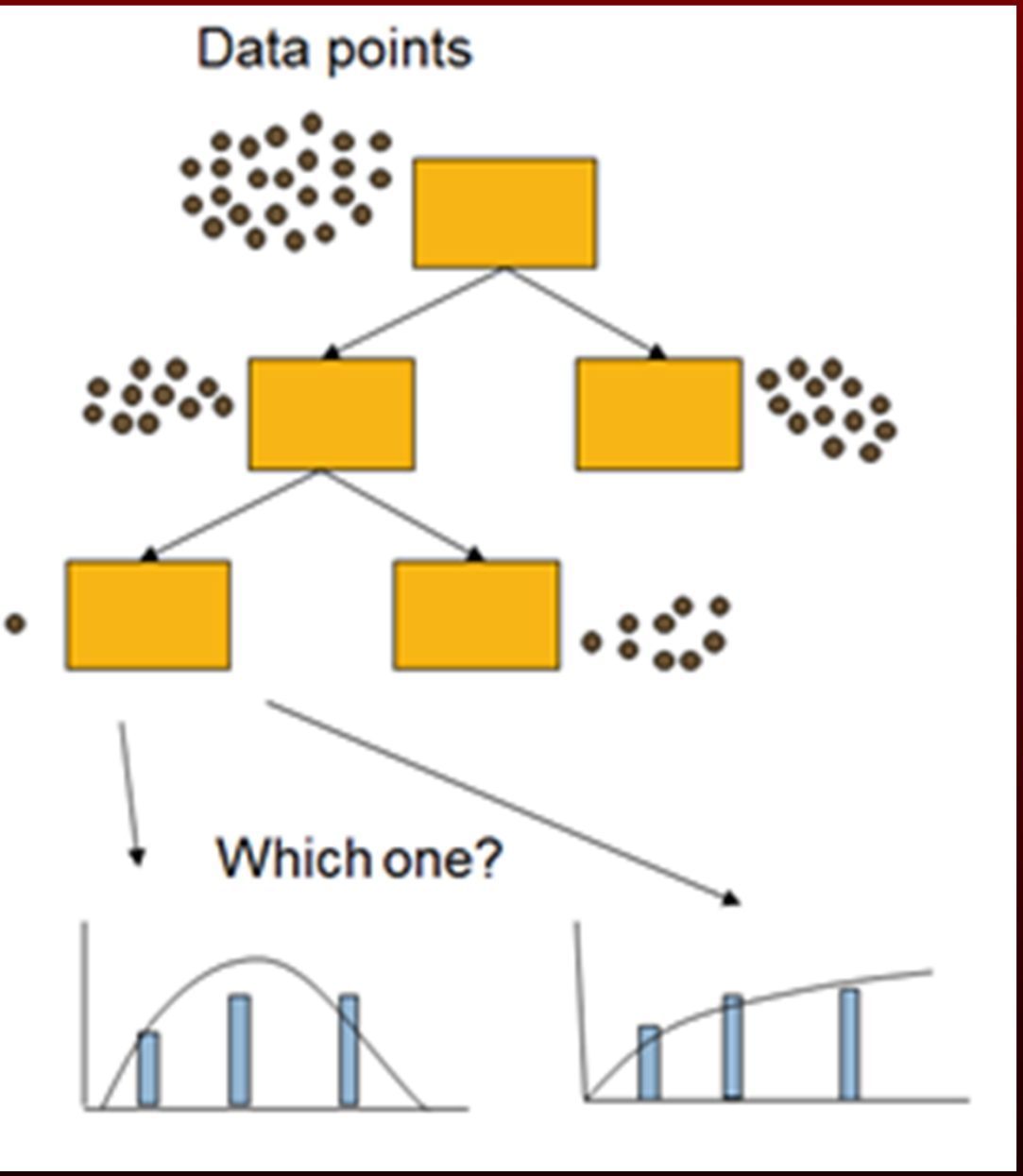


# *Multi-Variable Sensitivity Analysis (circa 2006 !)*

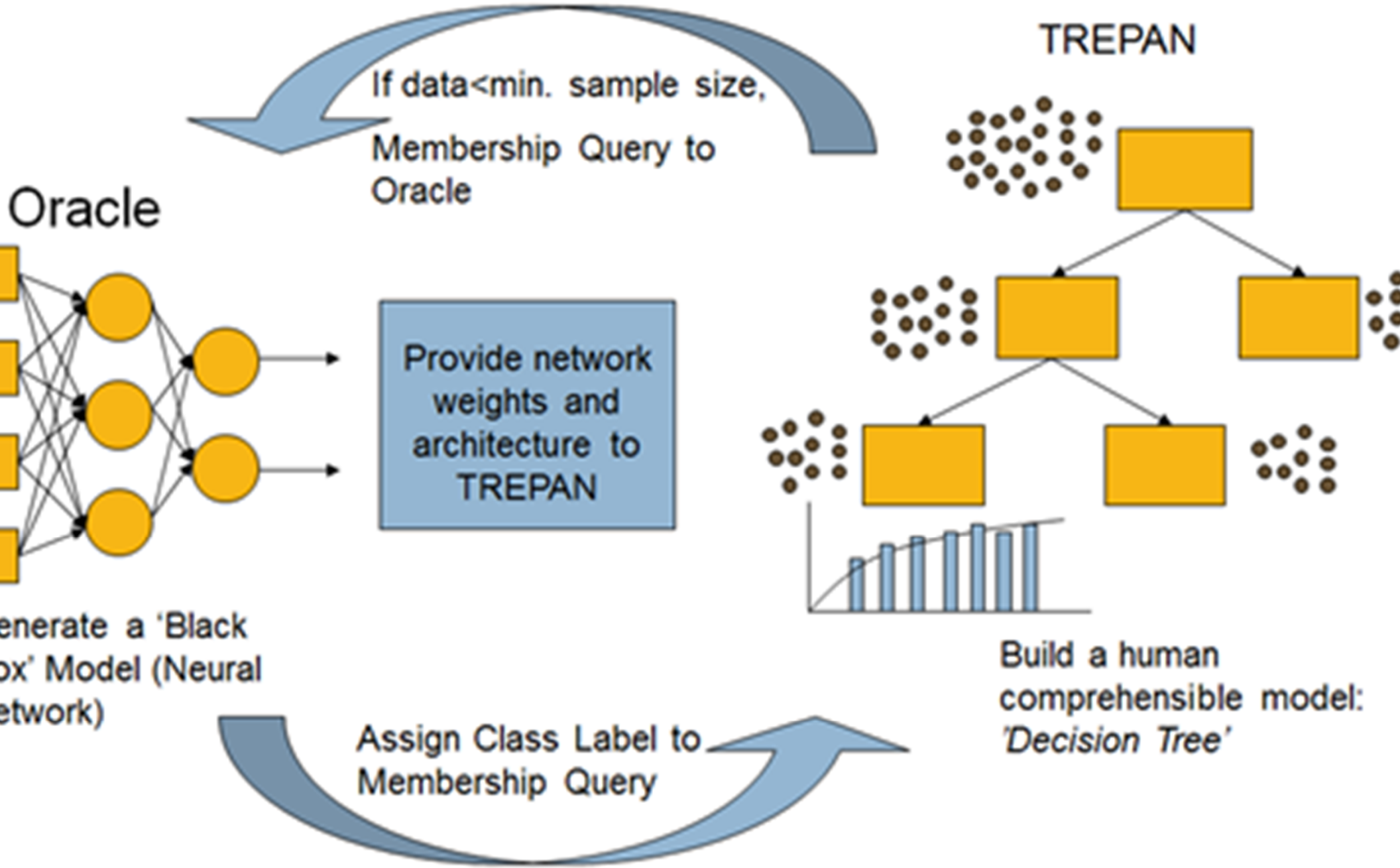


# Decision Trees




- Symbolic Knowledge Extr
  Technique
- Most commonly used dec
  tree induction algorithm –
  (Quinlan)
- Recursive partitioning of t
  data
- Drawback: Amount of dat
  reaching each node decre
  with the depth of the tree
- Alternative: TREPAN


If data<min. sample size,
Membership Query to Oracle
Oracle
Provide network weights and architecture to TREPAN
TREPAN
enerate a 'Black ox' Model (Neural etwork)
Assign Class Label to Membership Query
Build a human comprehensible model: 'Decision Tree'

# *Different Project: Crude Oil Impact*

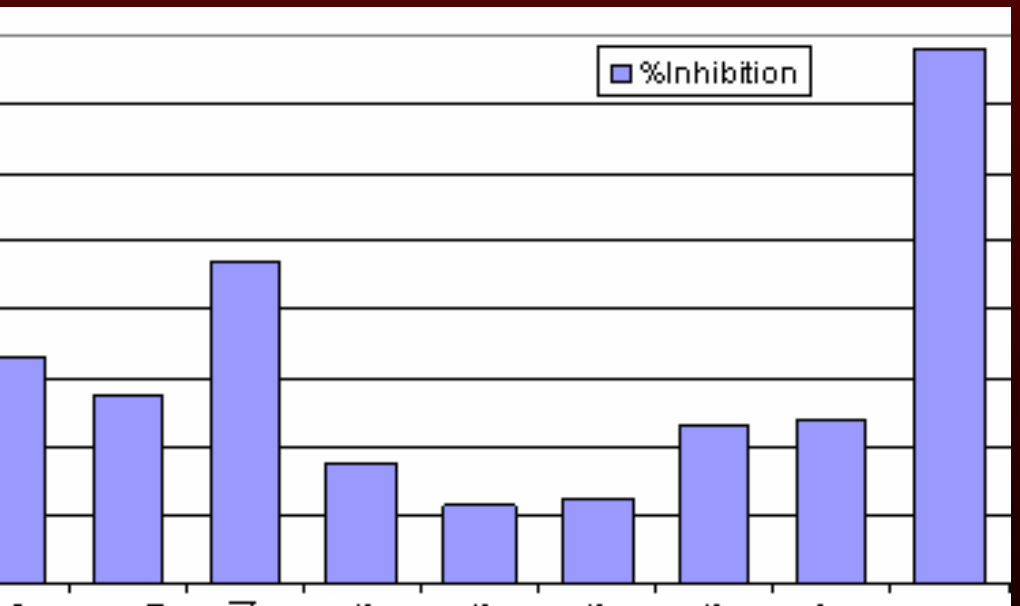
## d New Set of Tools:

## itations to Sensitivity:

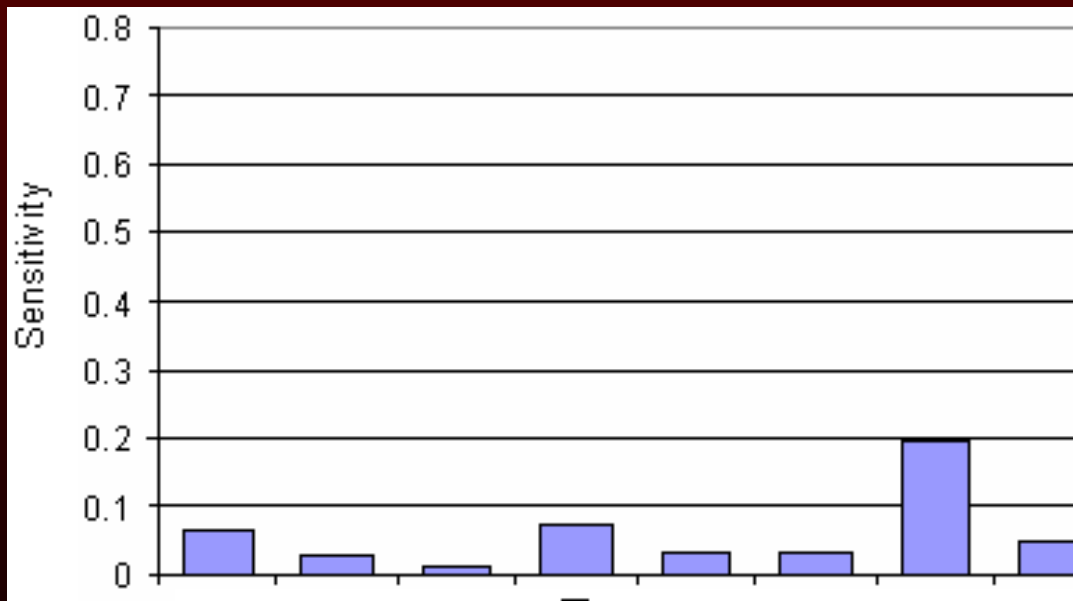
## ANNs were created for “high” and “low” %Crude Oi

