Model Quantification

GETTIN' MATHY WITH IT.

Outline

Why quantify?

What type of math?

What's the time step?

Functional forms

Parameter estimates

Quick, dirty, yet scientifically defensible tricks to generate patterns

Pitfalls

Don't have to be good at math!

Biologists and ecologists have deep understanding of their systems, but generally aren't exposed to advanced mathematical techniques

Elegant mathematical solutions are neat, but they're not the only approach

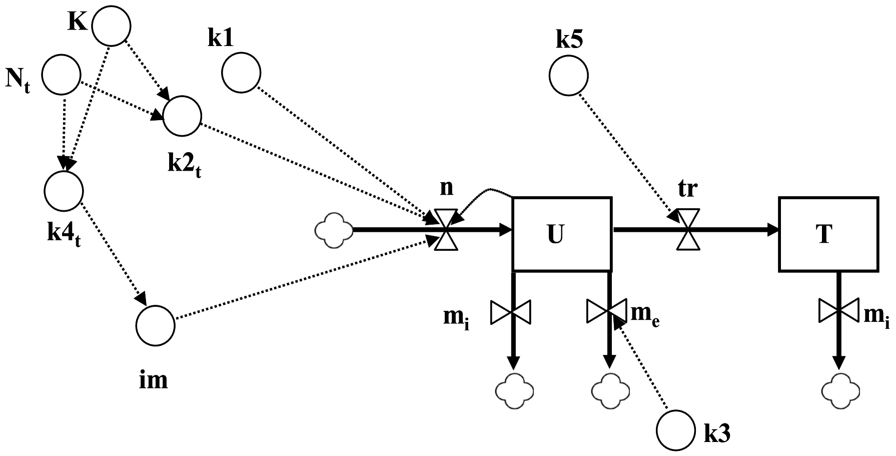
If you have an understanding of your system, you can model it

Quantification

Quantifying models provides ability to understand numerical consequences of ideas, scenarios, system dynamics, etc..

Conceptual Model should be used a templat

- Equations should be tightly coupled with conceptual model
- Helps with communication and transparency
- Don’t hide behind the math/code



Choosing appropriate mathematics & software

In Theory:

- Results should not depend on software or advanced math
- What is important is that the critical processes are captured

In Practice:

- Software/Mathematics affect efficiency and computation time
- Need to identify up-front how model will be quantified
- Mechanistic (process-based) models aren't developed that often for USACE planning
  - Statistical equations (correlations) used as proxies

How do you choose the approach?

Experience

Comfort-level

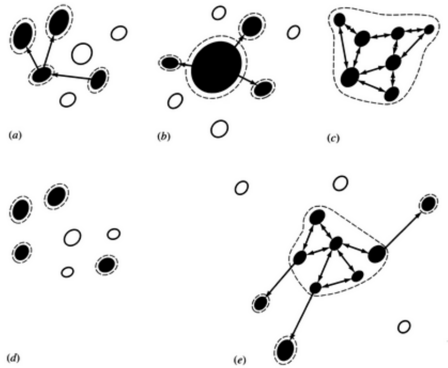
Deadlines

Question being asked

$$N_{t+1} = N_t + (births\ - deaths)$$

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right)$$

$$\mathbf{M_A} = \begin{bmatrix} F_0 & F_1 & F_2 & \cdots & F_{m-1} & F_m \\ P_0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & P_1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & P_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & P_{m-1} & 0 \end{bmatrix} \quad \mathbf{n_A} = \begin{bmatrix} N_0 \\ N_1 \\ N_2 \\ \vdots \\ N_{m-1} \\ N_m \end{bmatrix}$$



Simpler is better – Don’t make it too complicated!

Choosing the time step

Have to choose how often the model is updated, and how long to run it.

- We plan for 50 yr horizon, but how often do you need to calculate changes in order to get an accurate idea?

  - What processes are you interested in? How often do they occur?
  - Time step needs to reflect what's happening in nature, not what's convenient
  - Don't have to choose familiar units
    - Can use 12 seconds, 3 days, 14 months, 50 yrs, etc…

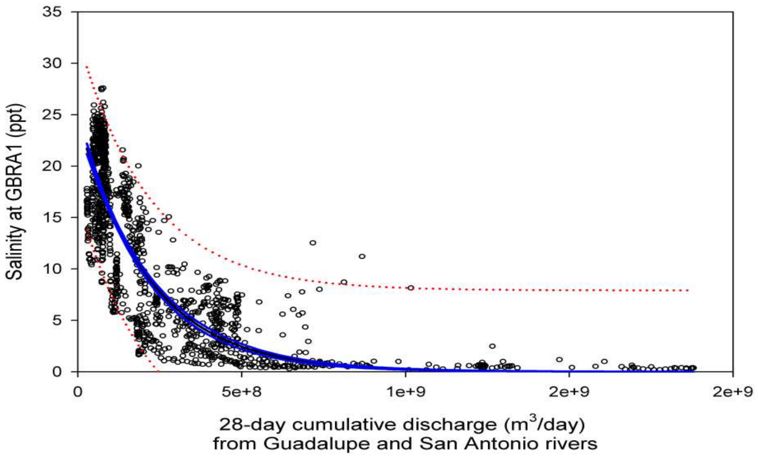
Time step, con't.

