DUE TODAY Assgn#5

Assignment # 5 Steady State Numerical Models: Create two steady state MODFLOW simulations of groundwater flow in your system. One of the simulations should represent the flow system without the stress and the other should simulate the steady state condition with the stress. Build on your work from assignments 1 through 3. Using the MODFLOW manuals and class examples, create a name file, then build each of the input files. When you have them all, execute the model, look at the output or error messages and revise the file until you have models that "run". Be sure to save your files because you will want to use them later in the semester. **Compare your results to the result of your spreadsheet and analytical modeling.** Be sure to save your files because you will want to use them later in the semester.

Suggested Steady State Modeling Report Outline
Title
Introduction
- **objective**
- **problem description**

Geohydrologic Setting
Results of analytical and spreadsheet modeling
Numerical Model setup
- **geometry**
- **boundary conditions**
- **initial conditions**
- **parameter value ranges**
- **stresses**
- **special considerations**

Uncalibrated model results
- **predictions**
- **problems encountered, if any**

Comparison with Analytical/Spreadsheet results
Assessment of future work needed, if appropriate
Summary/Conclusions
References

submit the paper as hard copy and include it in your zip file of model input and output
submit the model files (input and output for both simulations) in a zip file labeled:
ASSGN5_LASTNAME.ZIP

Calibration

(Parameter Estimation, Optimization, Inversion, Regression)

adjusting parameter values, boundary conditions, model conceptualization, and/or model construction until the model simulation matches field observations

We calibrate because

1. the field measurements are not accurate reflections of the model scale properties, and

2. calibration provides integrated interpretation of the available data
(e.g. the dependent observations tell us about the independent properties)

Calibrated model ~ provides minimized residuals (Observed – Simulated) without bias (N indicates the number of observations)

Global measures of error:
Mean Error: (Sum(Obs-Sim))/N
Mean Absolute Error: (Sum(ABS(Obs-Sim)))/N
Root Mean Squared Error: $((\text{Sum}((\text{Obs}-\text{Sim})^2))/N)^{0.5}$
Sum-of-Squared Weighted Residuals: $\text{Sum}(\text{weight}(\text{Obs}-\text{Sim})^2)$

Graphical measures of error
observed vs. simulated should form a 45° line passing through the origin
residual vs. simulated should form a uniform horizontal band around zero
ordered residuals on normal probability graph should form a straight line

Spatial and Temporal Distribution of Residuals
Map (obs-sim) in x, y, z space should exhibit a random pattern of positive and negative, as well as large and small, residuals

Graph of (obs-sim) vs. time OR vs. observation # should form a uniform horizontal band centered on zero

ALSO USE COMMON SENSE to spot errors

Optimal Parameter Values are the result of the calibration
They should correspond with field measured values

If they differ significantly carefully consider whether:
- such a difference is reasonable due to scale issues
- the conceptual model is in error, or
- there errors in the field data

Have expectations, question all aspects of the situation when calculations do not match expectations

We will use automated calibration (here nonlinear regression), it is a valuable tool for:
- finding the best fit to your field observations
- identifying the type and location of additional data that will be most helpful
- differentiating conceptual models
- identifying models that are most representative of the field

Unfortunately, many practicing ground-water professionals are still using trial-and-error but it is changing rapidly

Our objective is to minimize the sum of squared weighted residuals:

$$S(\underline{b}) = \sum_{i=1}^{ND} \omega_i \left[y_i - y_i'(\underline{b}) \right]^2$$ $\boxed{\text{Objective Function}}$

$\underline{b}$ vector of estimated parameter values 1xNP

ND number of observations

NP number of parameters being estimated

y_i i^{th} observation (head, flux, concentration)

$y_i'(\underline{b})$ modeled equivalent of the i^{th} observation

ω_i weight of the i^{th} observation

Weighting Squared Residuals because Observations are:

1. **not equally reliable** (some heads may have been determined from survey TOC (top of casing) while other TOCs were estimated from a topographic map)
2. **have different units** (a difference of 1 foot in head may not have the same importance as a difference of 1cfs flow rate)
3. **have true errors that are correlated** (e.g. many h obs @ one well but elevation of well or position of well is in error)

We deal with these issues through weighting observations. Research has indicated that ignoring the correlation of error between observations does not significantly influence the regression, but we can include them if we wish.

Using 1/variance of the measurement as the weight renders the weighted squared residuals unitless and assigns high weights to more accurate observations. THEREFORE we can sum weighted squared residuals and regardless of the units or magnitudes, they are of equal importance, except for their measurement certainty.

