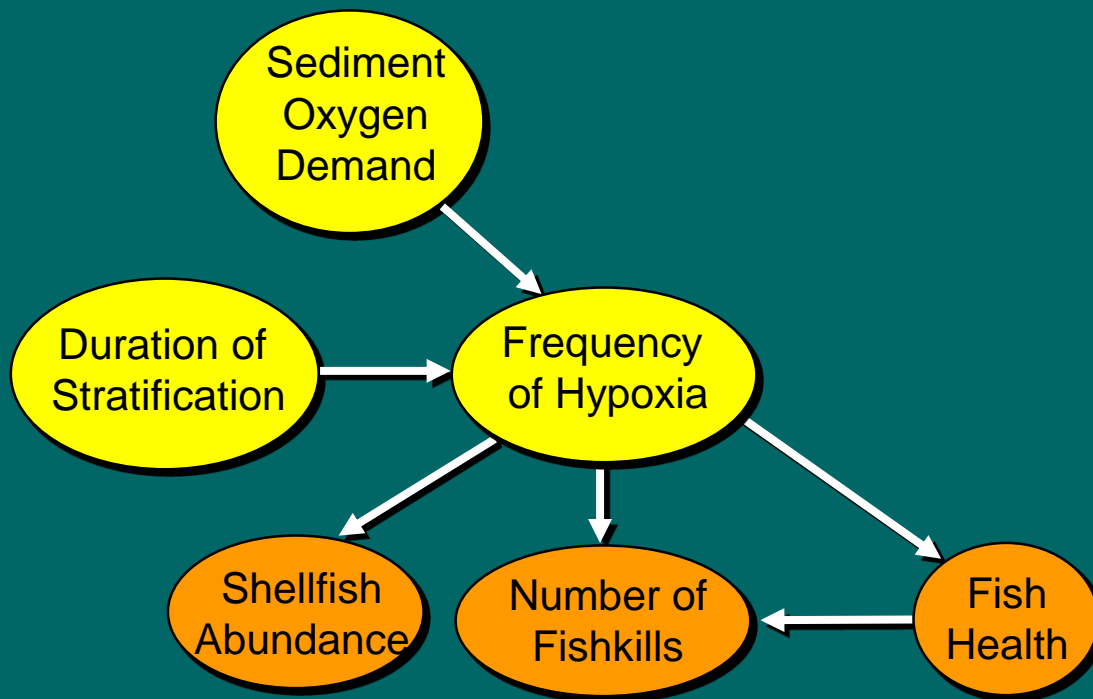


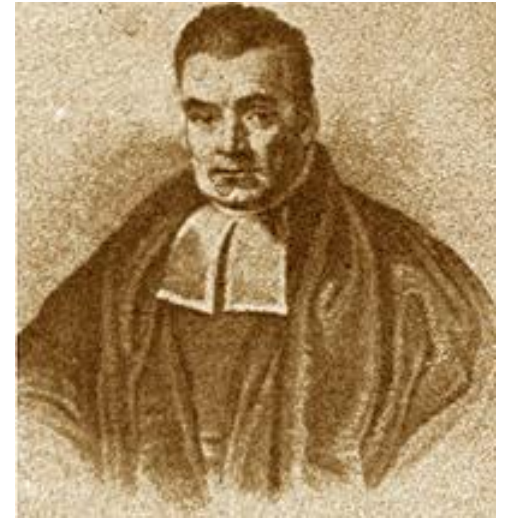
Bayesian Networks for Inference and Decision in Water Quality Assessment

Kenneth H. Reckhow
Duke University

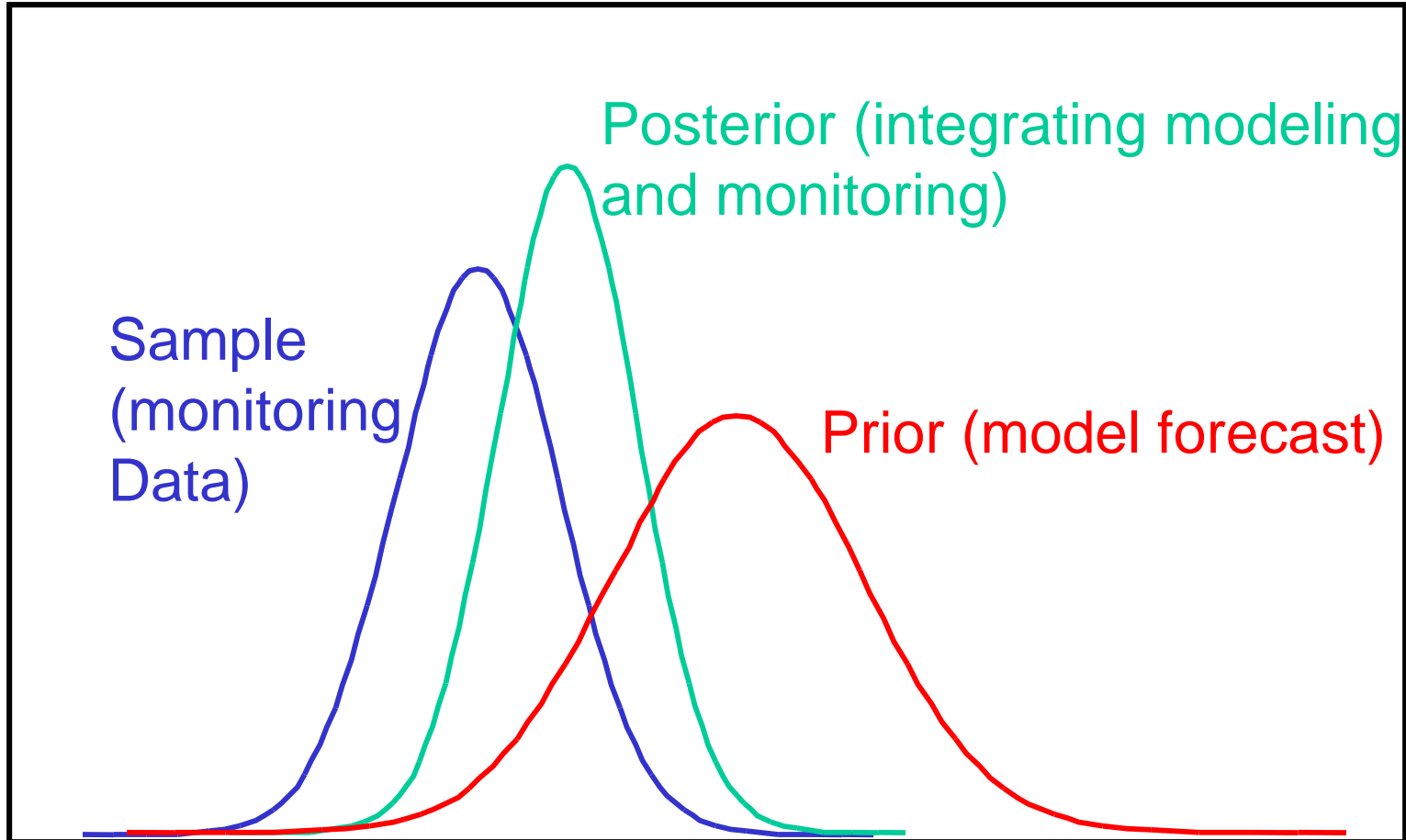


Bayesian Inference

- Why is Bayes' Theorem useful?
 - For combining information
- Why isn't it widely taught and used?
 - influence of Fisher, Neyman & Pearson; convenience of “procedures”
 - requirement for prior knowledge
 - formal use of subjective information (judgment)