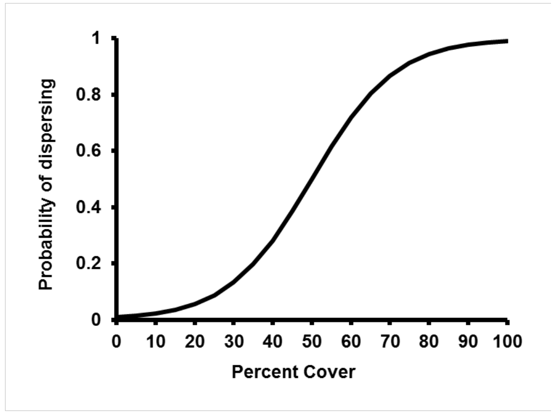
Can have nested time-steps within a model

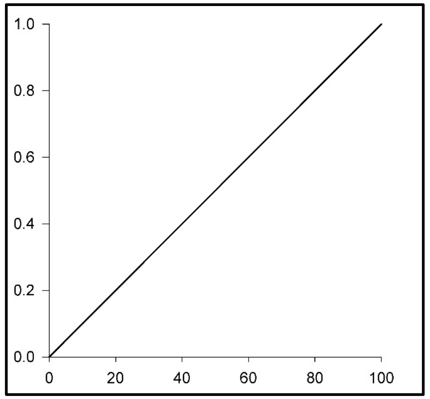
What level of precision is necessary?

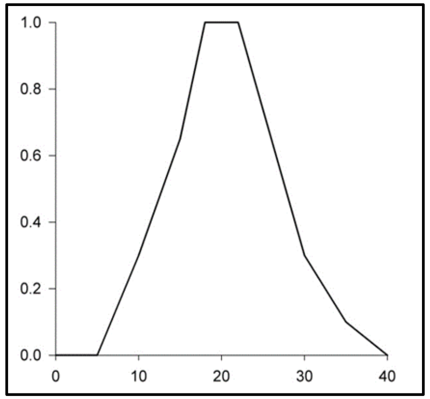
Functional Forms of Equations

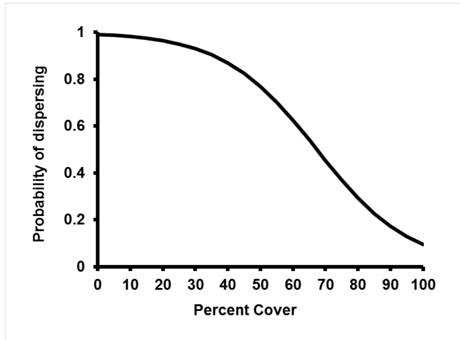
How should relationships be quantified?











$$p(j) = \frac{e^{V_j/\theta_k} \cdot \left(\sum_{i \epsilon C_k} e^{V_i/\theta_k}\right)^{\delta_k - 1}}{\sum_{k'}\left(\sum_{i \epsilon C_{k'}} e^{V_i/\theta_{k'}}\right)^{\delta_{k'}}}$$

$$I_A + I_B - (I_A * I_B), \text{ if } I_A > 0, I_B > 0$$

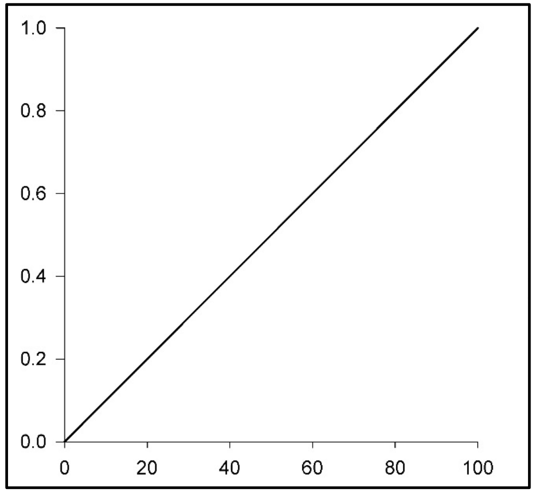
$$I_A + I_B + (I_A * I_B), \text{ if } I_A < 0, I_B < 0$$

$$\frac{I_A + I_B}{1 - \min[|I_A|, |I_B|]}, \textit{ otherwise}$$

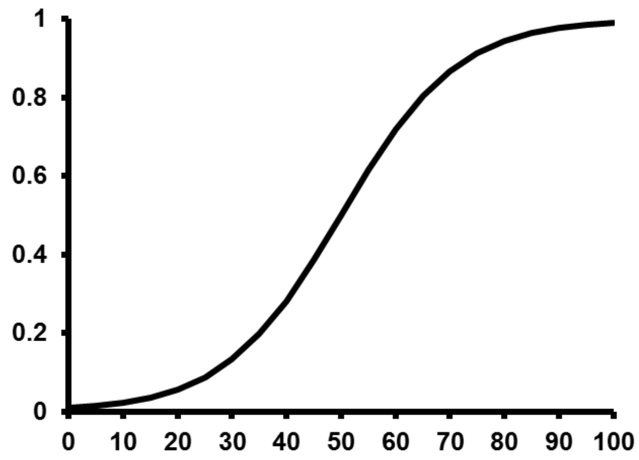
What if functional forms are unknown?

Use verbal descriptions and graphical functions

- Try to explain the relationship in a minute, then draw a picture
- Graphical representations provide an intermediate step between verbal and mathematical representations.



