Types of Models: WHAT's in the BOX

Conceptual.............Mathematical

Static…………..Dynamic :*TIME*

Lumped.............Spatially Distributed: *SPACE*

Stochastic.............Deterministic

Abstract…………..Physically / Process Based

any of these could be placed in an calibration, optimization context!

Model Performance

Defining objective functions - choose a function or set of functions that evaluate how well model and observed fit each other

- type of metric (R2, for, percent error, “fuzzy”)
- time-space scale, extent of the comparison

Multiplicative approach

Metric A * Metric B …

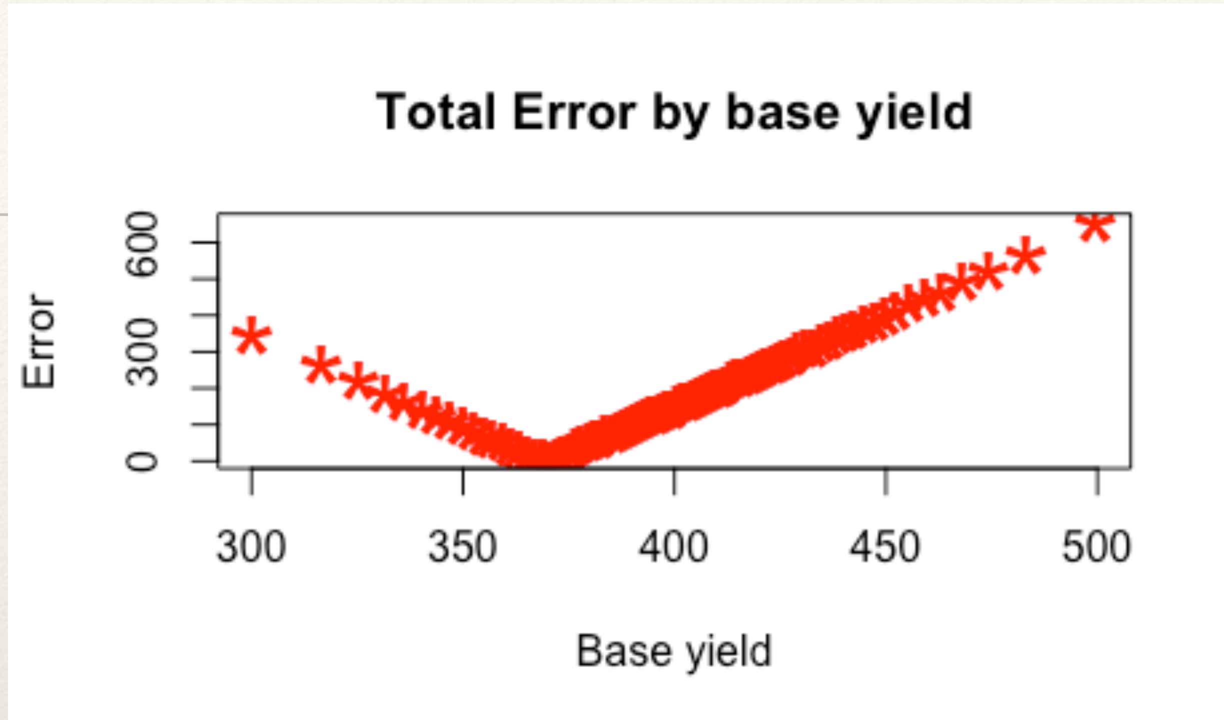
Metric A * weighting A + Metric B * weighting B…

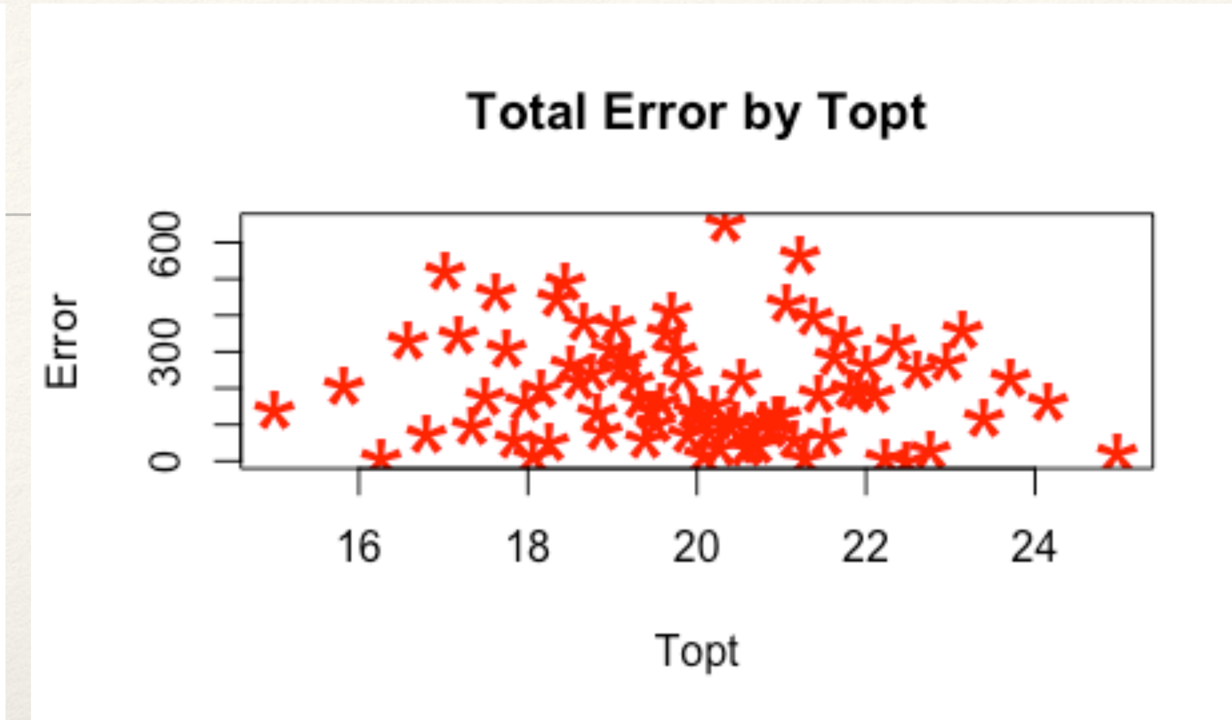
Calibration-Optimization

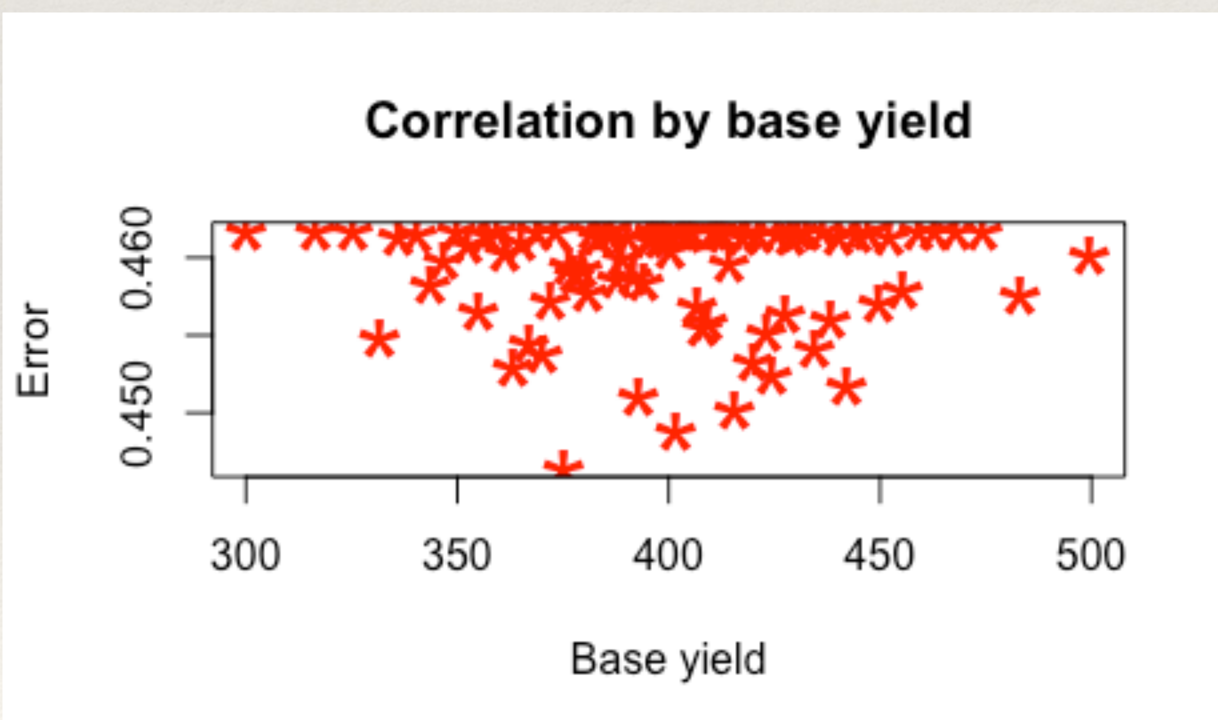
- choosing parameter sets to use based on comparison with observed data
- calibration is very similar to sensitivity analysis
  - we could use LHS/Sobel function to generate parameter sets and model runs
  - compute performance metrics for each run
  - graph and decide on a ‘cut off point’ of ‘acceptable parameters
- optimization
  - a way to to calibration - search procedure
  - solving problems

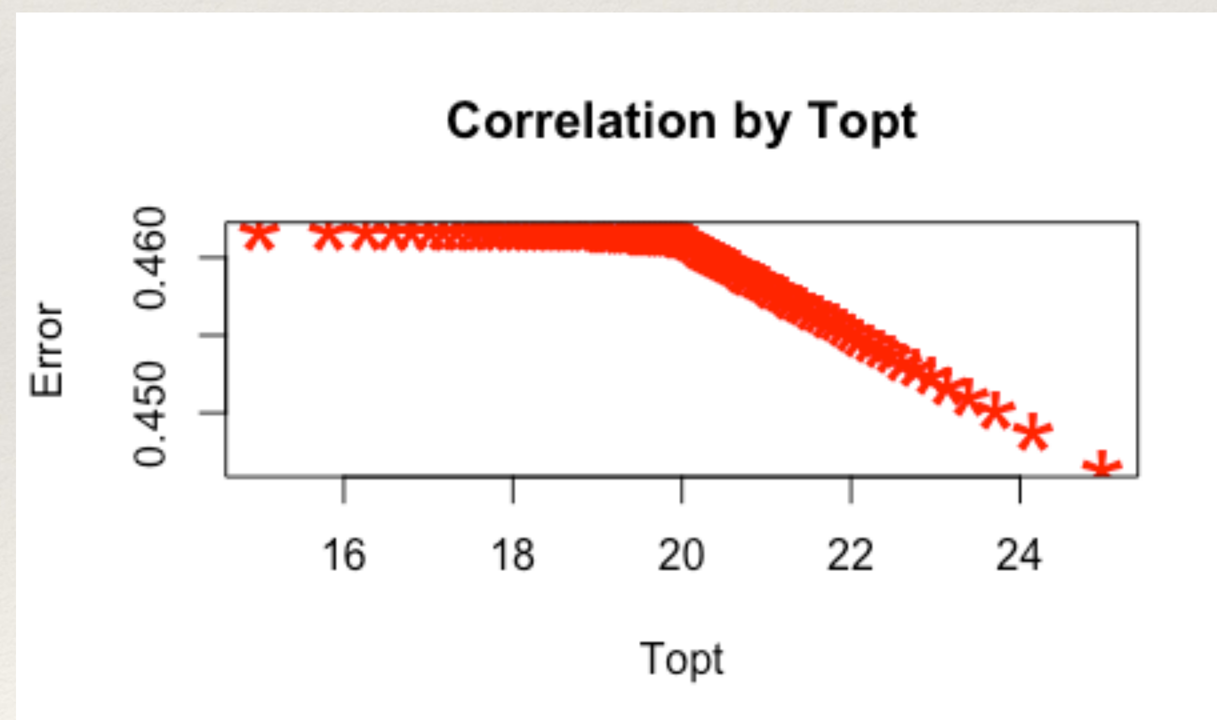
Calibration

- Collect observed data, and known parameters and inputs
- Determine how you will evaluate the function (e.g percent error, ranked correlation?) and time/space scales for evaluation
- Write a function to evaluate your model results over a range of parameter values
- Generate parameter sets (LHS, Sobel, other)
- Run evaluation function
- Look at sensitivity of performance measures to parameter values
- Decide which parameters are ‘functional”