## ensitive results were very different

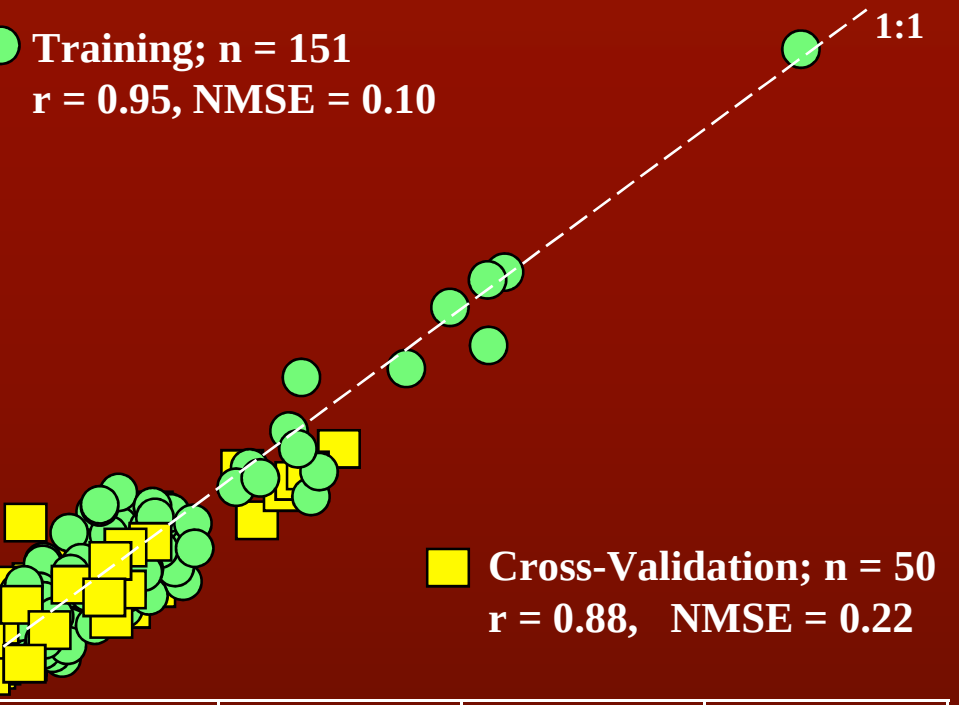
**Crude Oil <=20%**

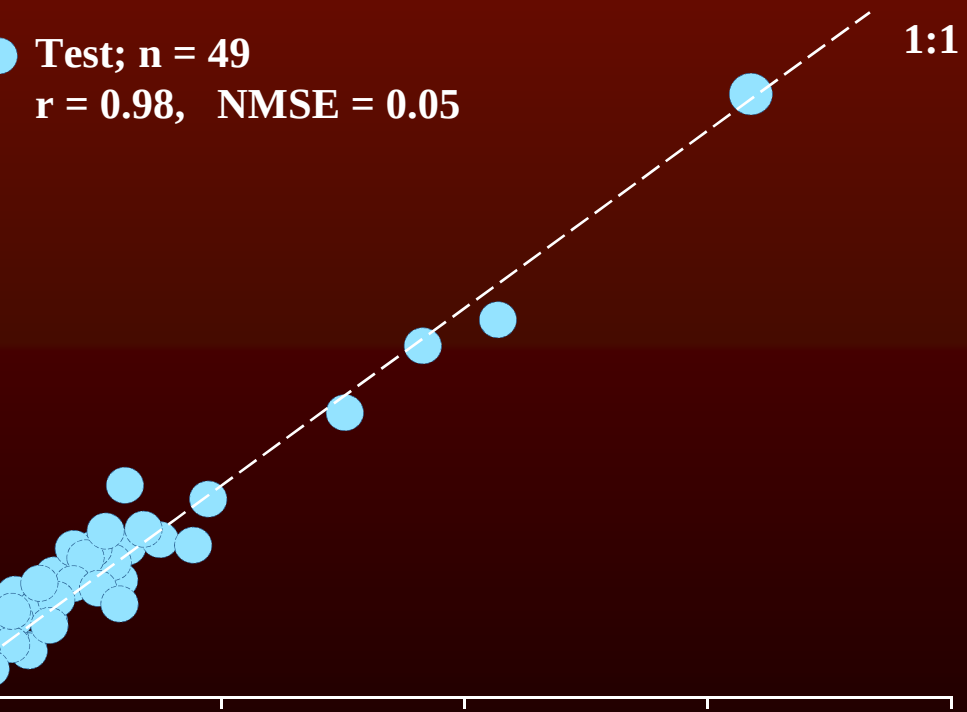


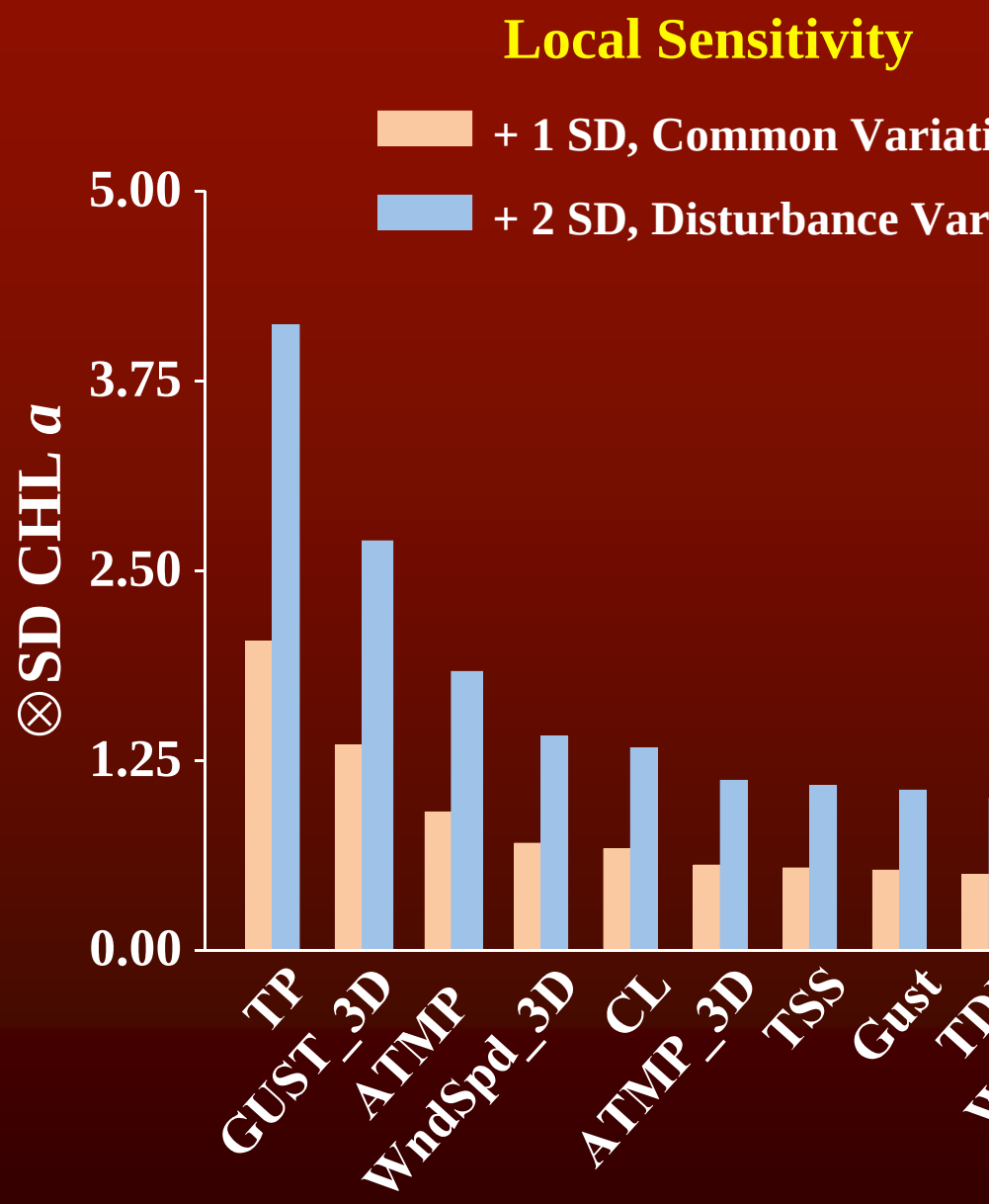

**Crude Oil >=50%**



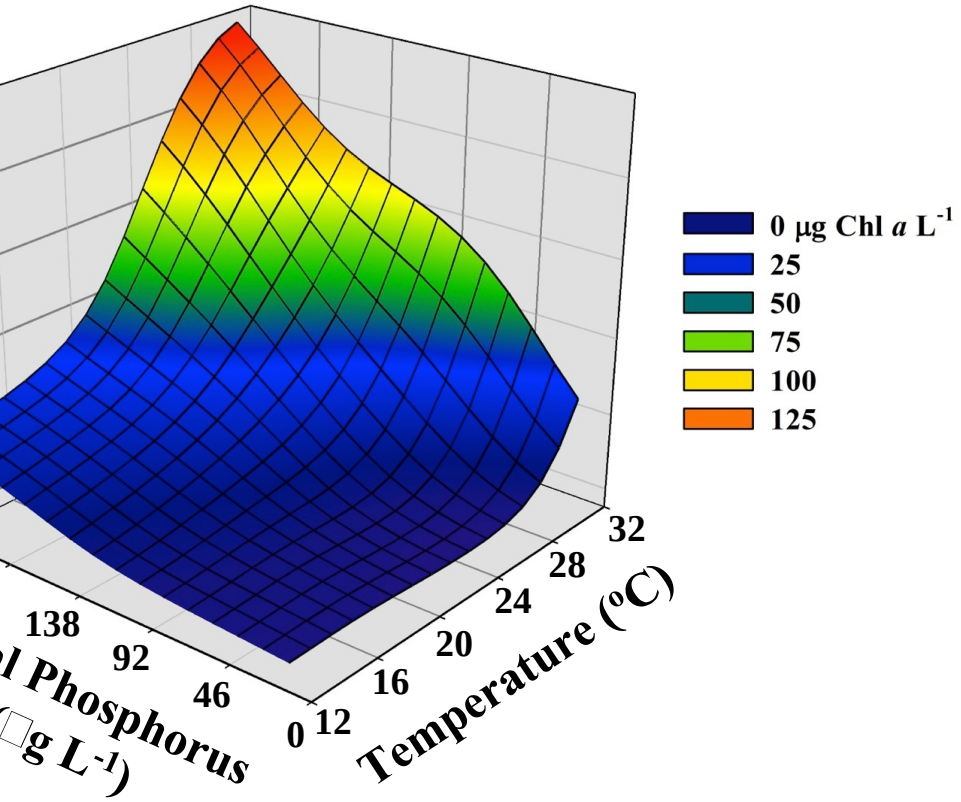
# range of Sensitivity



## HL as a function of TP & TEMP

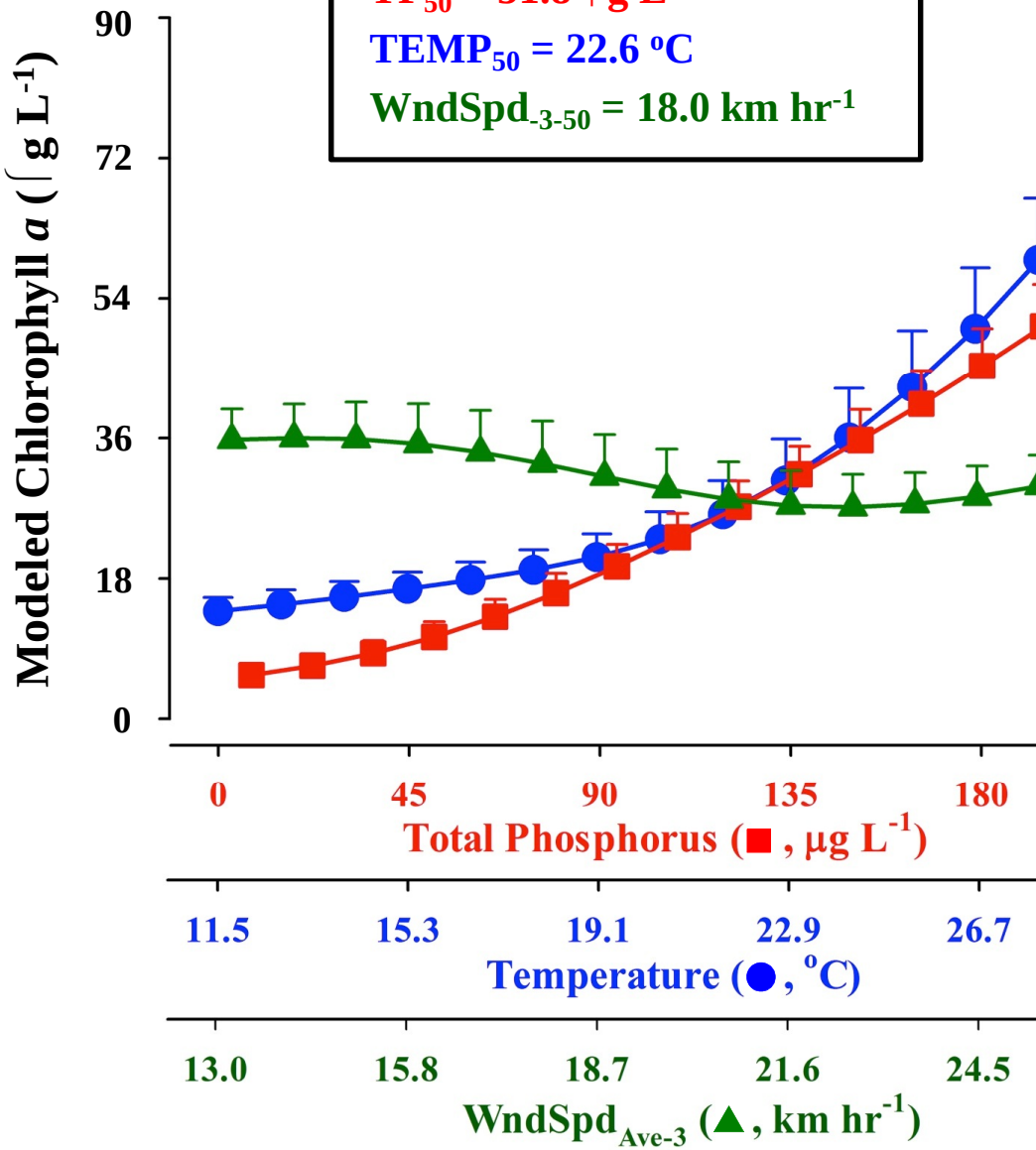



**Half-Maximal Abundance Concentrations / Conditions:**

$TP_{50}$ = 51.8 µg $L^{-1}$

$TEMP_{50}$ = 22.6 °C

$WndSpd_{3\text{-}50}$ = 18.0 km $hr^{-1}$