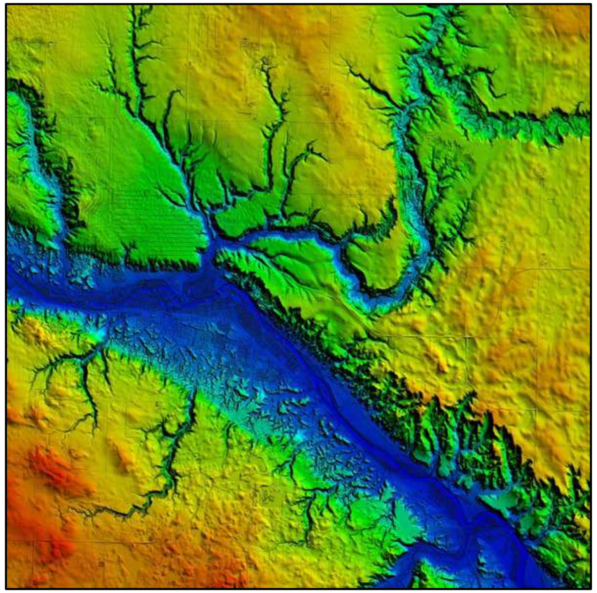
Linear functions: simplest relationship; the general relationship between two variables is understood (e.g., variable A increases when variable B decreases), but the exact form is not



ogistic functions: ıore complex; llows threshold ffects, and eriods of stasis nd rapid change

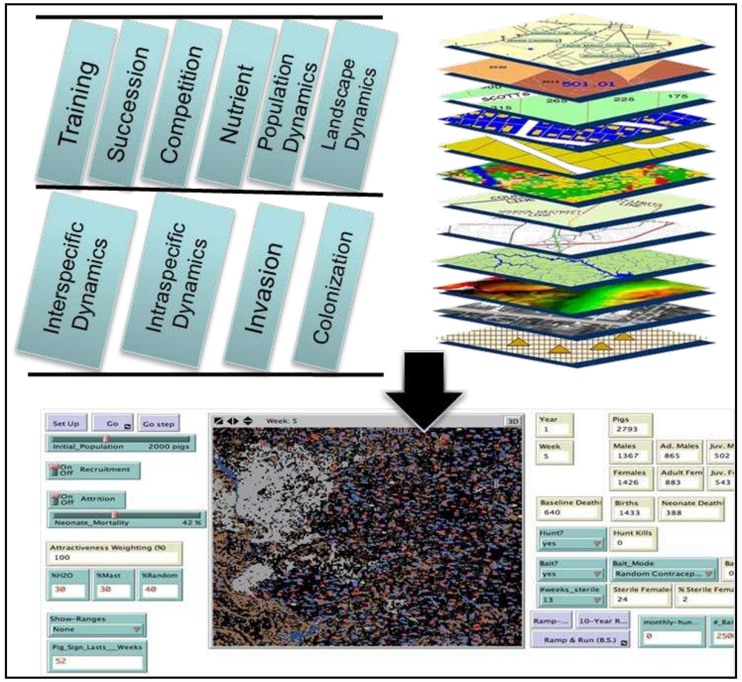
Types of data and parameterizations

Quantitative Data
- Field work
- Remotely sensed
- Other models
- Literature
- Theory



Qualitative data
- Expert opinion
- Hypotheses

The model itself
- Experimenting with model can reveal trends and patterns



Capturing feedbacks and thresholds

All environmental systems have feedback (positive/negative) and thresholds

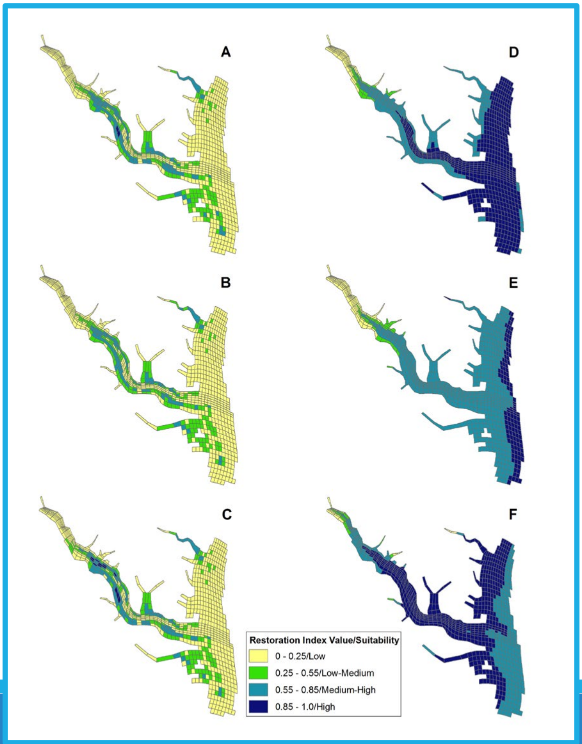
- E.g., crowding in populations is a negative feedback
- Species viability changes under different environmental conditions

These effects are often difficult to determine precisely in nature

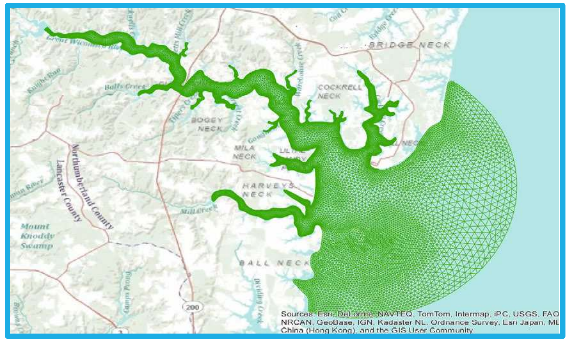
Integrated Models

Using Hydrodynamic Models

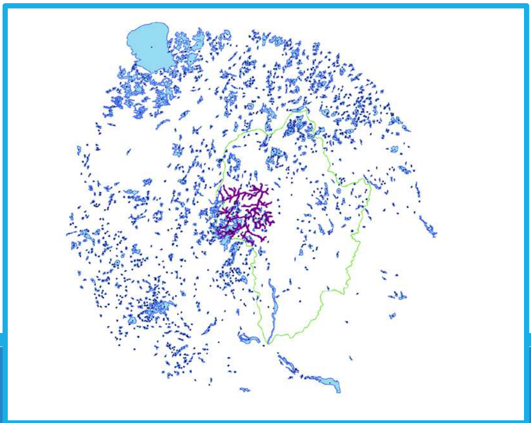
- *Integrated Models* are models composed of multiple models
- USACE modeling generally combines hydrodynamic and ecological models



CH3D temperature, salinity, used as inputs for Oyster Restoration model



Oyster metapopulation used ADH velocity, flow, and WQ data



SWAT (watershed model) calculated flow, and velocity to determine Zebra mussel larval mortality

Quantifying inputs from hydrodynamic models

Important considerations

- What time scale is important to the ecological components of the system?

