Quantification boot camp

Turning concepts into equations

Todd M. Swannack, Ph.D.

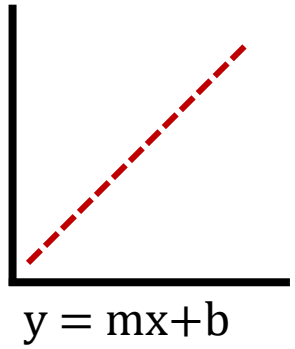
US Army ERDC

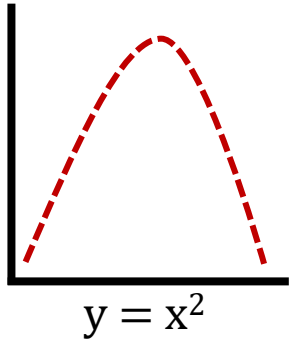
Homer Navigation Improvement Study

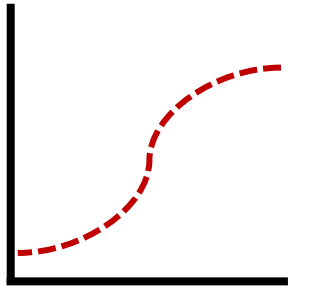
April 2024

Outline

- Why quantify?
- What type of math?
- Selecting a time step
- Functional forms
- Parameter estimates
- Quick, dirty, but scientifically defensible tricks to generate patterns
- Pitfalls







$$S(x) = \frac{1}{1 + e^{-x}}$$

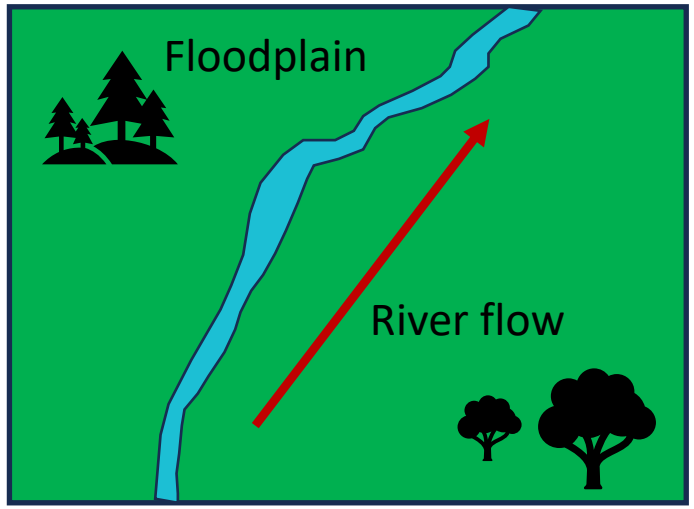
Quantifying environmental systems

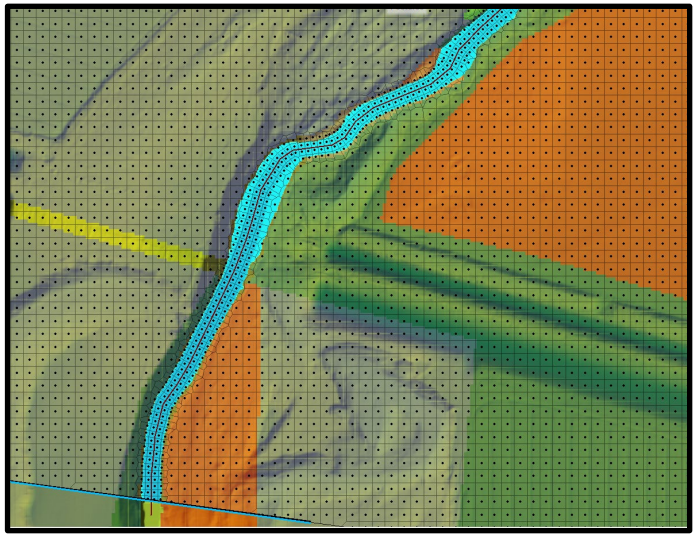
- Project delivery teams usually have a deep understanding of their systems, but not all members of the team may be familiar/comfortable with advanced math
- Elegant mathematical solutions are not the only approach
- If you understand **something** about your system, you can model it
- **Everyone** brings some knowledge to the table

Quantification

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

- **Conceptual models should be used a template**
  - Equations should be tightly coupled with conceptual models
  - Helps with communication and transparency
  - Document where equations come from and how they were chosen
  - Don’t hide behind the math/code

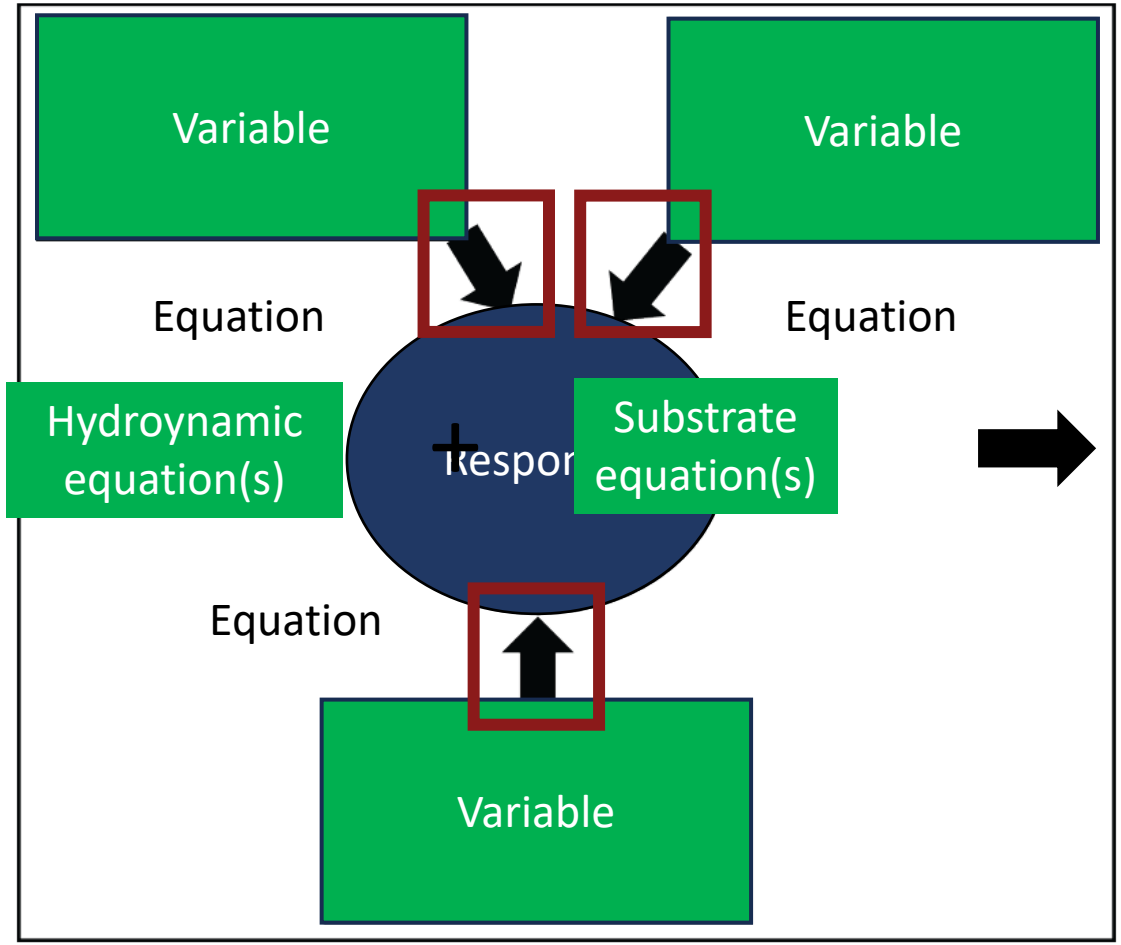




Conceptual models as templates for quantification

Boxes represent **variables** in the quan[illegible]del

Water quality equation(s) +



Conceptual model

Arrows represent **equations** in the quan[illegible]e model

Mussel habitat index

Herman et al. 2021

Choosing appropriate mathematics and software

In theory:

- Results should not depend on software or advanced math
- What is important is that 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) can be used as proxies

How do you choose an approach?