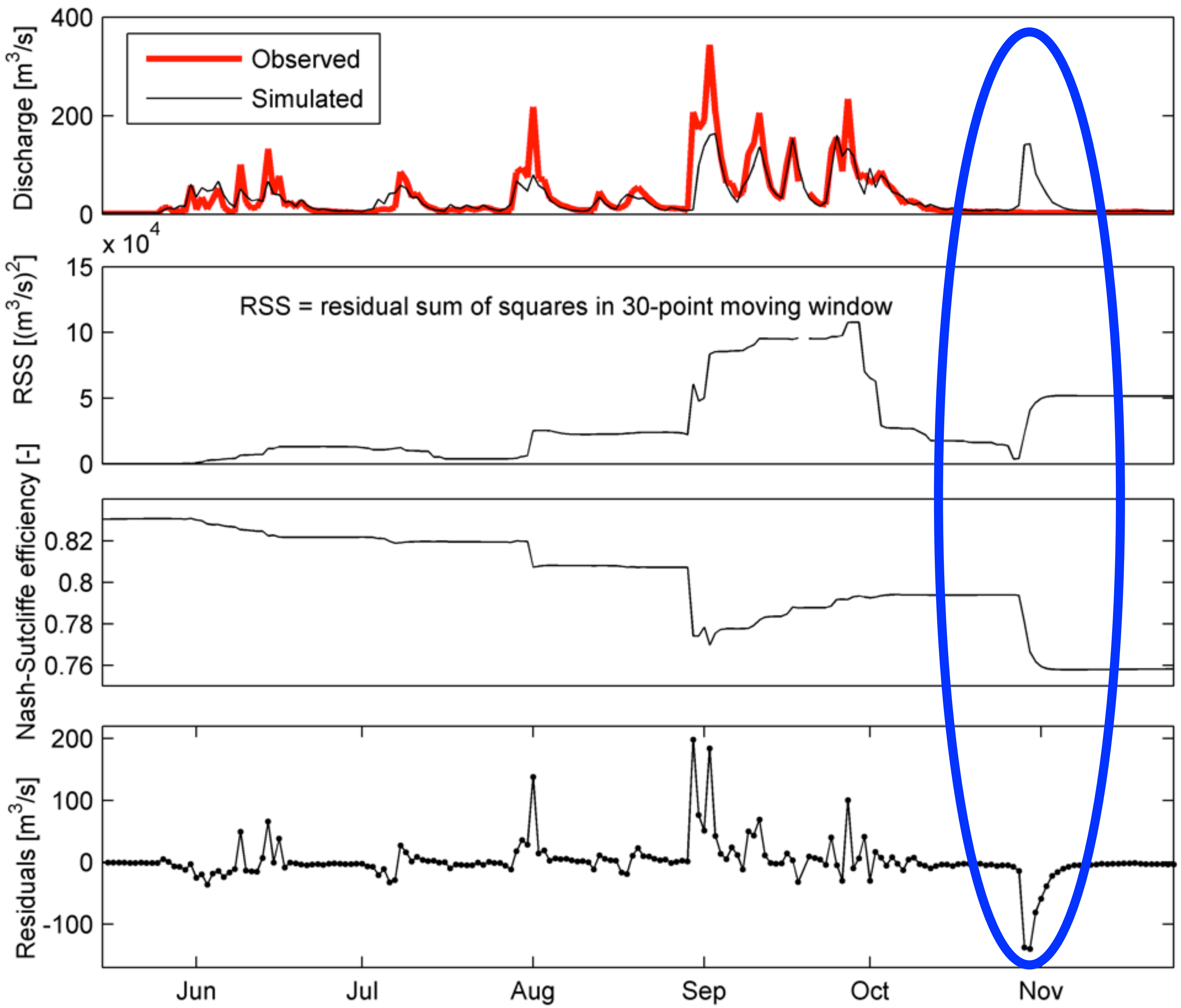


Note how different metrics help to define different parameters

Equifinality

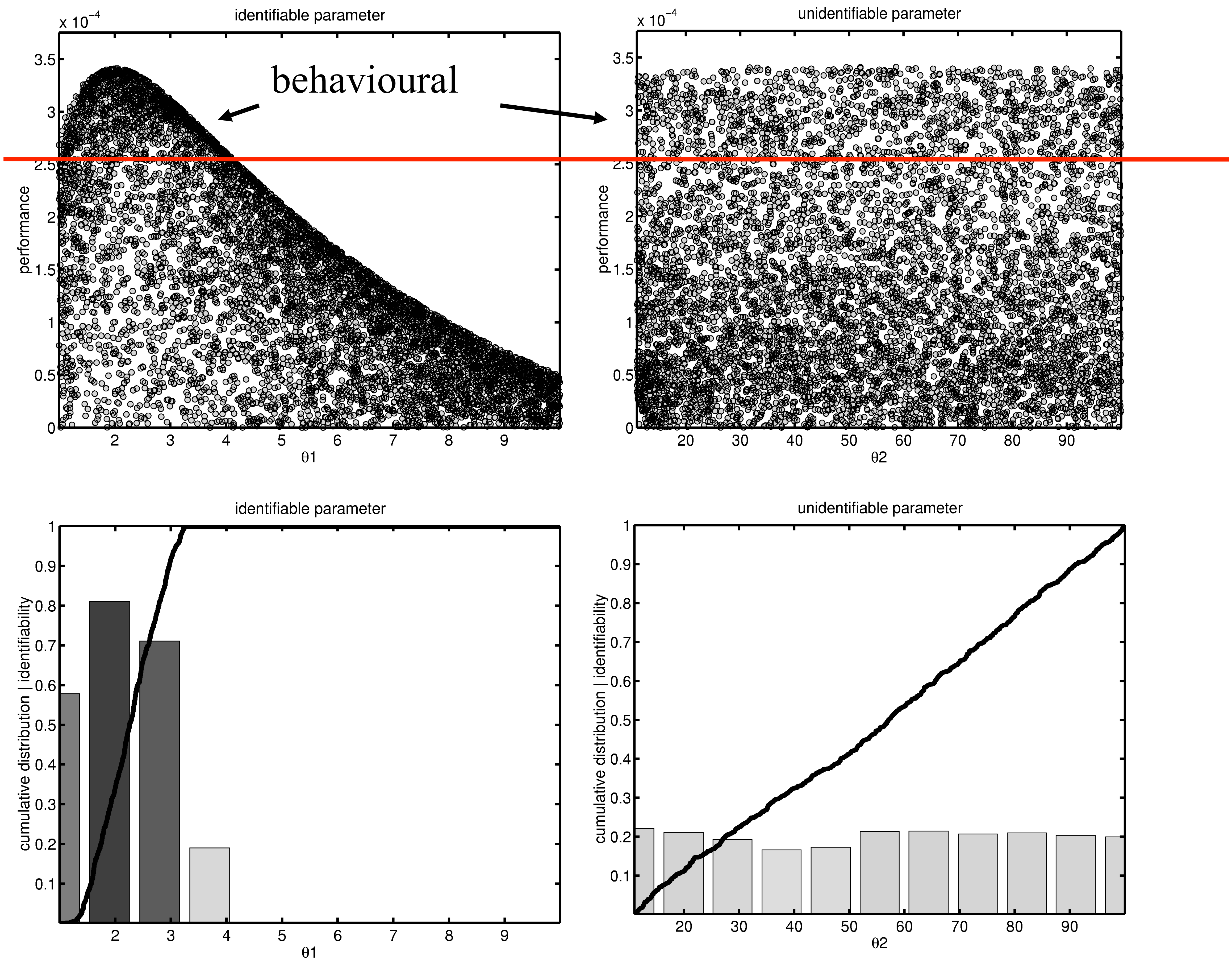
- Many parameter values can give equal performance
- And comparison with observed data is limited
    - errors in observed data
    - over-fitting to the particular time/place/set of observations available

Disinformation in calibration data

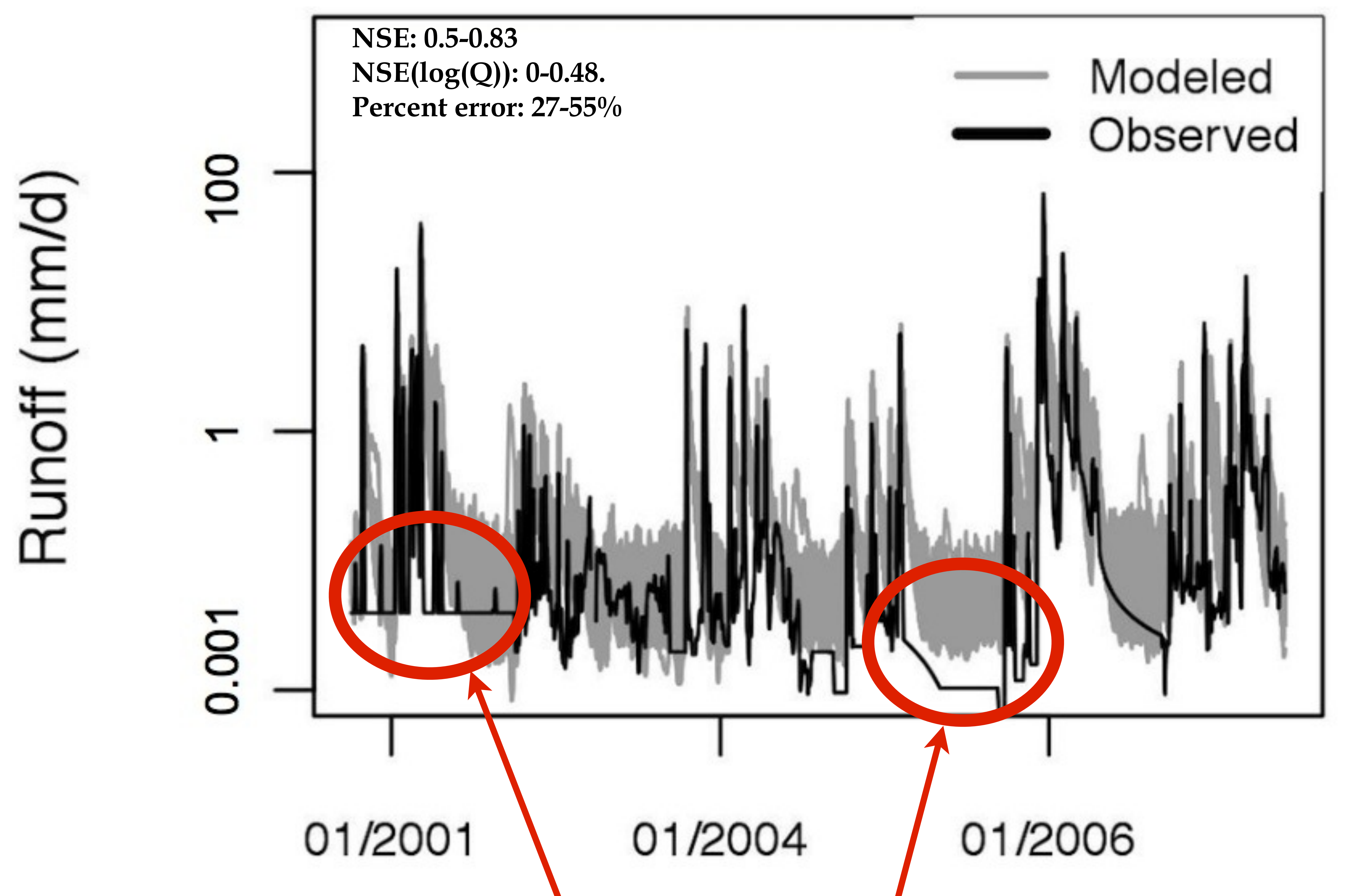


Application of WASMOD to Pasa La Ceiba, Honduras
(from Ida Westerberg, Uppsala)