  - Hydrodynamic models can run at small time steps that might not be link well with ecological processes.
    - Cumulative effects are more important (E.g., does seagrass care what happens every 30 seconds, or are exposures over weeks or months more important?

  - Requires aggregation of Hydro model data to reasonable scales

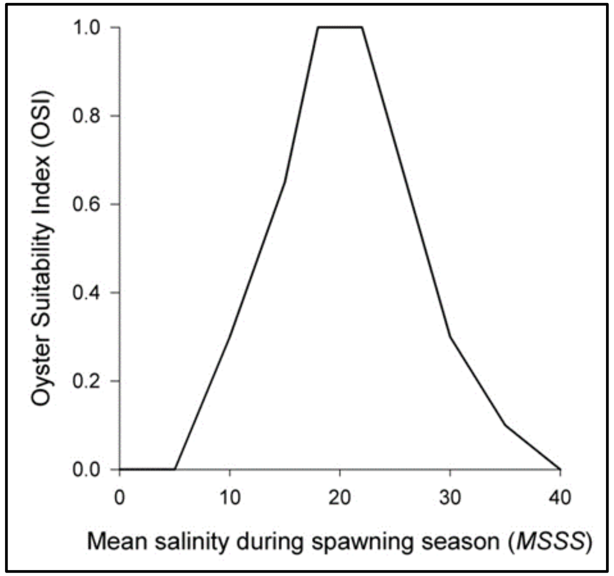
Quantifying inputs from hydrodynamic models

- Are you interested in historical patterns or future patterns?
- Does the hydro data contain a range of values that can show wide range of ecological response? If not, inference from eco model limited.
- Critical to organize with Hydro modeling team to ensure data will be delivered in appropriate format/scales
  - Costly and time-consuming to redo simulations

Quantifying Thresholds

Quickest way is with step-functions or if-then statements

- Equations are almost never reported

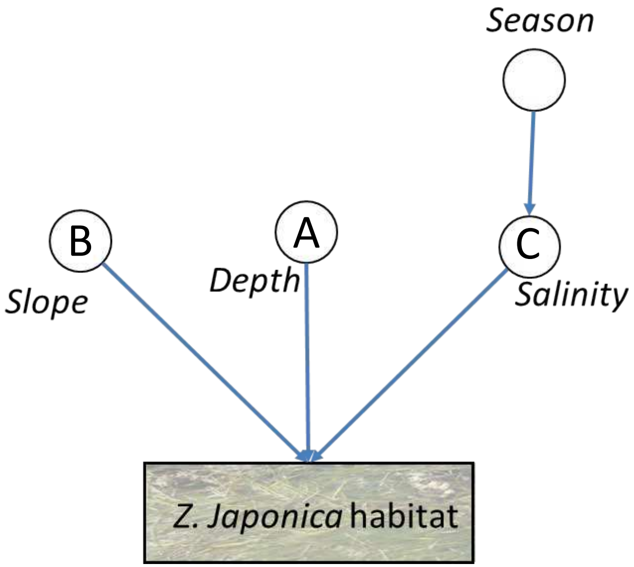


Typical HSI Representation

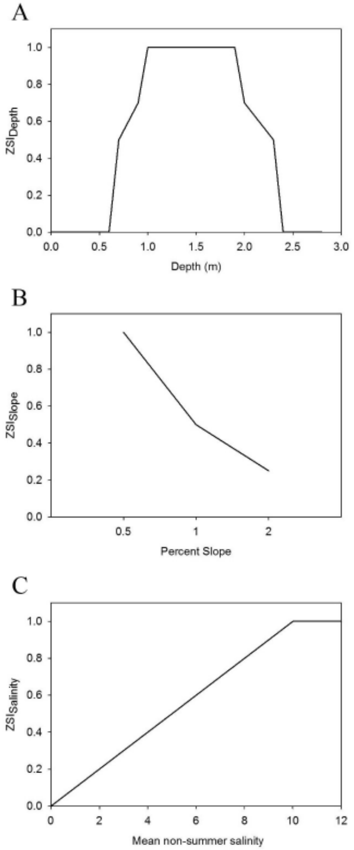
$MSSS \leq 5$ or $MSSS > 40$	$OSI_{MSSS} = 0$
$5 < MSSS \leq 10$	$OSI_{MSSS} = -0.3 + (0.06 * MSSS)$
$10 < MSSS \leq 15$	$OSI_{MSSS} = -0.4 + (0.07 * MSSS)$
$15 < MSSS < 18$	$OSI_{MSSS} = -1.1 + (0.1167 * MSSS)$
$18 \leq MSSS \leq 22$	$OSI_{MSSS} = 1$
$22 < MSSS \leq 30$	$OSI_{MSSS} = 2.925 - (0.0875 * MSSS)$
$30 < MSSS \leq 35$	$OSI_{MSSS} = 1.5 - (0.04 * MSSS)$
$35 < MSSS \leq 40$	$OSI_{MSSS} = 0.8 - (0.02 * MSSS)$

Take advantage of the math!
(equations look smarter)