**1.98 + (0.03*TP)**

it SE = 0.41,Fstat = 29857.36

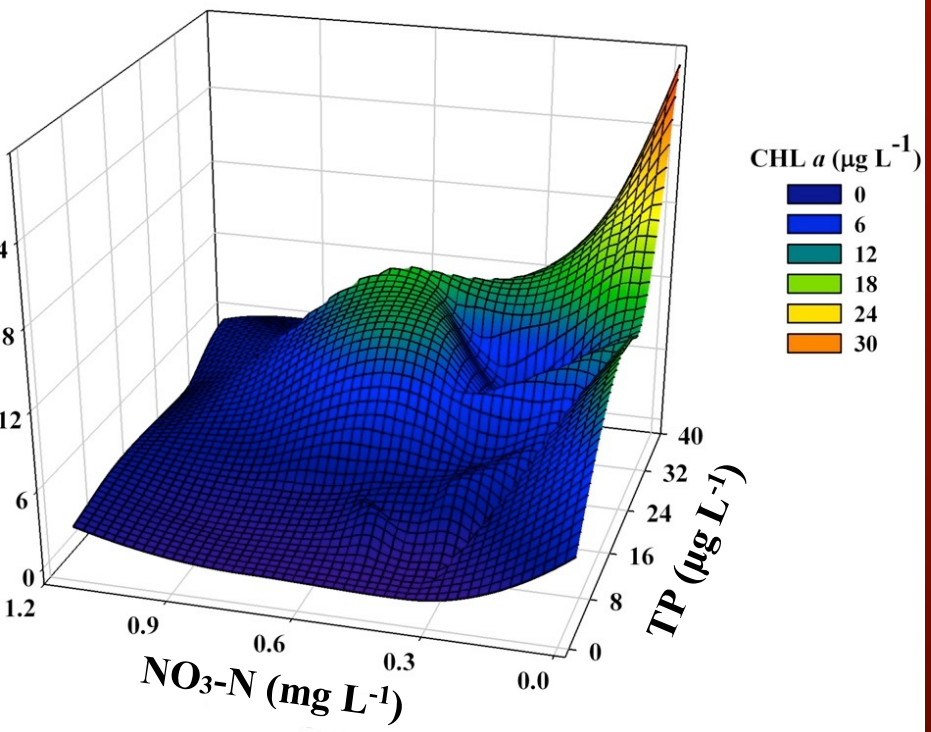
**2.23 + (0.002*$TEMP^2$)**

it SE = 1.03,Fstat = 6323.88

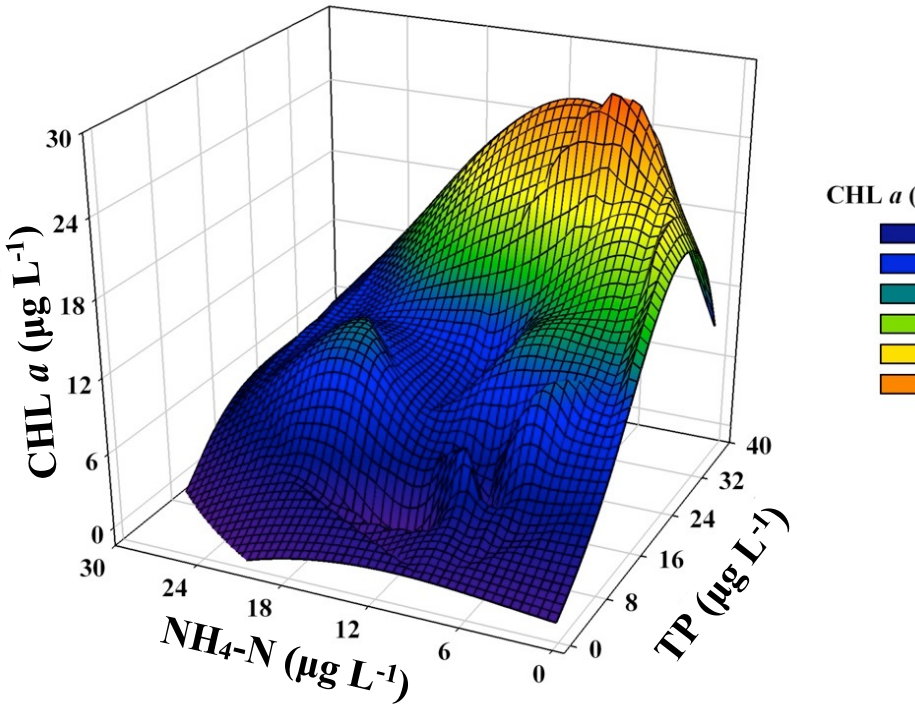
**62.16 + (473.88*$WndSpd_{Ave\text{-}3}$) - (103.65*$WndSpd_{Ave\text{-}3}^2$)**
**ndSpd$_{Ave\text{-}3}^3$) - (0.82*$WndSpd_{Ave\text{-}3}^4$) + (0.03*$WndSpd_{Ave\text{-}}$**
**$WndSpd_{Ave\text{-}3}^6$) + (5.80e-6*$WndSpd_{Ave\text{-}3}^7$)**

*...king Into Account the Interactions and/or Synergisms of Co-Limiting Nutrients*