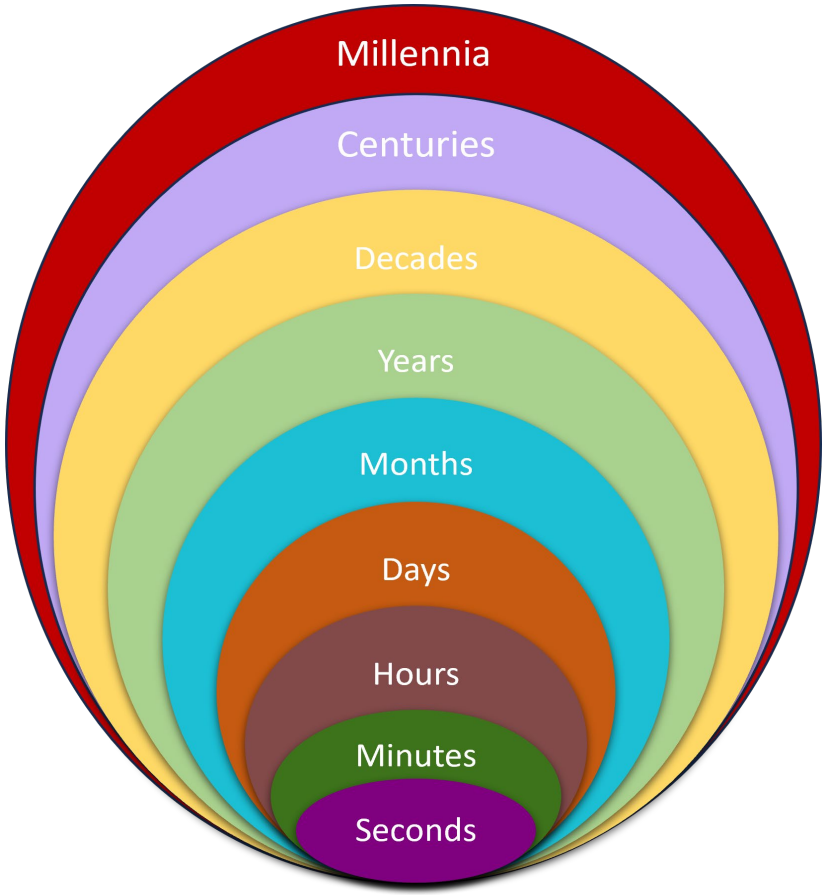
Key considerations:

- Experience
- Comfort-level
- Deadlines
- Question being asked
- Desired level of complexity for project stage and goals

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

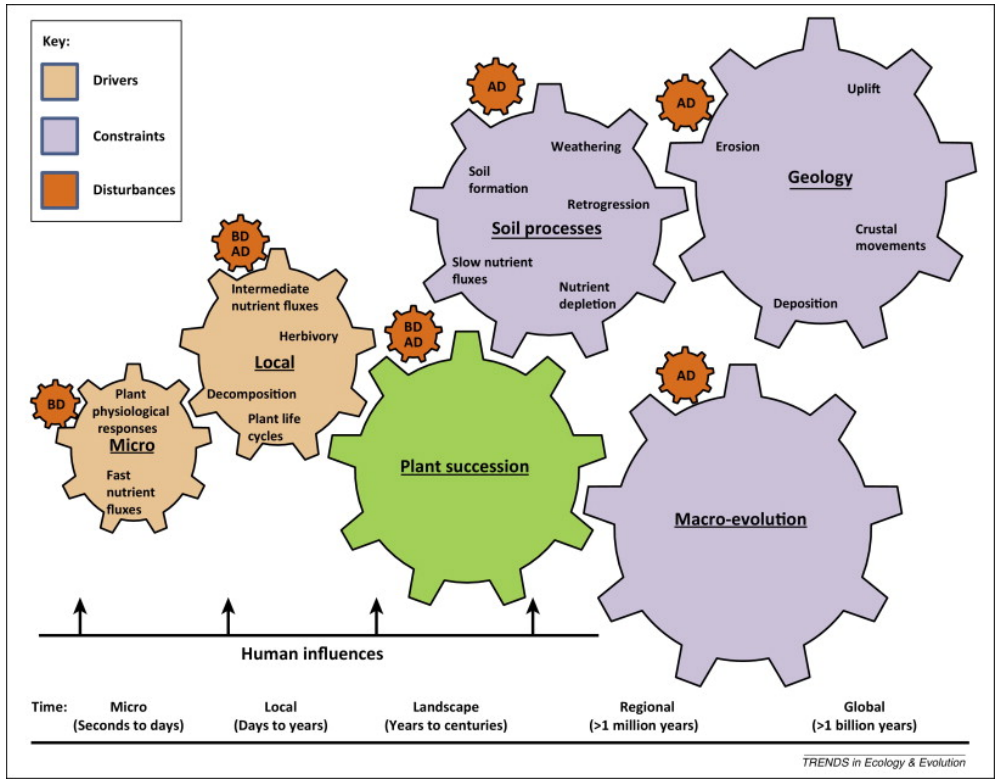
Selecting an appropriate temporal scale

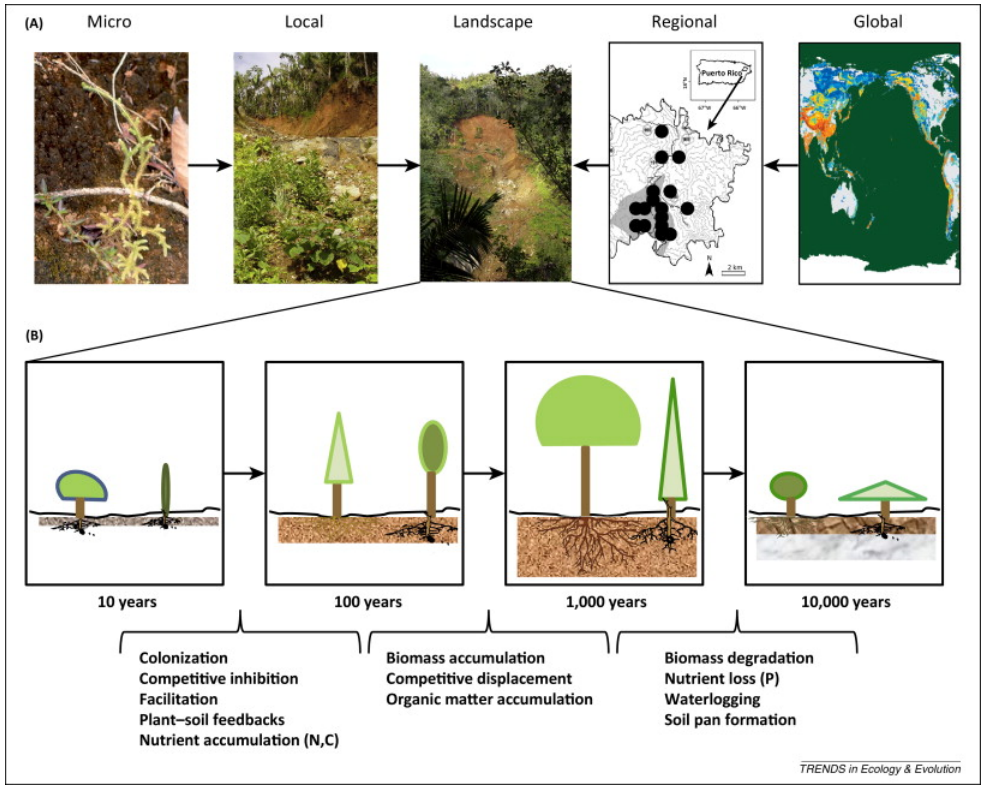
- **How often will the model be updated, and how long will it run?**
  - USACE plans for a 50 yr horizon, but how often do you need to calculate changes in order to get an accurate idea?
- **What processes are you interested in? How often do they occur? When are species present?**
  - Temporal scale needs to reflect what's happening in nature, not what's convenient
  - Familiar units aren't necessary
    - Can use 12 sec, 3 days, 14 months, 50 yrs, etc…



More on temporal scale

- Can have nested time scales within a model
- What level of precision is necessary?