Seagrass quantification (Yaquina Bay, OR)



Conceptual

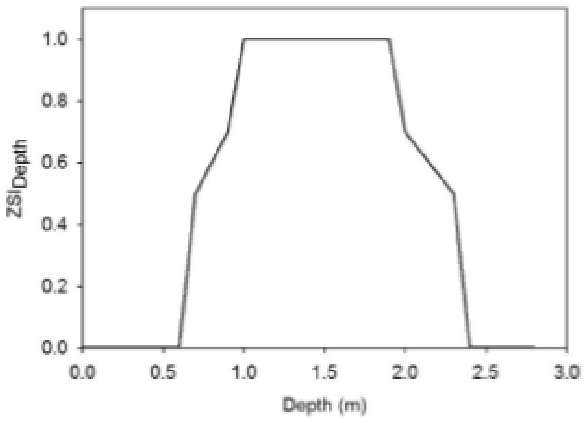


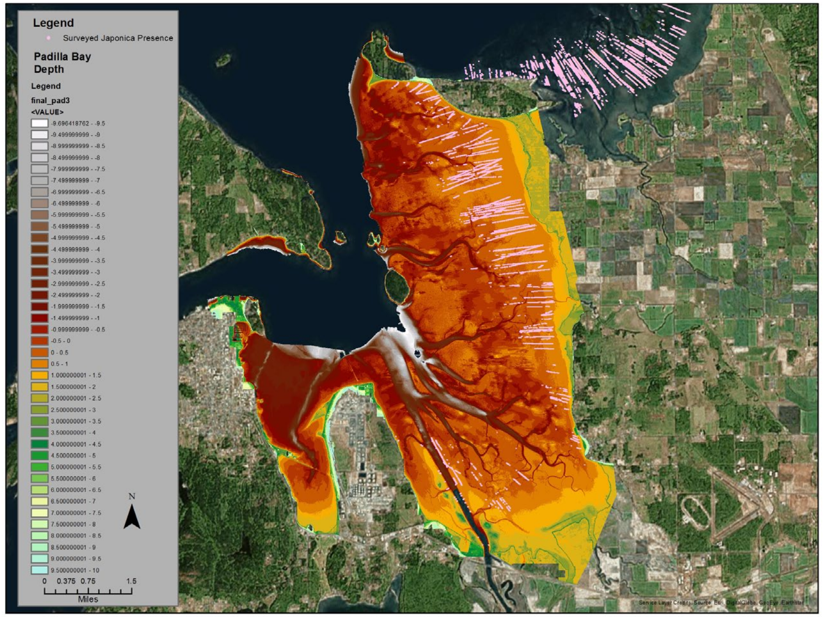
Functional

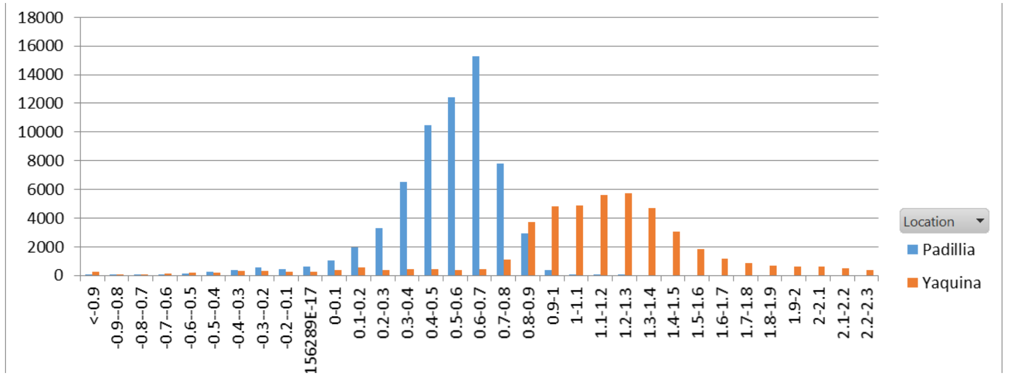
	Variable	Equation	Eq#
	Depth (m)		
A	$Depth \leq 0.6$ or $Depth > 2.4$	$ZSI_{Depth} = 0$	(1)
	$0.6 < Depth \leq 0.7$	$ZSI_{Depth} = -3 + (5 * Depth)$	(2)
	$0.7 < Depth \leq 0.9$	$ZSI_{Depth} = -0.2 + (Depth)$	(3)
	$0.9 < Depth \leq 1.0$	$ZSI_{Depth} = -2 + (3 * Depth)$	(4)
	$1.0 < Depth \leq 1.9$	$ZSI_{Depth} = 1$	(5)
	$1.9 < Depth \leq 2.0$	$ZSI_{Depth} = 6.7 - (3 * Depth)$	(6)
	$2.0 < Depth \leq 2.3$	$ZSI_{Depth} = 2.033 - (0.67 * Depth)$	(7)
	$2.3 < Depth \leq 2.4$	$ZSI_{Depth} = 12 - (5 * Depth)$	(8)
	Slope		
B	$Slope \leq 0.5\%$	$ZSI_{\%\ Slope} = 1$	(9)
	$0.5\% < Slope \leq 1\%$	$ZSI_{\%\ Slope} = 1.5 - Slope$	(10)
	$1\% < Slope \leq 2\%$	$ZSI_{\%\ Slope} = 1.5 - (0.25 * Slope)$	(11)
	$2.0\% < Slope$	$ZSI_{\%\ Slope} = 0.25$	(12)
	Salinity		
C	$0 \leq Salinity \leq 10$	$ZSI_{Salinity} = 0.1 * Salinity$	(13)
	$Salinity > 10$	$ZSI_{Salinity} = 1$	(14)

Mathematical

Re-quantifying model for new area







In first application (Yaquina Bayt), depth worked well. In second (Puget Sound), *Depth* relationship didn't capture the relationship as well. Have to re-quantify, or re-conceptualize functional form based on data.