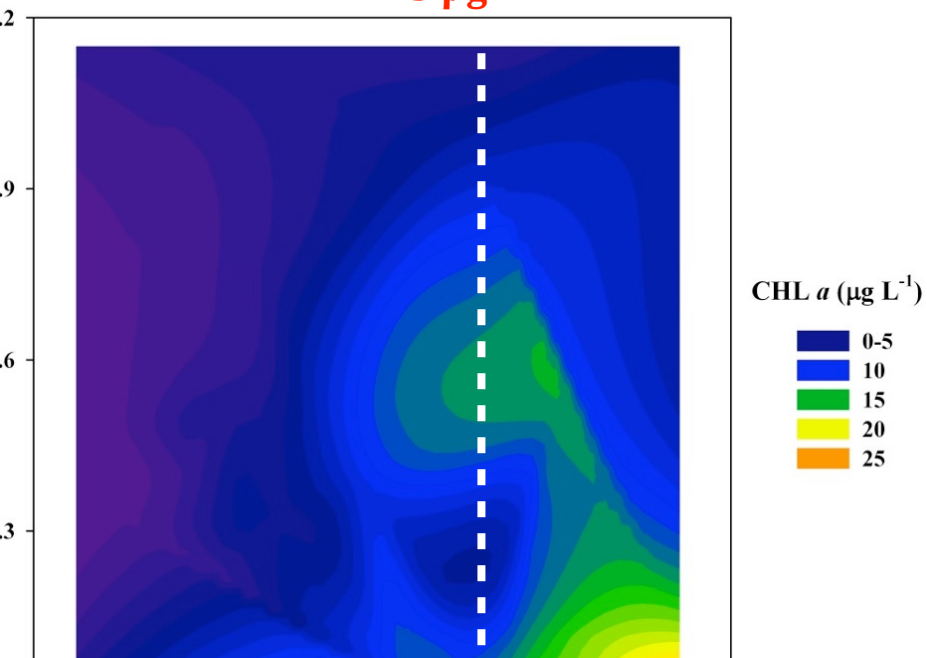
## L *a* as a function of TP & $NO_3$-N



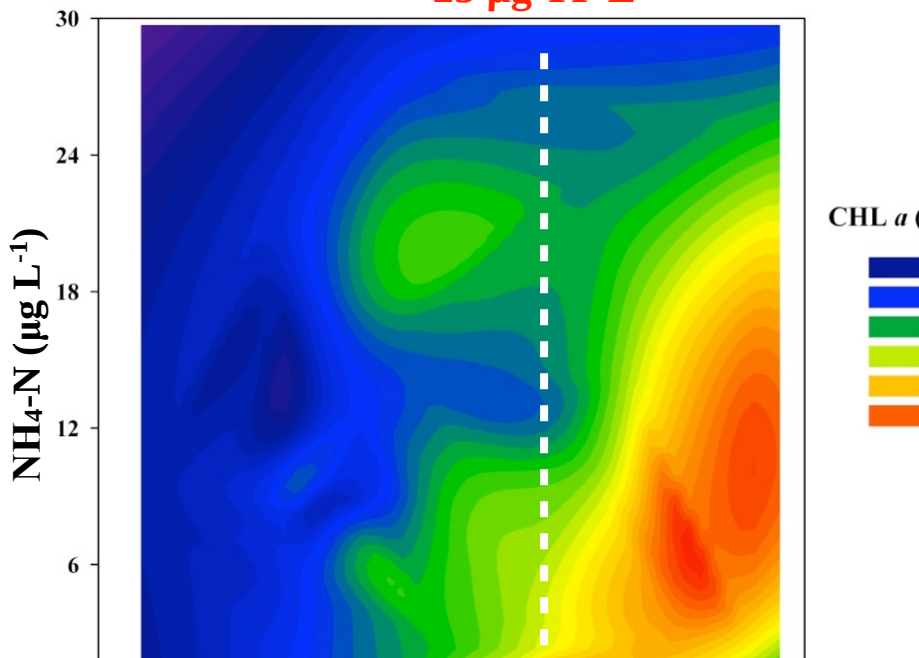

## CHL *a* as a function of TP & $NH_4$-N




## 25 µg TP $L^{-1}$



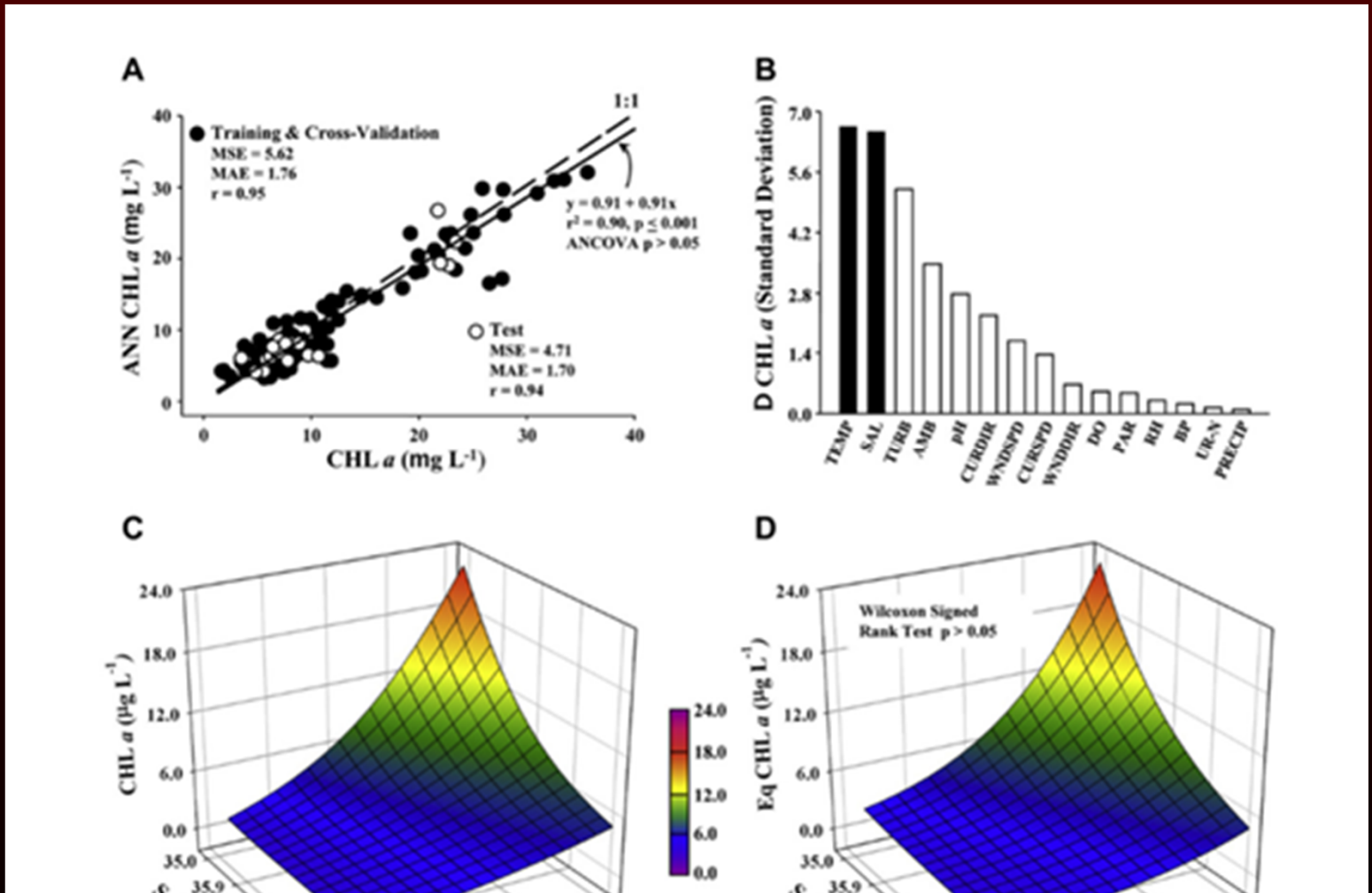

## 25 µg TP $L^{-1}$



# evelopment of Grey Box Technique

$$[\mathrm{CHL}\,a] = w_1 \cdot f(x_1, y_1) + r_1,\ r_1 = w_2 \cdot f(x_2, y_2) + r_2,$$
$$r_2 = w_3 \cdot f(x_3, y_3) + r_3,\ \text{and}\ r_{n-1} = w_n \cdot f(x_n, y_n) + r_n$$

**Generalized Equ for 2 variable int with output (CH**

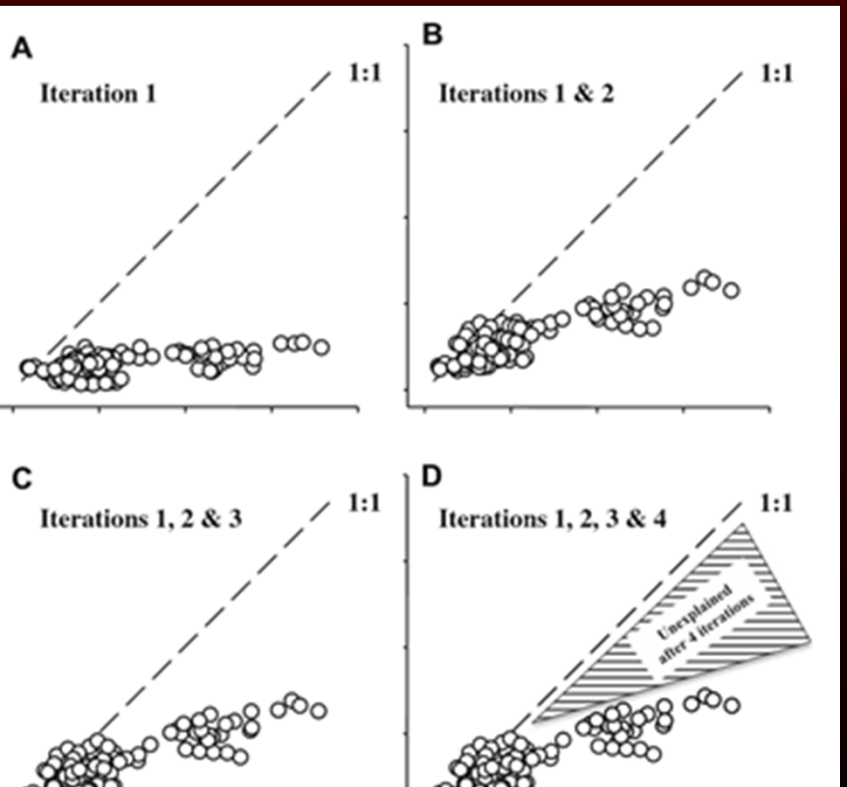


# Iterations : ANNs Models

$$[\text{CHL } a]_{\text{Grey–Box}} = [\text{CHL } a]_{\text{1st iteration}} + [\text{CHL } a]_{\text{2nd iteration...}} + [\text{CHL } a]_{\text{nth iteration}} + r_n$$

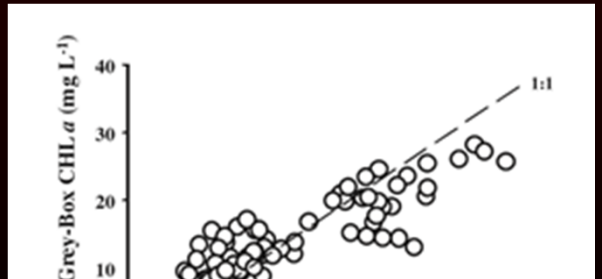
Multiple ANN mode
utilizing 2 variables
time to predict Outp

ons: Additive Models




Finalized Combined Model

$$[\text{CHL } a_{\text{Grey–Box}} = -13.3 + 2.25\left(\begin{array}{l}(0.41*1.16 + 5.0\text{E}17*(\text{TEMP}^{6.592}*\text{SA} \\ +\left(0.35*\frac{(0.72 - 0.01*\text{pH} + 0.115*}{(1 - 1.08*\log_{10}\text{pH} + e*\log} \\ +\left(0.12*\frac{(70.89 + 12.69*\log_{10}\text{CURSPD} - 12}{(1 - 13.43*\text{CURSPD} + 0.0} \\ +\left(0.06*\frac{(8.01 - 1.98*\log_{10}\text{CURDIR} - 2.69*}{(1 - 0.01*\text{CURDIR} + 0.03*\text{W}}\end{array}\right.$$



# Global Sensitivity

ensitivity about Means

- Local Sensitivity
- Does not consider variable interactions as states change

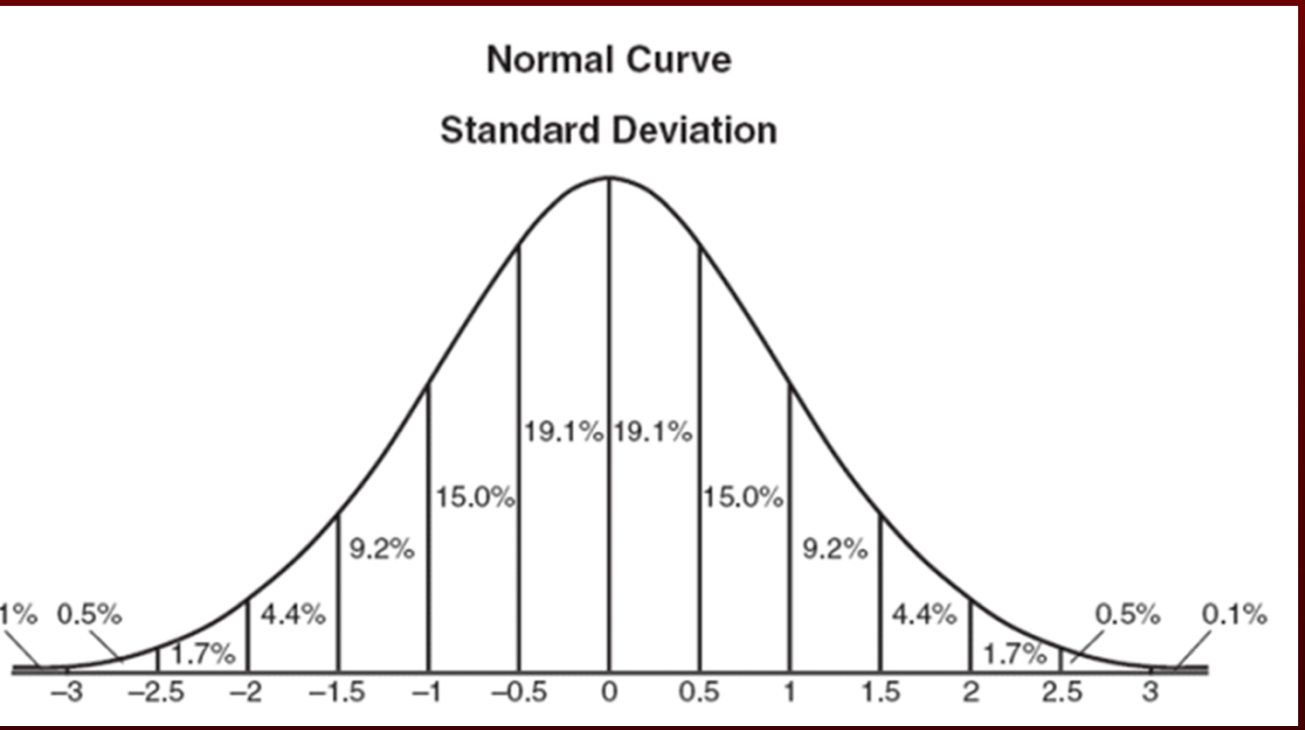
eveloped Global Sensitivity

- Looks at how variables interact as their state change!