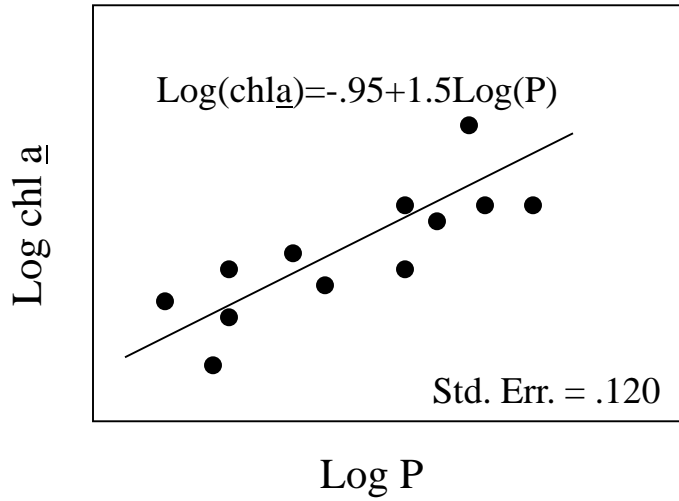
Bayesian Analysis: combining information



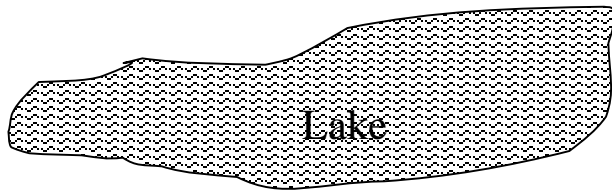
Why do I like, and use, Bayesian analysis?

- For the opportunity to formally incorporate subjective assessments into an analysis (Is science ever truly objective? How often do we use judgment in water quality assessment and modeling?)
- To permit quantitative updating of analyses/modeling as new information is acquired.

