MODEL EVALUATION

DOES YOUR MODEL MAKE SENSE TO SOMEONE ELSE?

OVERVIEW

What is evaluation?

Why is it useful?

How do you evaluate environmental models?

Dealing with uncertainty.

Practical Evaluation Techniques

Pitfalls

EVALUATION

- **Process of rigorously assessing model components, structure, parameter values, assumptions, but not scenario results**
- **Commonly called *model validation***
  - Models represent a point of view of a system. Validation probably not the best term b/c it indicates a model can be true. Are opinions true?
  - Evaluation captures the essence of validation without connoting that the model is true
- **Process needs to ensure scientific defensibility and transparency**

EVALUATION CON'T.

- **Is the model useful for its intended purpose?**
  - Given the assumptions, structure, and assumptions, can the model be used for what the developers intended.

- **What are its limits and weaknesses?**
  - Under what conditions does the model break?
  - Should you try to break it?

- **Is it re-creatable?**

EVALUATION IS OFTEN NOT RIGOROUS

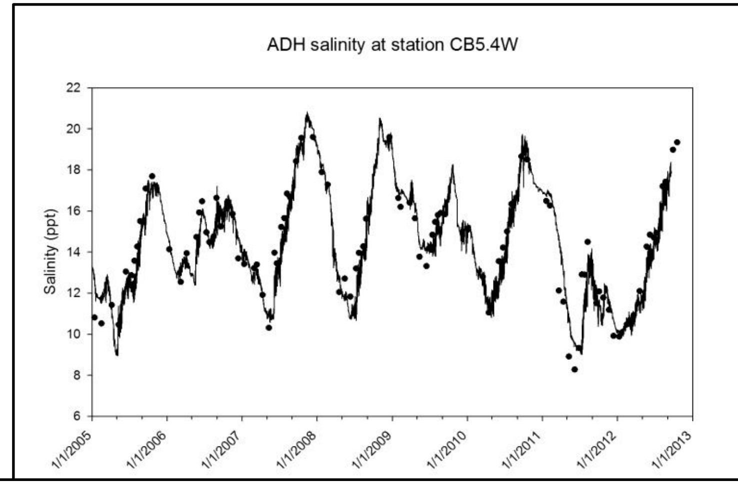
Detailed evaluation is rare

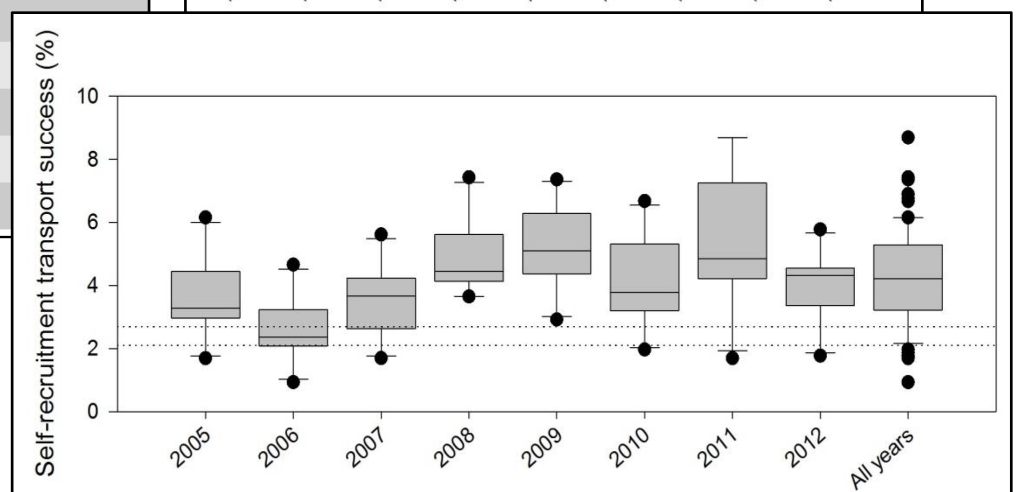
- Overly rely on software
- Don’t have time
- Aren’t concerned with recreatability
- Discipline hasn’t required it
    - Small field & modeling was esoteric
    - But most agencies rely on models now
    - Increased need for scientifically-defensible and detailed documentation
    - TRACE (Transparent and comprehensive model evaluation)

PROBLEMS WITH EVALUATION

Different disciplines have different expectations of model performance

	Hydrological Modeling	Ecological Modeling
Main Focus	Water	Persistence of species
Sub focus	Chemistry of water	Dynamic relationships
Environmental Hierarchy Target	Landscape	Molecular (Genetics) Organ systems / tissues Individuals Populations Landscape Ecosystem Biome
First principles?	Sometimes	Never
Model confidence	High	Low
Science/Art	99/1	25/75
Knowledge of dynamics	High	Low

