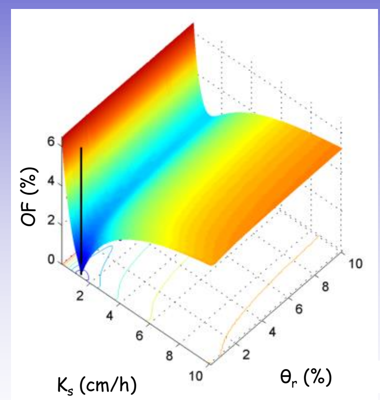
PARAMETER OPTIMIZATION USING INVERSE MODELING



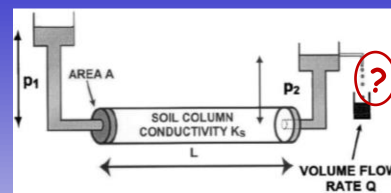
Contact: vcouvreur@ucdavis.edu

OUTLINE

- Direct vs inverse problems
  - Ex.1: Darcy's experiment
  - Analytical solutions
  - Ex.2: Multi-step outflow experiment
- Inverse modeling scheme for parameter optimization
  - Objective function
  - Optimization algorithms
  - Ex.3: Orchard soil water status monitoring
- Uncertainty Analysis
- Overview

EXAMPLE OF DIRECT PROBLEM

Model: $Q = \dfrac{K_s\, A\, \Delta P}{L}$



Initial cond.: $\theta(t=0,x) = \theta_s$

Boundary cond.: $\Delta P(t>0) = \Delta P_{obs}(t)$

System properties: K_s, A and L are *known*

Response: Q(t) is *unknown*

Typical application:

Direction and speed of pollutant plume propagation

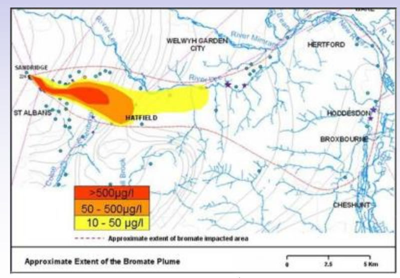
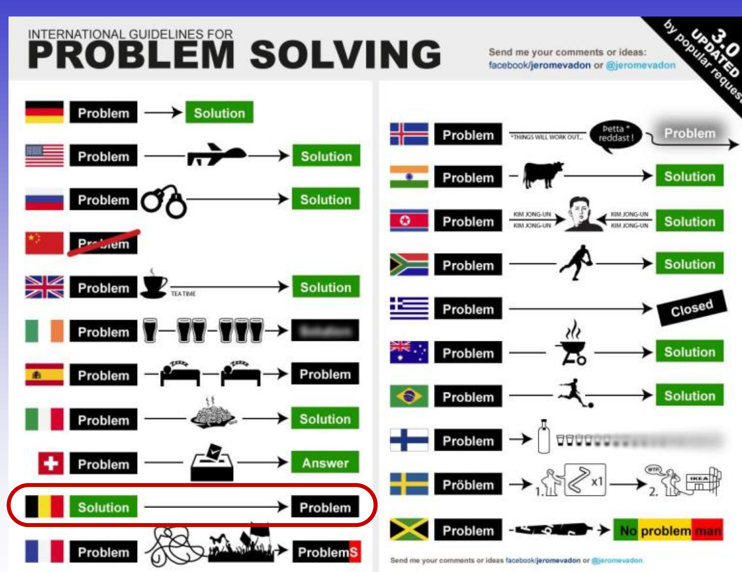


Image: www.groundwateruk.org

INVERSE PROBLEM ANALOGY



Response is *known* **BUT** system properties are *unknown*

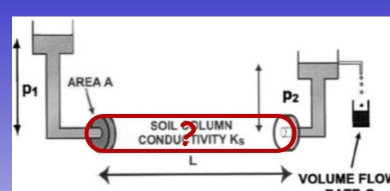
INVERSE PROBLEM DEFINITION

Solving an **inverse problem** is the process of calculating from a set of observations the causal factors that produced them. It is called an inverse problem because it starts with the results and then calculates the causes. This is in contrast to the corresponding **direct problem**, whose solution involves finding effects based on the complete description of their causes.

"**Inverse modeling** is a formal approach for estimating the variables driving the evolution of a system by taking measurements of the observable manifestations of that system, and using our physical understanding to relate these observations to the driving variables." (*Lectures on inverse modeling*, D. J. Jacob, 2007).

INVERSE PROBLEM EXAMPLE 1

Forward model: $Q = \dfrac{K_s\, A\, \Delta P}{L}$



Initial cond.: $\theta(t{=}0,x) = \theta_s$

Boundary cond.: $\Delta P(t{>}0) = \Delta P_{obs}(t)$

System properties: A, L are <u>*known*</u>

System response: $Q(t)$ is <u>*known*</u>

System property: K_s is <u>*unknown*</u> → **But straightforward K_s calculation thanks to the available analytical relation "$Q(K_s)$" in saturated conditions**

Typical application:

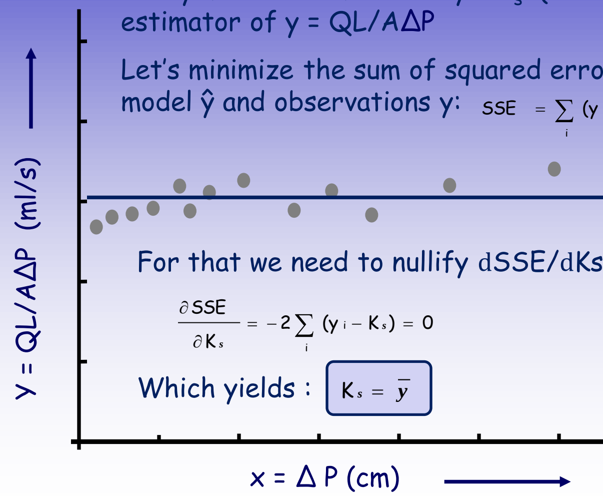
Characterization of system properties and/or system state

ANALYTICAL SOLUTIONS

Single parameter solution:

Darcy model assumes that $\hat{y} = K_s$ (i.e. a constant) is the estimator of $y = QL/A\Delta P$

Let's minimize the sum of squared errors between the model $\hat{y}$ and observations y: $SSE = \sum_i (y_i - \hat{y})^2 = \sum_i (y_i - K_s)^2$