Prior

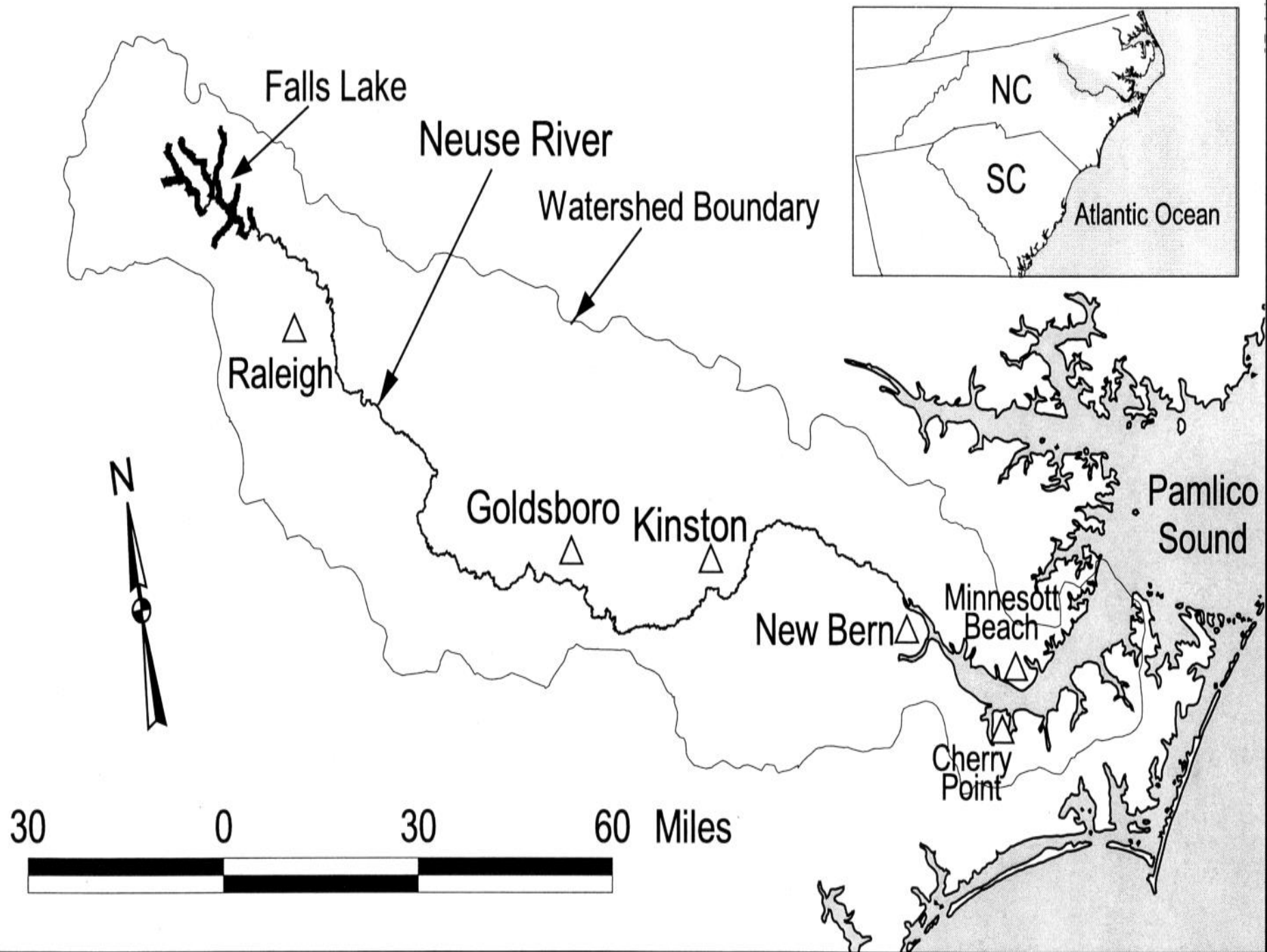


Sample

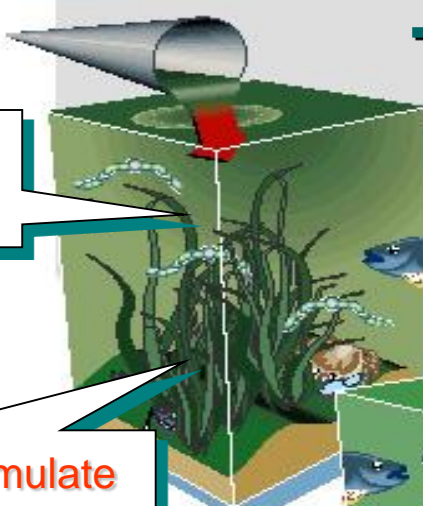


Bayes Theorem

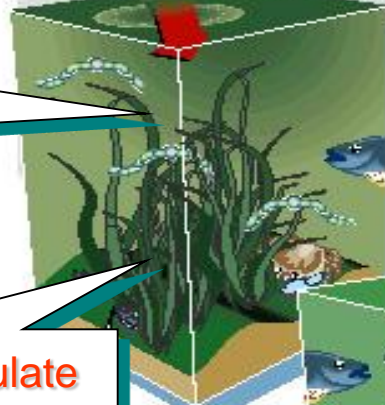
Posterior Probability




The Negative Effects of Excessive Nitrogen in an Estuary




Nitrogen stimulates the growth of algae.



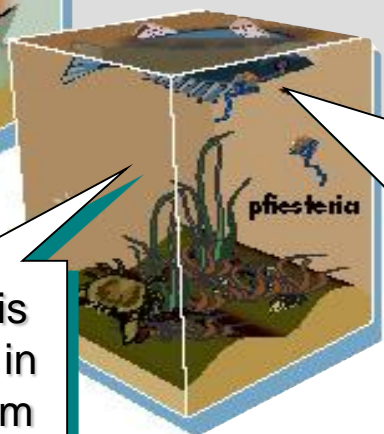
Algae die and accumulate on the bottom where they are consumed by bacteria.



Under calm wind conditions, density stratification occurs.



Oxygen is depleted in the bottom water.

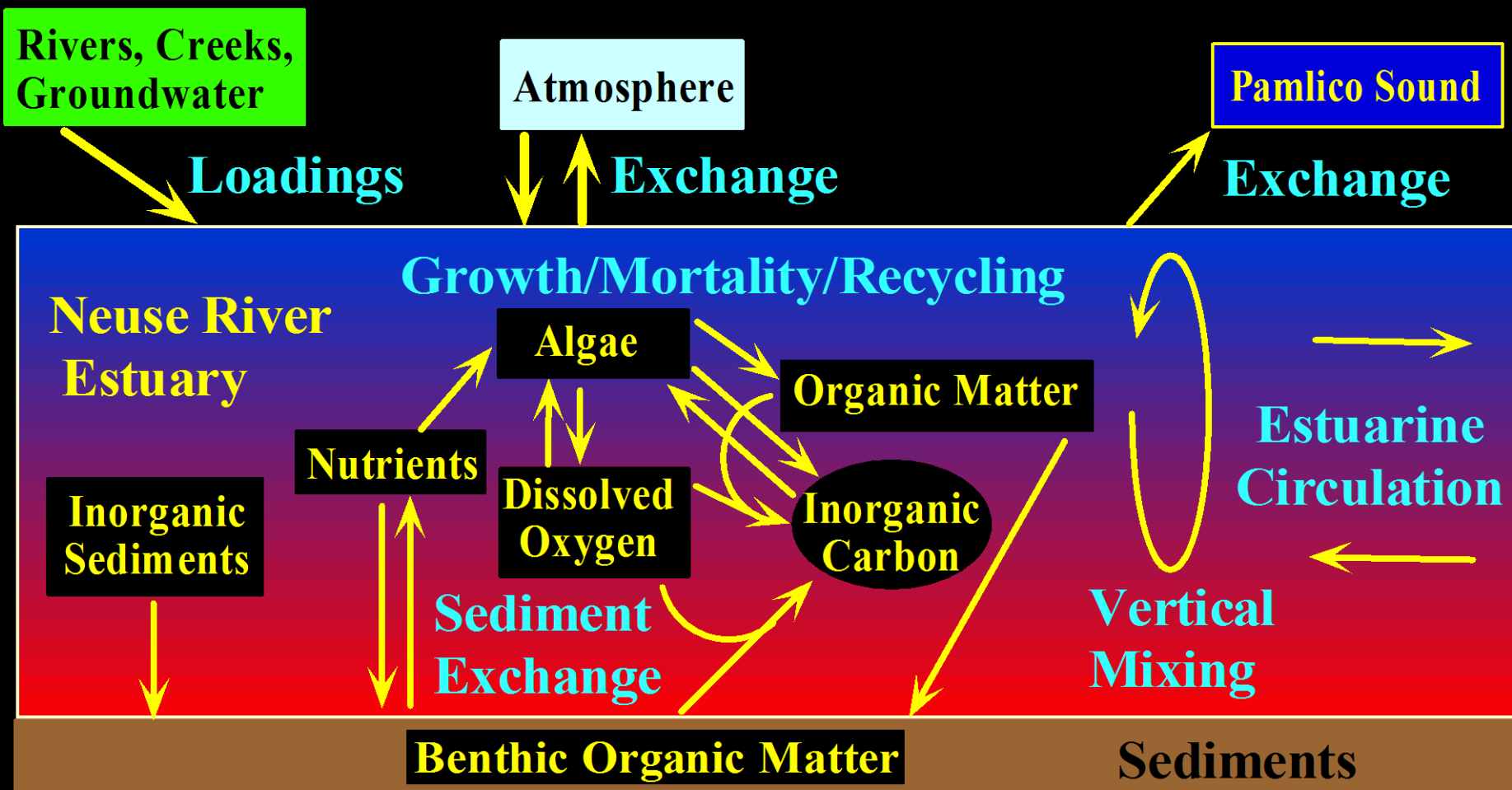


Fish and shellfish may die or become weakened and vulnerable to disease.

Three Different Models were Applied

- CE-QUAL-W2 (NEEM; 2-dimensional)
- EFDC-WASP (3-dimensional)
- A Probability Network Model (Neu-BERN)