# Culprit: Correlation

| | Tot_Par | Ave_WndSpd | Ave_WndDir | Ave_BP | Tot_Precip | Ave_AmbC | Ave_RH% | Ave_CurrentSpd | Ave_CurrentDir | Ave_Urea | Ave_WatC | Ave_PSU | Ave_pH | Ave_Turb |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 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 | | | | | | | | | | |
| Tot_Precip | -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 |

# Global Sensitivity

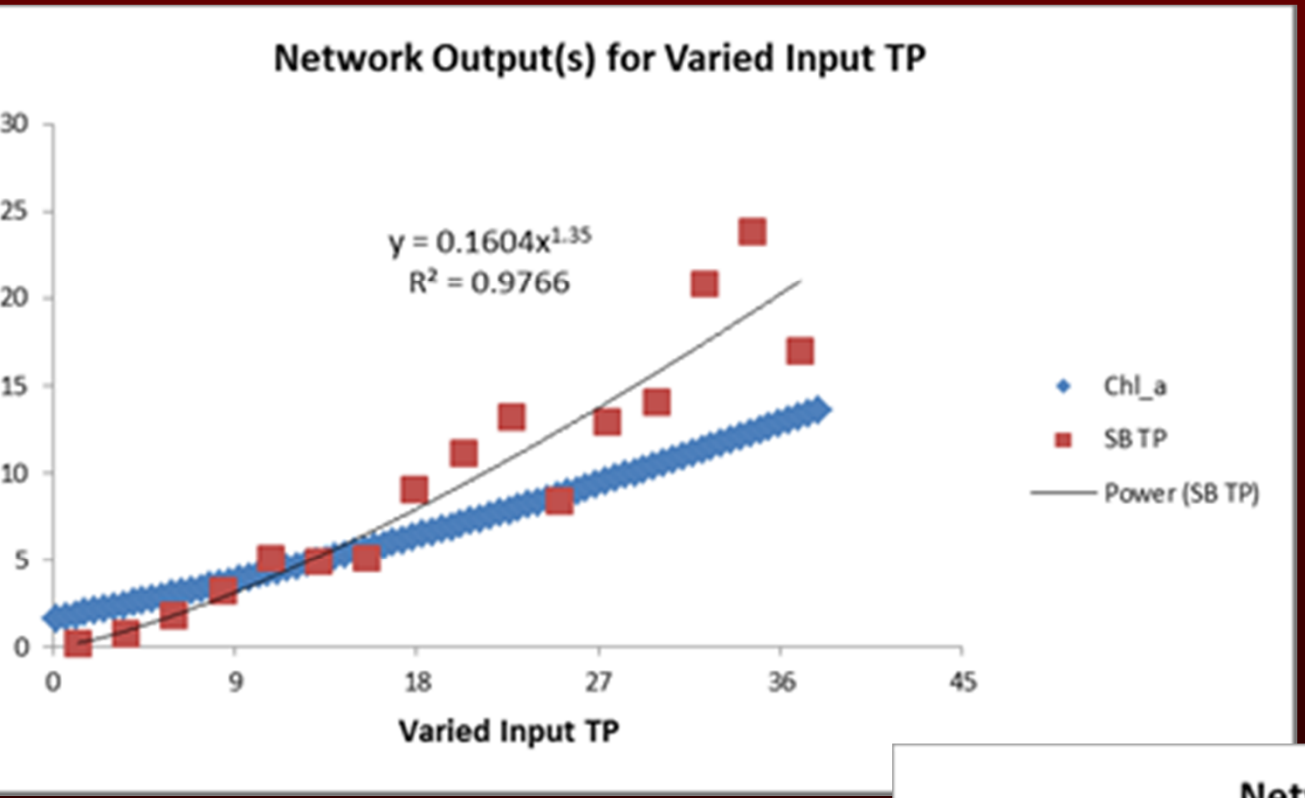



**Each Variable has its o**
**distribution of values (S**

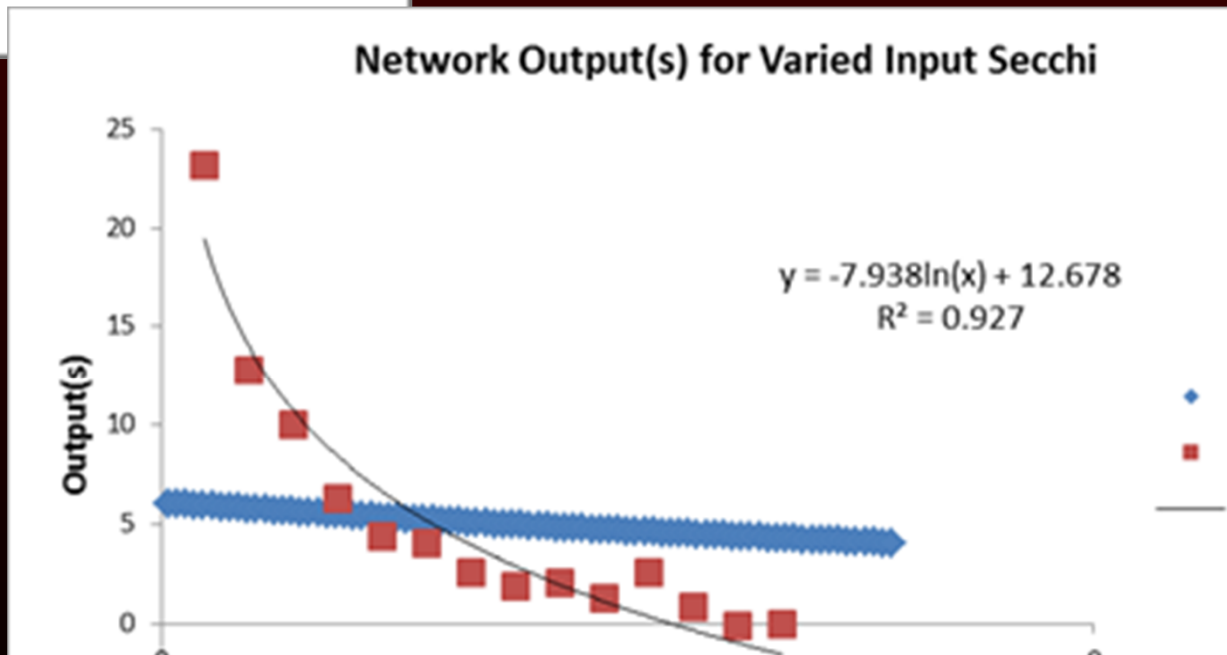
**Impact of Correlation o**
**Behavior**

| N | Secchi | TSS | TP | TDP | SRP | NH4 | NO3 | CL | Sol_Si | POC | DOC |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 σ | 1.57 | -0.70 | -0.98 | -0.48 | -0.25 | 0.02 | -0.02 | -0.57 | -0.16 | -1.16 | -0.80 |
| 5 σ | 0.53 | -0.67 | -0.59 | -0.04 | -0.14 | 0.09 | 0.41 | -0.02 | -0.40 | -0.79 | 0.04 |
| 5 σ | -0.17 | -0.08 | -0.16 | -0.11 | -0.09 | -0.09 | -0.04 | -0.14 | -0.09 | -0.26 | -0.14 |
| 5 σ | -0.40 | -0.02 | 0.14 | 0.13 | 0.04 | -0.26 | -0.24 | -0.16 | 0.39 | 0.35 | -0.06 |
| 5 σ | -0.68 | 0.50 | 0.31 | -0.37 | -0.06 | -0.49 | -0.35 | -0.06 | 0.20 | 0.87 | 0.15 |