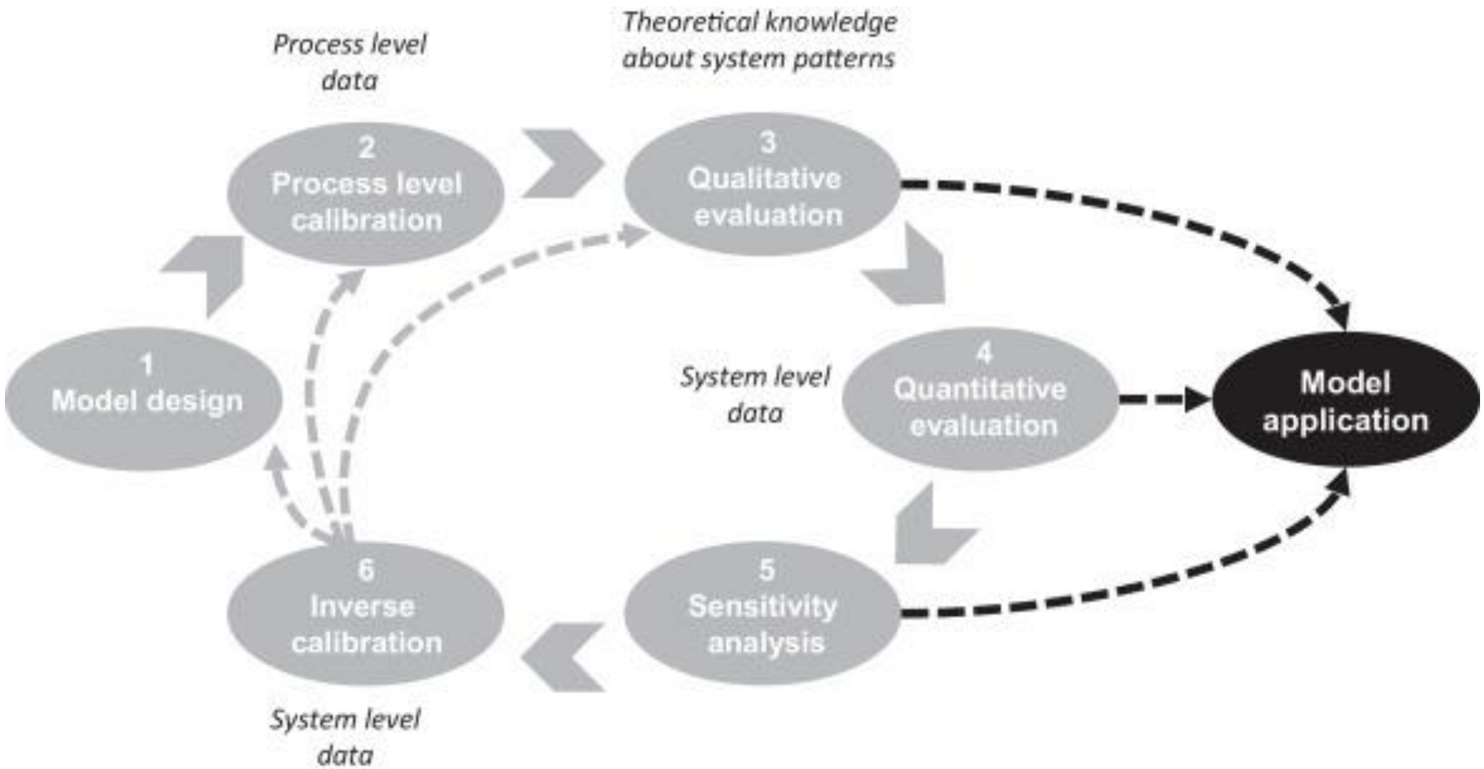




PROBLEMS WITH EVALUATION

Biggest issue is failure to document entire evaluation process

- Iterative approach
- Evaluation occurs throughout the modeling process, but is rarely documented thoroughly. Need to document each cycle.



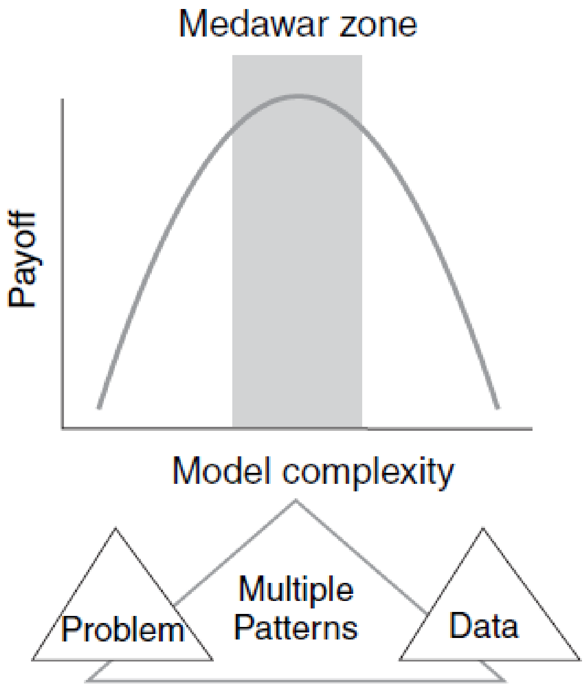
PATTERN-ORIENTED MODELING

Technique for analyzing ecological models with emergent properties

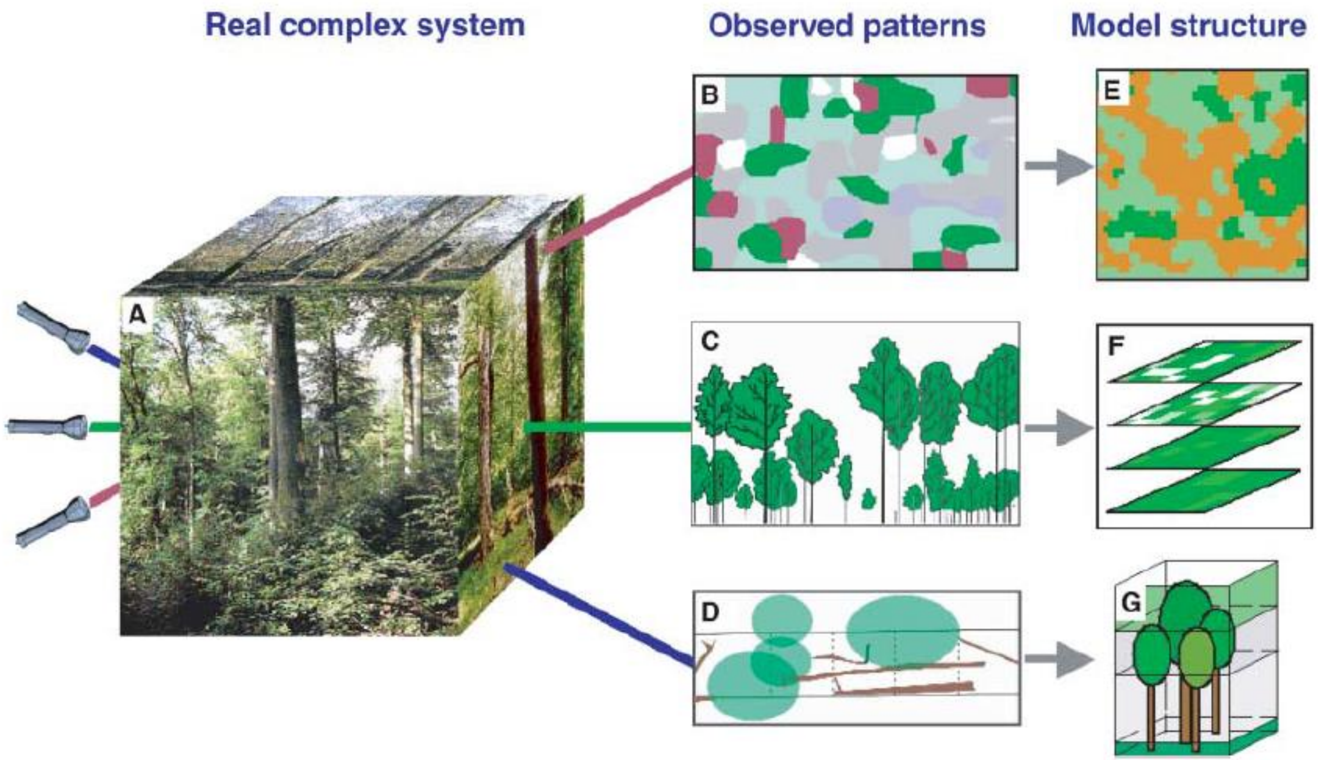
- Comparison of multiple, simulated system-level patterns to multiple observed and/or hypothesized patterns from real system
- Moves away from p-values and statistics towards weight of evidence and critical thinking