Neuse Estuary Eutrophication Model

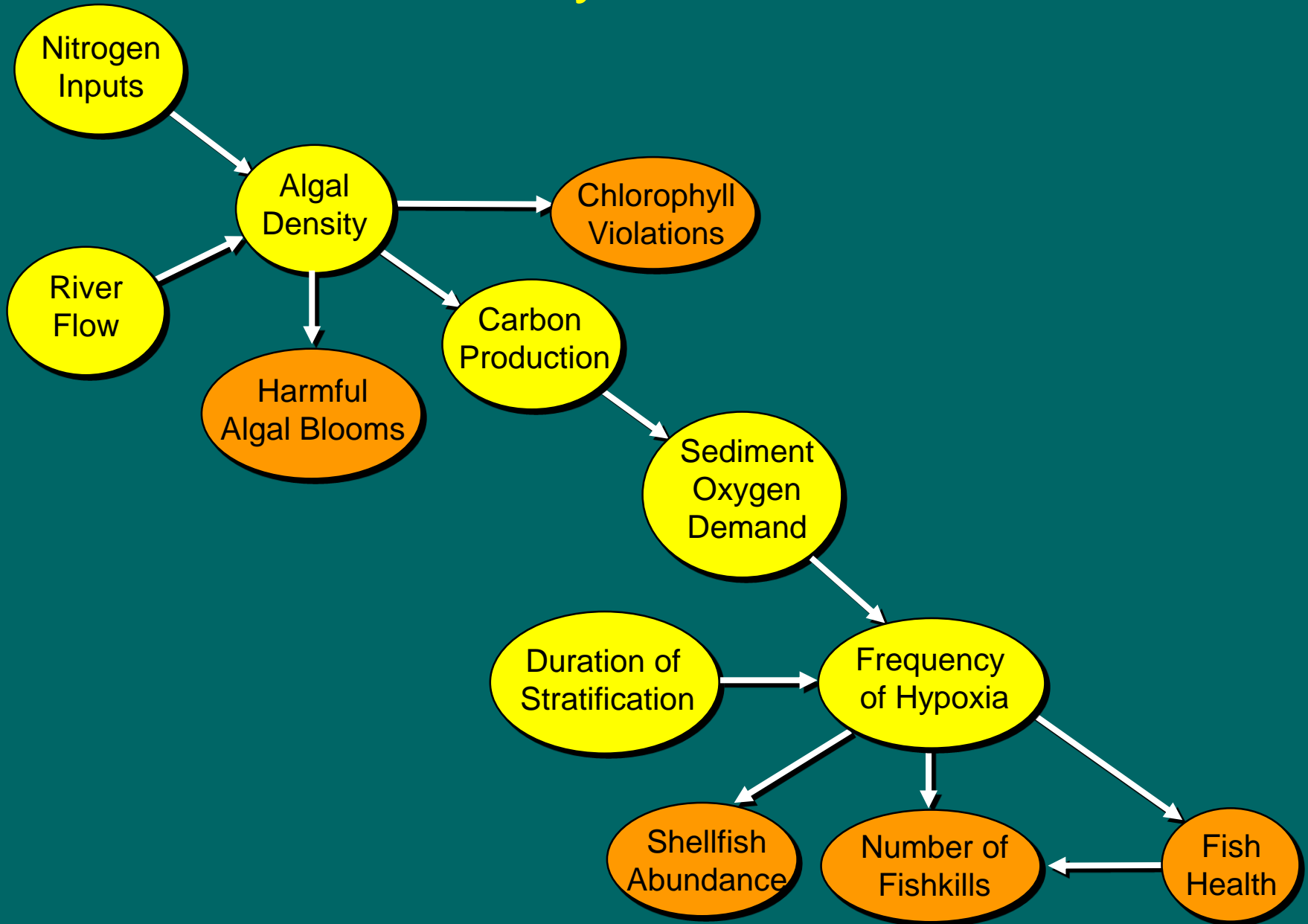


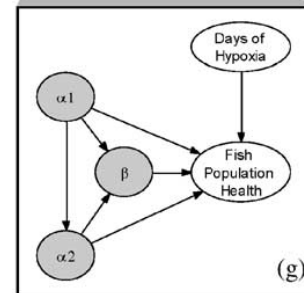
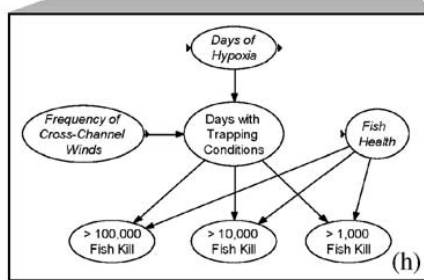
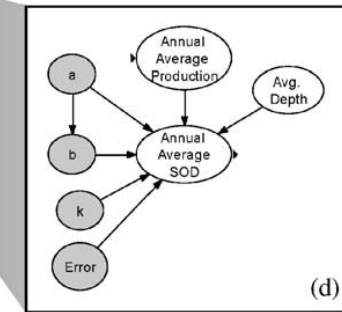
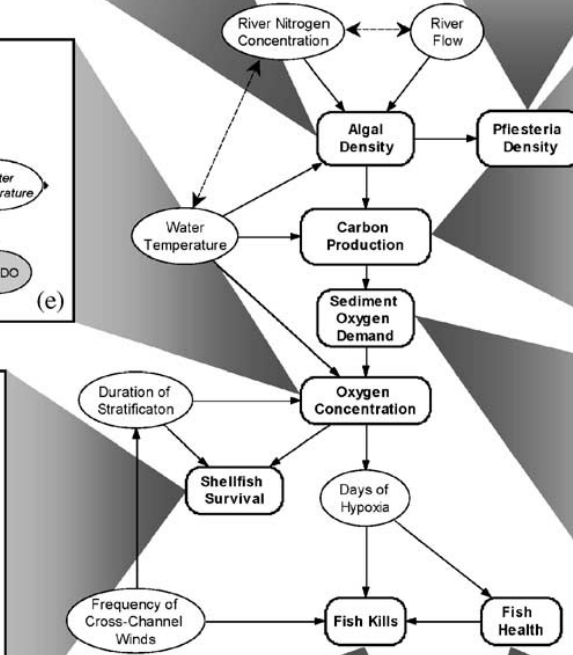
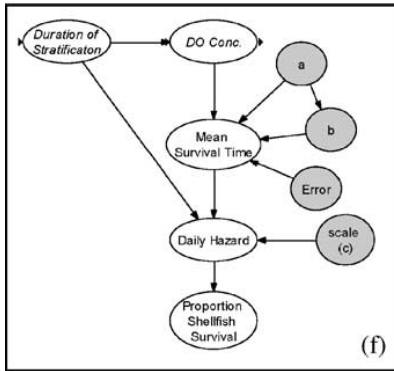
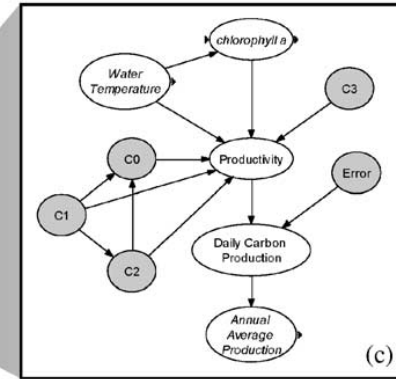
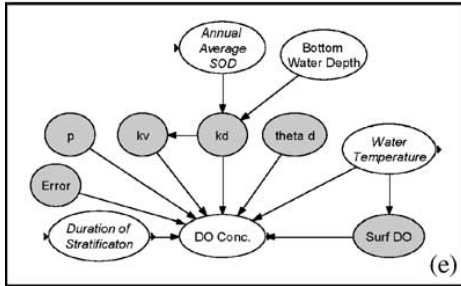
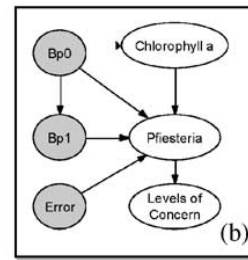
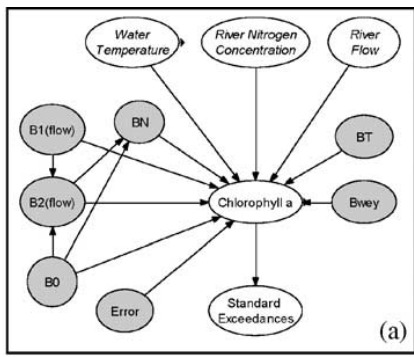
Bayes (Probability) Networks