For that we need to nullify dSSE/dKs :

$$\frac{\partial SSE}{\partial K_s} = -2\sum_i (y_i - K_s) = 0$$

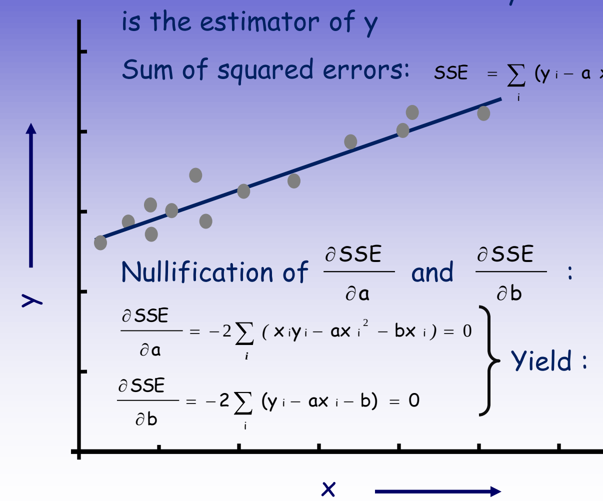
Which yields : $K_s = \bar{y}$

ANALYTICAL SOLUTIONS

Generalization for 2 parameters:

The linear model assumes that $\hat{y} = a\,x + b$ is the estimator of y

Sum of squared errors: $SSE = \sum_i (y_i - a\,x_i - b)^2$



Nullification of $\frac{\partial SSE}{\partial a}$ and $\frac{\partial SSE}{\partial b}$:

$$\frac{\partial SSE}{\partial a} = -2\sum_i (x_i y_i - a x_i^2 - b x_i) = 0$$

$$\frac{\partial SSE}{\partial b} = -2\sum_i (y_i - a x_i - b) = 0$$

Yield :

$$a = \frac{\sum_i (x_i y_i - \overline{x}\,\overline{y})}{\sum_i (x_i^2 - \overline{x}^2)}$$

$$b = \bar{y} - a\bar{x}$$

INVERSE PROBLEM EXAMPLE 2

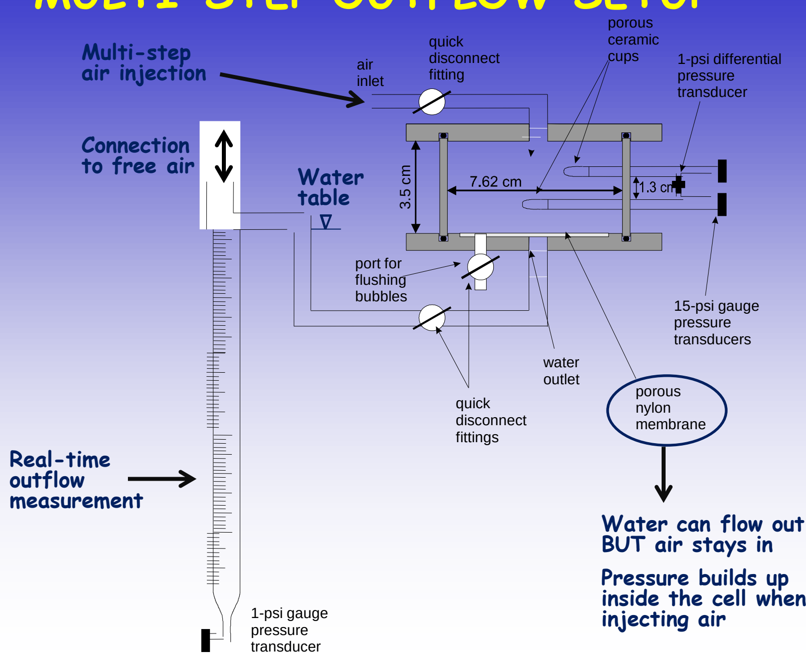
Unsaturated soil hydraulic properties estimation

Soil relative saturation index S_e: $S_e(h) = \left(\frac{1}{1+(\alpha h)^n}\right)^m$

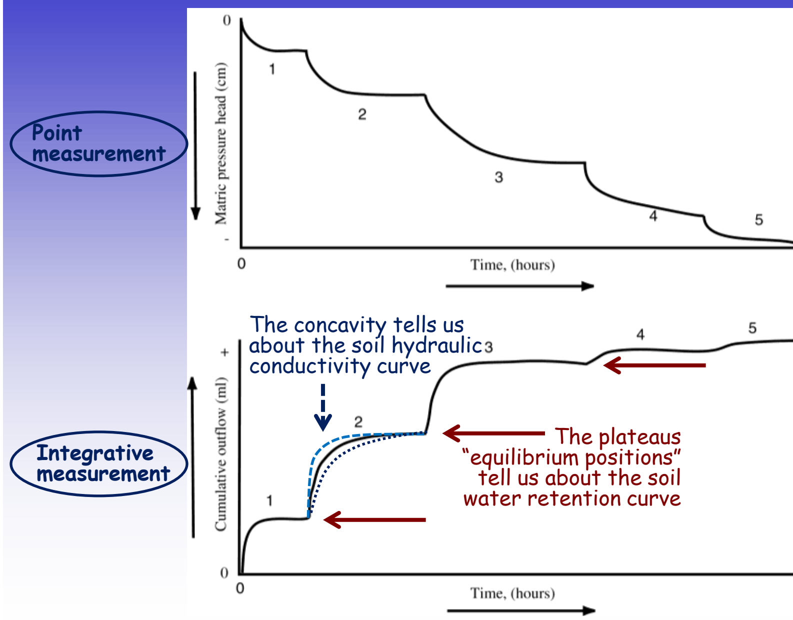
Soil water retention curve (WRC): $\theta(S_e) = \theta_r + S_e\,(\theta_s - \theta_r)$

Soil hydraulic conductivity curve (HCC): $K(S_e) = K_s\, S_e^{\wedge}\left(1-(1-S_e^{1/m})^m\right)^2$