Choosing patterns

- Pattern is anything beyond random variation
- Can be qualitative or quantitative
- Need to be diverse & not correlated



PATTERN-ORIENTED MODELING



Mosaic of development stages

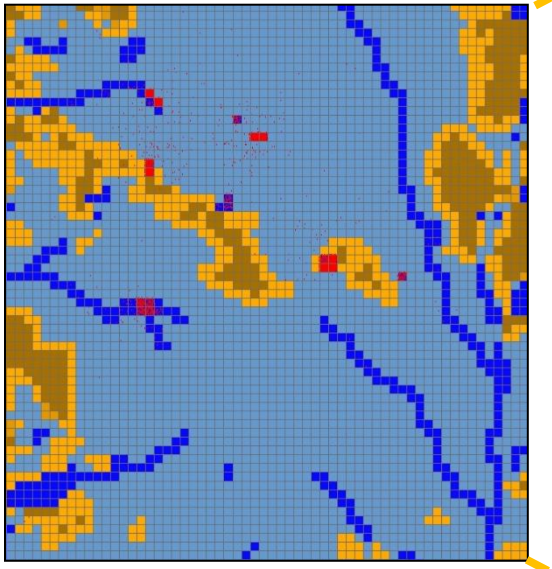
Vertical patterns of tree size

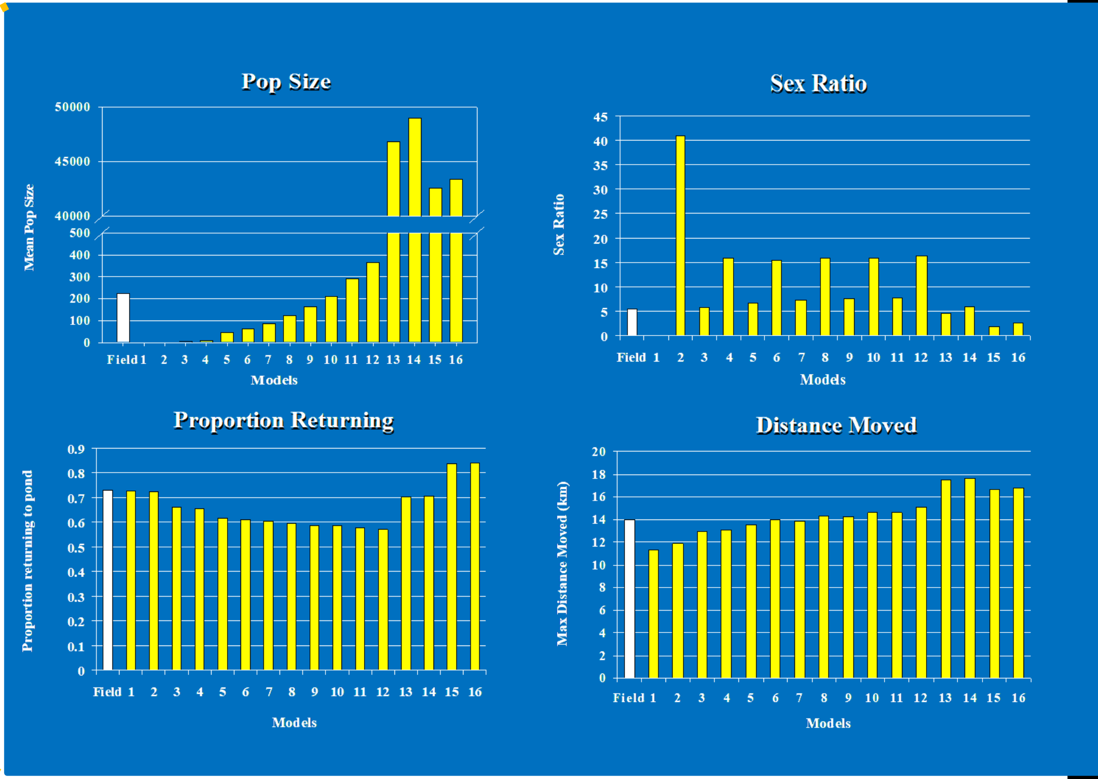
Fallen wood

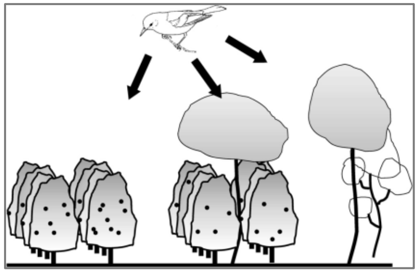
Grimm et al., 2005

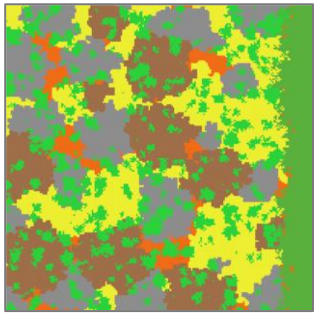
PATTERN-ORIENTED MODELING

What level of juvenile survivorship leads to sustainable toad populations?









Patterns for bird foraging and pest control of the coffee borer beetle (CBB)

1) Infestation rates increase over coffee season
2) In absence of birds, infestation rates are higher in low shade
3) Bird densities are higher in high-shade than low-shade
4) Positive relationship infestation rates in exclosures and % by which birds reduce infestation rate
5) Birds respond quickly to local irruptions on their normal food supply
6) Change in bird density within season related to food supply
7) Weak relationship between vegetation characteristics and bird density at intermediate scales
8) Bird movement distances (hourly) follow a log normal distribution
9) Bird consumption of CBB is episodic

What patterns tell us about structure

- Many of the patterns concern changes over time, so the model needs to be dynamic.
- Many patterns concern differences among habitat types, so the model must be spatial with resolution fine enough to represent the small, irregular patches of low- and high-shade coffee, forest, and forest fragments that characterize Jamaican coffee farms.
- Foraging habitat selection by birds is a key process driving the patterns, so it needs to be included in the model explicitly. Because this individual behavior is important, individual-based modeling is a natural approach. Birds move among the small habitat patches rapidly, so the model needs a very short time step.
- Many patterns involve CBB infestation rate, so the model must represent this rate and how it is affected by bird feeding. Infestation rate is not a direct measure of CBB population but is the measure used to quantify CBB in the field and calculate its economic impact.
- Patterns 1, 2, and 4 are results of exclosure experiments, so the model needs to represent exclosures.
- Reproducing pattern 5 requires the ability to simulate irruptions of arthropod prey.
- Pattern 6 includes increases and decreases in bird populations. The model does not include the reproductive season, so reproducing population increases would require including immigration. Reproducing population decreases requires including mortality or emigration.