Quantification for Kelp and Seagrass

Ideal conditions:
- clear water (light availability)
- nutrient-rich waters
- moderate water movement
- 20 C (68 F) temperatures
- typical depth range 25 – 90 feet

Ideal conditions:
- *Z. marina* near-shore, estuarine areas
- clear water (low turbidity and high PAR light)
- calmer waters
- optimal temperature 50-68 F
- typically 0-8 foot depths, max 20 feet
- 4 - 97% sand composition
- salinity 10-30 ppt

Missing data

There are often relationships that aren't defined quantitatively

- Have to rely on expert opinion
- Literature
- Interpolations

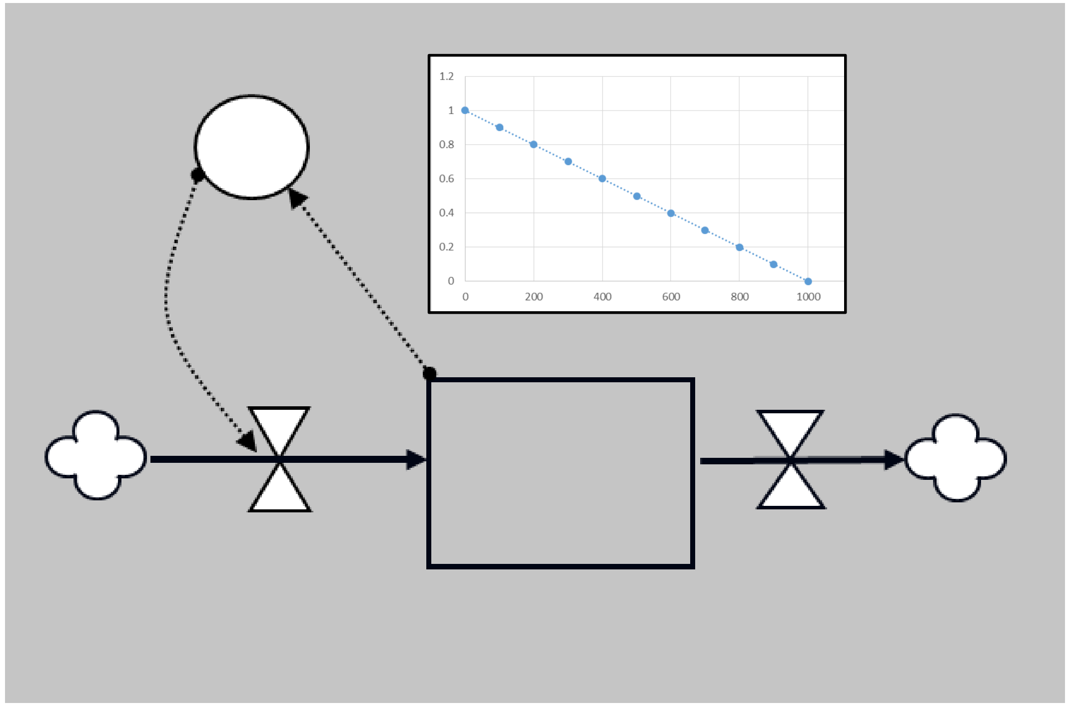
This is not less rigorous than quantitative data analysis, just less precise

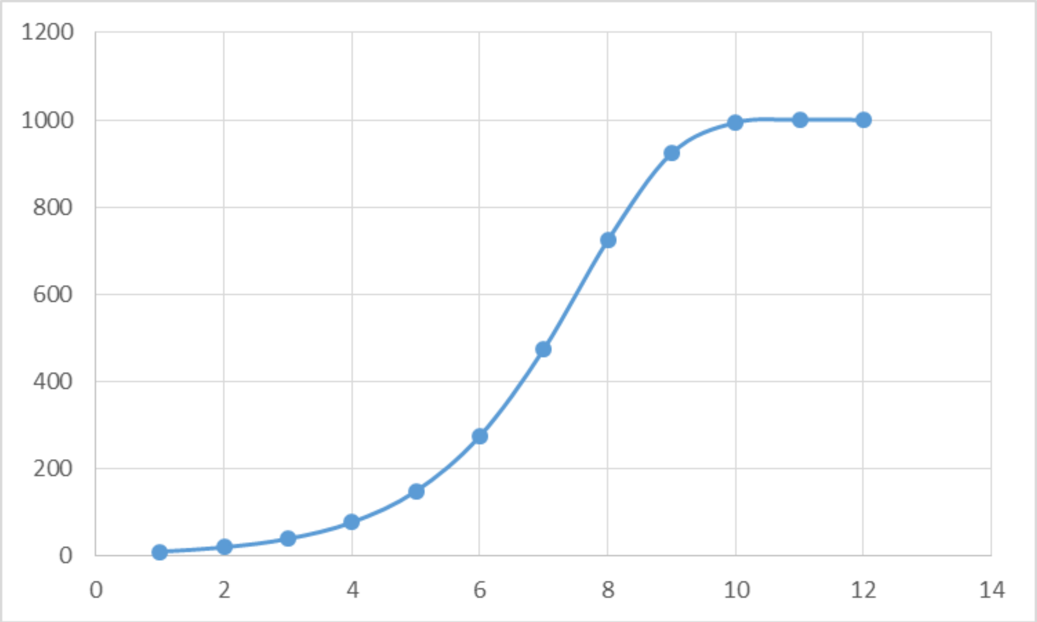
- Qualitative data requires increased attention in the documentation

Will make bigger mistake leaving out important relationships than hypothesizing about relationships

- Increased need for transparency

Quantifying feedbacks w/o data





Tell the story!

