Model evaluation - Day 5

CoSMo 2017
Gunnar Blohm

"All models are wrong, but some are useful" (Box 1976)

"The words *true model* represent an oxymoron" (Burnham & Anderson 2002)

Success?

- You have done everything right…
  - Chose right toolkit
  - Built model (s)
  - Fit to data
  - Selected the right number of parameters / the best model
  - Estimated parameters and bootstrapped 95% CIs
  - Cross-validated
  - Compared your best model to alternatives
  - …
- Does that mean you have a good model now?
- How can we know?

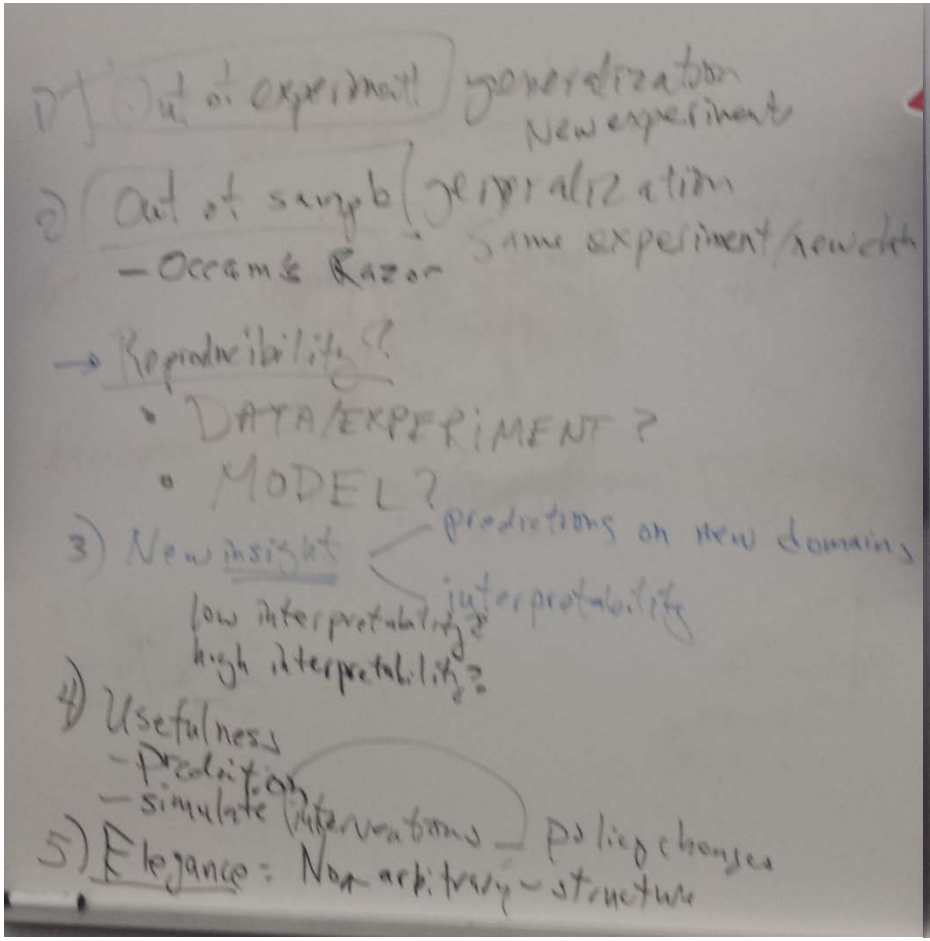
What's a good model? – see philosophy

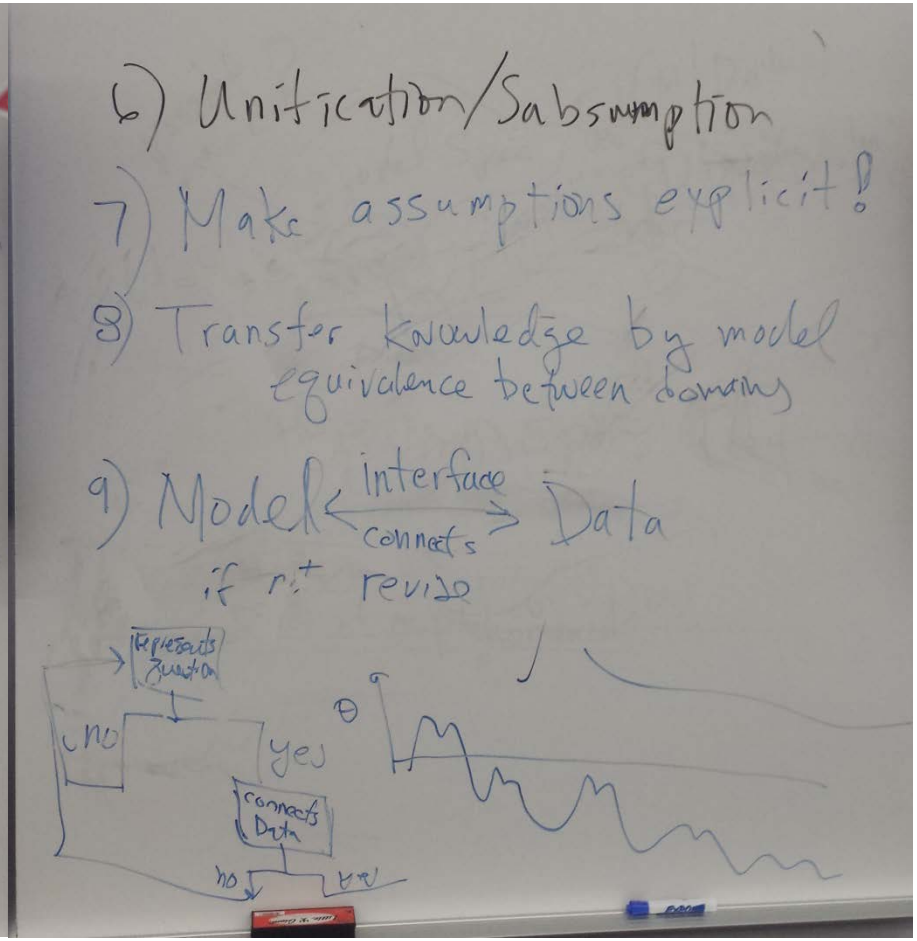
- Models are potentially good if they
  - Answer our research question (meet the goal) - pragmatism
  - Capture general principles
  - Advance our mechanistic understanding (beyond what's directly observable!)
  - Make testable predictions, i.e. reduce the space of potential answers to a specific question
  - Save animal lives (bc only a few experiments are needed)
  - Generalize! (many data sets)
  - Have a practical use (e.g. simulation of drug effects on behavior)
  - Are falsifiable (Popper)
  - …

What's a good model?

- Models should
  - Be as simple as possible and as complex as necessary – cf. Occam's razor ("shave away all that is unnecessary")
    - Principle of parsimony (Box & Jenkins 1970)
    - Bias-variance trade-off with # model parameters
  - Gain mechanistic understanding of an experimental phenomenon
  - Identify hypotheses, assumptions, unknowns
    - Make them explicit
  - Make quantitative predictions
    - E.g. parameters with biological interpretation
  - Build a theoretical brain as a model of the real brain (stroke lesions etc)
  - Inspire new technologies, facilitate translation
  - Models of diseases to help treatment, rehabilitation, quality of life

Brainstorming





"Everything should be made as simple as possible, but no simpler" (Einstein)

What's a good model?

- Interfacing with data
  - The model needs to be interpretable with respect to the data
  - Explicit representation of potentially measurable variables
  - Model needs to enable answering the questions at hand
  - (Data needs to be diverse enough!)

- Representing causal linkages
  - Bridging gap between neural properties, computational objectives and behaviour
  - Causal effect of manipulations (pharma, lesion, conditions, …)
  - Computational models are the only way!

Hard problems

- Model formulation / specification
  - How to find an appropriate set of candidate models?
  - The most original, innovative part of scientific work is the phase leading to the proper question
  - Requires critical thinking, intuition, creativity
- When is a model falsified?
  - When are contradictory experimental results due to experimental design problems?
  - How much evidence is needed to be sure?
    - Independent confirmation?
  - Can we compute p(model | all data)?
  - Credit assignment in hierarchical models – which part is wrong?