STEPS IN EVALUATING ENVIRONMENTAL MODELS

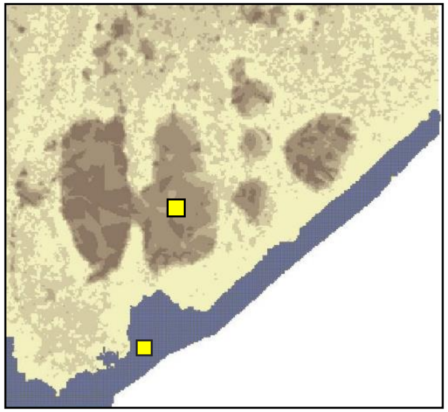
- Assess reasonableness of model structure
- Assess functional relationships & verify code
- Evaluate model behavior vs expected patterns
- Does model correspond well to data from real system? Does it matter?
- Analyze uncertainty

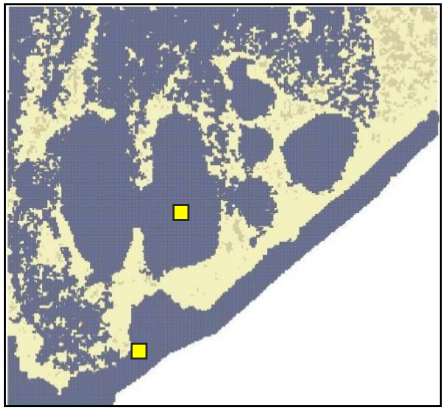
ASSESS REASONABLENESS OF MODEL STRUCTURE

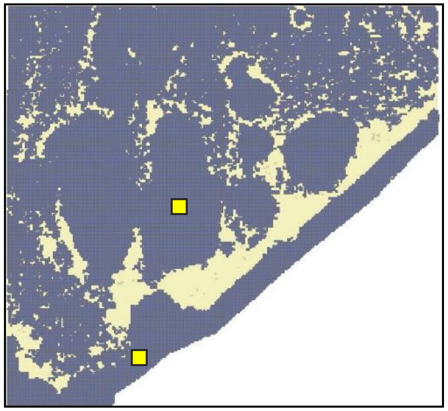
- **Does the structure make sense?**
  - Absolutely required for explanatory models (most environmental models), but not really for correlative models, which are less focused on capturing relationships b/w variables
- **Somewhat subjective**
  - Requires a priori hypotheses to test functional relationships
  - There are always simpler and more complex models