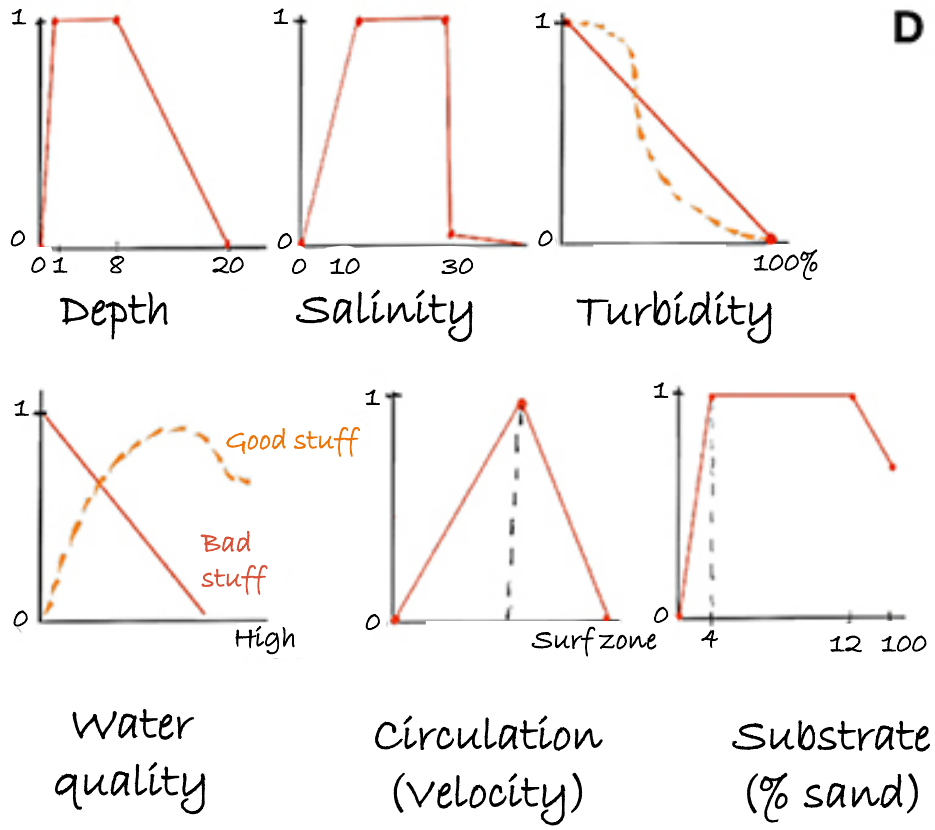




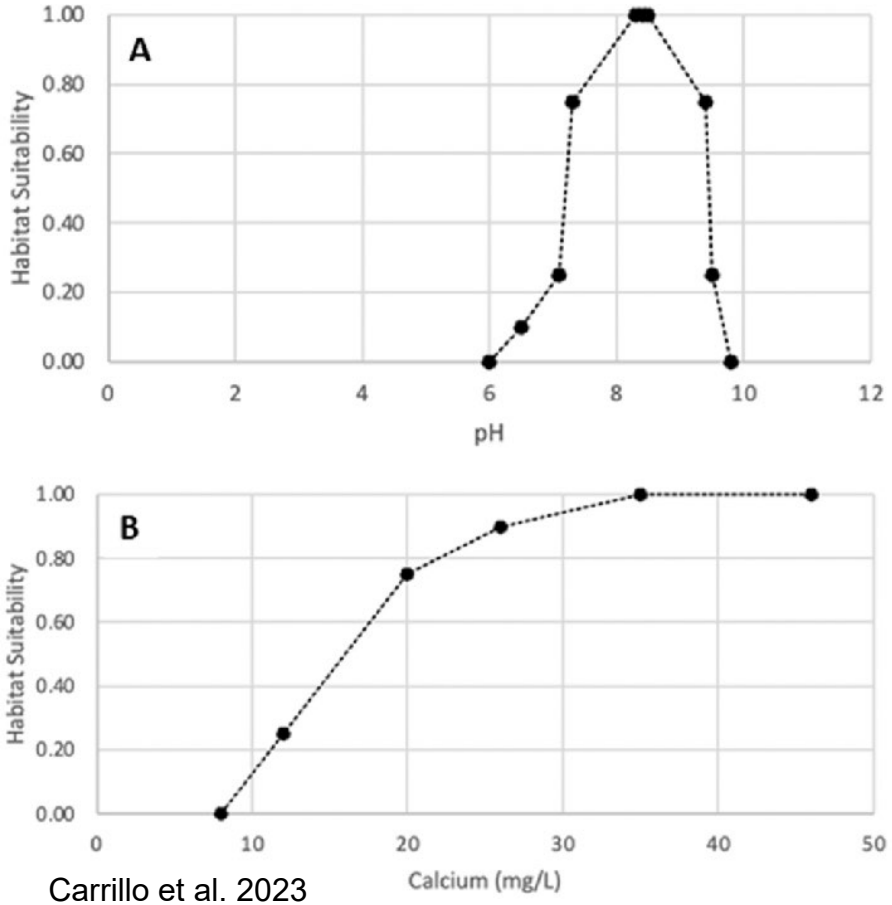
Walker and Wardle 2014

Functional forms of equations

How should relationships be quantified?



Herman et al. 2019

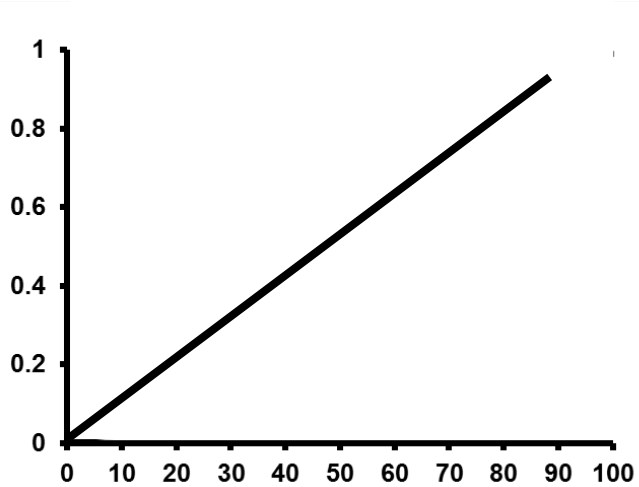


Carrillo et al. 2023

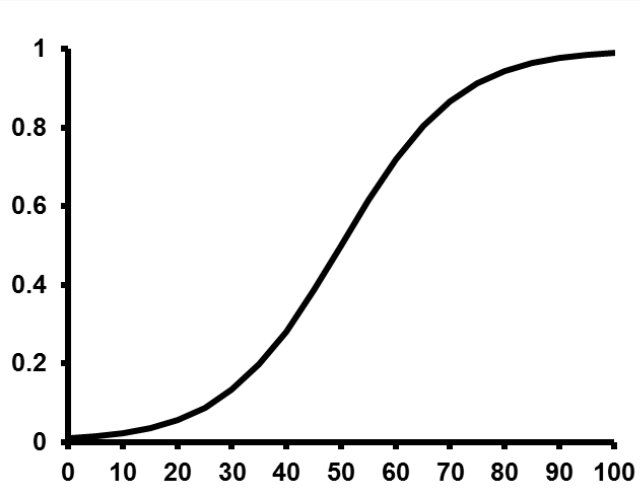
What if functional forms are unknown?