Dotty Plots and Identifiability Analysis



Other Issues:Observed Error



Error in observed streamflow can bias calibration:
Remove those periods from calibration

Problem:

model selection based on performance AND calibration

- Parameter optimization/evaluation: will not be robust
- calibration period
- performance measure
- input/measurement errors
- concept of equifinality

Beven, JH, 2006, Manifesto for the Equifinality Thesis

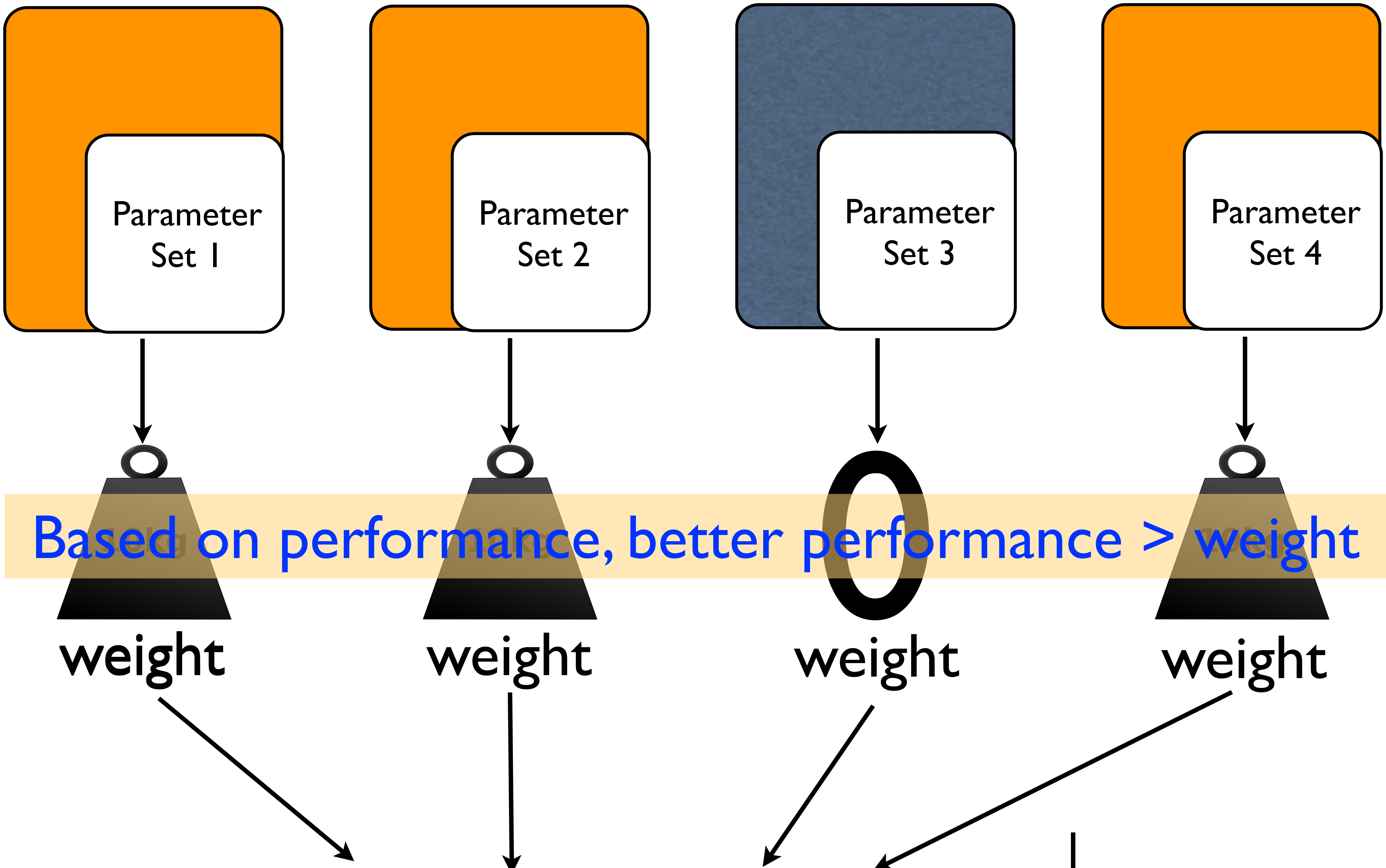
A possible solution:

Generalized Likelihood Uncertainty Estimation (GLUE)

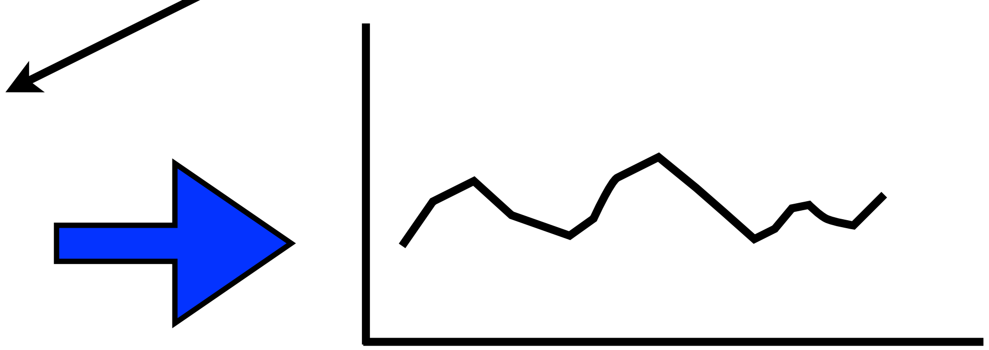
- assess the likelihood of different models + parameters being good predictors of the system of interest
- reject (give zero likelihood) those models that are clearly not good predictors of calibration data
- Can be done with different model structures as well as different parameter sets

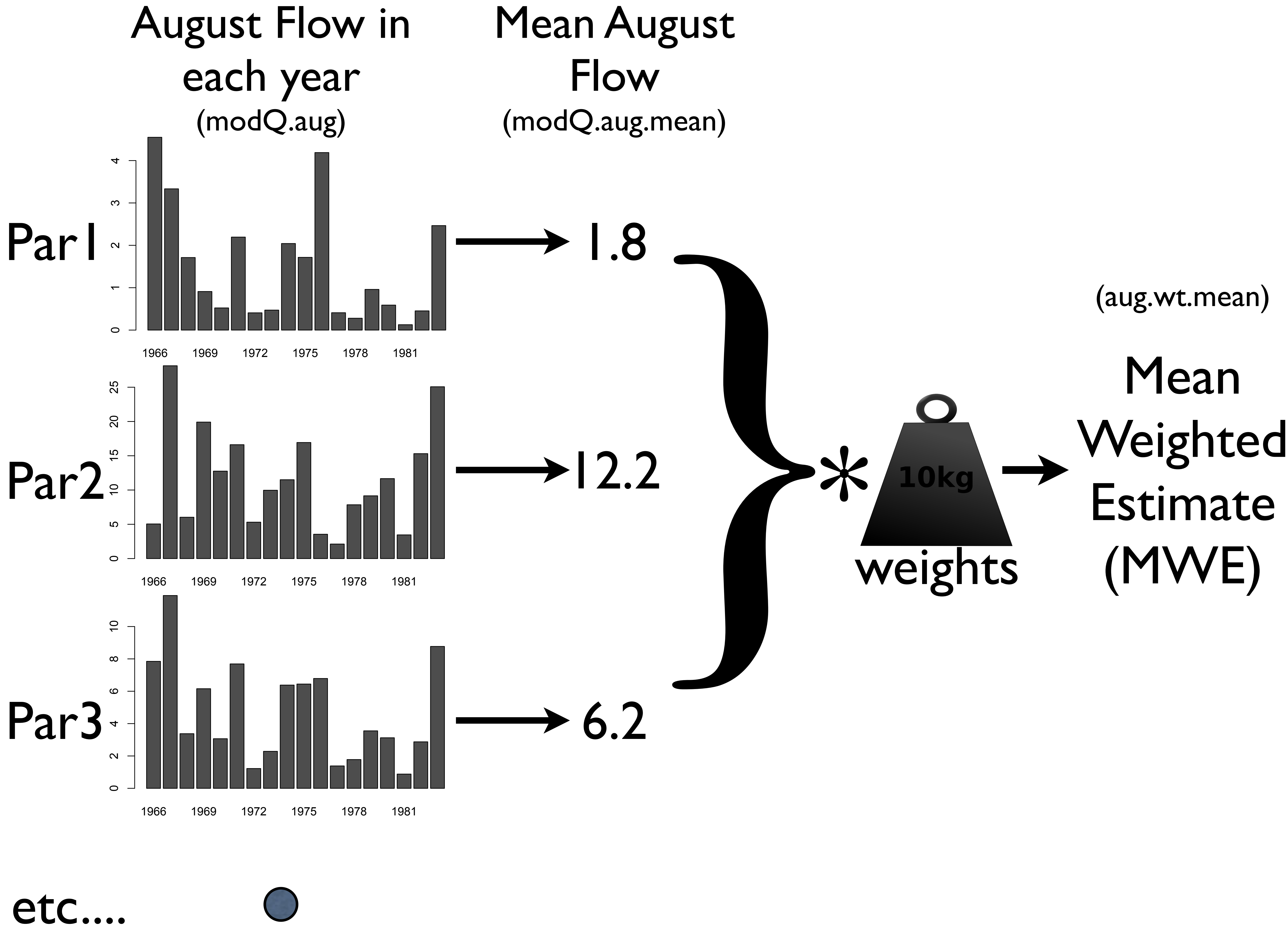
What to do about Equifinality

- Keep all parameter sets that are acceptable
  - acceptable: above some threshold of performance
  - always run the model for those parameter sets and use range of model output to define uncertainty bound
  - if you need a single model estimate:
    - combine results from all acceptable parameters
      - average
      - weight by performance



$$MWE = \frac{\sum output_i * weight_i}{\sum weights_i}$$




August Flow in each year
(modQ.aug)
Mean August Flow
(modQ.aug.mean)
Par1
1.8
Par2
12.2
Par3
6.2
etc....
*
10kg
weights
(aug.wt.mean)
Mean Weighted Estimate (MWE)

Calibration - Beyond GLUE:

Sampling the parameter space - Optimization and Formal Bayesian analysis

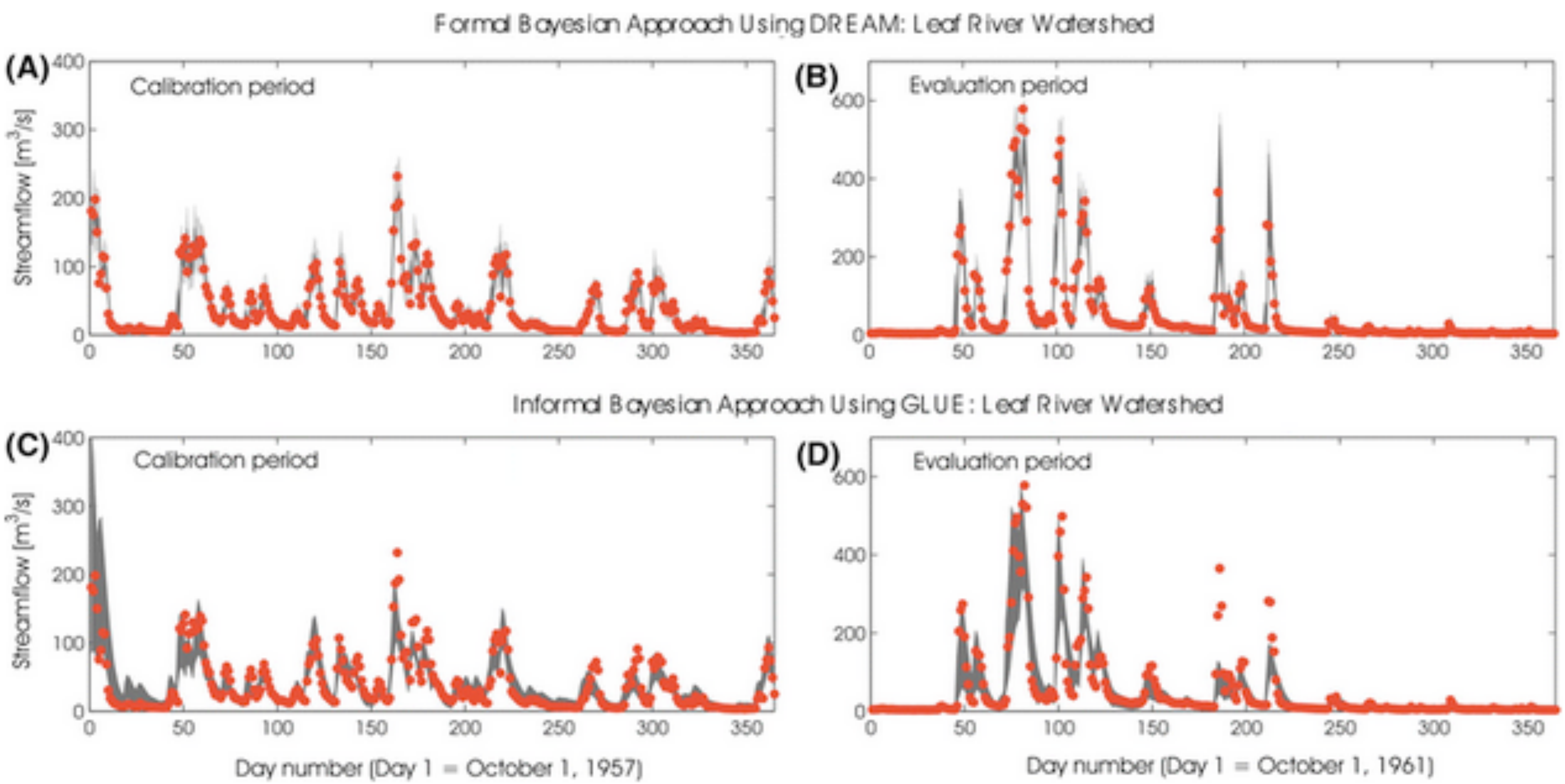


Fig. 5

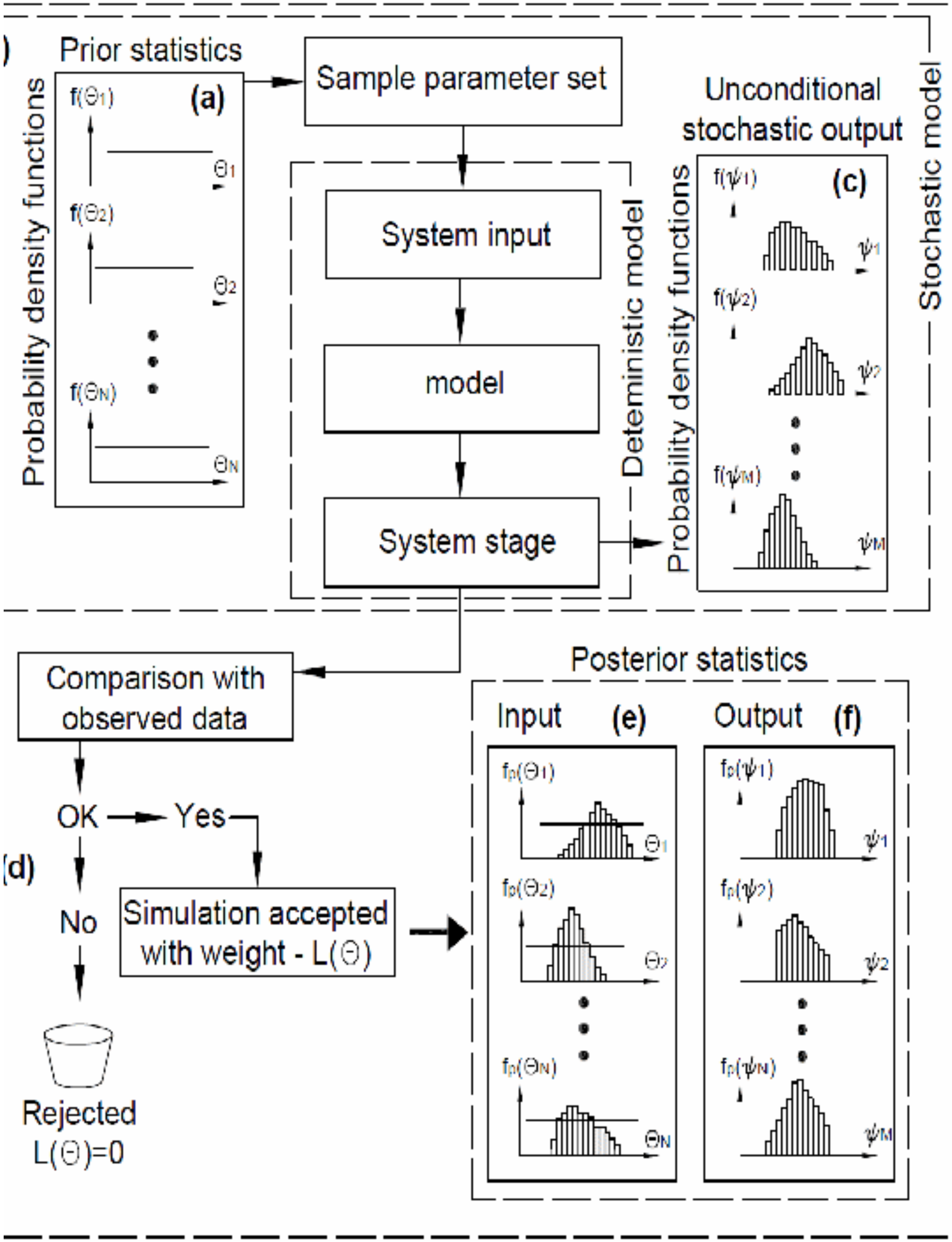
Streamflow prediction uncertainty ranges derived with DREAM (*top panels*) and GLUE (*bottom panels*) for a representative portion of the calibration (*left column*) and evaluation period (*right column*) for the Leaf River watershed. In each DREAM graph, the *dark gray* region represents the 95% confidence intervals of the output prediction due to parameter uncertainty, whereas the *light gray* region represents the additional 95% ranges of the prediction uncertainty. For GLUE the 95% prediction quantiles are presented. The *solid circles* denote the streamflow observations

J.A. Vrugt, C.J.F. ter Braak, H.V. Gupta, B.A. Robinson Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? Stochastic Environmental Research and Risk Assessment, 44 (2008)http://dx.doi.org.proxy.library.ucsb.edu:2048/10.1007/s00477-008-0274-

Generally found similar results using GLUE and formal MCMC calibration and uncertainty estimation

GLUE: Generalized Likelihood Uncertainty Estimation

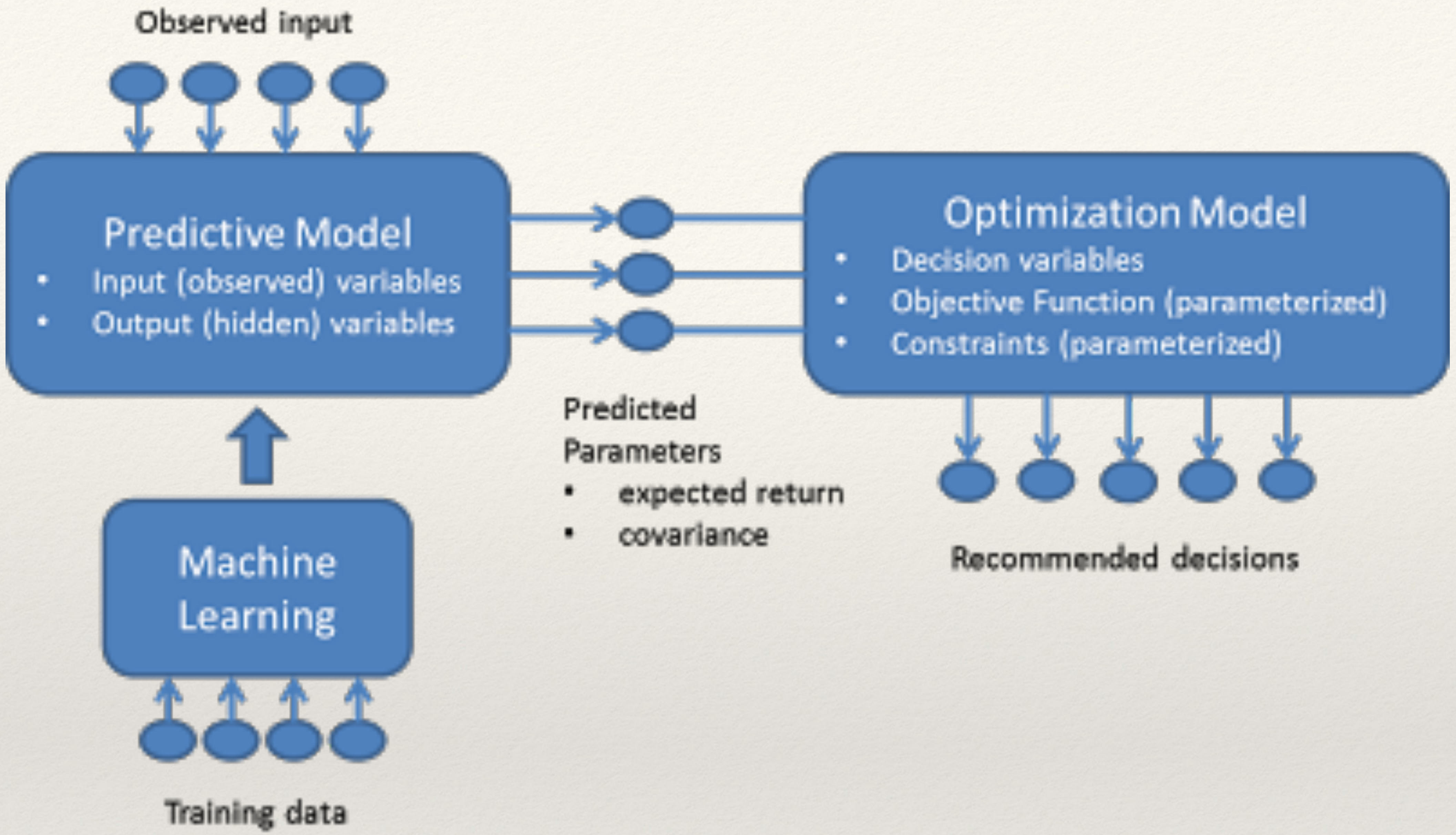
K. Beven and A. Binley, "The future of distributed models: model calibration and uncertainty prediction," Hydrological Processes, vol. 6, no. 3, pp. 279–298, 1992. View at Scopus