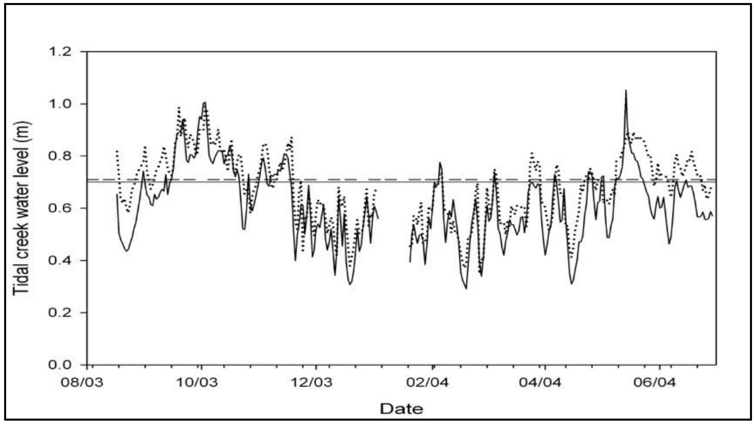
EXAMPLE OF MODEL STRUCTURE

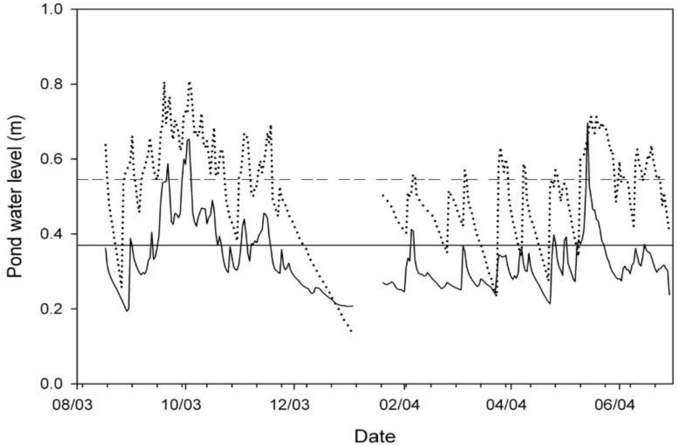
SLOSH: the does water run uphill model





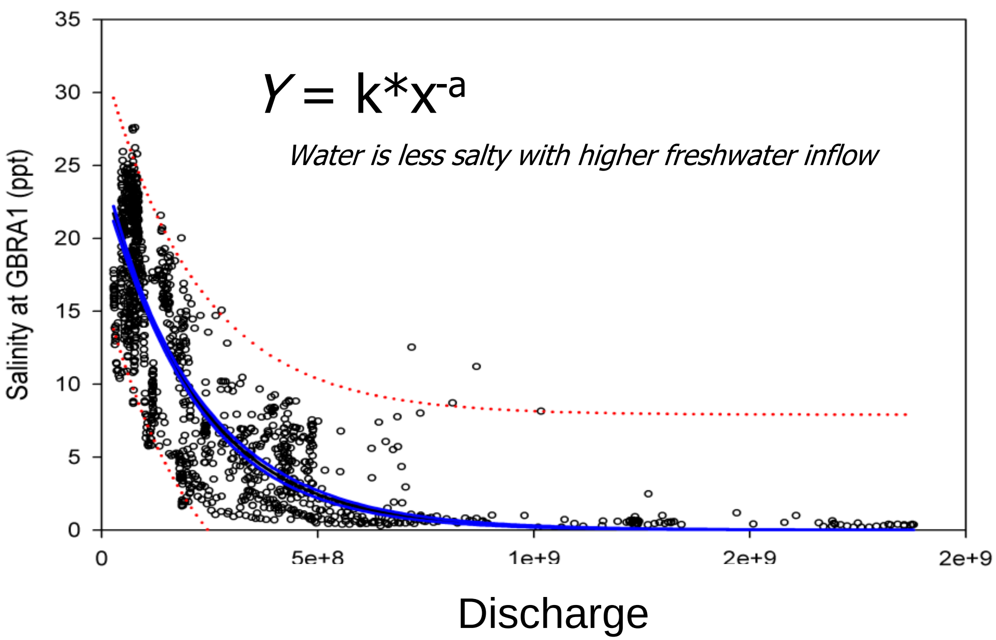


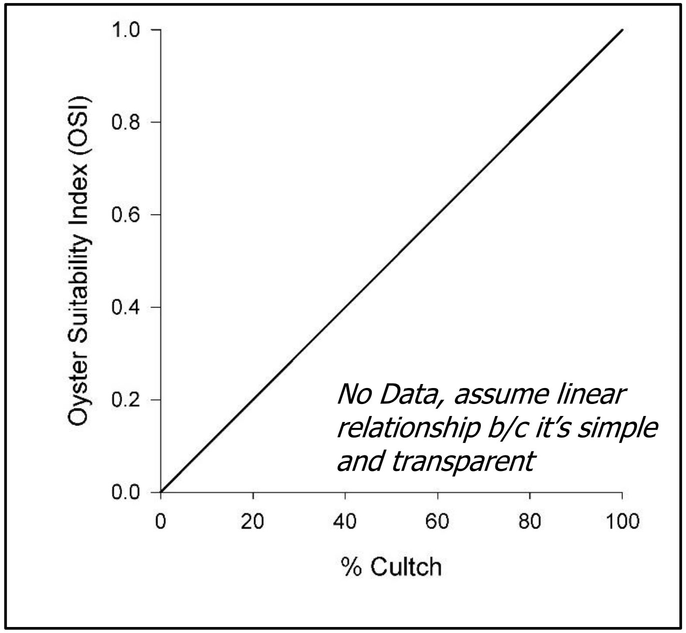




DO THE FUNCTIONAL FORMS MAKE SENSE?

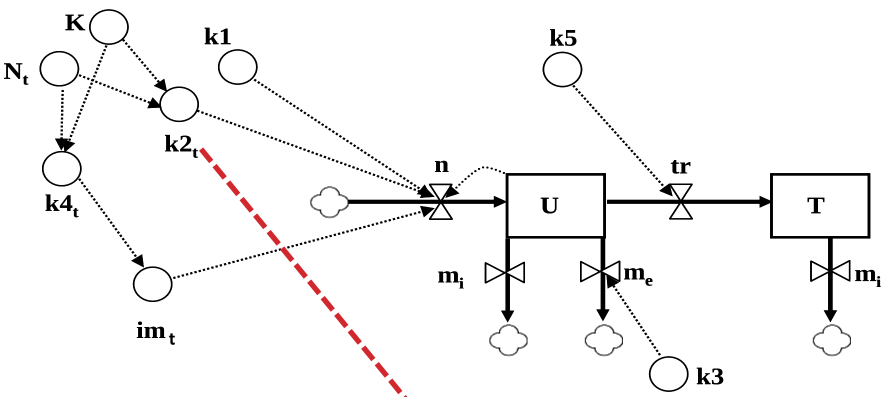
- Do the functional forms of the equations generate reasonable output given the other components in the model?
- Equations may not stand up after they're coupled with other model components





EVALUATING FUNCTIONS

Each function must be evaluated separately, then again when coupled to other components. **Document each step!**



$$k2_t = 1 - (1 / \mathrm{K})N_t$$

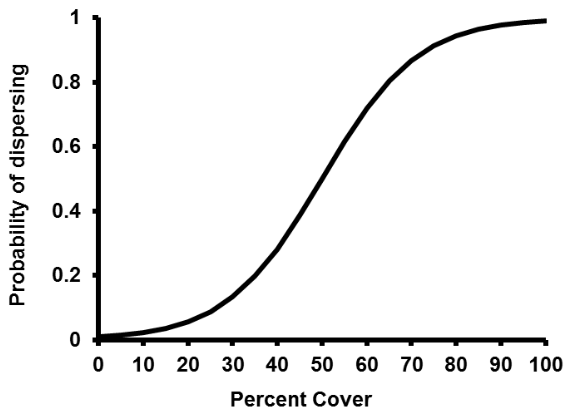
$$n_{i,\,t} = (U_{f,\,t})\,(k1)(k2_t)$$

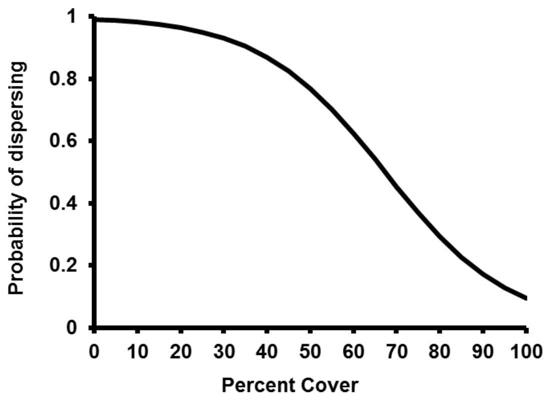
$$U_{i,\,t+1} = U_{i,\,t} + (n_{i,\,t} + i_{i,\,t} - m_{i,\,t} - me_{i,\,t} - tr_{i,\,t})$$

DOES MODEL ACT LIKE YOU THOUGHT IT WOULD?