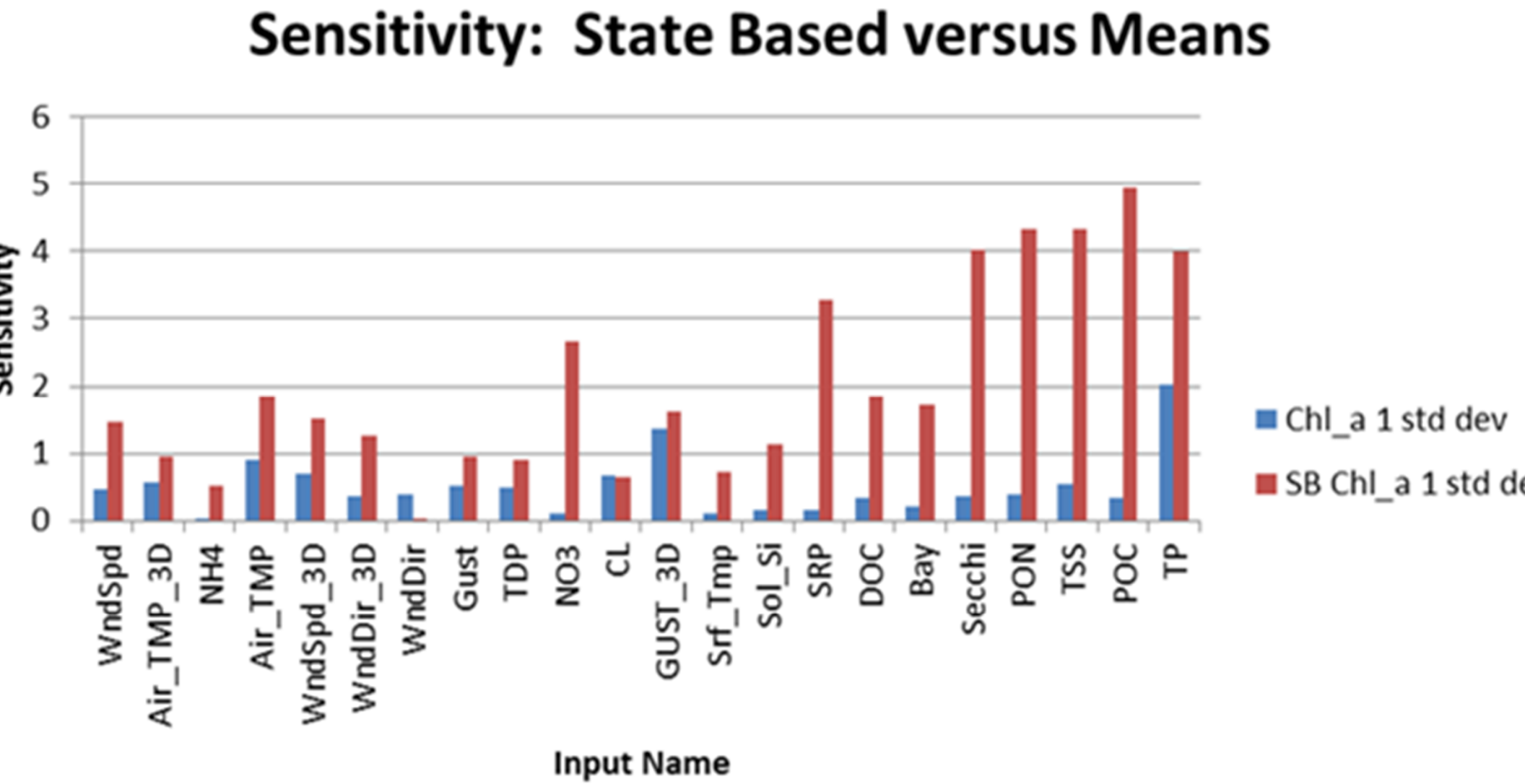
# Global Variation Across States




Significant difference Global versus Local Sensitivity



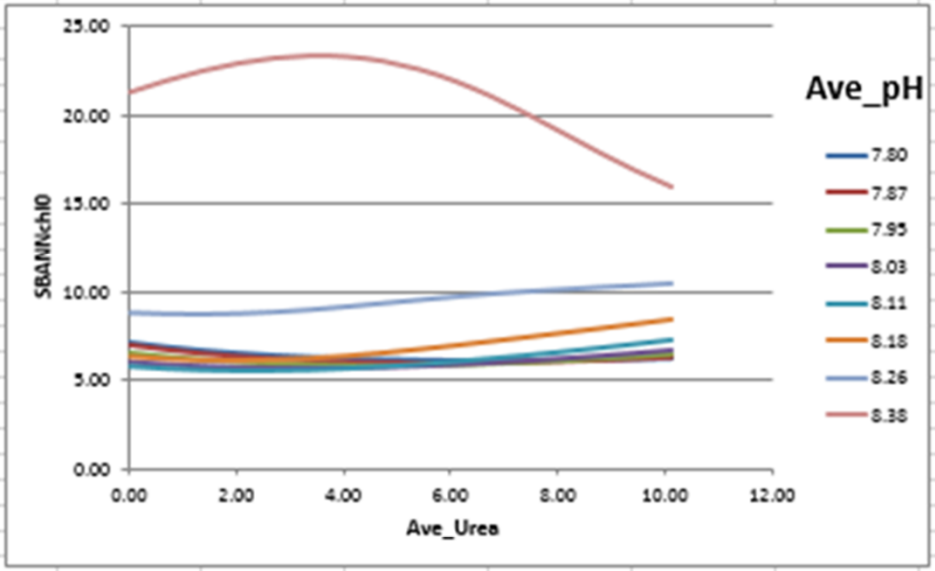
# Global (State Based) versus Local (Means) Sensitivity