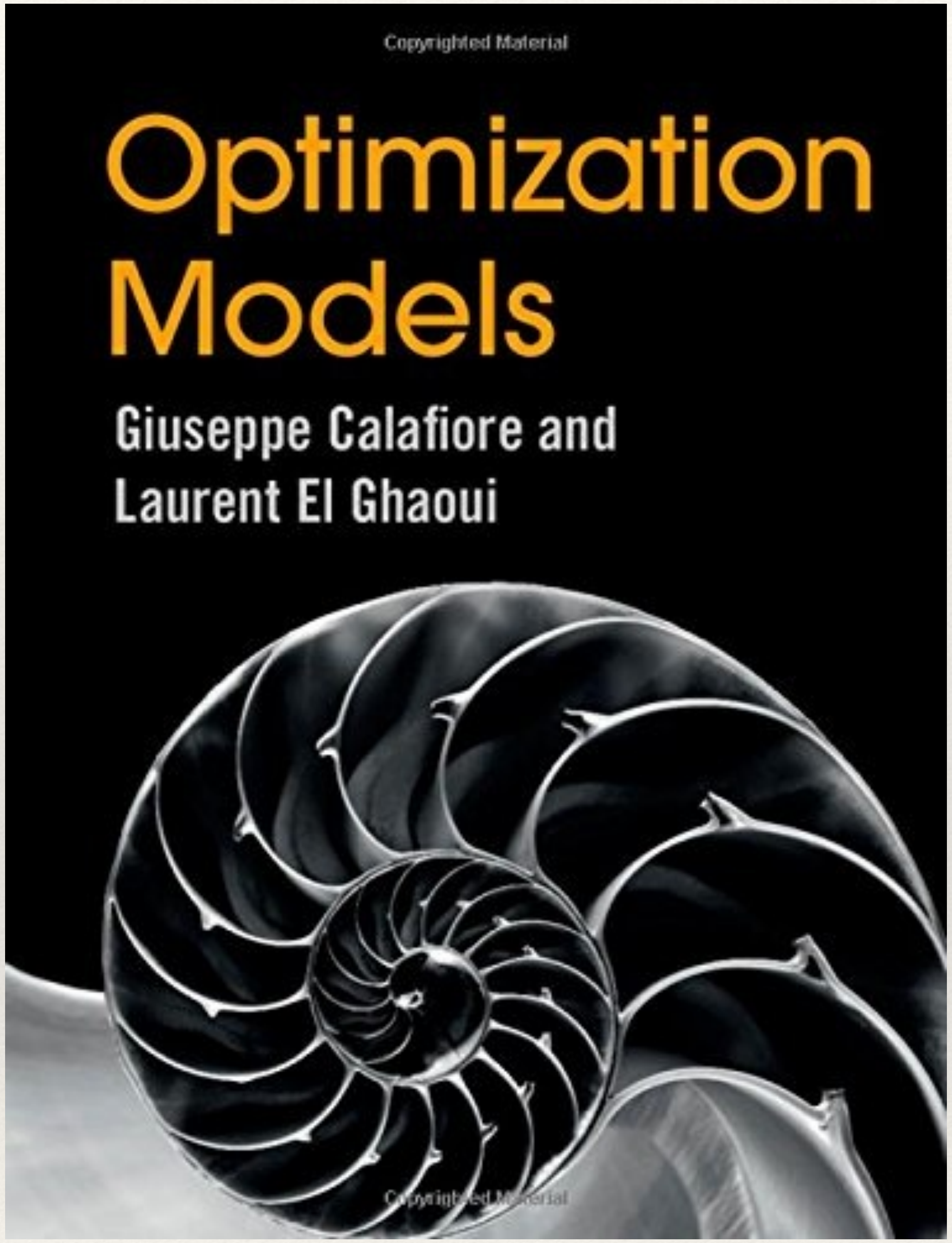
http://dc434.4shared.com/doc/R33bPROP/preview_html_m537a3256.gif

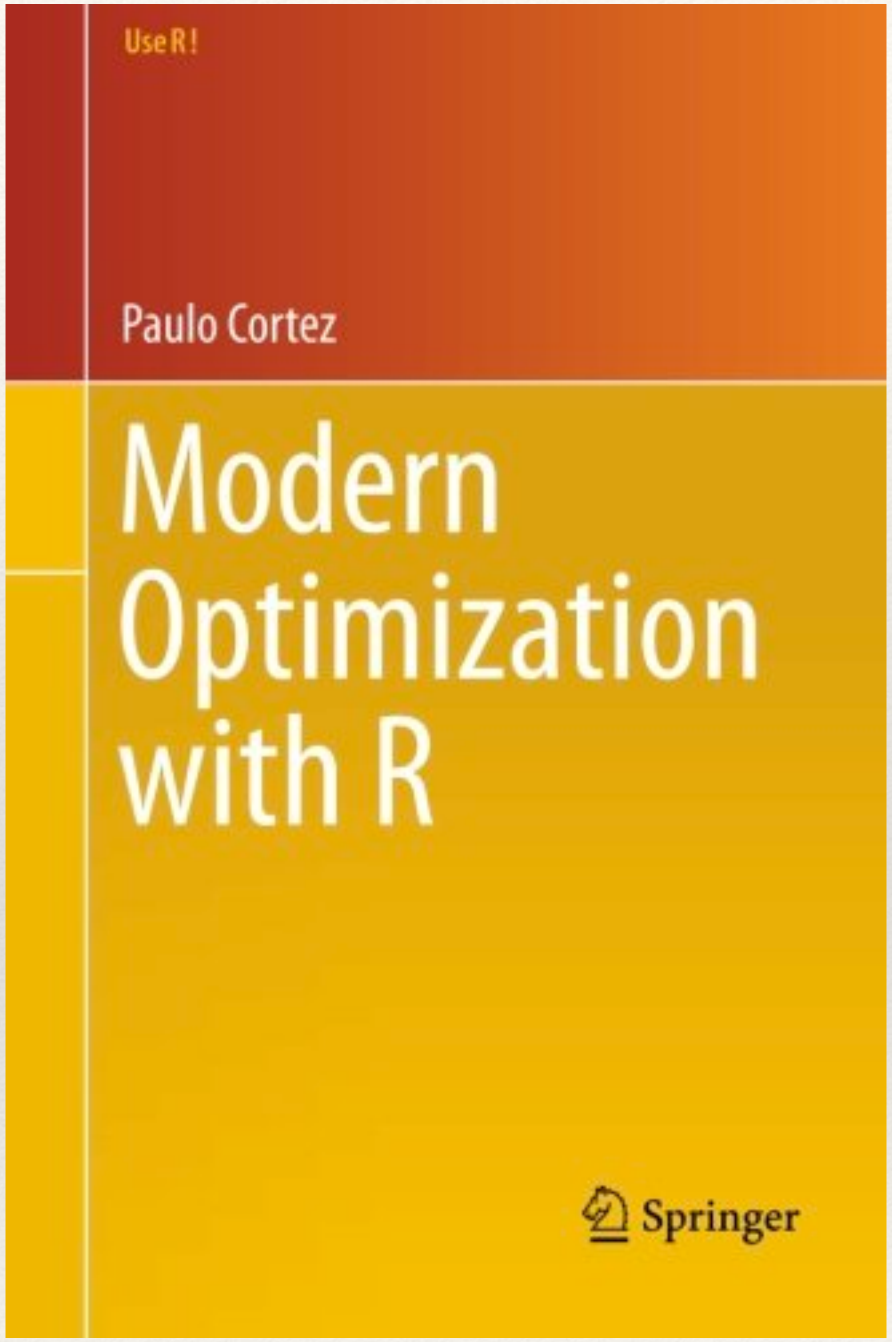
Optimization

- Minimizing (or maximizing) an objective
- Common goal in environmental problem solving
  - minimizing costs
  - minimizing material use
  - minimizing energy use
  - minimizing species loss …..



http://horicky.blogspot.com/2013/01/optimization-in-r.html


Copyrighted Material
Optimization
Models
Giuseppe Calafiore and
Laurent El Ghaoui
Copyrighted Material


Use R!
Paulo Cortez
Modern
Optimization
with R
Springer

Optimization

- Optimization in modeling, **identify**:
  - the *objective/output* that you want to minimize
  - the *parameters* (or inputs) that you can “play” with to change the output
  - the *model* (or set of models) that define the relationship between the output and the parameters

Optimization

- *Objective* - based on your output variable (or some transformation of it)
  - the output variable may be multi-dimensional
  - trying to minimize two or more things at the same time
    - minimize energy use and water use
    - maximize profit and minimize energy use

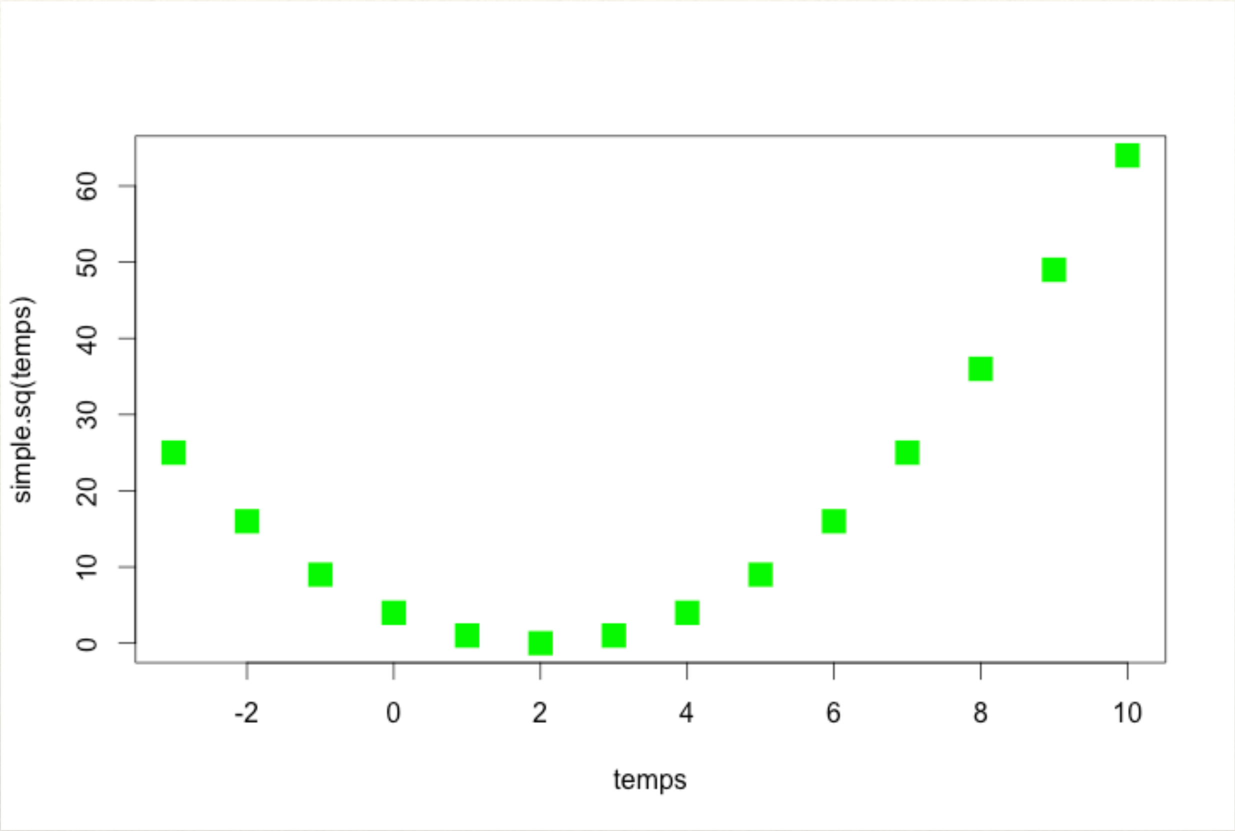
Optimization

- in modeling, identify:
  - the *objective* that you want to minimize
  - the *parameters* (or inputs) that you can "play" with to change the output - "free" parameters
  - the model (or set of models) that define the relationship between the output and the parameters