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



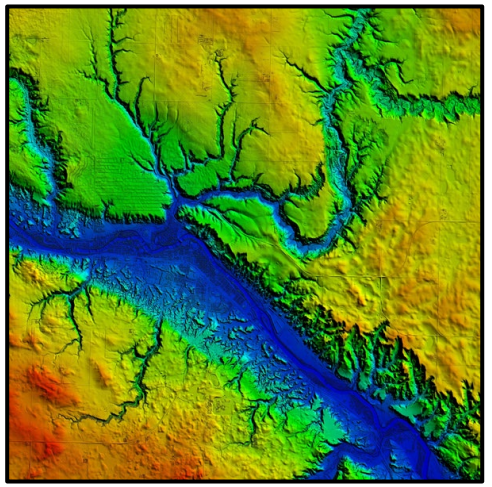
Linear functions: simplest relationship; the general relationship between two variables is understood (*e.g.,* variable A increases when variable B decreases), but the exact form is not



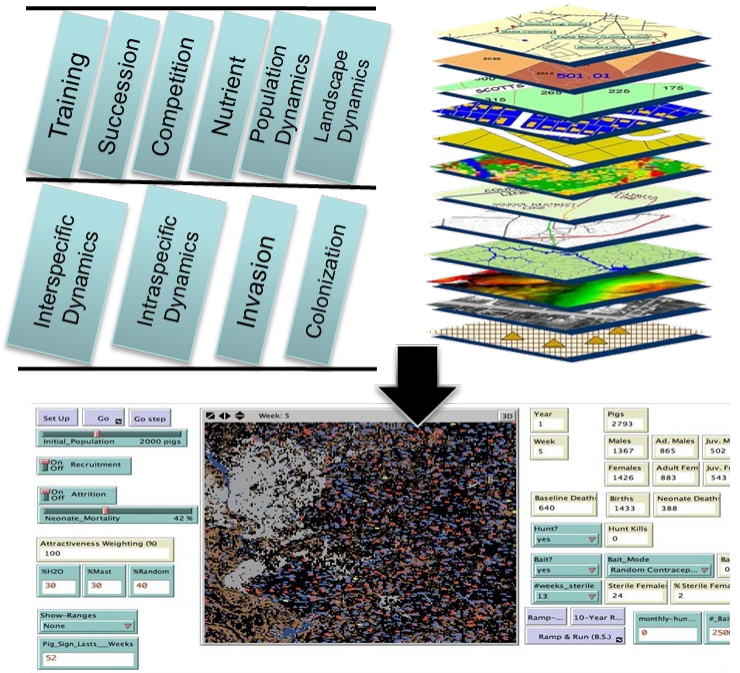
Logistic functions: more complex; allows threshold effects and periods of stasis and rapid change

Types of data and parameterizations

- **Quantitative data**
  - Field work
  - Remotely sensed data
  - Other models
  - Literature
  - Theory
- **Semi-quantitative data**
  - Ranked data
  - Indices
- **Qualitative data**
  - Expert opinion
  - Hypotheses
- **The model itself**
  - Experimenting with a model can reveal trends and patterns



