Writing Hierarchical Models

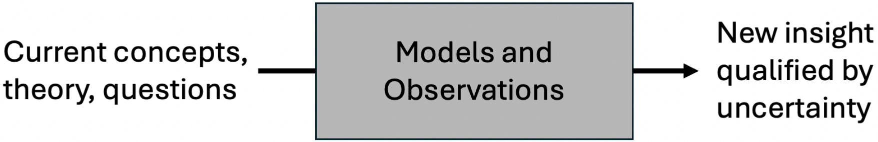
Bayesian Models for Ecologists

Becky Tang

June 09, 2024

DEB 2042028

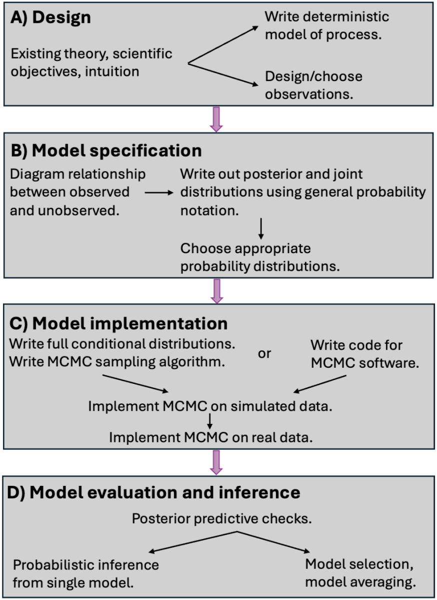
What is this course about?



All modeling problems have idiosyncracies

- Different types of data
- Different deterministic models
- Sampling error in the predictors or responses
- Calibration error for predictors or responses
- Prior knowledge of parameters
- Missing data
- Multiple scales of data (group level effects)
- Prediction and forecasting
- Spatial or temporal dependence
- Derived quantities

The Bayesian Method



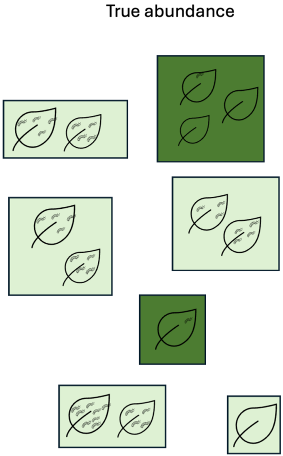
Steps in writing Bayesian models

1. Write your deterministic model. Be careful about support.
2. Draw Bayesian network (DAG) describing relationships between observed and unobserved quantities (including predictors, if it's useful to you).
3. Use the Bayesian network to write proportionality between posterior and joint distributions using bracket notation $[\ |\ ]$.
   a. Posterior distribution: $[\text{unobserved quantities} \mid \text{data}]$
   b. Joint distribution
      i. All nodes in Bayesian network at the heads of solid arrows (children) must be on the left hand side of a conditioning symbol $|$.
      ii. All nodes in Bayesian network at the tails of solid arrows (parents) must be on the right hand side of a conditioning symbol $|$.
      iii. All nodes at the end of an arrow with no arrow coming into them must be expressed unconditionally, i.e., they must have numeric arguments
4. Assign specific PDF or PMF to each of the brackets.
5. Choose numeric values for parameters of prior distributions. Do this sensibly! Do not default to vague priors. (Do as I say, not as I do.)

Sources of uncertainty

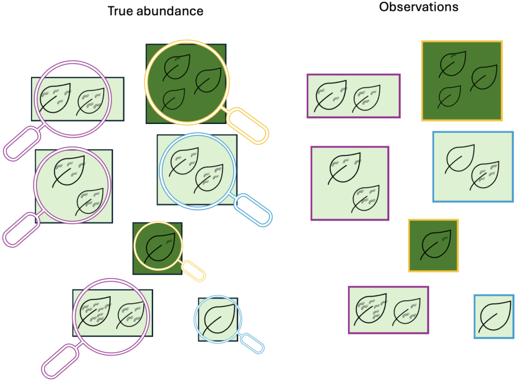
- **Process variance** arises because *our model* fails to represent all of the sources of variation in the quantity we seek to understand.
- **Observation variance** arises because *what we observe* does not perfectly represent the true, unobserved quantity we seek to understand.
  - Sampling error
  - Observer/calibration error (typically requires repeated measurements to estimate)

Example of sources of uncertainty



- True abundance per unit area z_i of herbivore on host plant i
- Background color corresponds to environmental condition
- Process model: $[z|g(x,\beta),\sigma_p^2]$

Example cont.



- y_{ij} = observed abundance per area at location i by observer j
  - Outline color denotes observer
- Maybe there is both sampling error σ_s^2 and observer error/variation σ_o^2

A note on uncertainty

- Incorporating uncertainty transforms a deterministic model into a probabilistic model.
- It's important to properly partition/attribute uncertainty into different sources, However, not all three sources are necessarily present.
    - Think about the design and assumptions of your study

Board work

Practice forming hierarchical models with different sources of error.