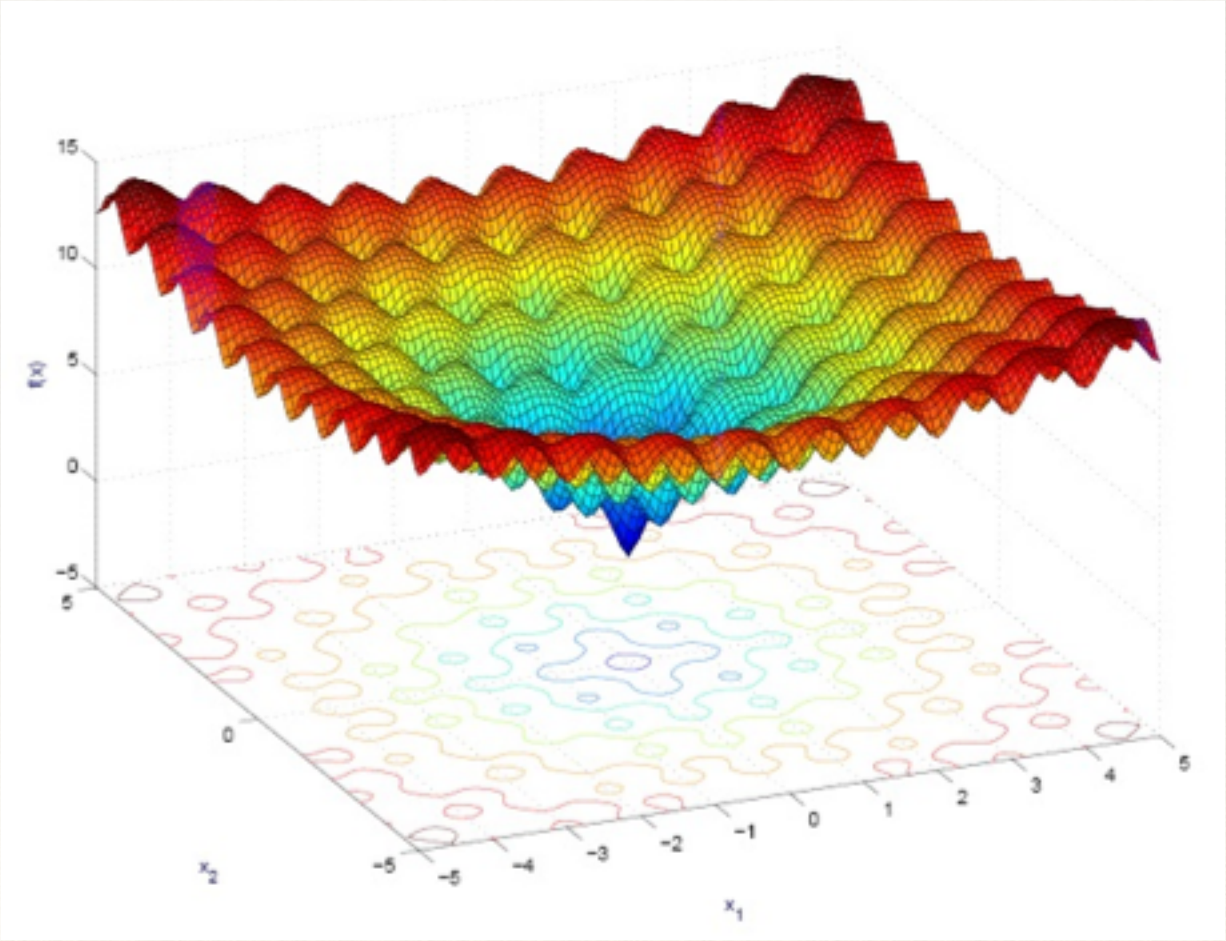
Optimization

- You can have more than one "free" parameters
  - minimizing energy use by changing both time of day that the plant functions and the operating temperature
  - minimizing air pollution by reducing the number of drivers and the time spent driving
  - maximizing species richness by increasing fire frequency and reducing fire severity

Optimization



Single free parameter - e.g minimize energy use as a function of temperature



Two free parameters - Minimize energy use as a function of temperature AND type of substrate/material

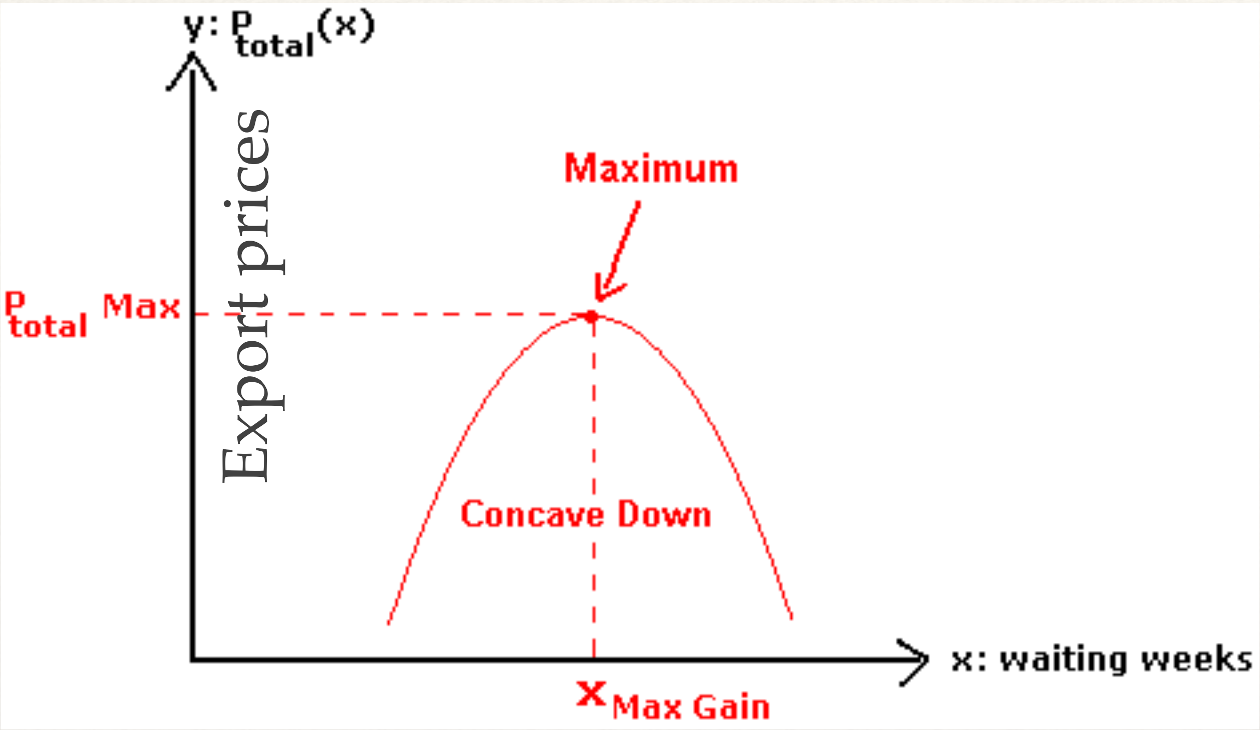
Optimization

- Objective - what you are trying to minimize
- Objectives using two or more variables
  - pareto optimality
    - point where any change in one variable will reduce objective for the other variable (balancing your multiple objectives)
  - define an objective function (single variable) that combines the two variables
    - useful because you can decide on the weighting
    - 0.5*energy.use + 0.5*cost
    - 0.8*energy.use + 0.2*cost
    - make sure that energy.use and costs are relative values to that it makes sense to add them

Optimization

- How easy the optimization is depends on the model - or the functional relationship between free parameters and output
- Three possibilities
  - the minimum can be derived mathematically
  - you can graph the relationships and "see" the minimum
  - you get there by trial and error - we have very sophisticated ways of doing the "trial and error"

Optimization

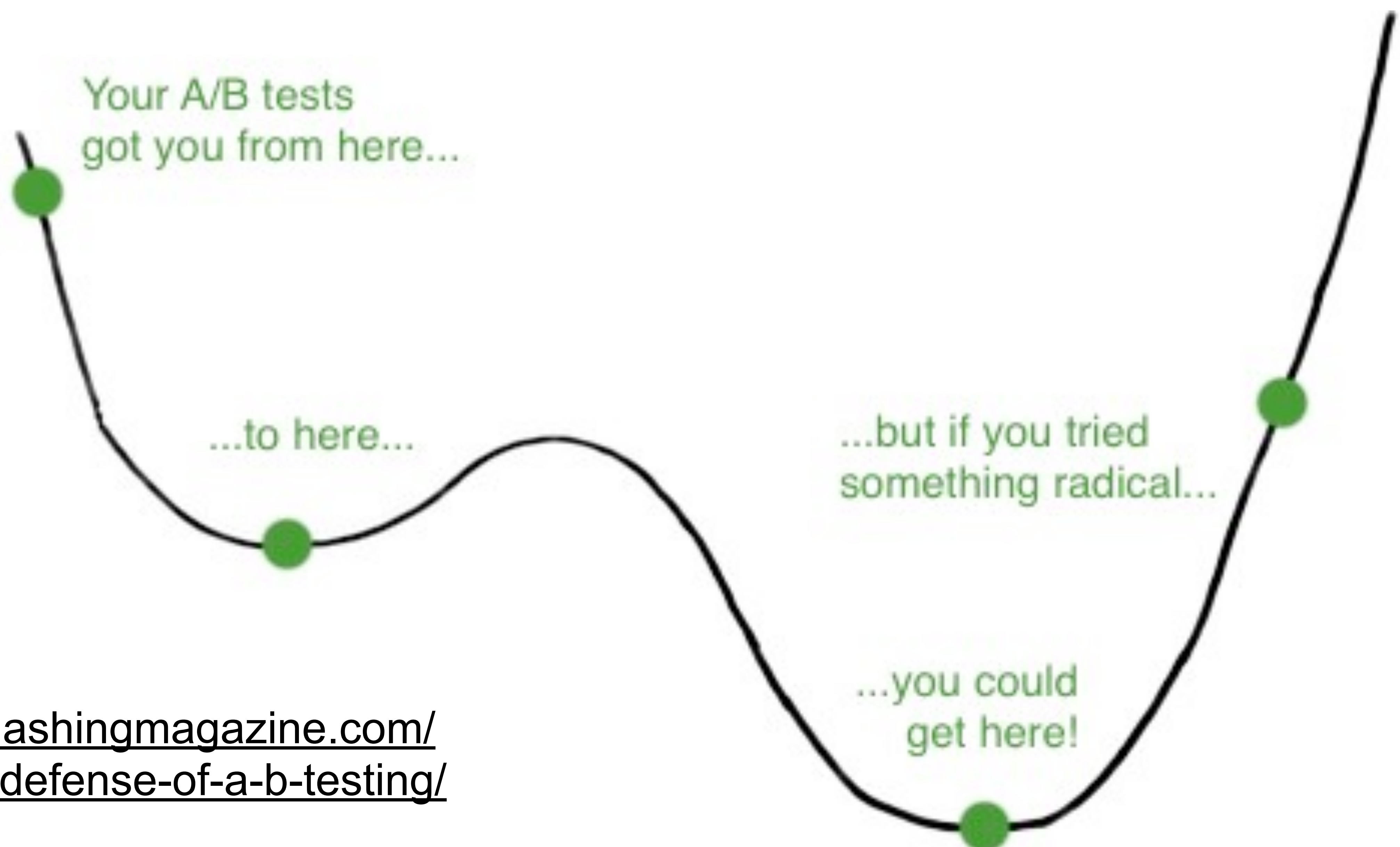


Finding the "optimal" value of the parameter (waiting weeks) and the value of the objective function at the optimum is visually easy to assess

http://derivative-functions.cours-de-math.eu/advanced-math-english14.php

Calibration:
Sampling the parameter space - Optimization

Optimization approaches vary in terms of
- ability to find non-local optima
- ability to handle multiple criteria