These are conditional probability models that:

- can be mechanistic, statistical, judgmental
- use probability to express uncertainty
- use Bayes theorem for adaptive implementation updating.

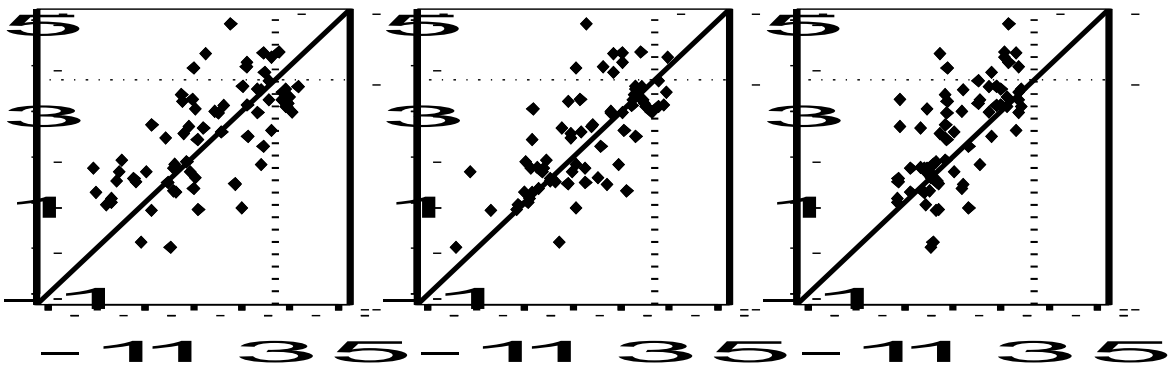
The Prior Bayes Net Model



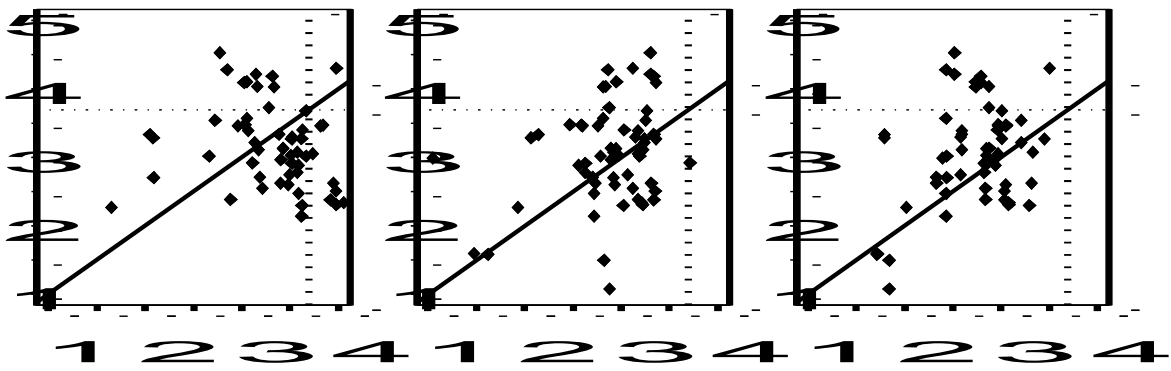


NEWMAN ET AL.