***A priori* expectations are critical for thorough evaluation**

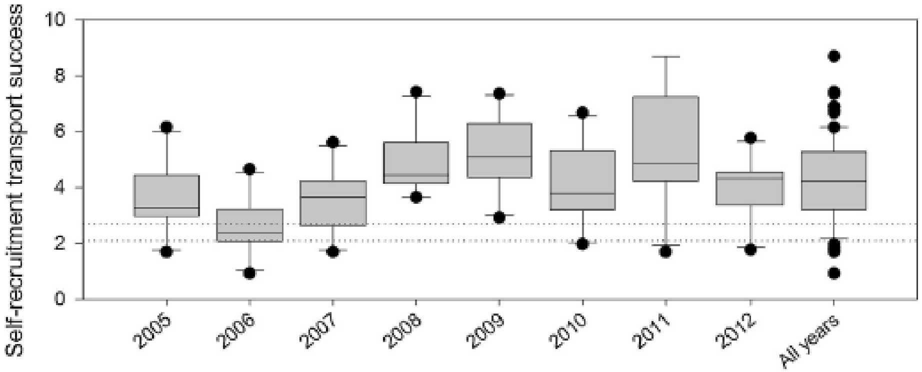
- Without documenting expected patterns of behavior, it becomes difficult to ascertain whether the model is producing the correct values
- Evaluate code and each function to make sure everything is being calculated correctly
- Practice iterative evaluation

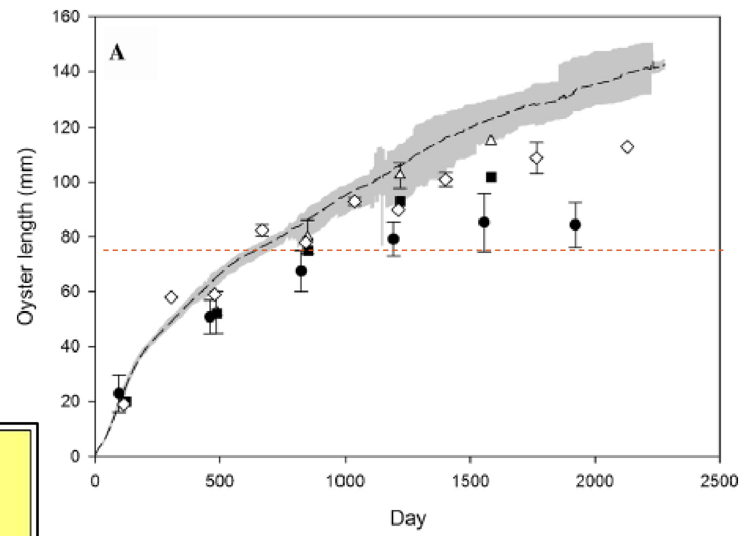


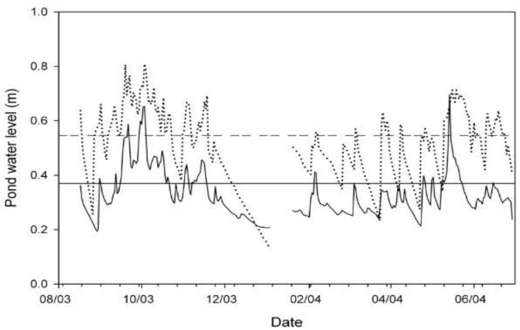


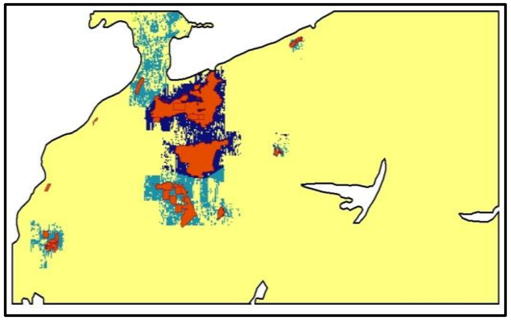
HOW DOES THE MODEL OUTPUT COMPARE TO REAL DATA?

In general, environmental models are used to project system dynamics, so some understanding of how output compares to real data can be useful.