MULTI-STEP OUTFLOW SETUP



SYSTEM RESPONSE TO AIR PRESSURE



INVERSE PROBLEM EXAMPLE 2

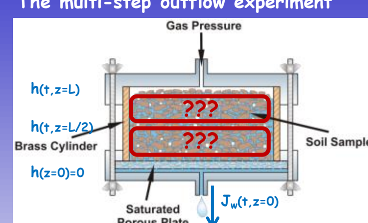
Model: $J_w(t,z) = K(h(t,z)) \dfrac{\Delta(h(t,z)+z)}{\Delta z}$

$$\Delta\theta(t,z) = \frac{\Delta J_w(t,z)\,\Delta t}{\Delta z}$$

$$\theta = f_1(h, \mathbf{WRC_{parms}})$$

$$K = f_2(h, \mathbf{HCC_{parms}})$$

The multi-step outflow experiment



No analytical relation available for "$J_w(t,z=0,WRC_{parms},HCC_{parms})$"

↓

Need a full inverse modeling scheme to find WRC_{parms} & HCC_{parms}

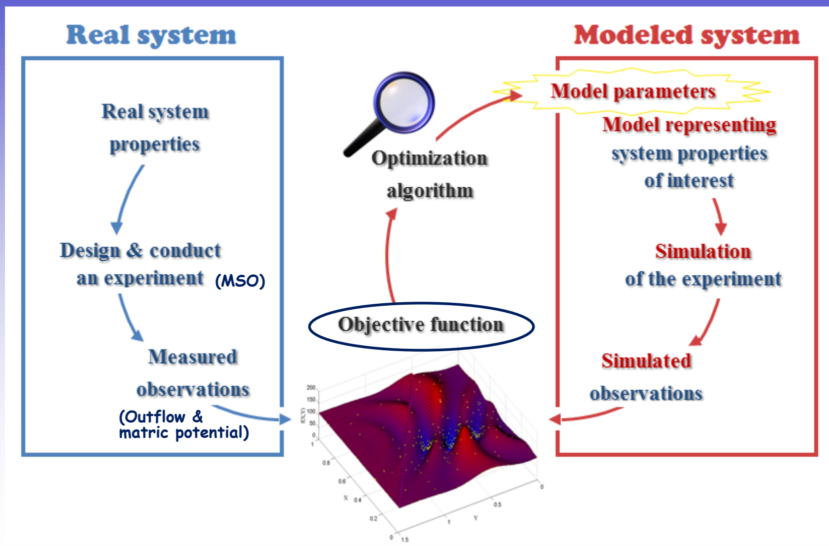
Initial cond.: $h(t=0,z) = z$

Boundary cond.: $h(t>0,z=0)$ & $h(t>0, z=L)$

System response: $J_w(t,z=0)$ is *known*

System properties: WRC_{parms} & HCC_{parms} are *unknown*

INVERSE MODELING SCHEME FOR PARAMETER OPTIMIZATION



OBJECTIVE FUNCTION

Definition: the objective function (OF) quantifies the quality of the fitness between measured and simulated observed system responses for any set of parameters.

- Example 1, weighted-average difference:

$$OF(parms) = \Sigma_{obs} (Sim_{obs}(parms) - Meas_{obs})\, w_{obs}$$

- Example 2, root weighted-average square difference:

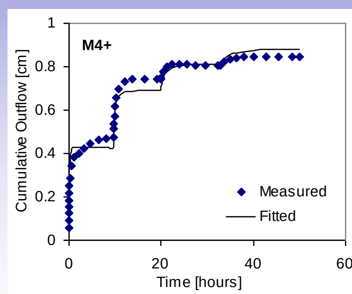
$$OF(parms) = sqrt(\Sigma_{obs} (Sim_{obs} (parms) - Meas_{obs})^2\, w_{obs})$$

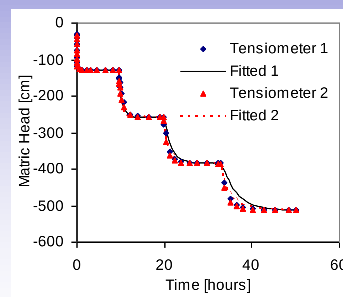
Individual weights « w_{obs} » can be attributed to each observation in space and time.

OBJECTIVE FUNCTION

Definition: the objective function (OF) quantifies the quality of the fitness between measured and simulated observed system responses for any set of parameters.

$$OF(parms) = \Sigma_{obs} (Sim_{obs}(parms) - Meas_{obs})\, w_{obs}$$



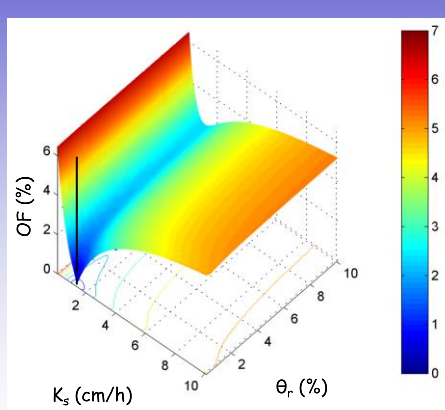


Errors for outflow & matric head ≈ 0.01 cm & 10 cm resp.

⟶ Adjust w_{obs} to the type of observation !

OBJECTIVE FUNCTION

The ensemble of all possible combinations of values for all the model parameters constitutes the "parametric space".



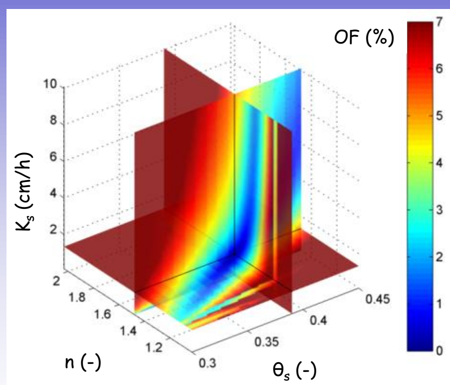
If the model has 2 parameters, the parametric space is a plan. Each point in this plan is a "parameter set" X.

Displaying the objective function value for all parameter sets in the parametric space is a way to visualize the OF "topography" (only possible for 2 parameters at a time).

The type of topography determines what type of optimization algorithm is more likely to find the optimal parameter set.

OBJECTIVE FUNCTION

In a 3-dimensional parametric space, the third dimension cannot be used to display the OF topography.