Manitoba Land Initiative

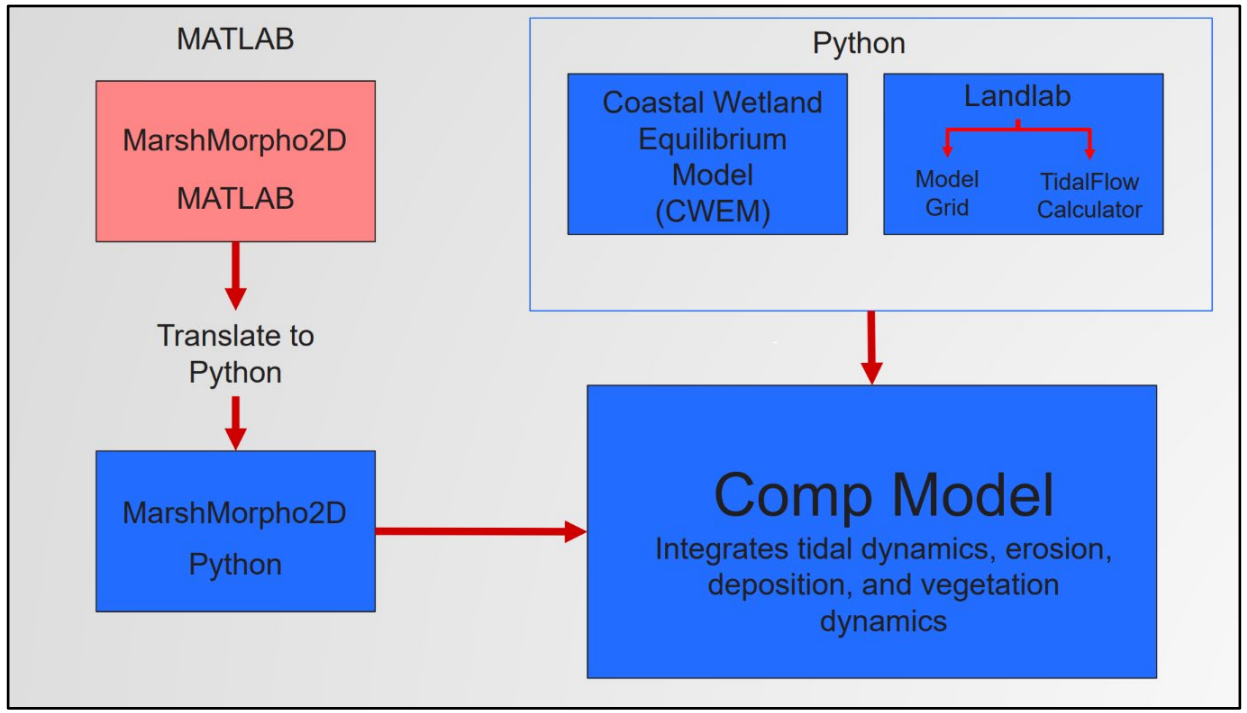


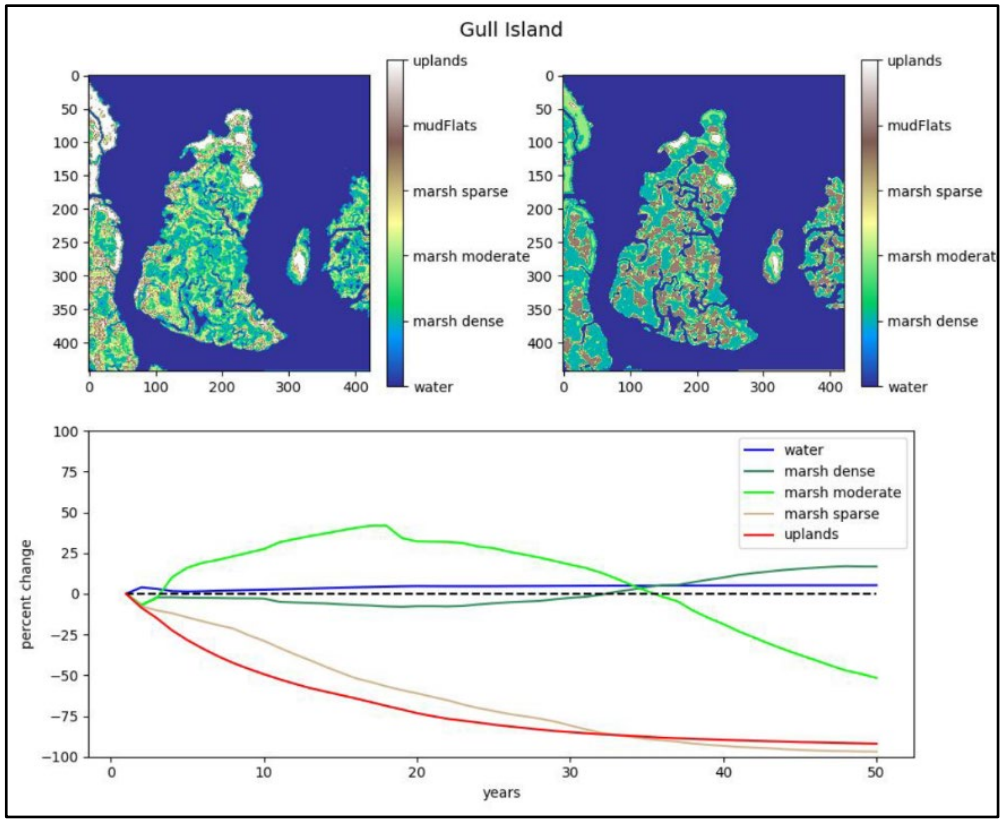
Swannack and Westervelt 2011

Integrated models

Integrated models are models composed of multiple models

- USACE modeling generally combines hydrodynamic & ecological models

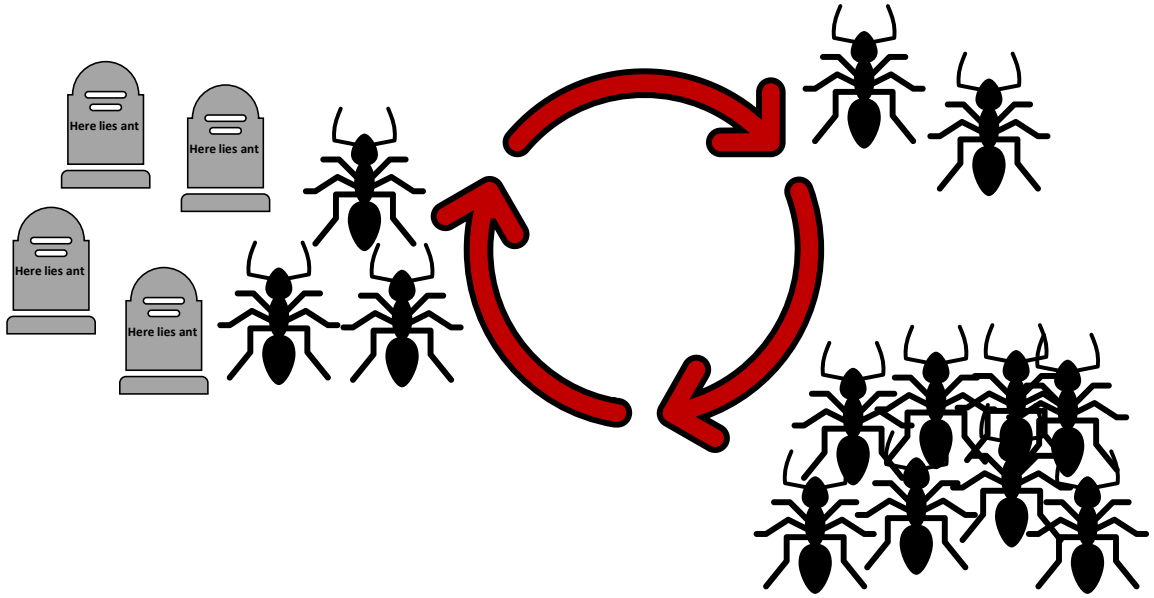




Huff et al. 2024, EMRRP Webinar

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



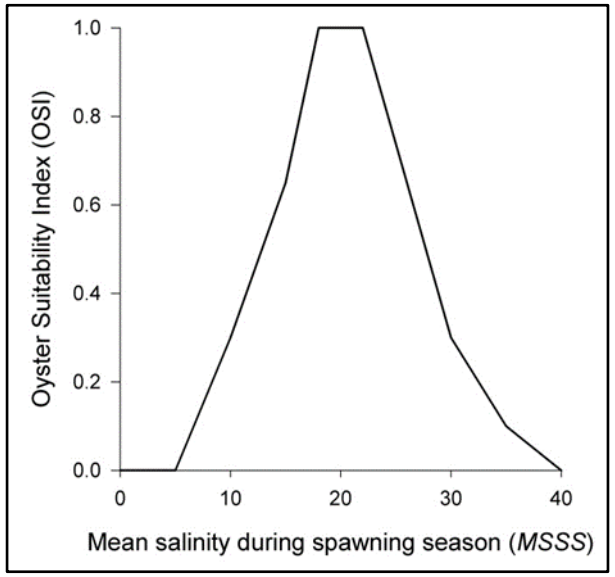
These effects are often difficult to determine precisely in nature

Quantifying thresholds

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

- Equations are almost never reported, but are needed for transparency

Swannack et al. 2014



Typical HSI step function used in planning models

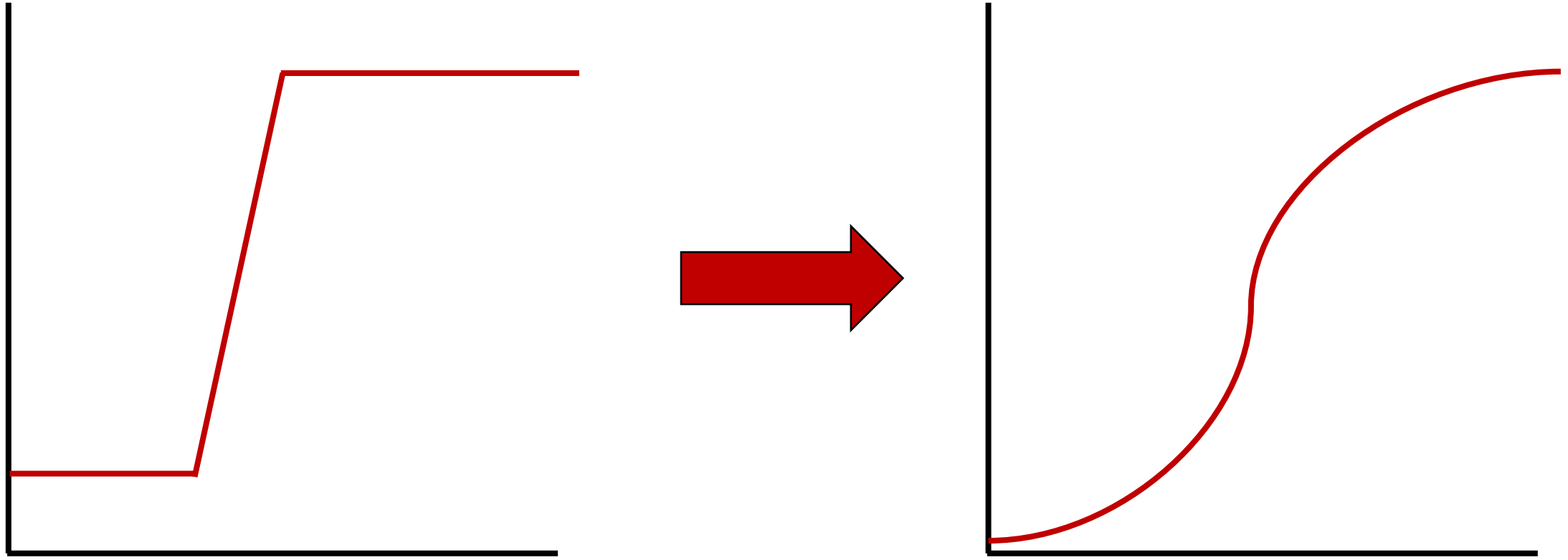
$MSSS \leq 5$ or $MSSS > 40$	$OSI_{MSSS} = 0$
$5 < MSSS \leq 10$	$OSI_{MSSS} = -0.3 + (0.06 * MSSS)$
$10 < MSSS \leq 15$	$OSI_{MSSS} = -0.4 + (0.07 * MSSS)$
$15 < MSSS < 18$	$OSI_{MSSS} = -1.1 + (0.1167 * MSSS)$
$18 \leq MSSS \leq 22$	$OSI_{MSSS} = 1$
$22 < MSSS \leq 30$	$OSI_{MSSS} = 2.925 - (0.0875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