http://www.smashingmagazine.com/2010/08/26/in-defense-of-a-b-testing/

Optimization

- An interactive approach - keep “bracketing” the function

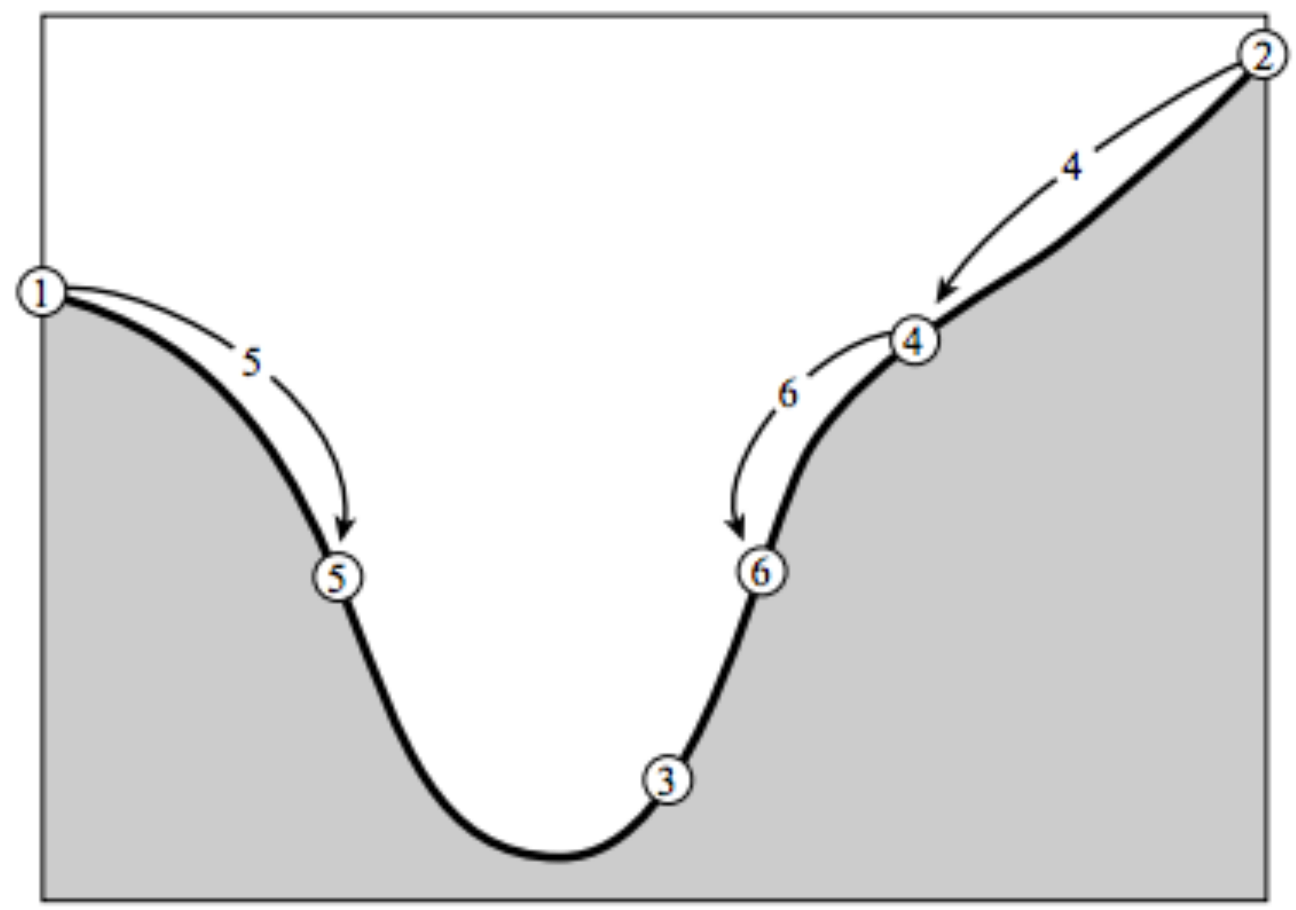


Figure 10.1.1. Successive bracketing of a minimum. The minimum is originally bracketed by points 1,3,2. The function is evaluated at 4, which replaces 2; then at 5, which replaces 1; then at 6, which replaces 4. The rule at each stage is to keep a center point that is lower than the two outside points. After the steps shown, the minimum is bracketed by points 5,3,6.

https://www.uam.es/personal_pdi/ciencias/ppou/CNC/TEMA6/f10.pdf

Optimization

- But this might get you to a “non” local minimum

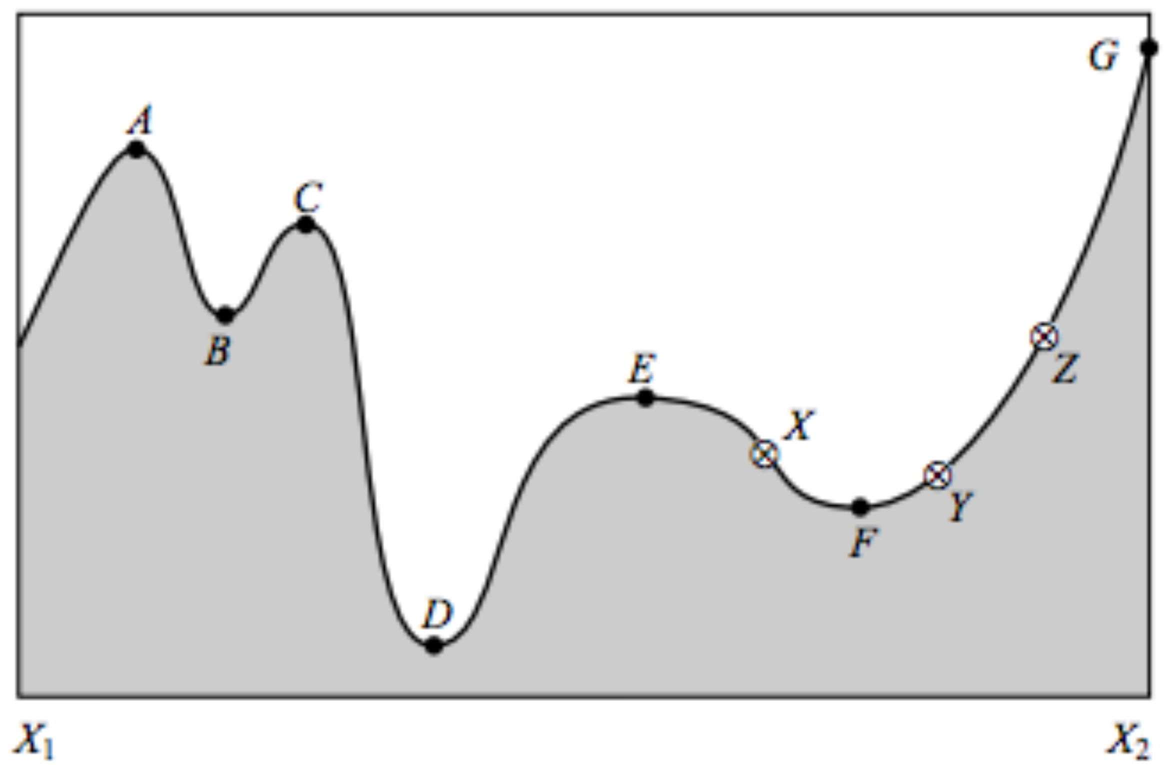
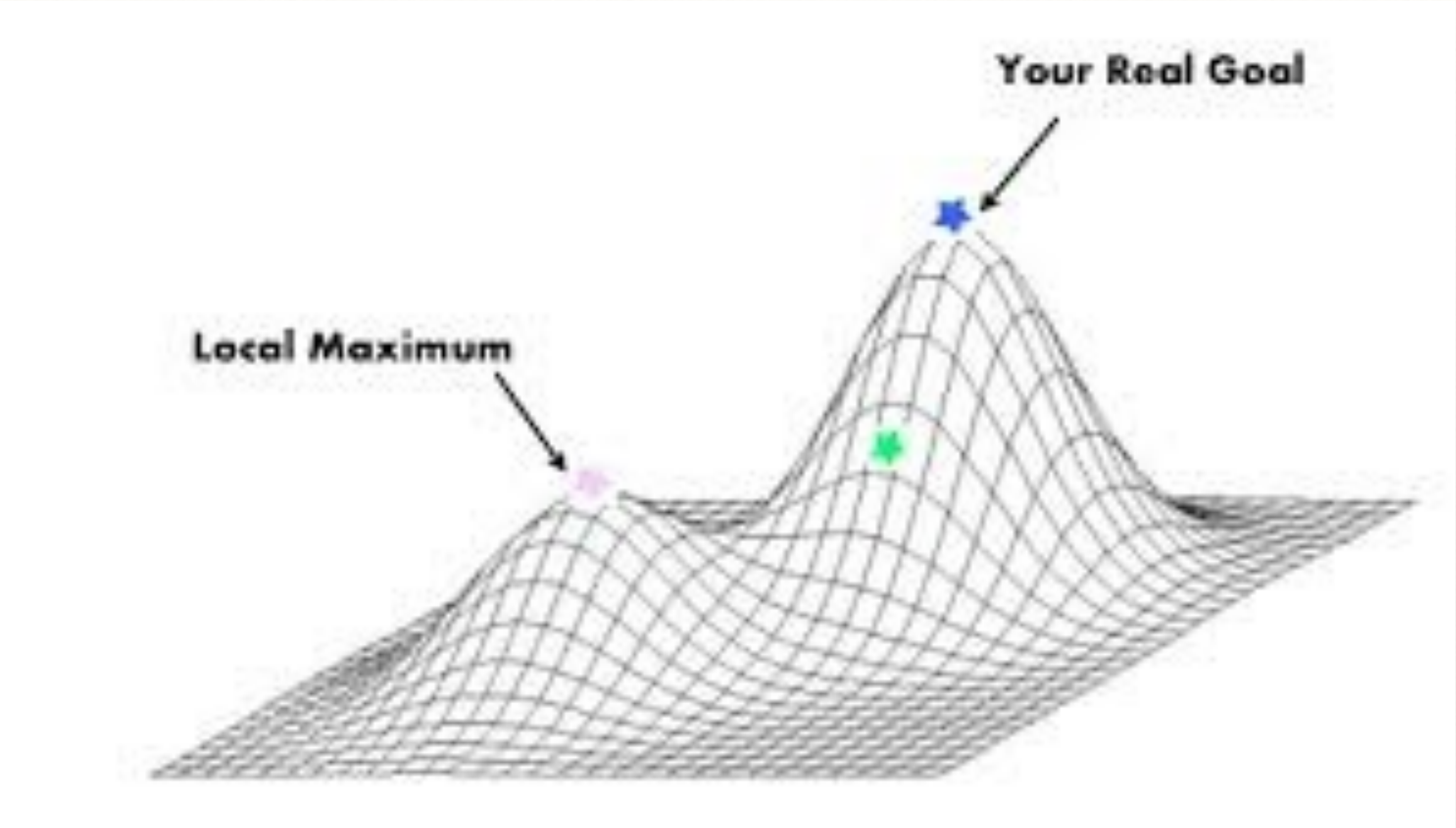
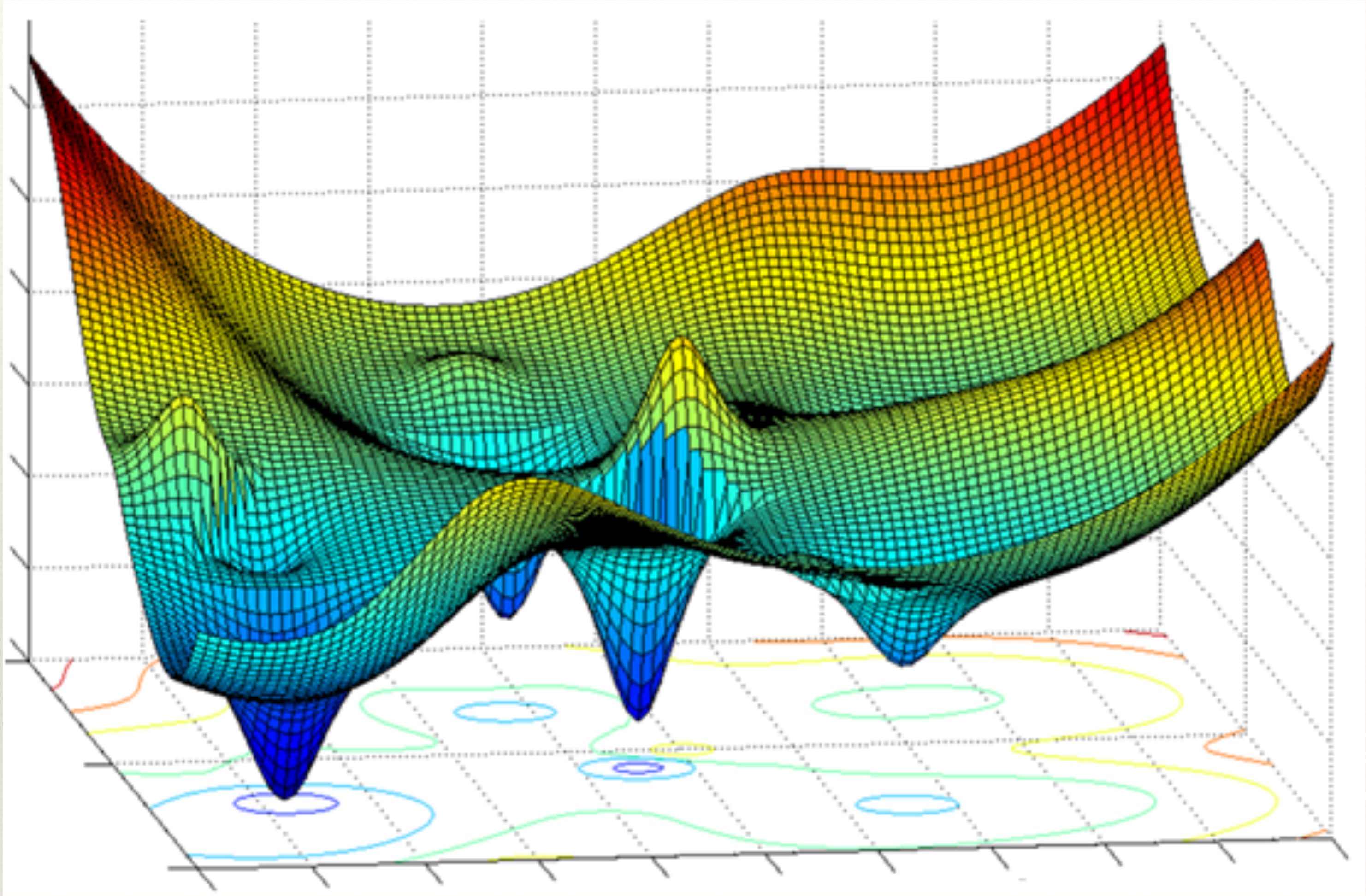