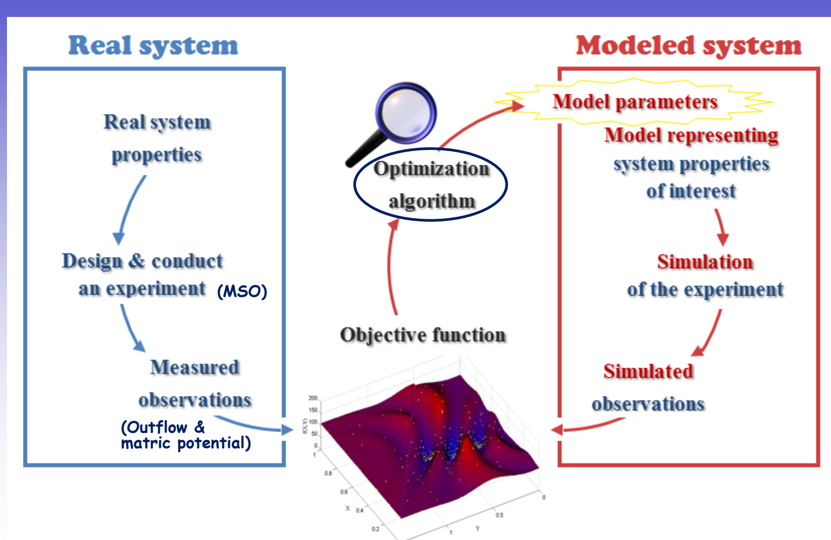
Color-code needed to display OF values on 2-D slices of the parametric space.

Discretization of each slice is 50 x 50 => 7500 Hydrus runs necessary to evaluate the OF at each point and get this image (not quite efficient to search for the minimum).

Warning: Incomplete view of the OF topography.

If more parameters need to be optimized, this process becomes less and less efficient.

INVERSE MODELING SCHEME FOR PARAMETER OPTIMIZATION



OPTIMIZATION ALGORITHMS

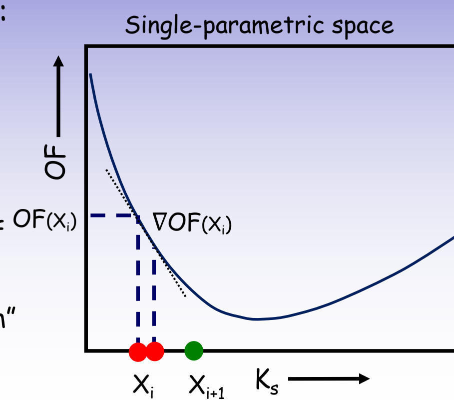
Definition: the optimization algorithm selects parameter values X_{i+1} based on prior information on the objective function score $(\ldots, OF(X_{i-1}), OF(X_i))$ of previous simulations. The way this information is used differs among optimizers:

- Gradient descent method:

$$X_{i+1} = X_i - p\ \nabla OF(X_i)$$

where X_i is a vector containing the parameter set of the i^{th} iteration and p is a property of the optimizer.

Type: "Sliding search algorithm" (see also Levenberg-Marquadt)



OPTIMIZATION ALGORITHMS

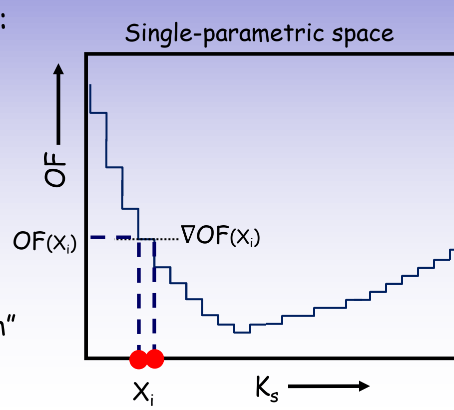
Definition: the optimization algorithm selects parameter values X_{i+1} based on prior information on the objective function score $(\ldots, OF(X_{i-1}), OF(X_i))$ of previous simulations. The way this information is used differs among optimizers:

- Gradient descent method:

$$X_{i+1} = X_i - p\ \nabla OF(X_i)$$

where X_i is a vector containing the parameter set of the i^{th} iteration and p is a property of the optimizer.

Type: "Sliding search algorithm" (see also Levenberg-Marquadt)



OPTIMIZATION ALGORITHMS

- Simplex Search Method (local):

Construct simplex using best *N* + 1 points (*N* = number of parameters)

$$X_{test} = X_m + \mathbf{p}\,(X_m - X_w)$$

where X_w is the worst parameter set in the simplex and X_m is the mean of the best *N* parameter sets

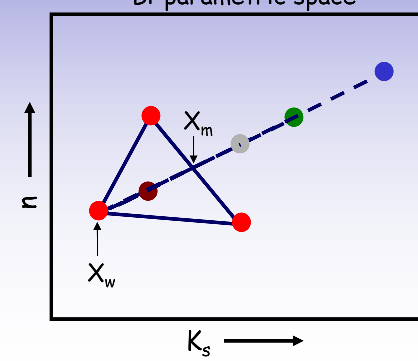
<u>Tested p values in an iteration:</u>

p_r = reflection
p_e = expansion
p_{c+} = positive contraction
p_{c-} = negative contraction

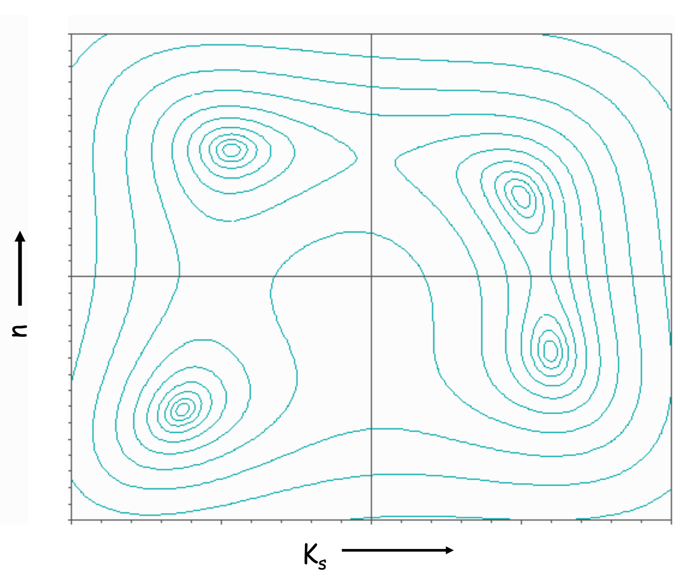
X_w is replaced by the best X_{test} and a new simplex is formed.