In this case, negative feedback causes incremental decrease in growth

Quick way to generate patterns

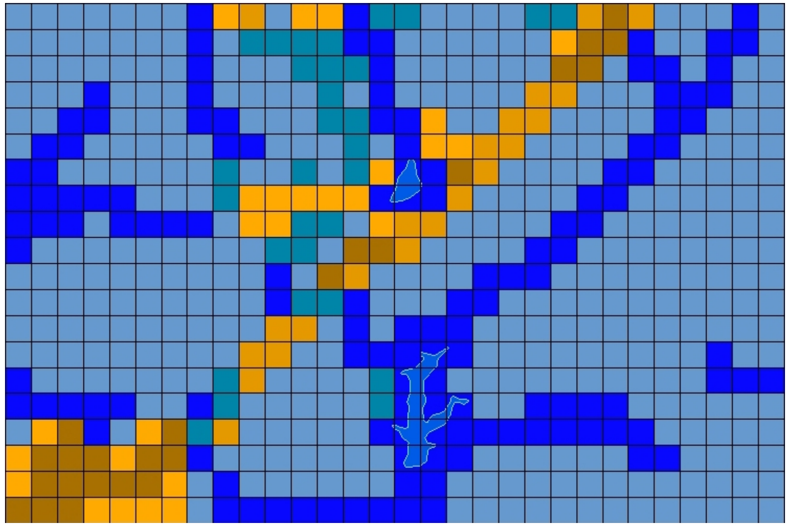
Quantifying expert opinion: Habitat

Landscape

Expert Opinion

+1.00	ABSOLUTELY
+0.75	STRONGLY
+0.50	MODERATELY
+0.25	SLIGHTLY
0.00	INCONCLUSIVE
-0.25	SLIGHTLY
-0.50	MODERATELY
-0.75	STRONGLY
-1.00	ABSOLUTELY

Model



$$I_A + I_B - (I_A * I_B), \text{ if } I_A > 0, I_B > 0$$

$$I_A + I_B + (I_A * I_B), \text{ if } I_A < 0, I_B < 0$$

$$\frac{I_A + I_B}{1 - \min[|I_A|, |I_B|]}, \ otherwise$$

Modeling without data

Decisions will need to be made, regardless of data availability

Need for transparency

Simple functions can help identify magnitude and general trends in absence of data

Expert opinion can be used to parameterize equations until other datasets are available

Considering Scale

System properties emerge as scale changes

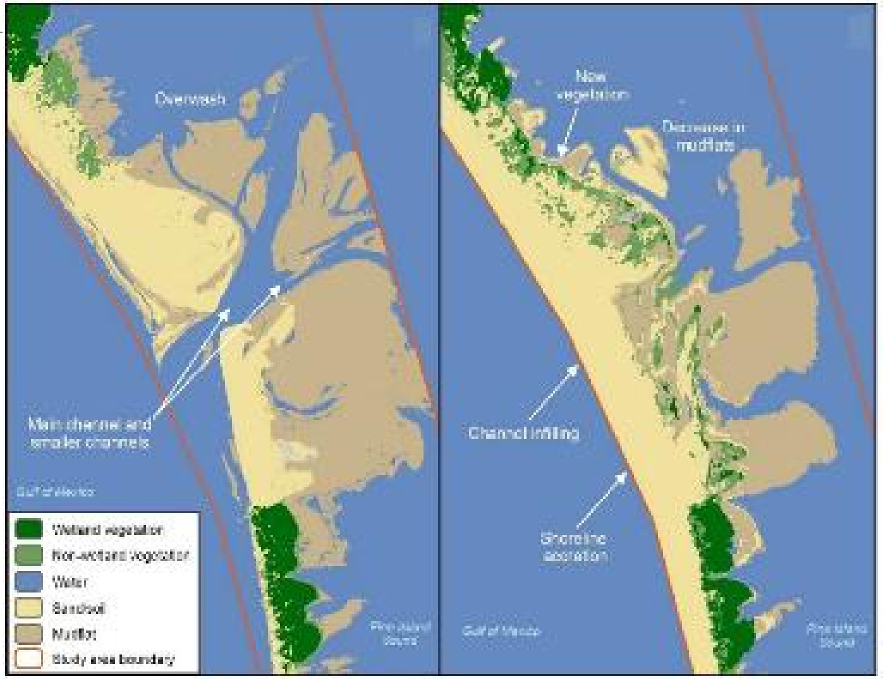
Spatial Modeling

Incorporating topographic, geomorphic, and/or land use patterns into models to understand how changes in spatial configurations affect ecological dynamics

Space matters

- Configuration and composition of landscapes can affect ecological structure and function