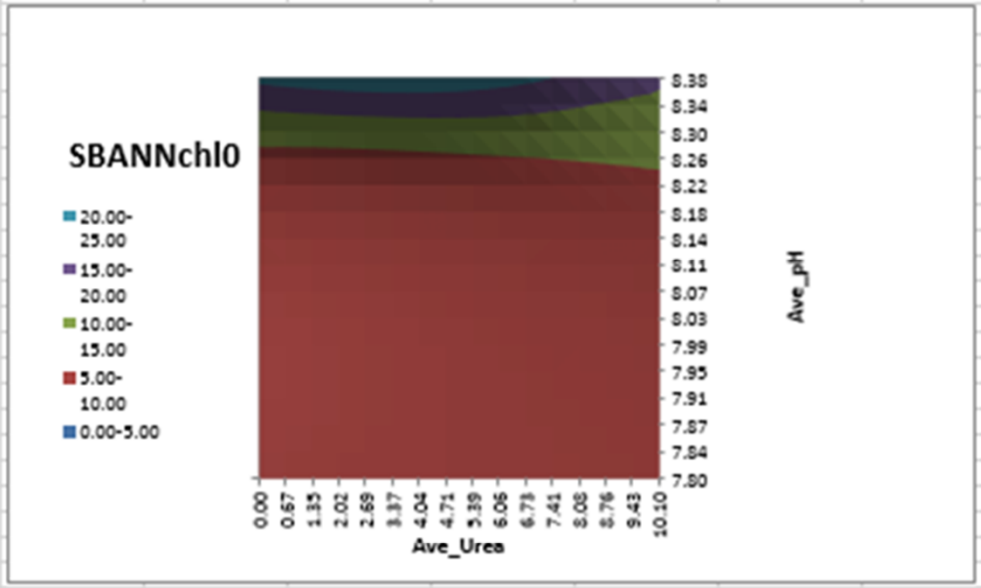


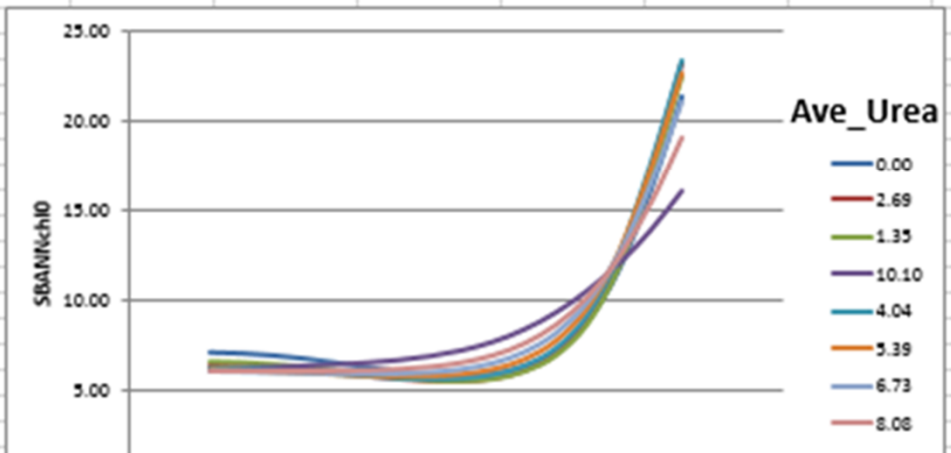
# New Grey Box

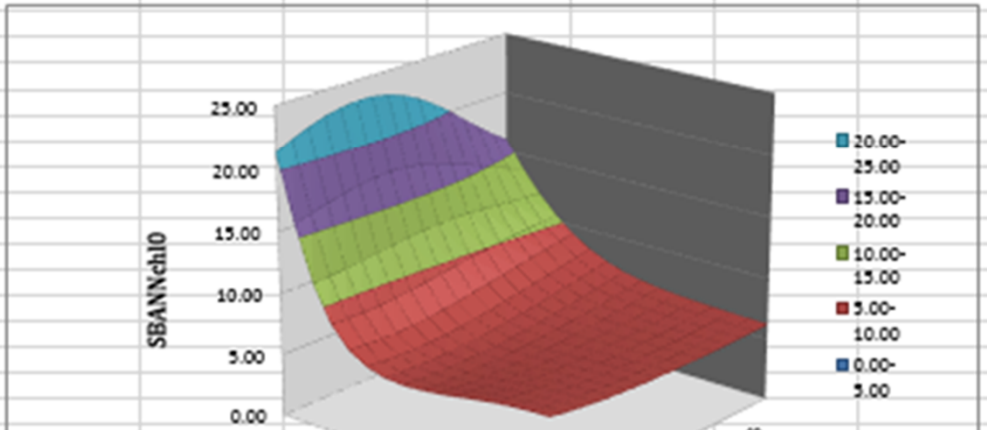
## NN Model Output

## $^{t}$ Iteration

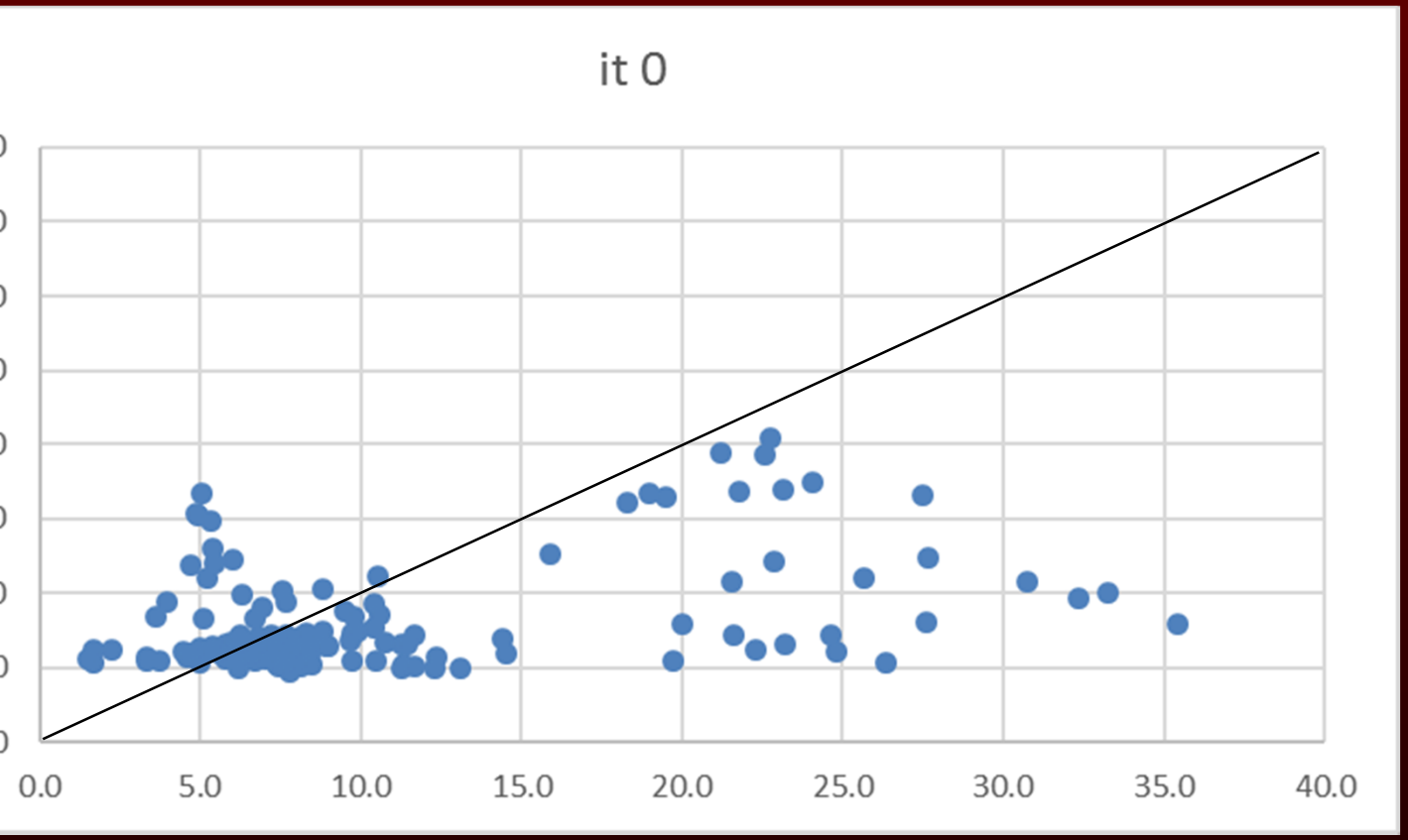


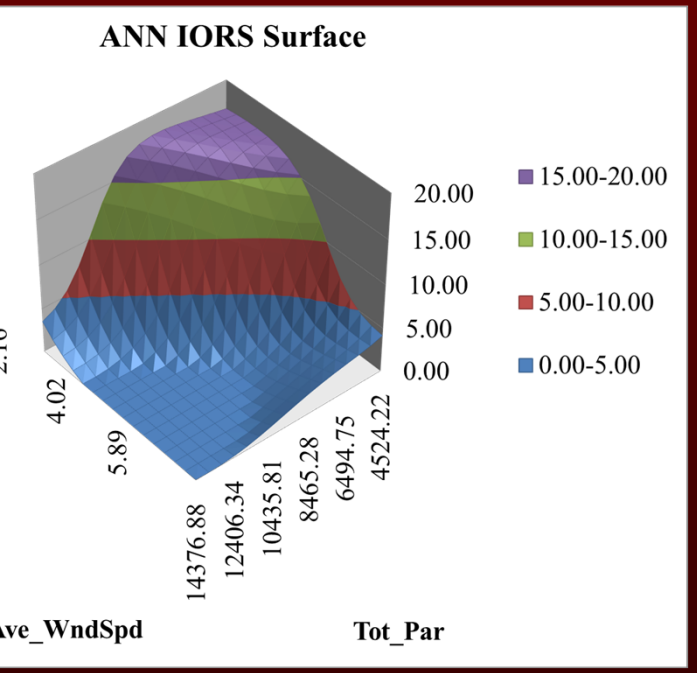



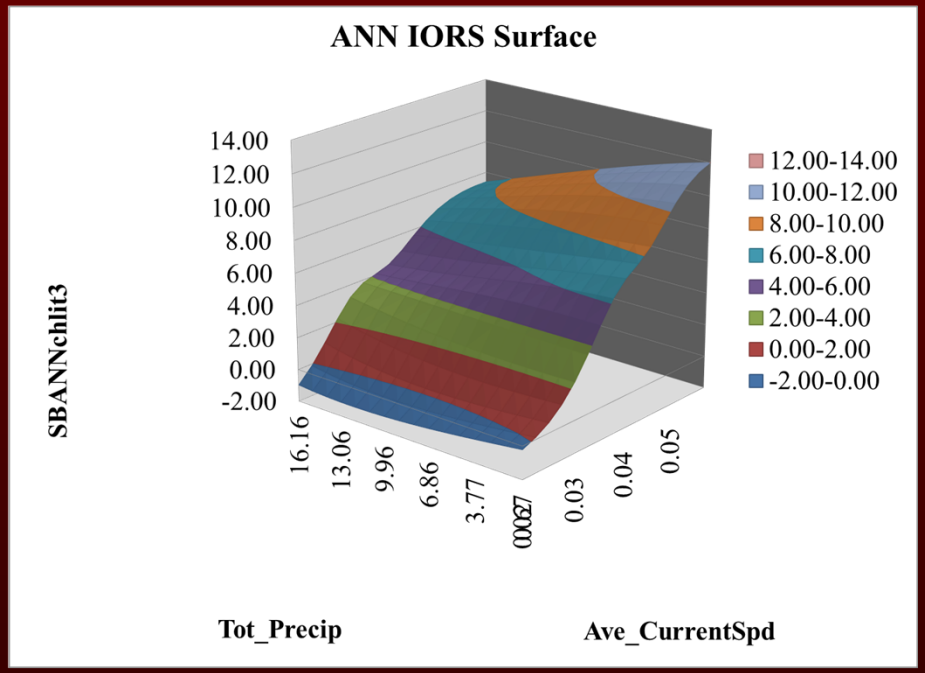
# impact 1st Iteration

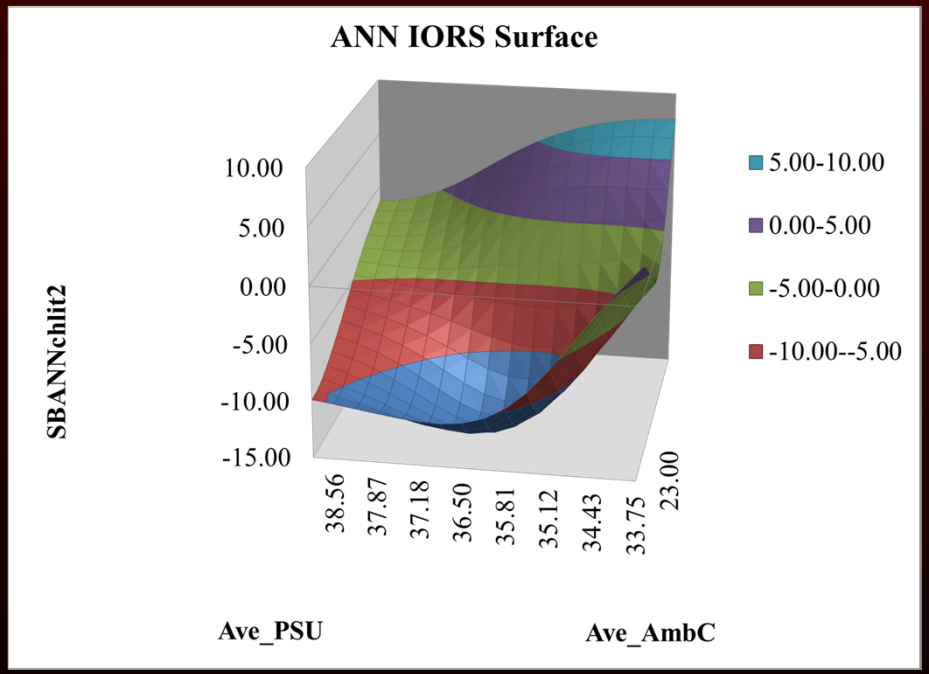


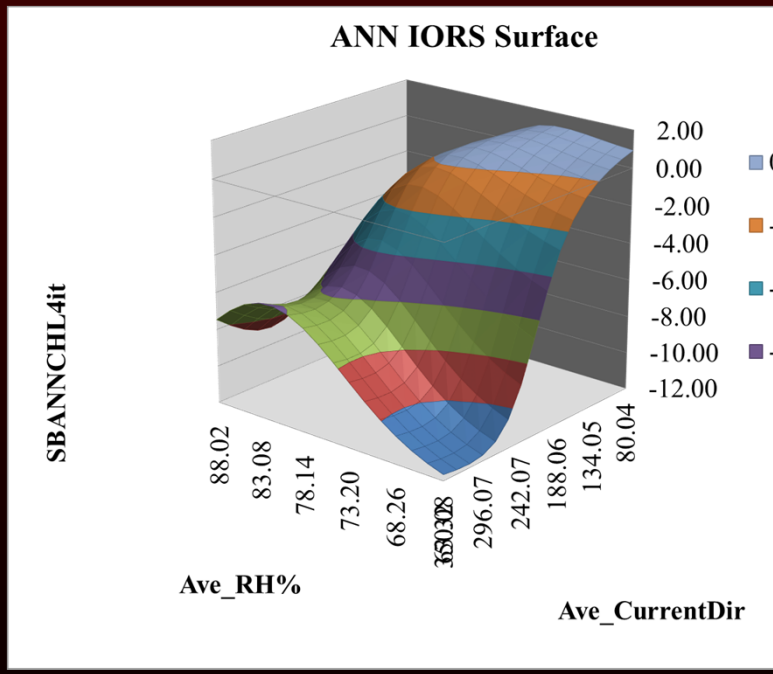

| P | Q | R |
|---|---|---|
| Chl | CHL EQN ph Urea | CHL-EQNw |
| ***6.2*** | 4.9932 | 1.2085 |
| ***6.5*** | 6.6976 | -0.2309 |
| ***15.9*** | 12.6408 | 3.2301 |
| ***5.8*** | 5.5323 | 0.2552 |
| ***5.4*** | 12.9518 | -7.5435 |
| ***1.7*** | 5.3084 | -3.6438 |
| ***6.6*** | 5.7440 | 0.8257 |
| ***22.3*** | 6.2293 | 16.0415 |
| ***12.3*** | 5.6523 | 6.6811 |
| ***19.7*** | 5.4937 | 14.2521 |
| ***7.7*** | 7.1795 | 0.4705 |
| ***14.5*** | 6.0184 | 8.4858 |