Type: "Jumping search algorithm"

Bi-parametric space

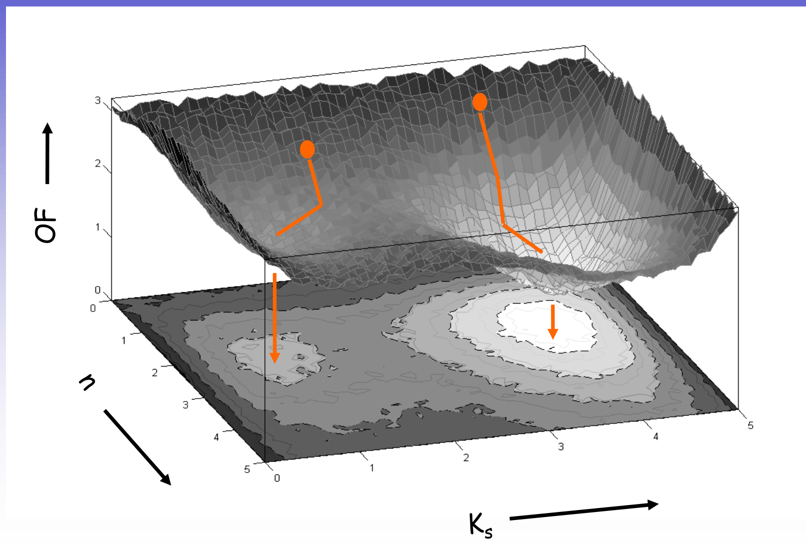


OPTIMIZATION ALGORITHMS



OPTIMIZATION ALGORITHMS

PROBLEMS WITH LOCAL SEARCH METHODS …



OPTIMIZATION ALGORITHMS

- The Shuffled Complex Evolution Method (global):

Uses multiple simplex searches & complex shuffling

(1) Generate sample: Sample M parameter sets $\{X_1, …, X_M\}$ and compute the OF of each of these points;

(2) Rank points: Sort the M points in order of decreasing fitness;

(3) Partition into complexes: Partition the M points into complexes, each containing one of the best ranked points and at least N+1 points on total (N = number of parameters);

(4) Evolve each complex: Evolve each complex using the Simplex Search Method;

(5) Shuffle complexes: Sort the points in order of decreasing fitness (i.e. increasing OF value);

(6) Check convergence: If convergence criteria are satisfied, stop; otherwise return to step 3;

OPTIMIZATION ALGORITHMS

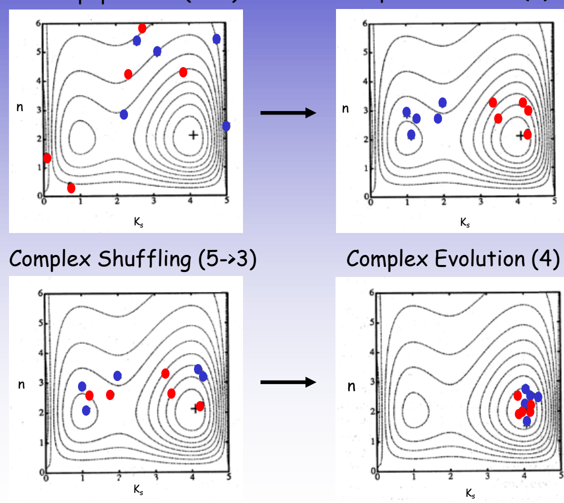
- The Shuffled Complex Evolution Method (global):

Initial population (1->3) | Complex evolution (4)

Iteration 1

Complex Shuffling (5->3) | Complex Evolution (4)

Iteration 2



Evaluate OF

Sort points based on OF

Partition them into red and blue complexes

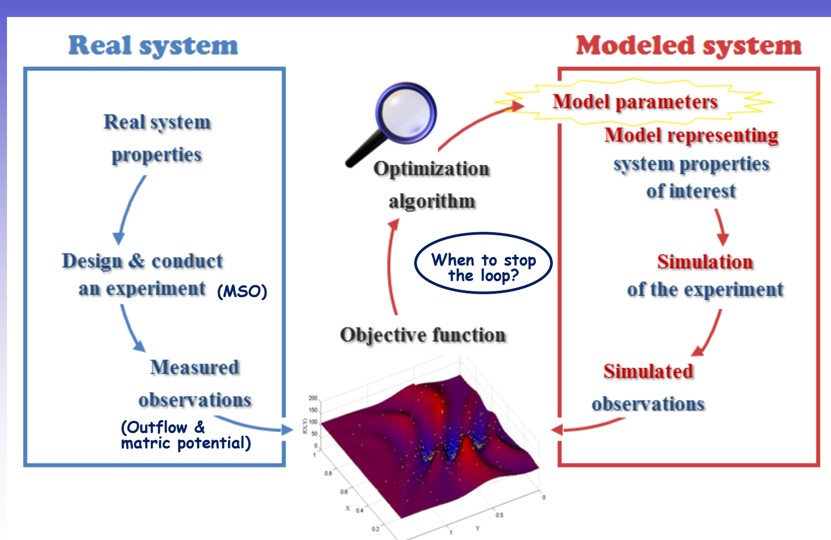
Convergence?

No: go to next iteration

Yes: end optimization

See also "Genetic Algorithm", "DREAM", ...

INVERSE MODELING SCHEME FOR PARAMETER OPTIMIZATION



OPTIMIZATION ALGORITHMS

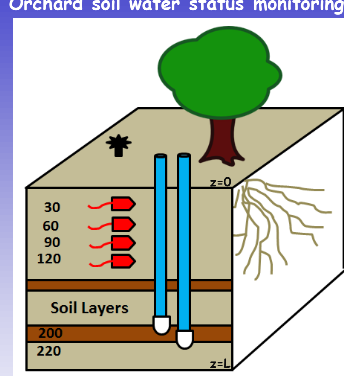
When do we want to stop the parameter optimization loop ?