Considering Space

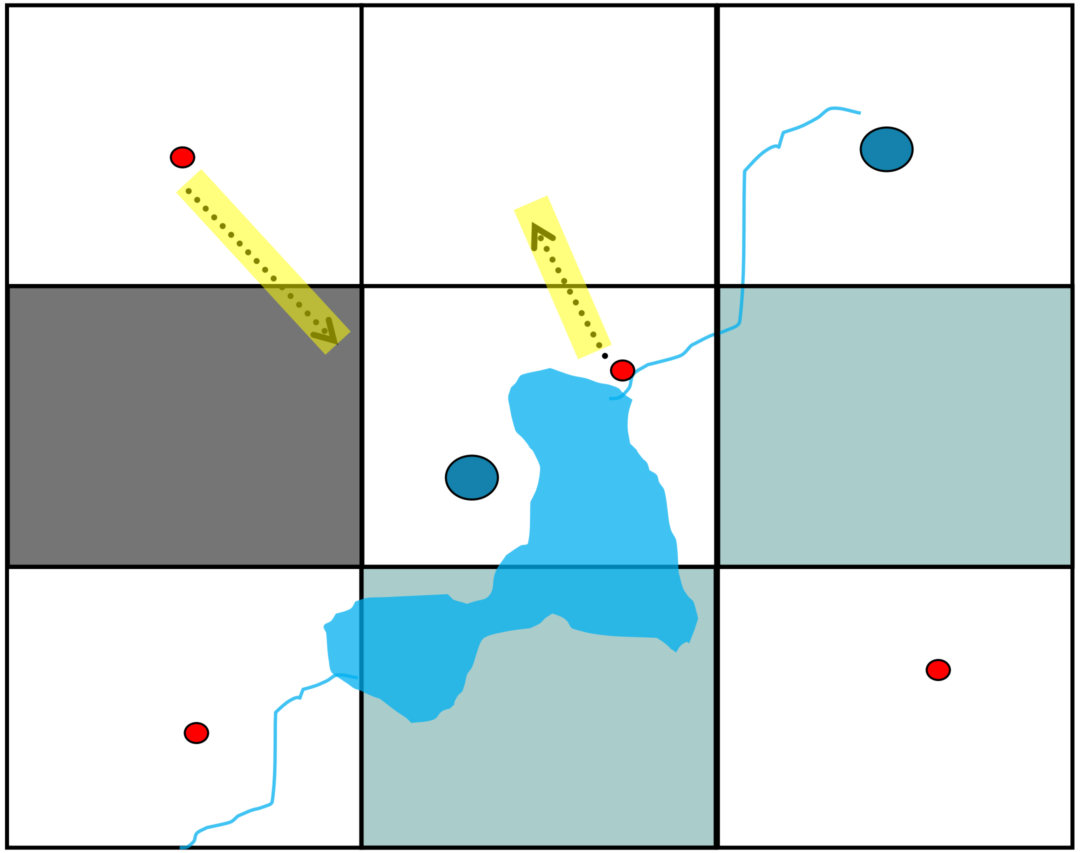


Working with spatial models

Considerations:

- Location-based differences across the project area

- What spatial scale is relevant?
  - Link ecological processes to a spatial scale (i.e., the grid/DEM/etc)



interactive Toolkit for Applied Modeling (TAM)

Platform developed for rapid model development

Quantifies threshold-based datasets

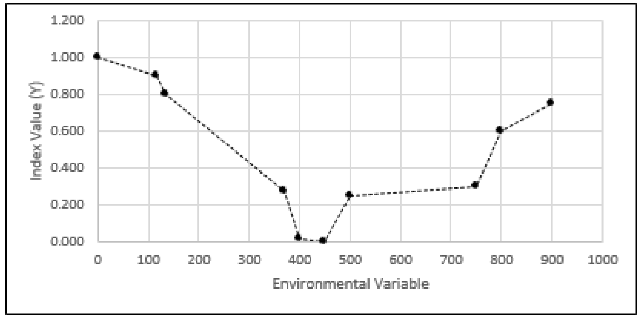
Certified for USACE

ENTER DATA INTO HIGHLIGHTED CELLS

Breakpoint #	Environmental Variable	Index Value (Y)
1	0	1.000
2	118	0.900
3	136	0.800
4	368.4	0.275
5	400	0.01
6	450	0
7	500	0.25
8	750	0.3
9	800	0.6
10	900	0.75

Values	Intercept	Slope	Equation
0 -118	1.00	-0.0008	Y= 1 + (-0.0008 * Environmental Variable)
118 -136	1.56	-0.0056	Y= 1.56 + (-0.0056 * Environmental Variable)
136 -368.4	1.11	-0.0023	Y= 1.11 + (-0.0023 * Environmental Variable)
368.4 -400	3.36	-0.0084	Y= 3.36 + (-0.0084 * Environmental Variable)
400 -450	0.09	-0.0002	Y= 0.09 + (-0.0002 * Environmental Variable)
450 -500	-2.25	0.0050	Y= -2.25 + (0.005 * Environmental Variable)
500 -750	0.15	0.0002	Y= 0.15 + (0.0002 * Environmental Variable)
750 -800	-4.20	0.0060	Y= -4.2 + (0.006 * Environmental Variable)
800 -900	-0.60	0.0015	Y= -0.6 + (0.0015 * Environmental Variable)
900 -			Y = 0.75

References



Pitfalls

Choosing inappropriate mathematics & software

- Not all formats lend themselves to a given problem
- Can get trapped by constraints of approach

Not choosing appropriate time step

- Too long: violate assumption that change is system is constant b/w time steps
- Too short: lose interpretability, longer simulation time

Pitfalls

Relying on automated parameterization techniques

- Processes that test every possible combination of parameter values can quickly turn the model into a black box

Using overly sophisticated equations

- It’s easy to rely on fancy stats, but make sure they are appropriate for the objective of the model

Pitfalls

Uninterpretable functional relationships/coefficients without meaning

- Functional relationships should make sense (within your discipline)
- Coefficients should reflect magnitude of process occurring in nature

Not paying attention to units of measure

- Can violate assumptions and create nonsensical results

Pitfalls

No clear verbal description

- If you can’t explain it clearly, you can’t math it correctly
- Try to explain it in one minute – where you get hung up can help identify problem areas

Don’t consider graphical relationships

- Intermediate step b/w verbal and mathematical model
- Can serve as proxy for formalized equations

Reluctance to use qualitative information

- Specific numbers can be difficult to find. Stories aren’t.

Removing functional relationships due to lack of data