Figure 10.0.1. Extrema of a function in an interval. Points A, C, and E are local, but not global maxima. Points B and F are local, but not global minima. The global maximum occurs at G, which is on the boundary of the interval so that the derivative of the function need not vanish there. The global minimum is at D. At point E, derivatives higher than the first vanish, a situation which can cause difficulty for some algorithms. The points X, Y, and Z are said to “bracket” the minimum F, since Y is less than both X and Z.

https://www.uam.es/personal_pdi/ciencias/ppou/CNC/TEMA6/f10.pdf

Optimization

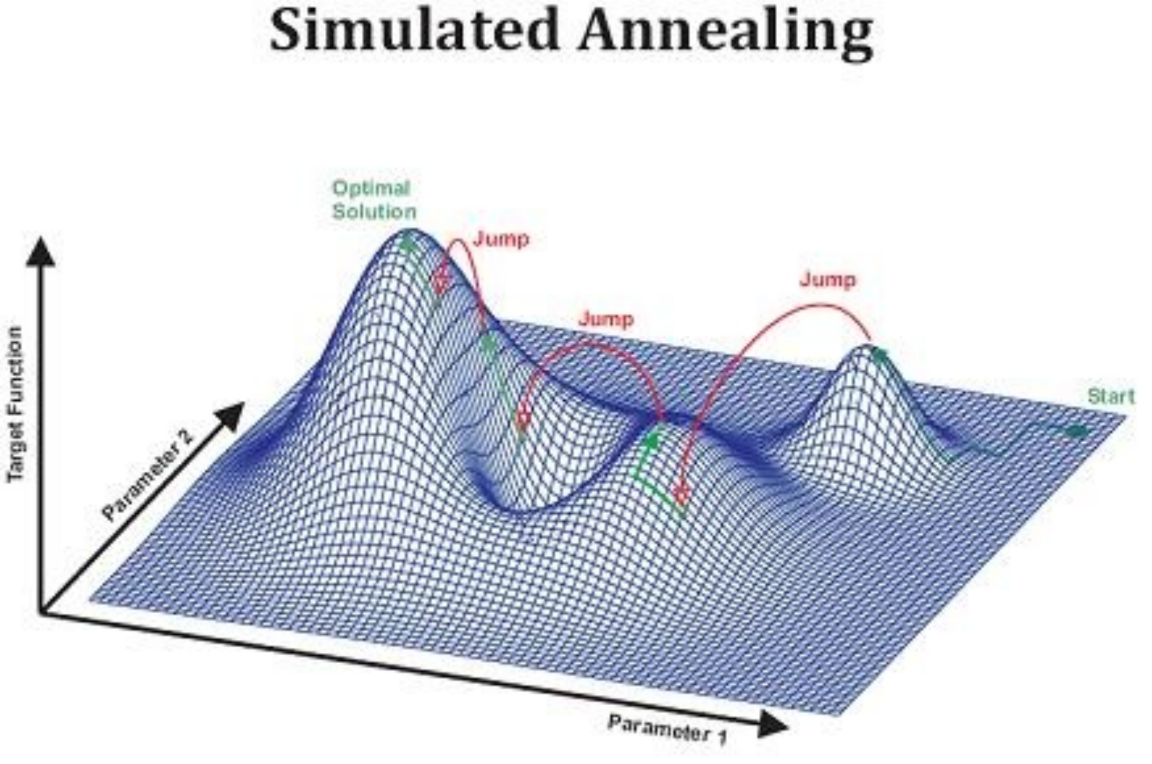




www.mathworks.com

Optimization

Math geeks and engineers have developed all kinds of strategies for search parameter spaces to find minimum, that deal with complex spaces with lots non-local minimum



www.frankfurt-consulting.de

Optimization

- One more distinction
- Constrained vs unconstrained
  - unconstrained …what we've been looking at
  - constrained….add rules about what parameters can be selected (e.g you must purchase at least 20 units of x)

Optimization

Optimization of more complex models - almost always requires one of these "search" techniques

- think about coming up with the optimal sea wall size in the mangrove model

Optimization techniques are also used for calibrating models - where the objective is the fit between model estimates and observations - we will get to that

Optimization in R

- Different functions available that help with
  - automate the “searching” for you: smart searching!
  - dealing with multi-dimensional search spaces
  - dealing with complex (multi-min, multi-max) search spaces
  - constrained / unconstrained
    - optimiz() - one dimensional
    - optim(), nlm() , nlminb()

Optimization in R

- similar to using an *ode* solve a differential equation
- in some cases, you can solve it directly - you can use math/graphs to find the optimal value - other wise you iterate - R can help you do that
- two steps
  - code your model as a function
  - send function, and parameters to the optimizer
    - the optimizer search for a minimum value of the model output

Optimization in R

Use “optimize” in R,

Optimize(function to be minimized, lower bound, upper bound..)

Can also provide:

- tol - tolerance (how close do you have to get)
- maximum - if maximum=T, finds the maximum; default is finding the minimum

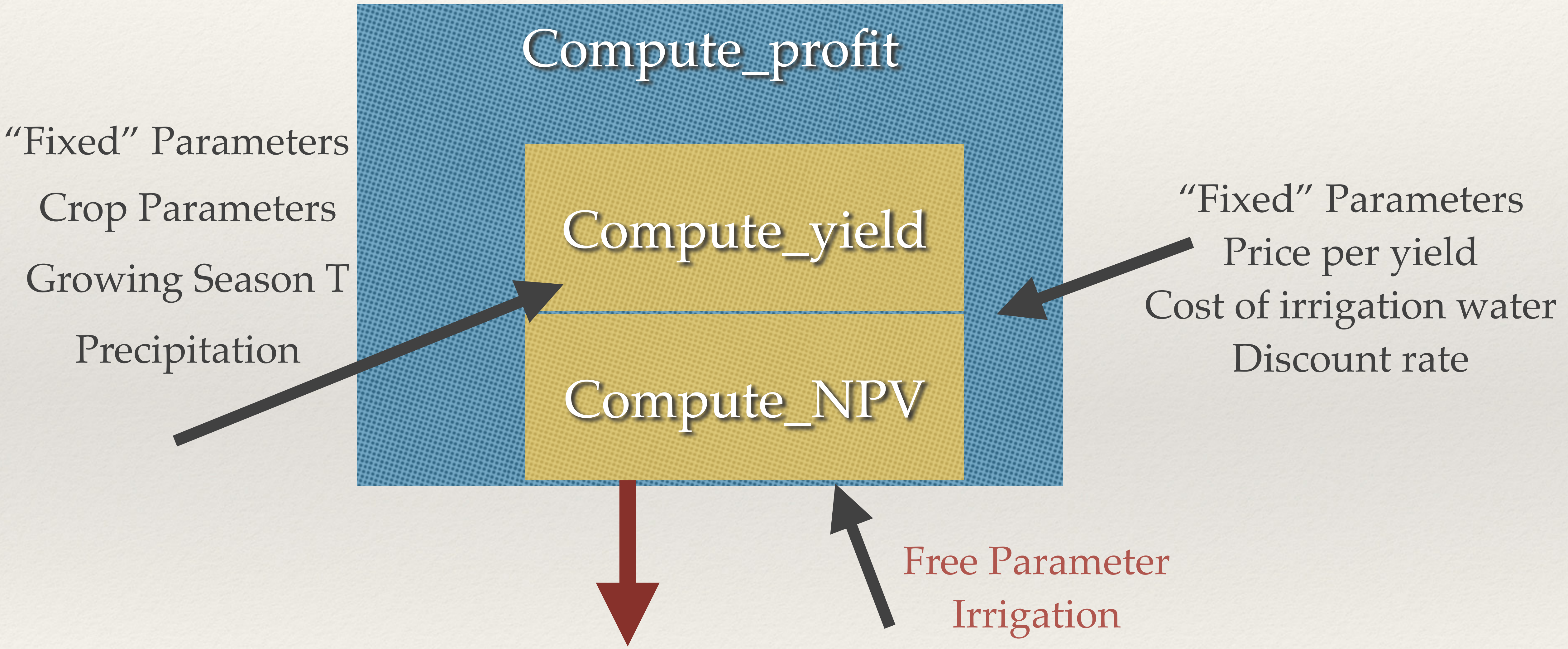
Optimization in R

- A slightly more complicated example….
- Lets say we have an option to purchase irrigation water; the contract requires us to commit to purchasing the irrigation water for a number of years in the future; we want to know how much irrigation water to purchase to maximize our profit from growing almonds (or some other crop)

Optimization in R

- What we have
  - we have a model of yield as a function of irrigation and climate
  - we have a cost of irrigation water and prices for yields
  - we have all of this for 10 years

Optimization in R



Optimize (compute profit, lower value for irrigation, upper value for irrigation, other fixed parameters, maximum=T)

Optimization in R

- Profit model
- output: total profit after n years
- inputs: inputs for yield model (growing season temperature (T) and annual precipitation (P), irrigation/year (irr), crop.pars), price of that crop gets (price), cost of irrigation water (cost)
    - calculate yield in each year- from our yield model
    - income is price*yield in each year
    - cost is irrigation * number of years
    - net = income - costs
    - find net present value of net

Optimize in R

Lets say we have 3 years of temperature and precipitation data, and price of almonds is $50/kg and irrigation water will cost us $150 per unit; and we know the crop specific parameters for almonds - how much irrigation water should we buy?