"Convergence criteria":

- When a parameter set attaining a threshold OF value is found
  Example: OF(X)=0 -> no difference between Sim_{obs} and $Meas_{obs}$
- When the best OF value has not significantly changed for long
  Example: less than 1% improvement for more than 50 iterations
- When too much time elapsed
  Example: optimization loop has been running for 2 days
- When too many iterations were ran
  Example: optimization loop has been running for 10^5 iterations
- Any combination of the previous criteria…

INVERSE PROBLEM EXAMPLE 3

Orchard soil water status monitoring



Model: $J_w(t,z) = K(h(t,z)) \dfrac{\Delta(h(t,z)+z)}{\Delta z}$

$$\Delta\theta(t,z) = \frac{\Delta J_w(t,z)\ \Delta t}{\Delta z}$$

$\theta = f_1(h,\mathbf{WRC_{parms}})$

$K = f_2(h,\mathbf{HCC_{parms}})$

$S(t,z) = ET(t)\ \alpha(h,\mathbf{P_2},\mathbf{P_3})\ \beta(z)$

Initial cond.: $h(\theta_{obs}(t=0,z^*))$ & $h_{obs}(t=0,z^*)$

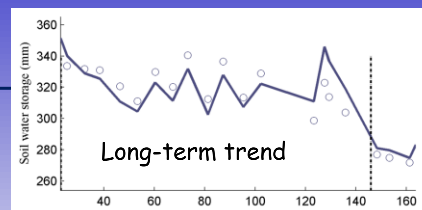
Boundary cond.: $J_w(t>0,z=0) = Irrig(t)$ & $\nabla h(t>0, z=L) = 0$ & $ET_{obs}(t)$

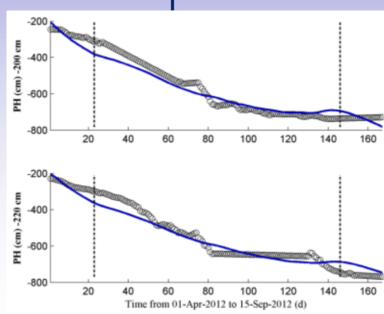
System response: $\theta_{obs}(t^*,z^*)$ & $h_{obs}(t^*,z^*)$ "*" for discrete obs.

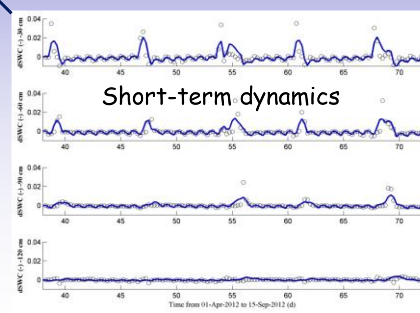
System properties: WRC_{parms} & HCC_{parms} *unknown* for 2 soil types
Feddes water stress parms P_2 & P_3 *unknown*

INVERSE PROBLEM EXAMPLE 3

Multi-objective optimization of 9 parameters (θ_{s1},a_1,n_1,K_{s1}, a_2,n_2,K_{s2},P_2,P_3) using the Genetic Algorithm (global)

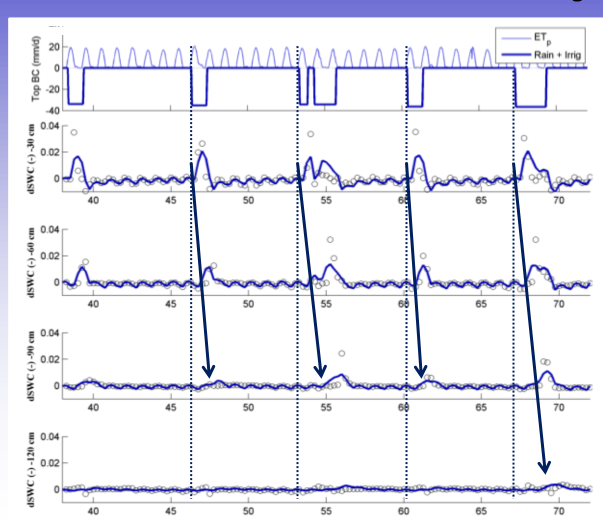






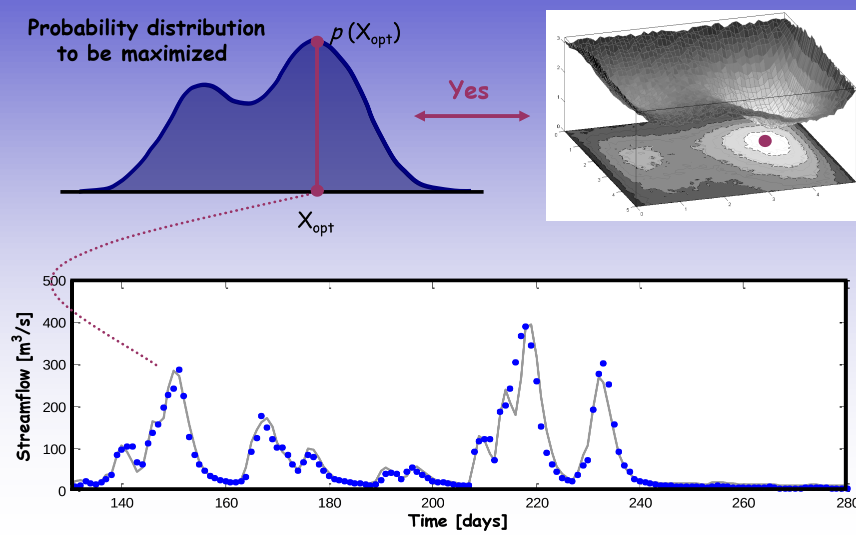
INVERSE PROBLEM EXAMPLE 3

Using the soil water content change "dSWC" to focus on the **amplitude** and **travel time** of the infiltration front signal



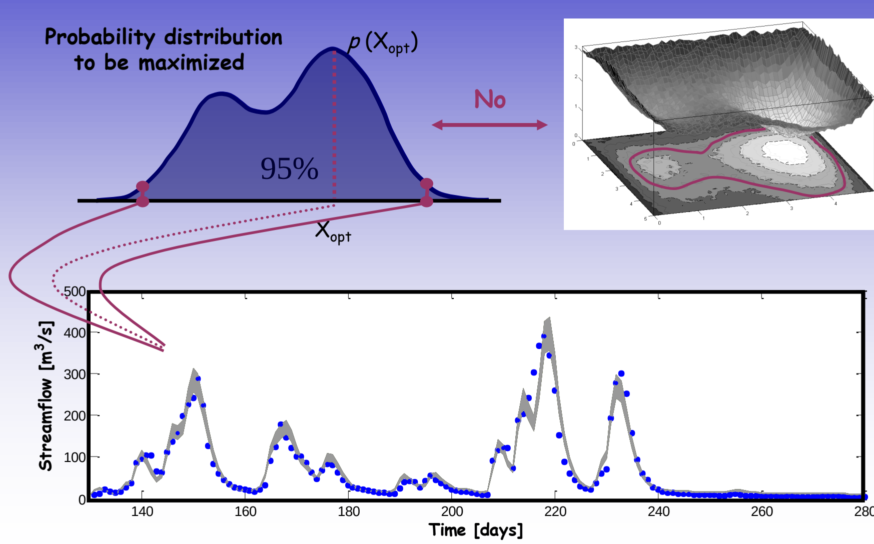
UNCERTAINTY ANALYSIS

Optimization identifies the mode… (most probable parameter set)