Take advantage of the math! (equations look smarter)

Quantifying thresholds

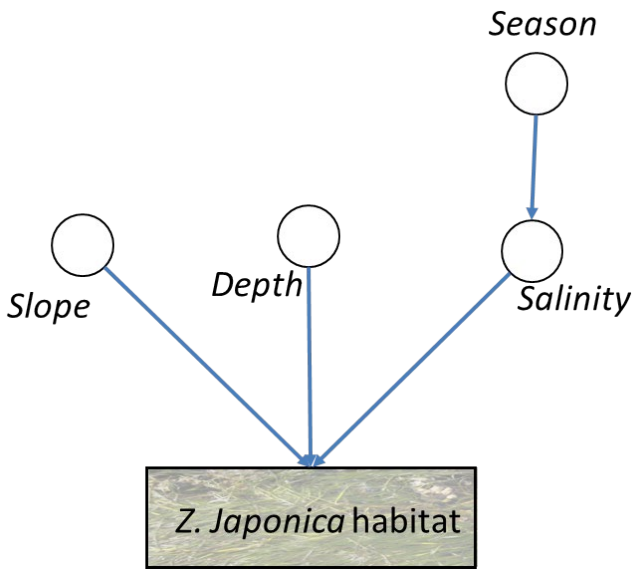
Quickest way is with step-functions or if-then statements

- Equations are almost never reported, but are needed for transparency

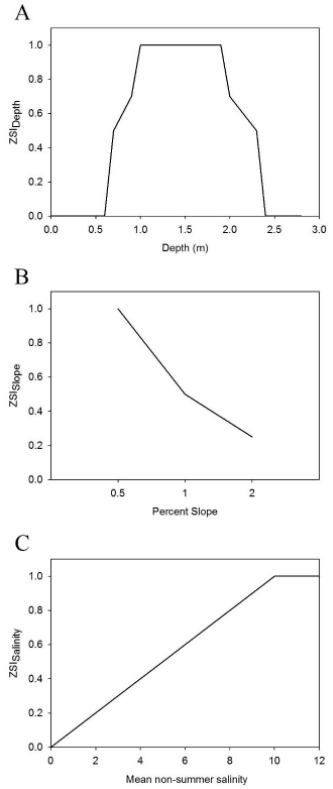


Example: seagrass quantification (Yaquina Bay, OR)

Conceptual (simple, 3 variables)



Functional, captures thresholds



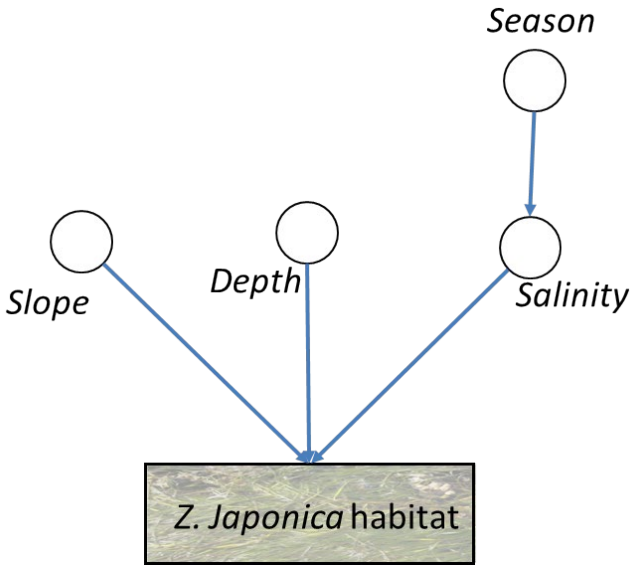
Mathematical, captures breadth of parameter space

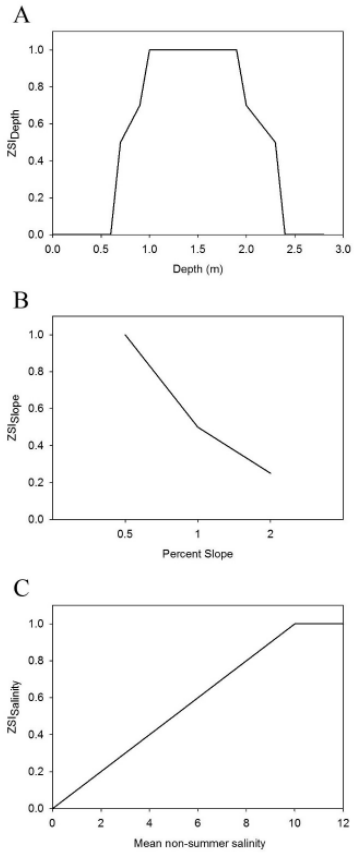
Variable	Equation	Eq#
Depth (m)		
$Depth \leq 0.6$ or $Depth > 2.4$	$ZSI_{Depth} = 0$	(1)
$0.6 < Depth \leq 0.7$	$ZSI_{Depth} = -3 + (5 * Depth)$	(2)
$0.7 < Depth \leq 0.9$	$ZSI_{Depth} = -0.2 + (Depth)$	(3)
$0.9 < Depth \leq 1.0$	$ZSI_{Depth} = -2 + (3 * Depth)$	(4)
$1.0 < Depth \leq 1.9$	$ZSI_{Depth} = 1$	(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		
$Slope \leq 0.5\%$	$ZSI_{\% Slope} = 1$	(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		
$0 \leq Salinity \leq 10$	$ZSI_{Salinity} = 0.1 * Salinity$	(13)
$Salinity > 10$	$ZSI_{Salinity} = 1$	(14)

Shafer et al. 2016

Breaking down models

Three parameter model=14 different equations





Variable	Equation	Eq#
Depth (m)		
$Depth \leq 0.6$ 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		
$Slope \leq 0.5\%$	$ZSI_{\% Slope} = 1$	(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		
$0 \leq Salinity \leq 10$	$ZSI_{Salinity} = 0.1 * Salinity$	(13)
$Salinity > 10$	$ZSI_{Salinity} = 1$	(14)

Conceptual **Functional** **Mathematical**

Shafer et al. 2016

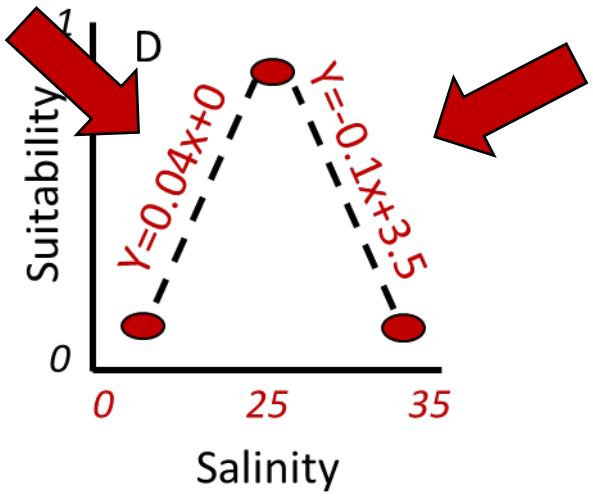
Missing data

- **There are often relationships that aren't defined quantitatively**
  - Must rely on expert opinion
  - Literature
  - Interpolations
- **This is not less rigorous than quantitative data analysis, just less precise**
  - Qualitative data requires increased attention during documentation
- **Will make bigger mistake leaving out important relationships than hypothesizing about relationships**
  - Increased need for transparency