R

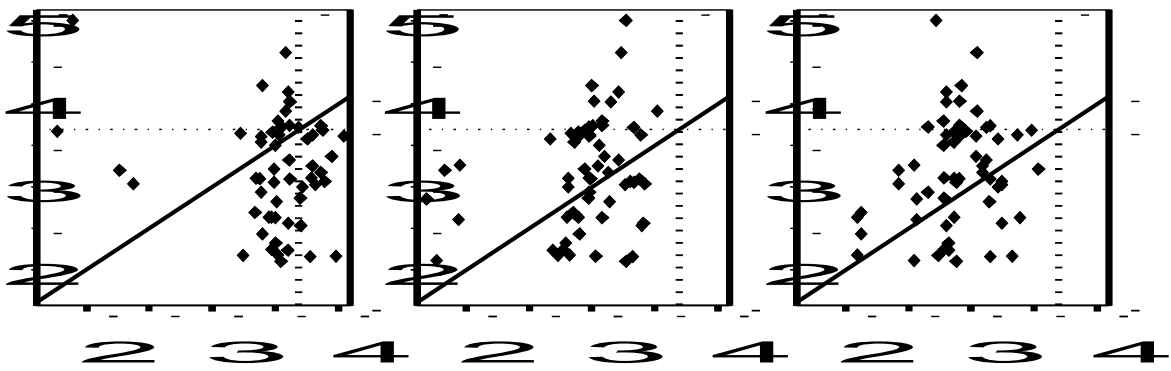


E



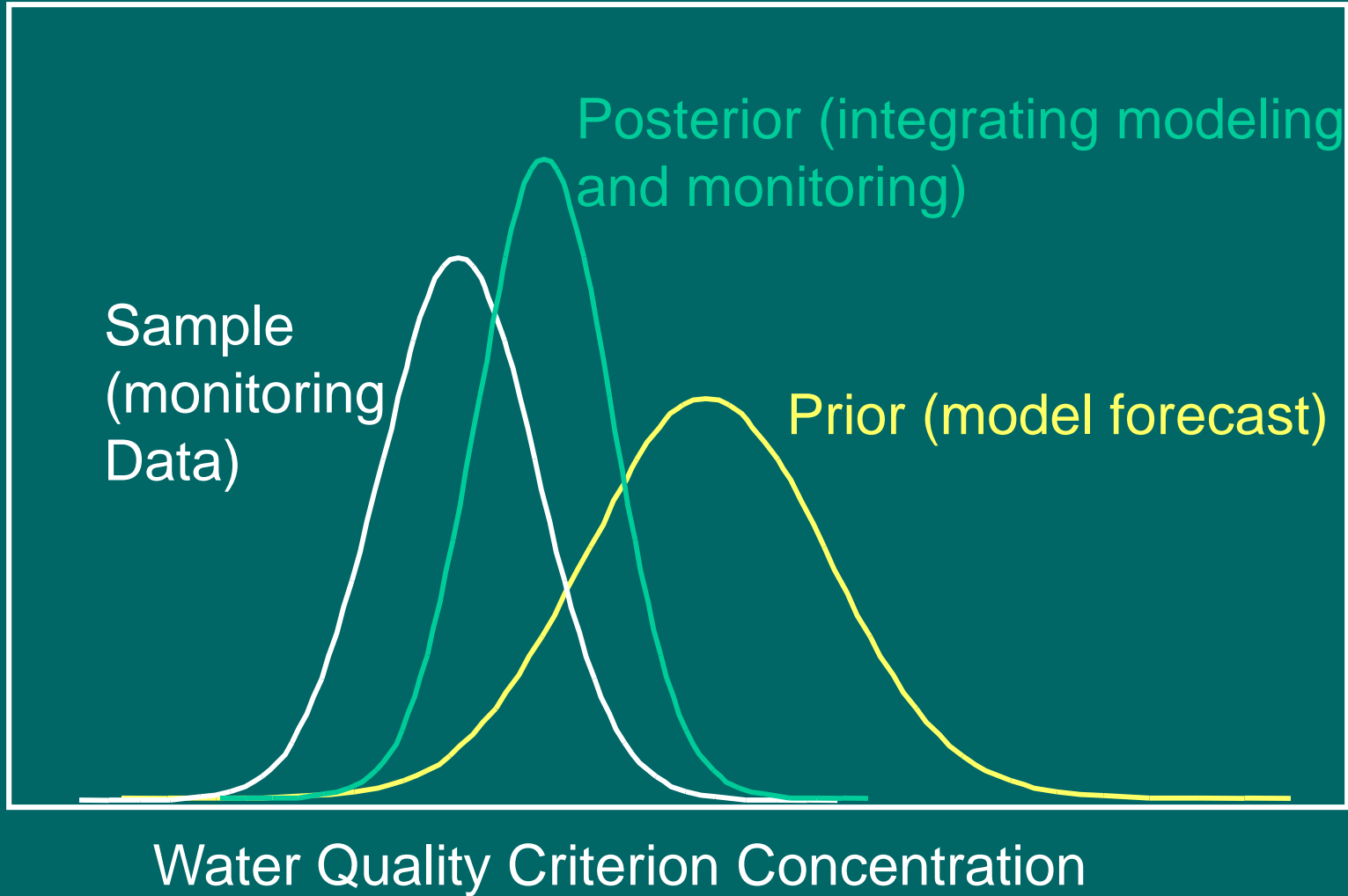
~~CONFIDENTIAL~~

B



Log P r

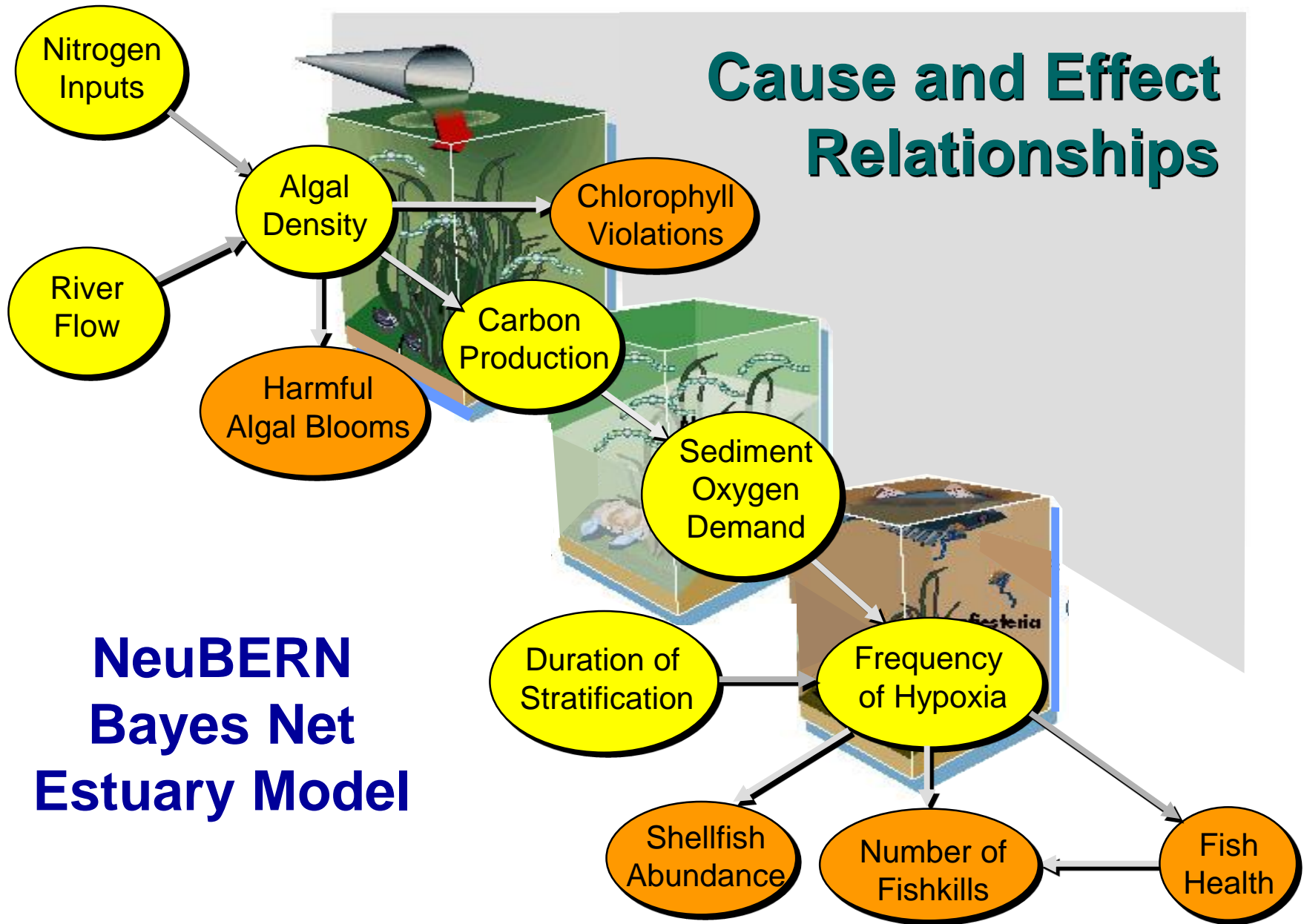
Adaptive Implementation: Bayesian Analysis