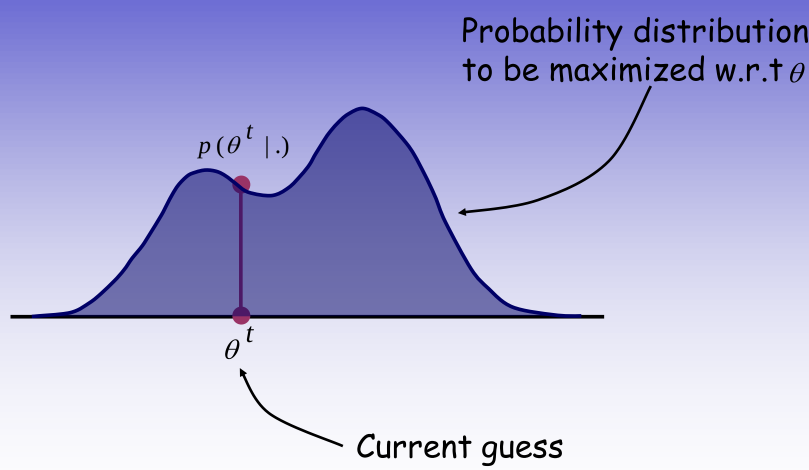
UNCERTAINTY ANALYSIS

Need estimates of uncertainty …



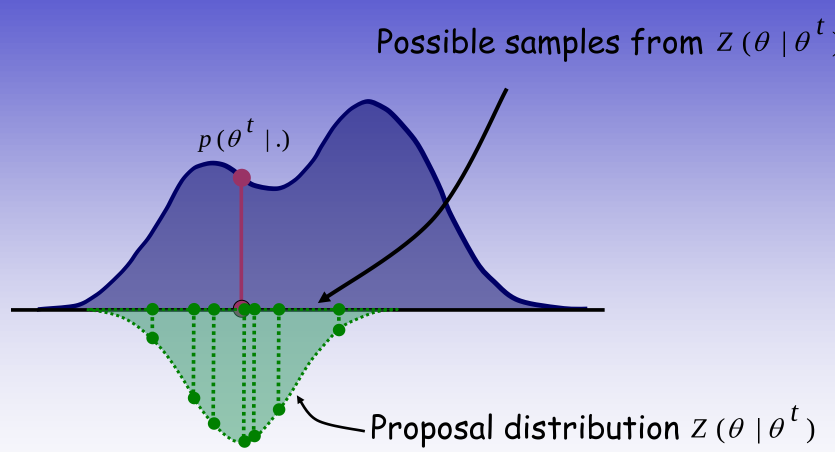
UNCERTAINTY ANALYSIS

Markov Chain Monte Carlo (MCMC) method



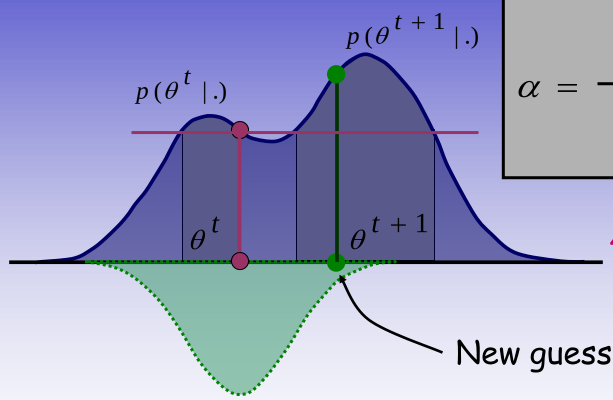
UNCERTAINTY ANALYSIS

MCMC proposal distribution z(.|.)



UNCERTAINTY ANALYSIS

MCMC acceptance rule:



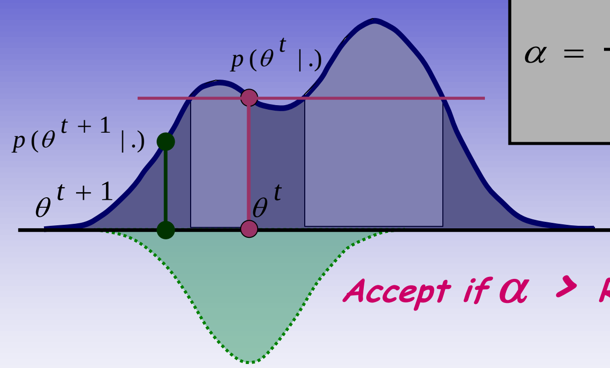
$$\alpha = \frac{p(\theta^{t+1}|.)}{p(\theta^{t}|.)} > 1$$

Always accept

MCMC – Acceptance of New Points Having Higher Probability than the Old Point is _100%_

UNCERTAINTY ANALYSIS

MCMC acceptance rule:



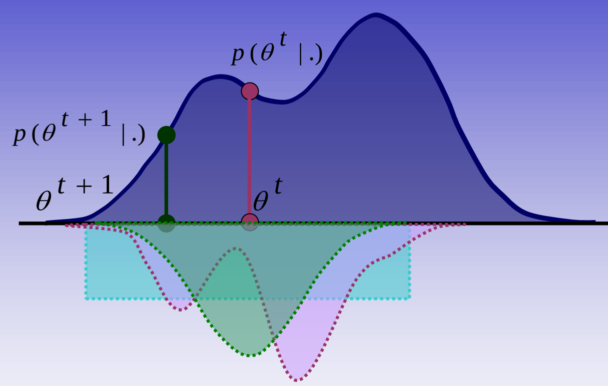
$$\alpha = \frac{p(\theta^{t+1}|.)}{p(\theta^{t}|.)} < 1$$