# Repeat Iterations

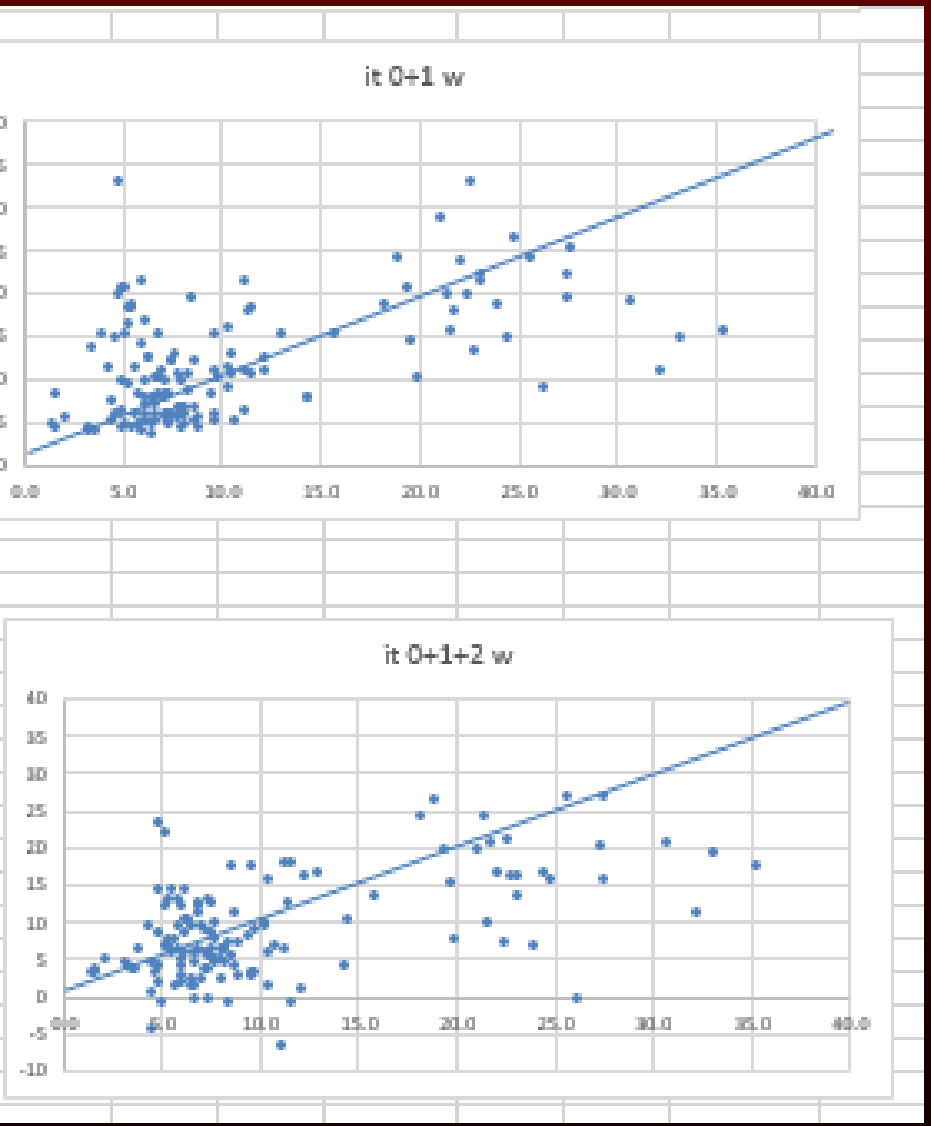


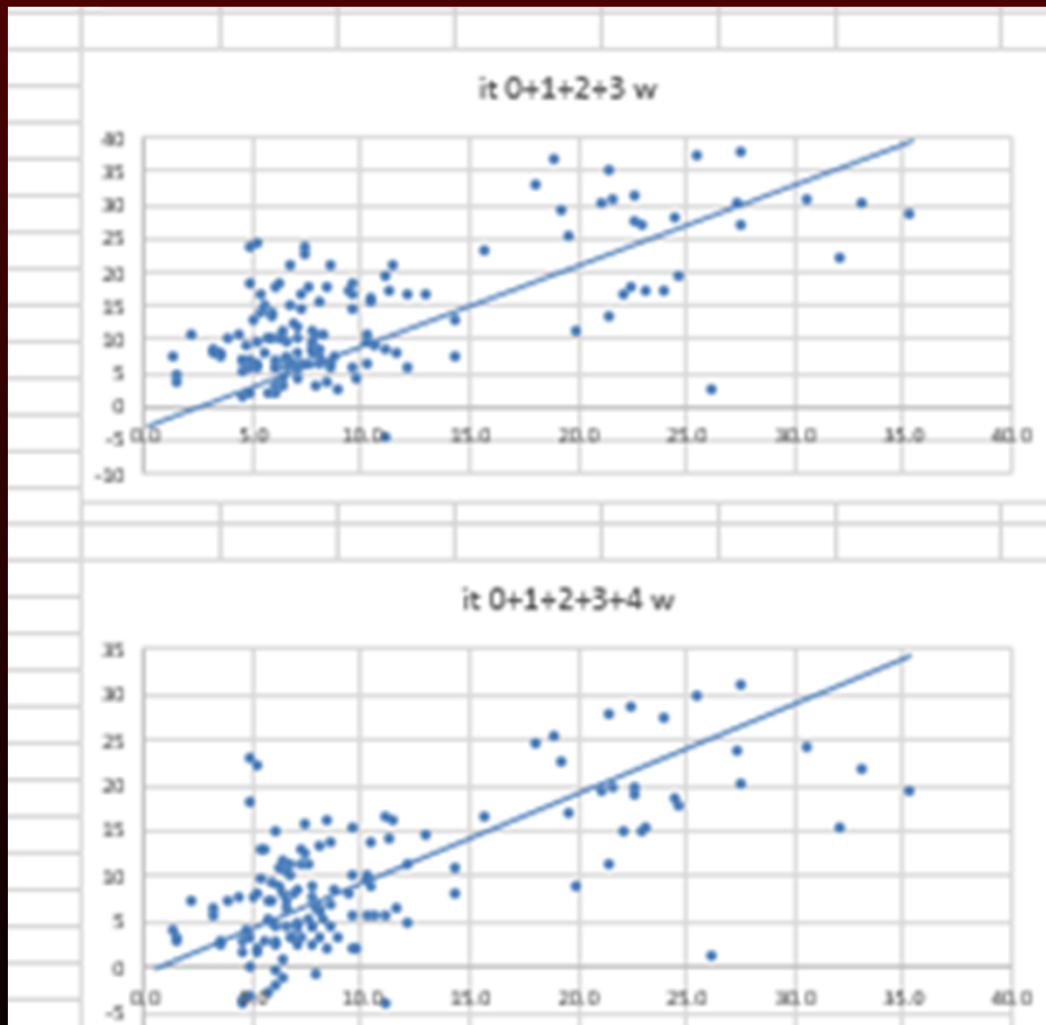



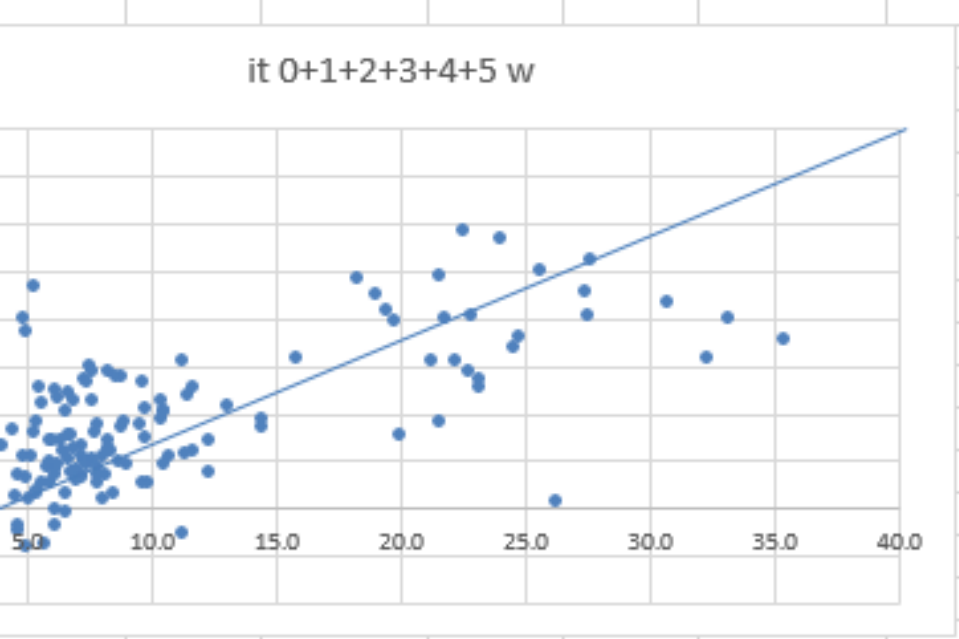
# Equation Model: summed after ea
# Iteration for remaining attributes

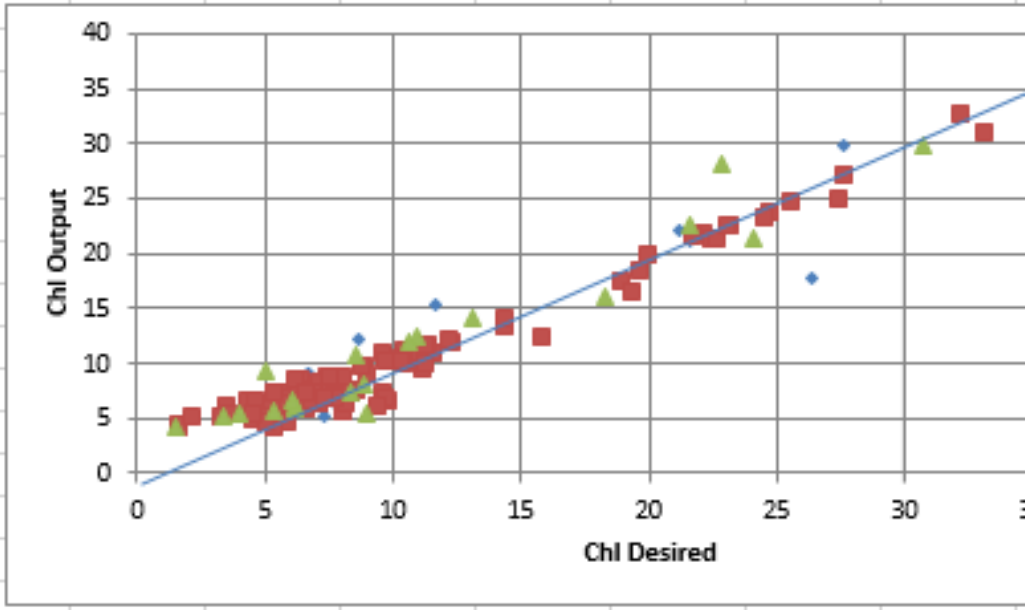


# Grey Box versus ANN

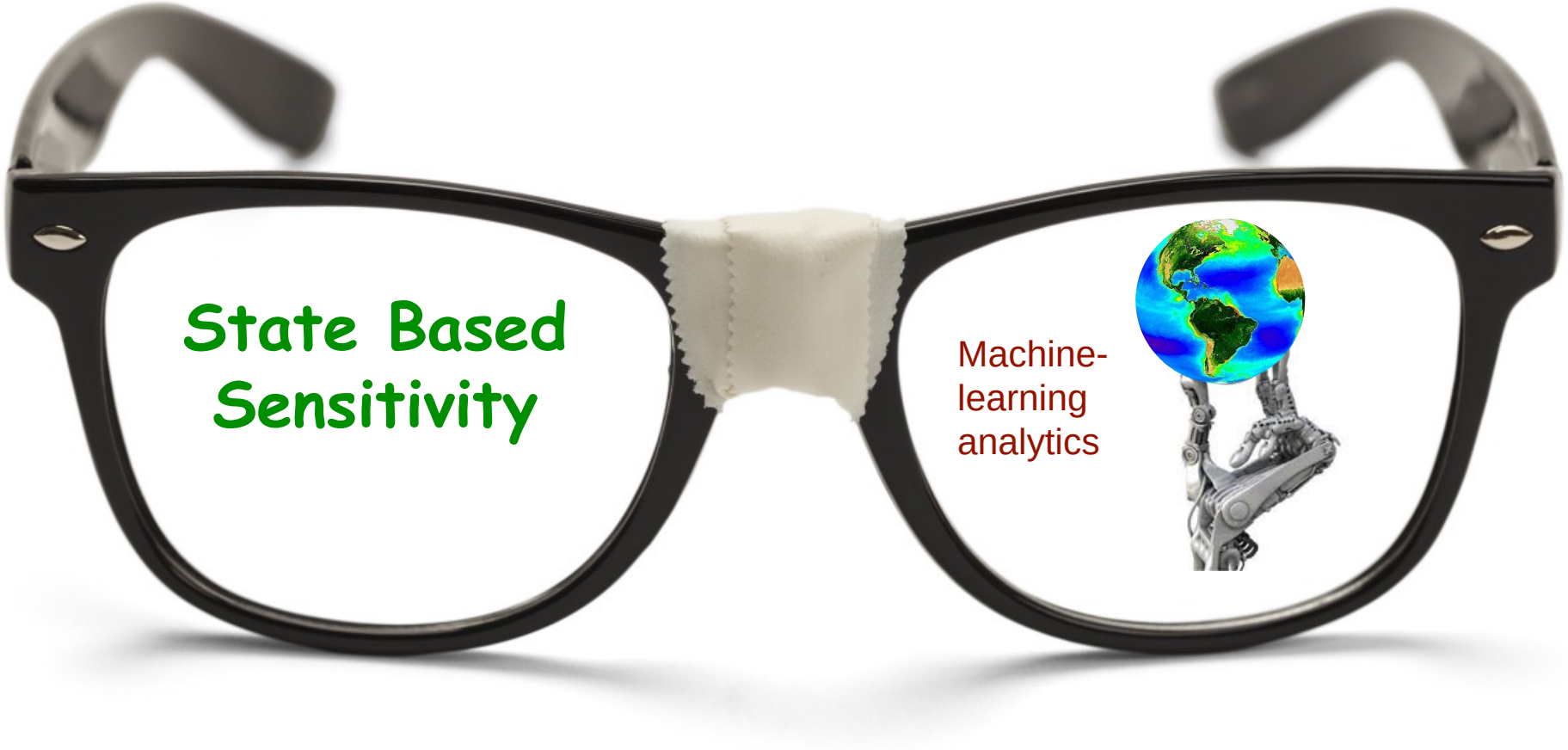
## Grey Box:  Deviations still high versus ANN




ssible Improvements – more detail breakdown i
viations

more effort to develop and investigate new id




achine-learning algorithms capable of autonomously unearthing and
eproducing complex patterns within sizeable data quantities afford gr
otential for fueling ecological hypothesis creation and ‘intelligent’
nowledge derivation