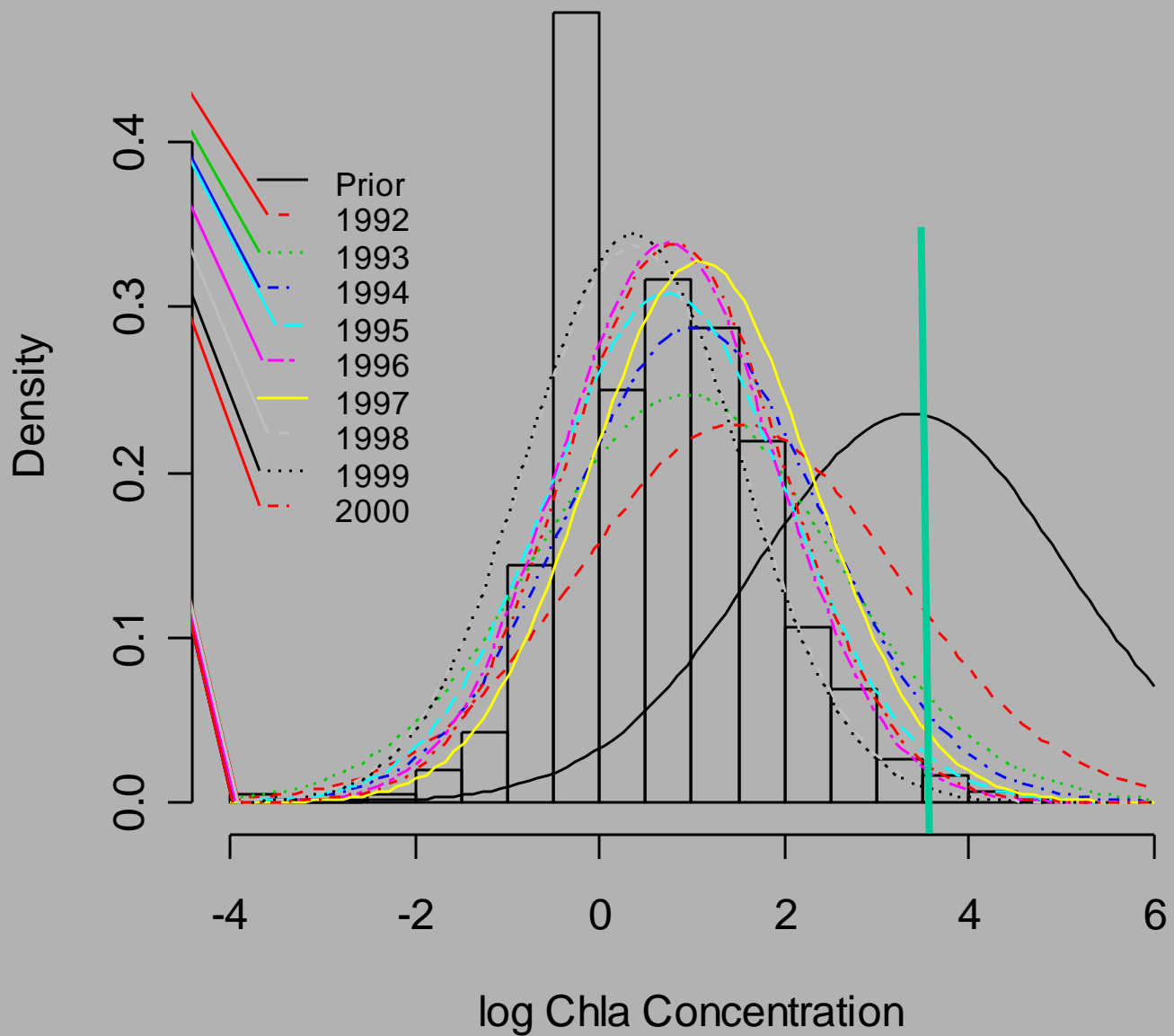
Sequential Updating

- Repeated use of the Bayes theorem
- Prior probability is generated for TN from Bayesian SPARROW
- Current posterior becomes prior when new data are available.