Accept if α > R ~ Uniform (0,1)

If the α ratio is small, then probability of acceptance is small

MCMC –Acceptance of New Points Having Lower Probability than the Old Point is _Probabilistic_

UNCERTAINTY ANALYSIS



Warning: Performance of a MCMC Sampler depends strongly on the choice of the *Proposal Distribution*

OVERVIEW

LOCAL SEARCH ALGORITHMS:

Start from a single randomly chosen point in the parameter space, and seek iterative improvement from this point

Examples: Gradient Descent, Levenberg-Marquardt, Simplex Search and Line Search

GLOBAL SEARCH ALGORITHMS:

Typically implement a number of different parameter combinations simultaneously (called a population) and use evolutionary principles of survival of the fittest to iteratively improve this population

Examples: Genetic algorithms, Differential Evolution, Particle Swarm Optimization and Evolutionary Strategies (Covariance Matrix Adaptation)

OVERVIEW

Choice of the optimization algorithm:

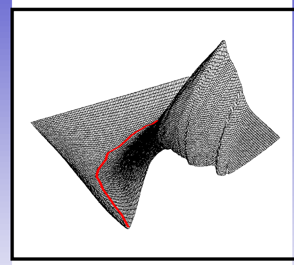
1 – Multiple regions of attraction

2 – UNCOUNTABLE local optima

3 – Discontinuous derivatives

4 – Long and curved ridges

5 – Poor sensitivity



Global search of space

Derivative-free method

OVERVIEW

Well-posed inverse problems:

- Test for global and local minima
- Test for unique solutions
- Independently measure parameters that are not sensitive to solution
- Do not estimate highly correlated parameters
- Include independently-measured information to objective function
- Minimize number of optimized parameters
- Minimize measurement errors
- Estimate model error
- Compare uncertainties of optimized parameters

OVERVIEW

Other applications of inverse modeling:

- Other soil hydraulic properties techniques, such as evaporation method, suction infiltrometer method and instantaneous profile method;
- Estimation of solute and heat transport properties;
- Estimation of root water and nutrient uptake parameters;
- Effective field soil properties, and in multi-layered systems;
- . . . .

OVERVIEW

LIMITATIONS:

- Inverse problems are not necessarily well-posed;
- Selection of weighting factors;
- Parameter estimates are valid for experimental range only;
- Method requires a lot of experience

Non-uniqueness

Instability

OVERVIEW

BENEFITS:

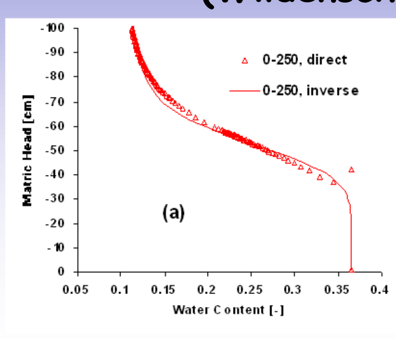
- Mandates marriage of experimentation with numerical modeling;
- Method is consistent, I.e. estimated hydraulic functions are used in model predictions;
- Uses transient measurements, as in real world;
- Relatively fast method, and lends itself for automation

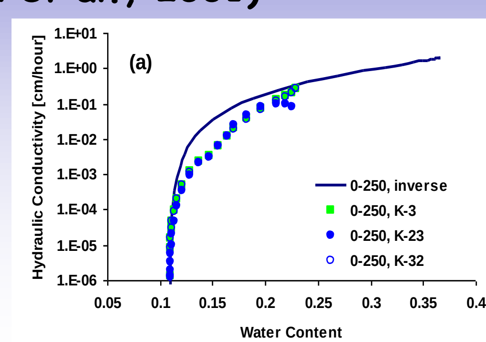
WHY NEED FOR MEASUREMENT OF SOIL HYDRAULIC PARAMETERS ?

- As input to water flow and contaminant transport models;
- To characterize soil physical characteristics, including their spatial and temporal variability;
- To correlate with other, more easily to measure soil physical properties, e.g. texture.

Multi-step outflow method to indirectly estimate soil water retention and unsaturated hydraulic conductivity functions

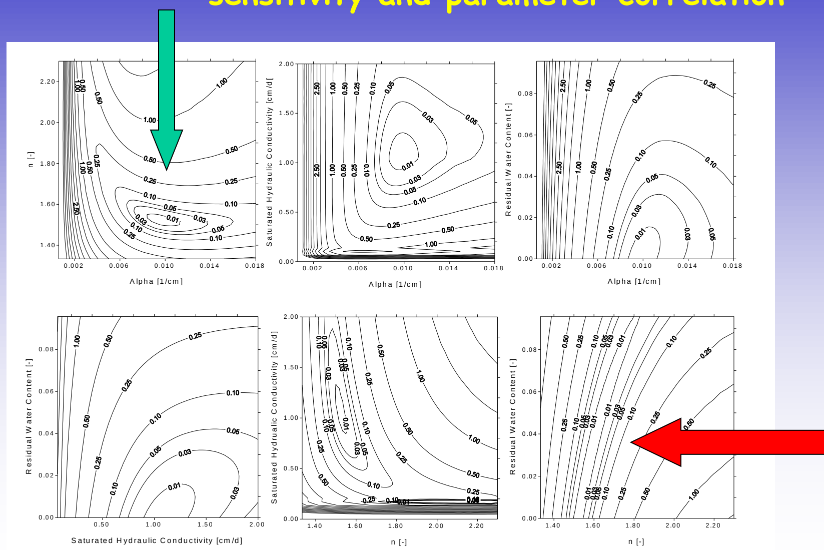
Lincoln sand
(Wildenschild et al., 2001)





RESPONSE SURFACE ANALYSIS:

Can also be used to investigate parameter sensitivity and parameter correlation



ITERATIVE METHODS FOR PARAMETER ESTIMATION

MANUAL PARAMETER ESTIMATION (HAND CALIBRATION)

Advantage: Simple to implement

Disadvantage: subjective, time-consuming and requires considerable experience

COMPUTERIZED ALGORITHMS (AUTOMATIC CALIBRATION)

Advantage: Objective and more efficient

Disadvantage: Complicated and typically requires programming experience and familiarity with technical jargon and cluster computers

SOME PRELIMINARY CONCLUSIONS …

Global search methods can handle complex response surfaces with multimodal optima and are therefore capable of handling a relatively large number of parameters