Quantifying expert opinion

Update the model to reflect new knowledge

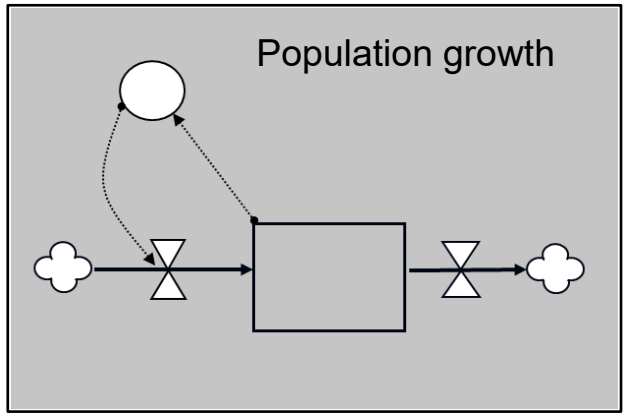
What if we collect more data and the lines aren't straight anymore?

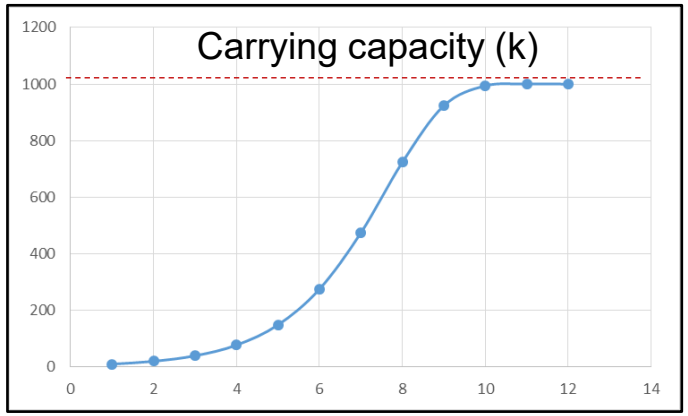


Adapted from Herman et al. 2019

Modeling without data

- Decisions will need to be made, regardless of data availability
- Transparency is important
- Simple functions can help identify magnitude and general trends in the absence of data
- Expert opinion can be used to parameterize equations until other datasets are available

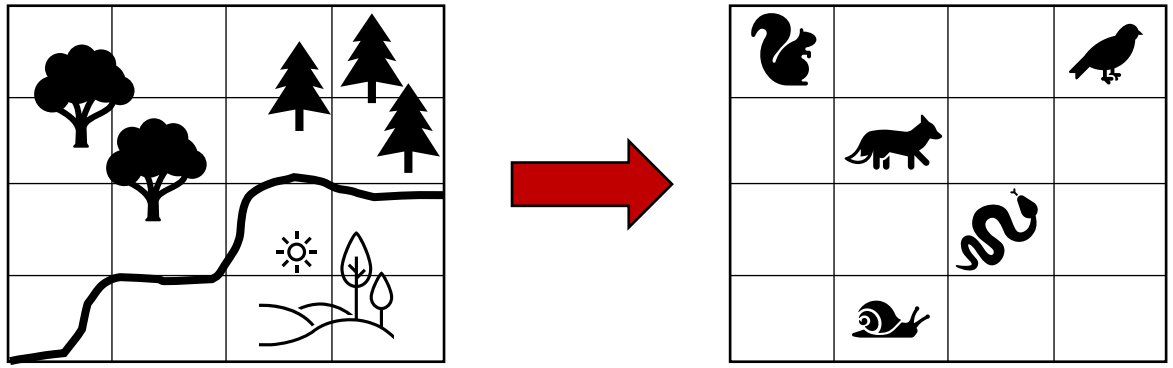




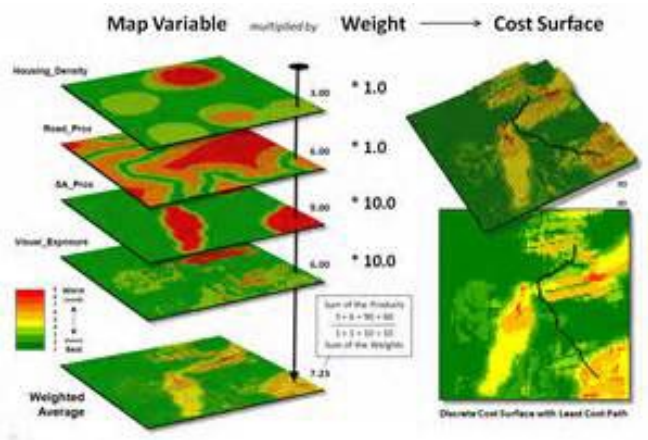
Spatial modeling

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

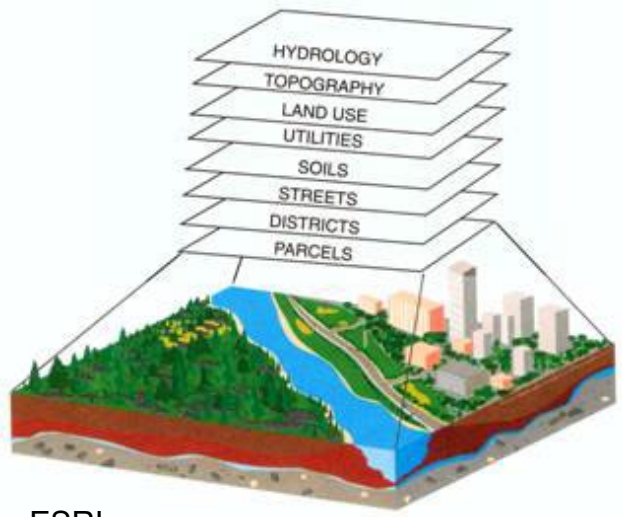
- **Space matters:** configuration and composition of landscapes can affect ecological structure and function