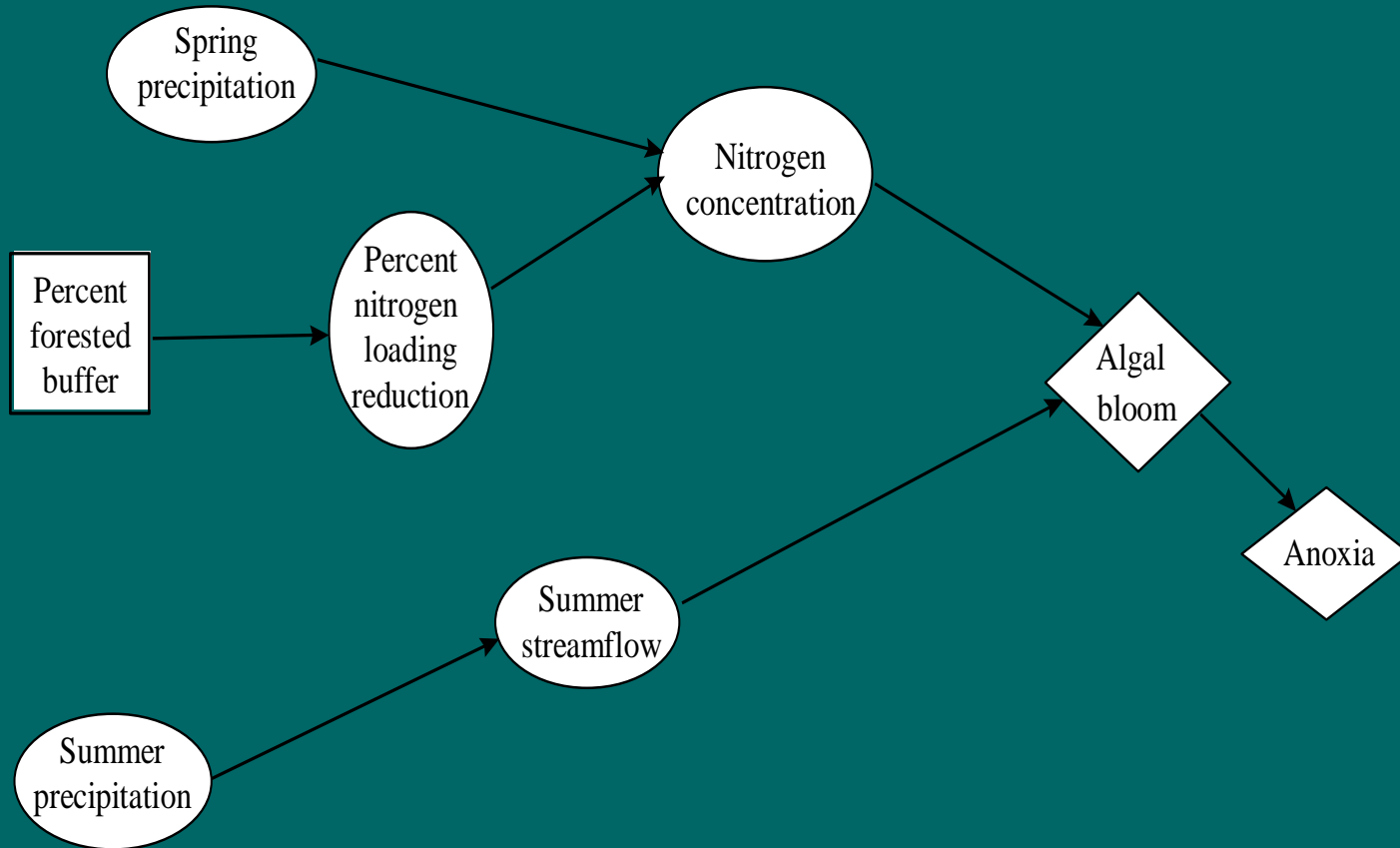
Cause and Effect Relationships



**NeuBERN
Bayes Net
Estuary Model**



Bayes Networks for Water Quality Decision Support



“Given (conditional on) a 30% reduction in nitrogen loading, what is the probability of an algal bloom?”

$p(\text{ALG}/30\% \text{ N-load reduction}) = ?$

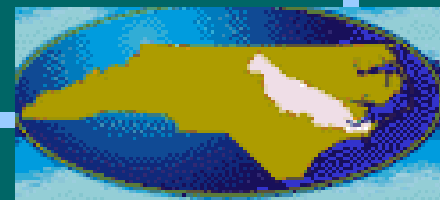


Table 5. Conditional probabilities for summer streamflow.

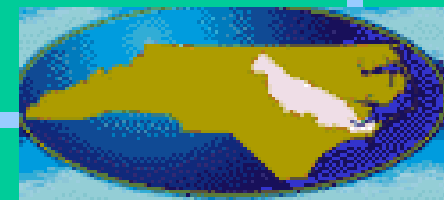
Summer Streamflow	Summer Precip Below Average	Summer Precip Average	Summer Precip Above Average
<500cfs	0.5	0.15	0.15
500-800cfs	0.35	0.7	0.35
>800cfs	0.15	0.15	0.5

Table 6. Conditional probabilities for an algal bloom.

Sum Flow	Total N < 1.4 mg/l			1.4 < Total N < 1.7			Total N > 1.7 mg/l		
	<500cfs	500-800	>800cfs	<500cfs	500-800	>800cfs	<500cfs	500-800	>800cfs
No Bloom	0.1	0.35	0.9	0	0.2	0.8	0	0.1	0.7
Slight	0.35	0.3	0.1	0.3	0.35	0.2	0.1	0.2	0.3
Moderate	0.35	0.25	0	0.4	0.25	0	0.3	0.3	0
Severe	0.2	0.1	0	0.3	0.2	0	0.6	0.4	0

Table 7. Conditional probabilities for anoxia.

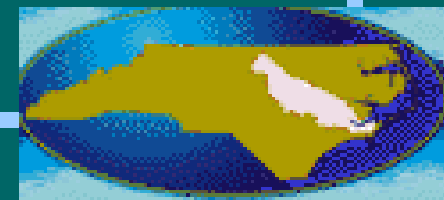
	No Bloom	Slight Bloom	Moderate Bloom	Severe Bloom
Anoxia	0.1	0.2	0.4	0.7
No Anoxia	0.9	0.8	0.6	0.3



Once we have estimated all conditional probabilities, we can calculate any joint or marginal probability. For example, using the probability estimates from Reckhow (*CJF&AS* 1999) for the Neuse, we estimate:

$$p(\textit{severe algal bloom}) = 0.18$$

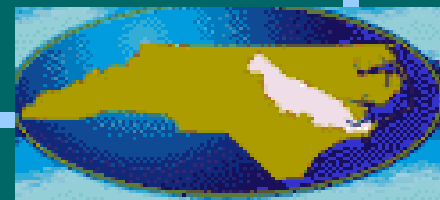
This is useful as it provides a probabilistic expression of knowledge, reflecting uncertainty.



However, using Bayes Theorem with the Bayes net, we can estimate the change in the “baseline” probability, once new (sample) information is acquired. This is the analytic approach for *adaptive implementation*.

Thus, if spring precipitation is observed to be in the “above average” category, the (conditional) probability becomes:

$$p(\textit{severe algal bloom} / \textit{above ave spring precip}) = 0.21$$

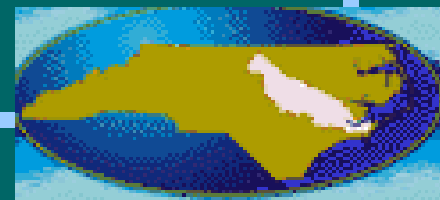


If instead, summer streamflow is extremely low (<500cfs), then:

$$p(\text{severe algal bloom} / \text{summer flow} < 500\text{cfs}) \\ = 0.33$$

Both events together yield:

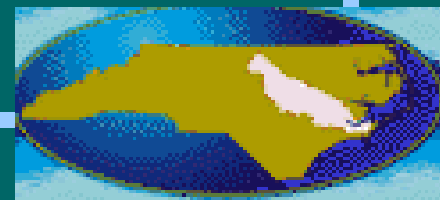
$$p(\text{severe algal bloom} / \text{above ave spring precip, summer flow} < 500\text{cfs}) = 0.37$$



Suppose the nitrogen concentration for next spring is found to be less than 1.4 mg/l. How does this realization of nitrogen change the probability of an algal bloom? This revised probability may be expressed and computed as:

$$p(ALG=severe|evidence\{nitrogen<1.4\})$$

$$= 0.159$$



Merits of The Neuse Estuary Multiple Models Approach

- The mechanistic models are structured to provide a deterministic prediction of compliance that is compatible with individual measurements of water quality criteria.
- The Bayes net model can provide an estimate of the uncertainty in compliance with criteria on a coarser spatial/temporal scale, as well as provide an estimate of the uncertainty in expended improvements in ecological endpoints of direct meaning to stakeholders.

The Bayes net probabilistic approach provides the opportunity to:

- Determine confidence intervals associated with achieving a given set of goals under a particular scenario.
- Link nutrient-reductions to ecologically meaningful endpoints in the estuary.
- Rapidly evaluate the impact of load reduction scenarios.
- Link modeling uncertainties with gaps in our understanding of the system and in monitoring programs.
- Combine the model predictions of Neuse Estuary status/improvements with assessments based on water quality measurements (data).

For the Neuse, the Bayes Network (BN) Model Complemented the Two Mechanistic Models.

- While the BN could not provide the space/time resolution of a detailed mechanistic model to evaluate dissolved oxygen and other small-scale outcomes, it is probabilistic and highly flexible in structure.
- The probabilistic nature of a BN means that prediction uncertainty could be estimated; also, the BN flexibility allows extension of the model for probabilistic prediction of endpoints concerning fish and shellfish.