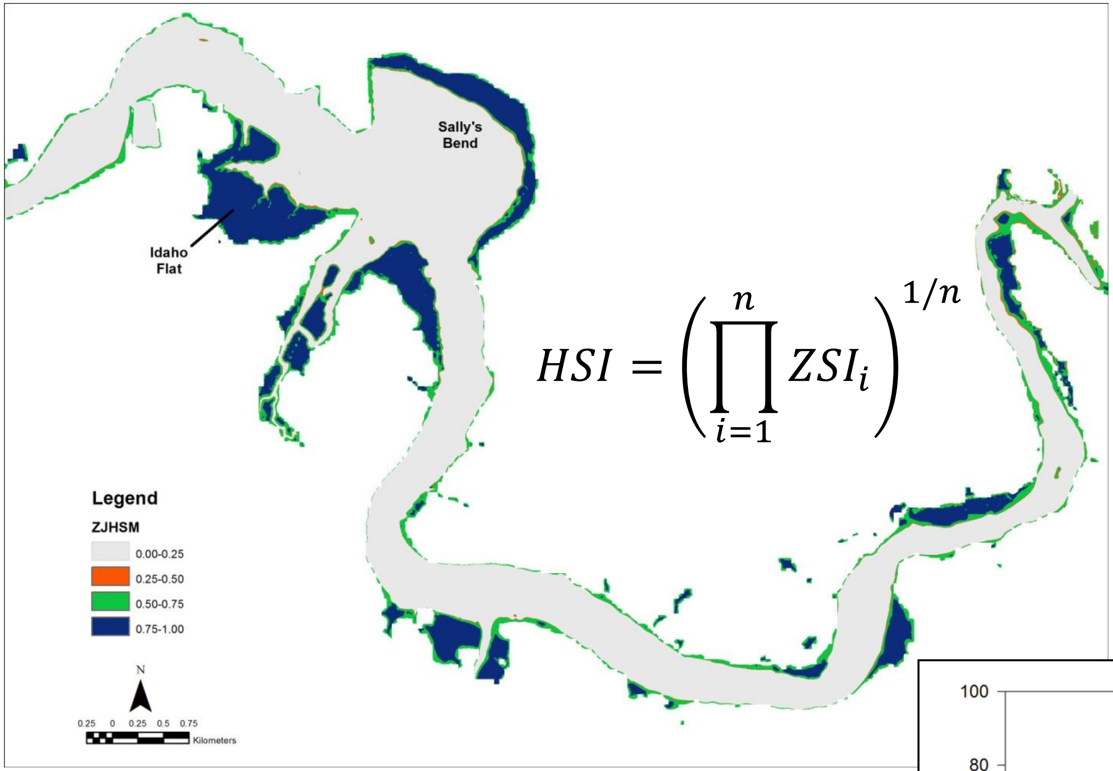






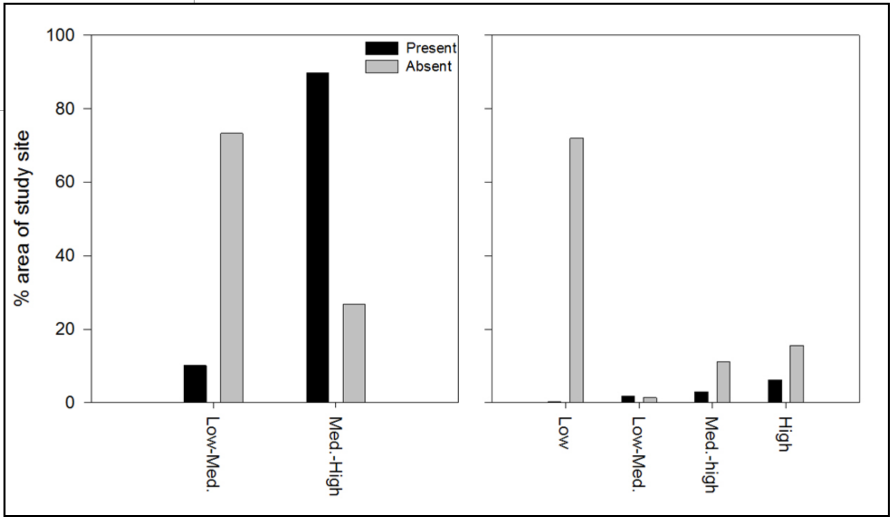


SEAGRASS EXAMPLE



$$HSI = \left(\prod_{i=1}^{n} ZSI_i\right)^{1/n}$$

- Model does a good job of predicting presence of *Z. japonica,* as well as predicting its absence.

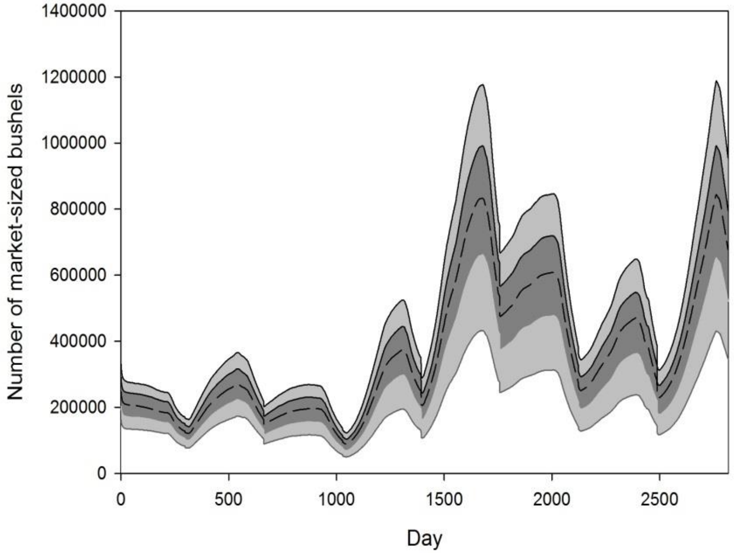


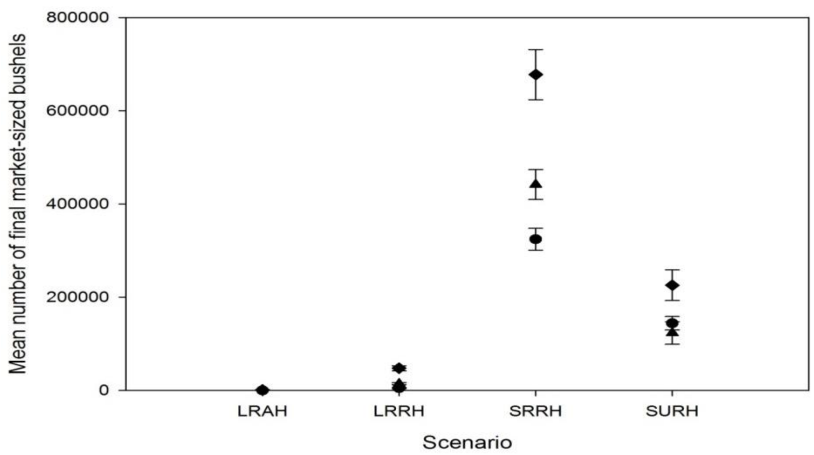
MODEL SENSITIVITY

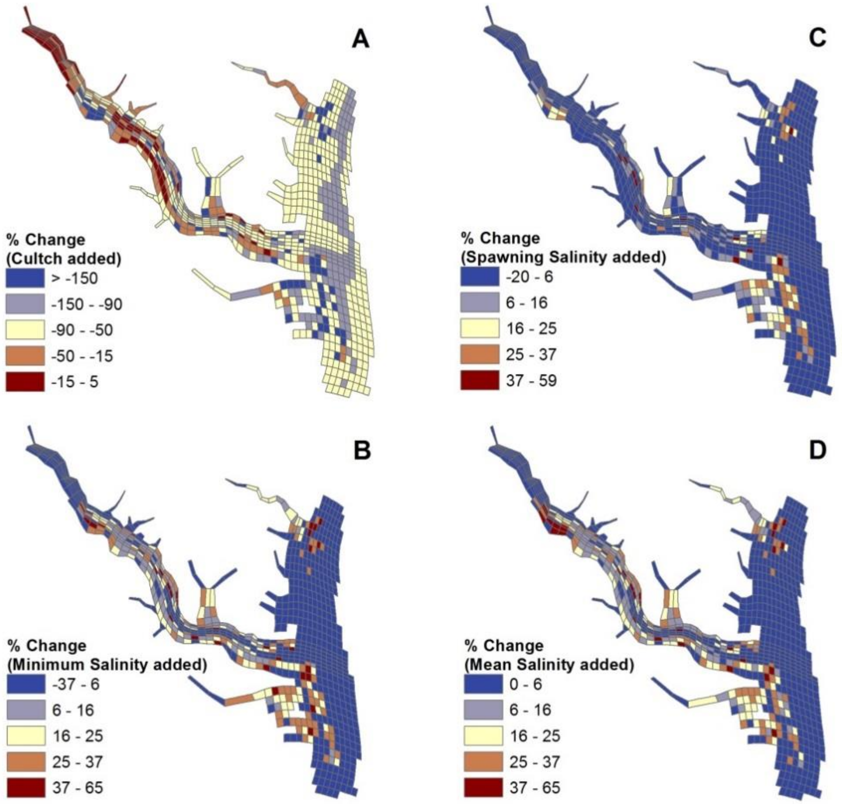
Determines degree of response of model behavior to changes in various components.

- Provides indication of relative precision of each parameter
- Run model over range of values representing degree of uncertainty
- Indicates level of confidence we have with model's ability to address question

MODEL SENSITIVITY







DISCOVERING THE MODEL'S LIMITS

How does the model perform under extreme circumstances?