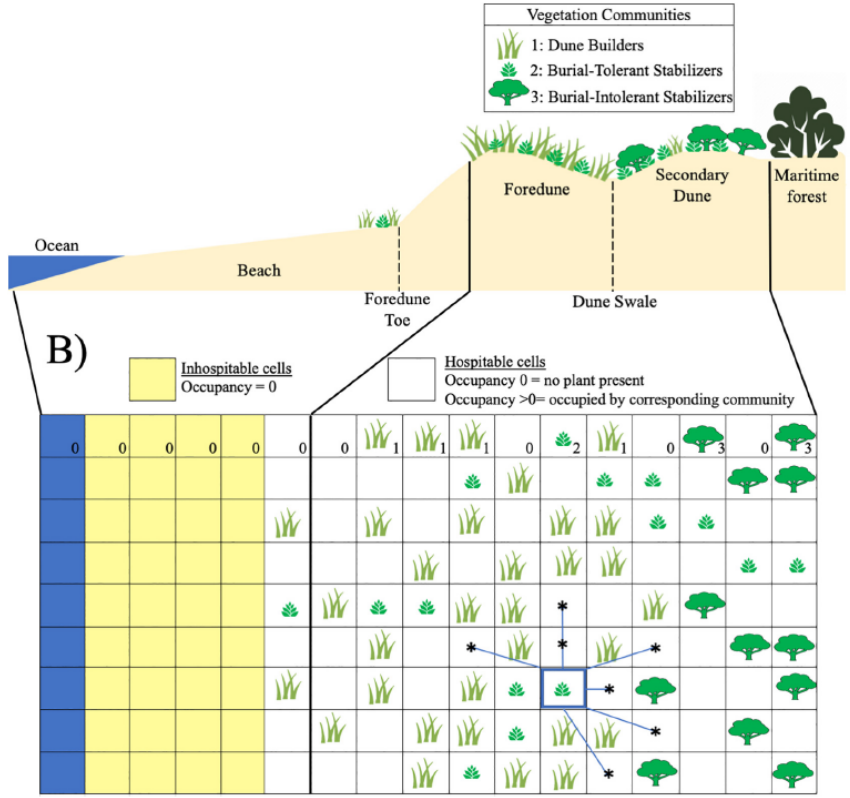
Considering space



Berry 2013



ESRI

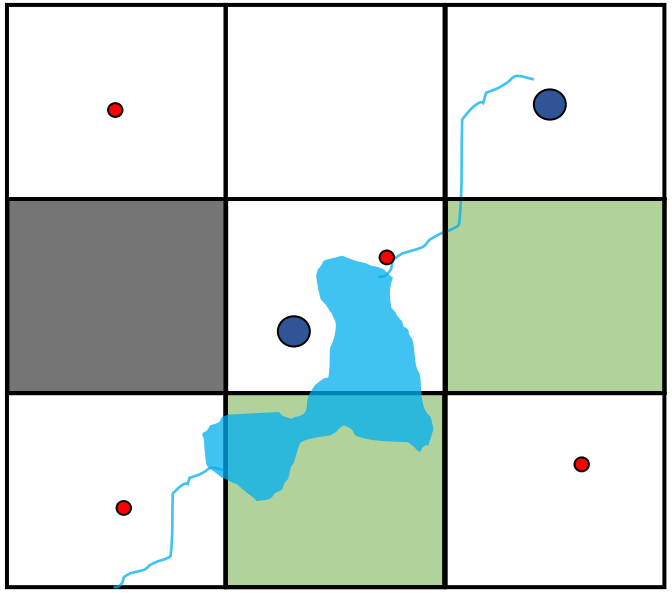


Charbonneau et al. 2022

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.,* grid/DEM/etc.)



Pitfalls of model quantification

- **Choosing inappropriate mathematics & software**
  - Not all formats lend themselves to a given problem
  - Can get trapped by constraints of approach

- **Failing to select an appropriate temporal scale**
  - Too long: violates assumption that change in system is constant b/w time steps
  - Too short: lose interpretability, longer simulation time

Pitfalls of model quantification

- **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 of model quantification

- **Uninterpretable functional relationships/coefficients without meaning**
  - Functional relationships should make sense (within your discipline)
  - Coefficients should reflect the magnitude of the process occurring in nature

- **Failing to consider units of measure**
  - Can violate assumptions and create nonsensical results

Pitfalls of model quantification

- **Lack of 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
- **Failing to 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**

Interactive toolkit for applied modeling (TAM)

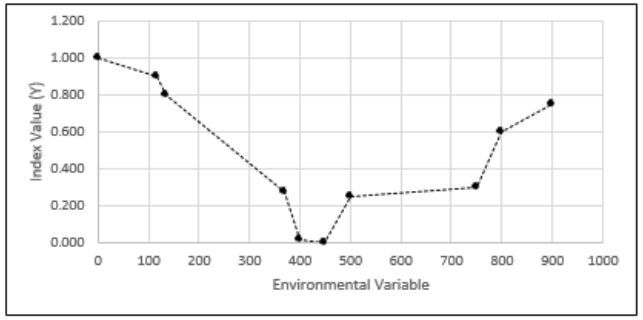
- Platform developed for rapid model development
- Quantifies threshold-based datasets
- Certified for USACE

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

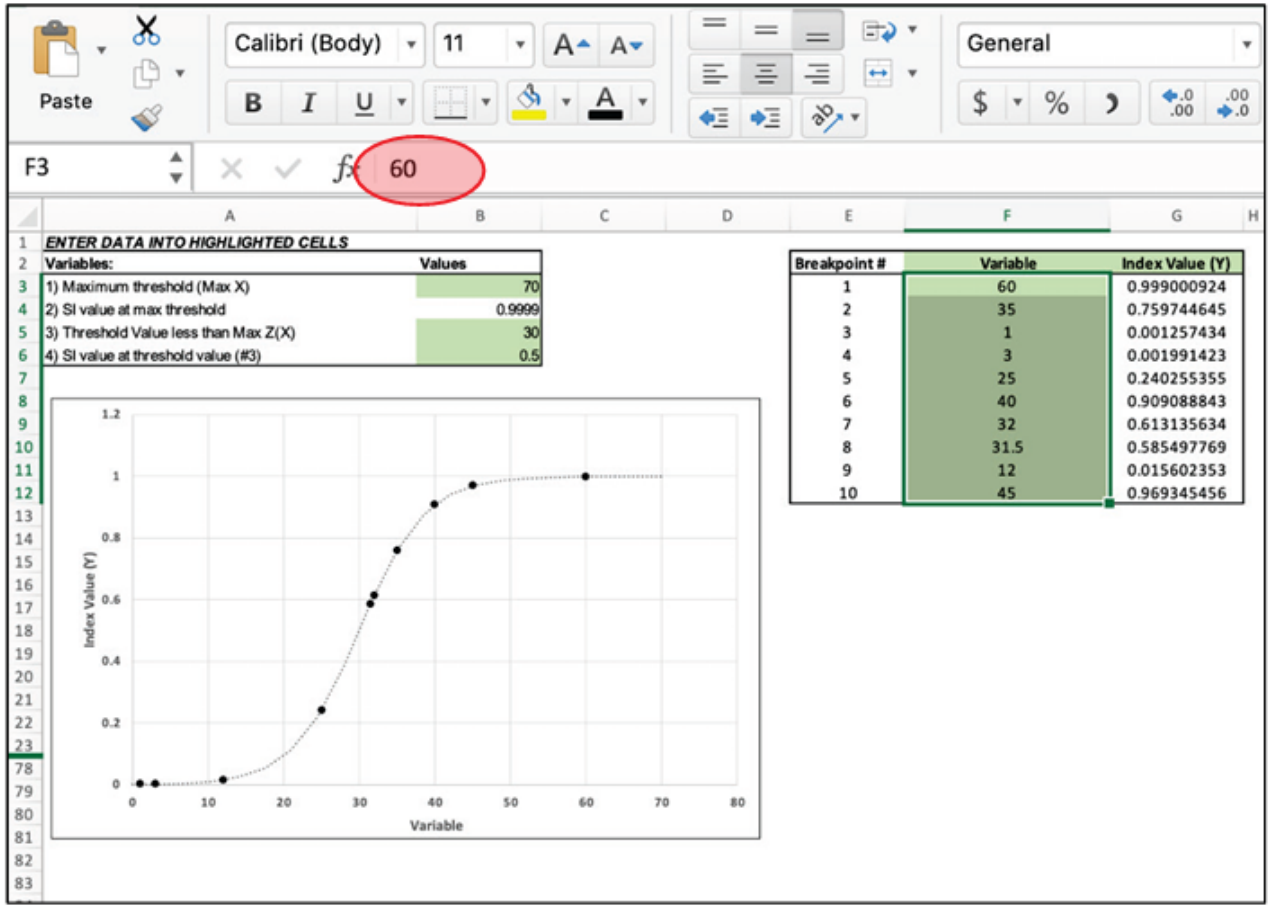
References



Carrillo et al. 2022

TAM example

Quantifying a conceptual relationship using TAM



Carrillo et al. 2022