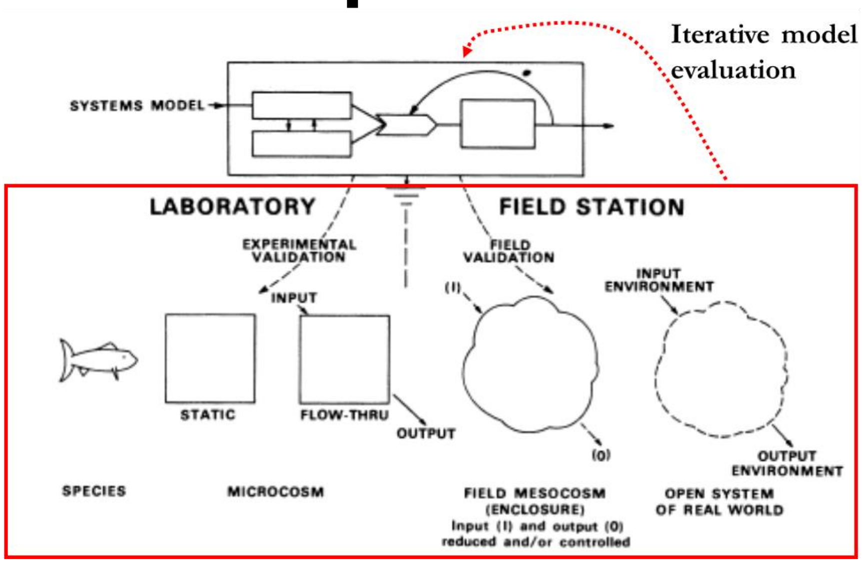
- Run model across wide range of values outside of range

Break the model!

Critical for understanding how model functions across a wide range of conditions.

PITFALLS

Failure to iteratively evaluate and document each step



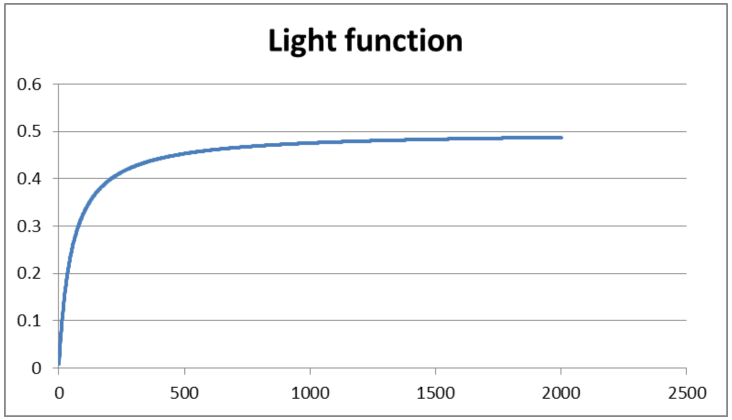
$$\Delta W = W_s P - W(R_m + M) \tag{1}$$

$$R_m = r_{20} * Q_{10}^{((T-20)/10)} \tag{2}$$

$$P = P_{max} * \frac{I}{I+H_I} * \frac{S*T^{pt}}{T^{pt}+H_T^{pt}} * \frac{H_D}{D+H_D} \tag{3}$$

$$I_{z,t} = I_0 * e^{-K_d - K_p * b_z} \tag{4}$$

$$I_{z,t} = I_0 * e^{-K_d} \tag{5}$$



PITFALLS

Underestimating importance of qualitative components of model

- Does it look right?

Accepting conceptually flawed functional relationships

- Immediately places you in position to defend

PITFALLS

Acceptance of surprising model results

- “Hmmm…that seems weird. Oh well.”
- Need to figure it out and document it!
    - Could be coding issue, flawed conceptualization, quantification, etc..

Interpreting initial results without letting model burn-in

- Initial behavior might not represent model patterns
- Resulting from initial conditions and parameter values

PITFALLS

Over-reliance on automated evaluation techniques

- Canned software cannot provide level of rigor needed for complex models

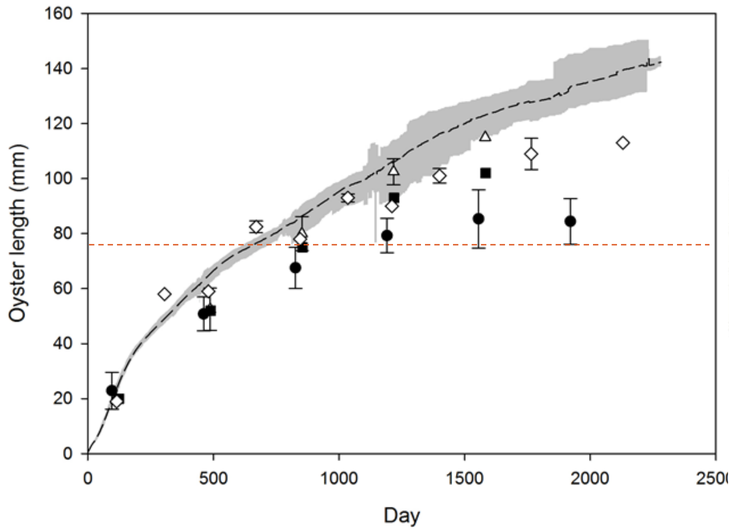
Over-reliance on statistics

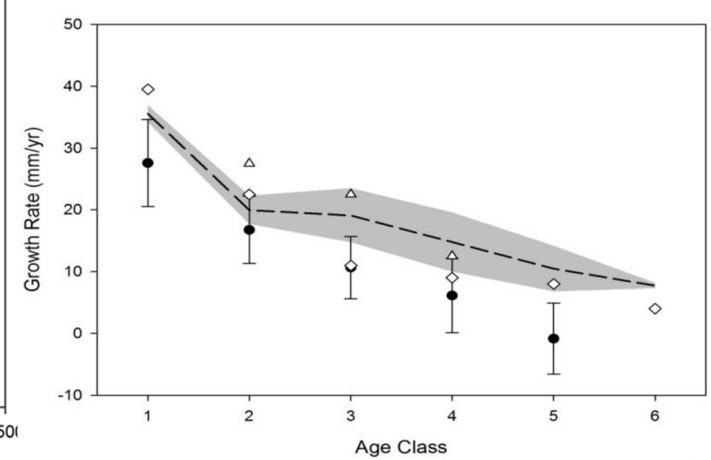
- Statistical and practical significance aren’t always the same thing

PITFALLS

Believing data from real system are correct

- Field data were collected over a specific period of time, under a particular set of environmental conditions.
- May not necessarily correlate with model projections





PITFALLS

Careless design of sensitivity analysis

- What parameters actually make sense for a sensitivity analysis?

Tendency to equate model sensitivity with model failure