A reasonable weight is the inverse of the measurement variance because more uncertain observations receive less weight and the weighted squared residuals are unitless and so can be summed

Observations have **units** (e.g. ft cfs mg/L etc)

Observations have uncertainty (e.g. measurement variance σ^2 which is the square of the standard deviation σ)

Standard deviation has same units as the observation (eg ft cfs mg/L etc)

Variance has observation units squared ft^2 $(cfs)^2$ $(mg/L)^2$ etc

Residuals have the same units as the observation ft cfs mg/L etc

Squared residuals have units that are squared observation units
ft^2 $(cfs)^2$ $(mg/L)^2$ etc

Weight = 1 / Variance units are the inverse of squared residuals
ft^{-2} $(cfs)^{-2}$ $(mg/L)^{-2}$ etc

Sum of Weighted Squared Residuals (∑wt*squared residual) unitless

for example:

say heads are accurate to within 1 ft of measurement

express this quantitatively as 95% confidence that head is within 1 ft of measurement

using a cumulative distribution of a standard normal distribution table we find a 95% confidence is 1.96 standard deviations, so

1.96 stddev = 1.0 ft

stddev = 0.51 ft

variance is the stddev squared

variance = 0.26 $(ft)^2$

and the weight is 1/variance

weight 3.85 $(ft)^{-2}$

NOTE

heads are derived from an elevation AND depth measurement

Variances can be summed (standard deviations cannot)

example

95% confidence elevation is 120ft +/- 10ft

1 stddev ~ 5.1 ft

variance ~ 26 ft^2

95% confidence depth to water 25ft +/- 0.1ft

1 stddev ~ 0.051 ft

variance ~ 0.0026 ft^2

Variance on the head measurement ~ 26 + 0.0026 ft^2

1 stddev ~ square root of 26.0026

Weight ~ 0.0385 ft^{-2}

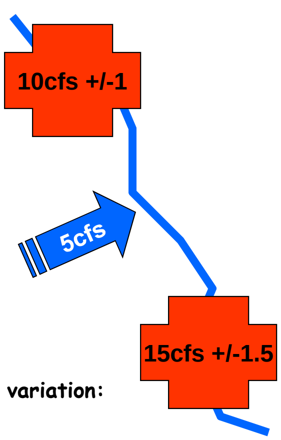
in the case of ground water flow measurements, 2 measurements are required and their variance must be combined (actually this is usually the case with head also because both top of casing and depth to water are needed):

Upstream Q = 10cfs   Downstream Q = 15cfs
95% certain upstream measurement is within 1 cfs
90% certain downstream measurement is within 1.5 cfs

upstream
95% confidence is   1.96stddev = 1.0 cfs
stddev = 0.51 cfs
variance is the stddev squared = 0.26 $(cfs)^2$

downstream
90% confidence is   1.65stddev = 1.5 cfs
stddev = 0.90 cfs
variance is the stddev squared = 0.81 $(cfs)^2$

the groundwater flux is 15cfs - 10cfs = 5 cfs
the variance is the sum of the individual variances
variance = 0.26 + 0.81 = 1.07 $(cfs)^2$
weight = 1/variance = 0.93 $(cfs)^{-2}$



Sometimes we express uncertainty as coefficient of variation:
stddev / mean
we use the measured value as the mean
coeff var = $(1.07\ (cfs)^2)^{-2}$ / 5 cfs = 0.21 cfs

The errors associated with observations that share measurements are correlated.

Generally this does not have a big influence on the regression or the associated statistics.

We can accommodate this with a full weight matrix.

In the case of stream flow observations with a shared gage, the off-diagonal variance is:

–(variance of the measurement at the shared gage)

The regression is not extremely sensitive to the weights, thus the casual approach to their definition is not a problem

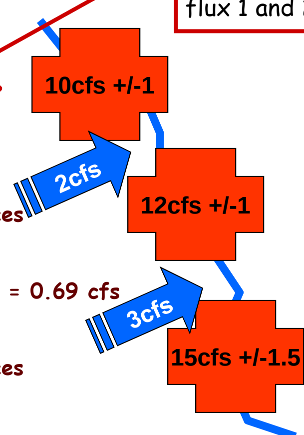
The weighting can be evaluated at the end of the regression by considering the cev (calculated error variance) more on that later

---

upstream = 10 cfs midstream = 12 downstream = 15 cfs
95% certain upstream measurement is within 1 cfs
95% certain midstream measurement is within 1 cfs
90% certain downstream measurement is within 1.5 cfs

upstream
variance is the stddev squared = 0.26 $(cfs)^2$
midstream
95% confidence is 1.96stddev = 1.0 cfs
stddev = 0.51 cfs
variance is the stddev squared = 0.26 $(cfs)^2$
downstream
variance is the stddev squared = 0.81 $(cfs)^2$

-0.26$(cfs)^2$ expresses the covariance between groundwater flux 1 and 2



the groundwater flux1 is 12cfs - 10cfs = 2 cfs
the variance is the sum of the individual variances
variance = 0.52 $(cfs)^2$
weight = 1/variance = 1.92 $(cfs)^{-2}$
coefficient of variation = $(1.92\ (cfs)^2)^{-2}$ / 2 cfs = 0.69 cfs

the groundwater flux2 is 15cfs - 12cfs = 3 cfs
the variance is the sum of the individual variances
variance = 1.07 $(cfs)^2$
weight = 1/variance = 0.93 $(cfs)^{-2}$
coefficient of variation = $(1.07\ (cfs)^2)^{-2}$ / 3 cfs = 0.34 cfs

Recall our objective is to minimize the sum of squared weighted residuals:

$$\mathbf{S(b)} = \sum_{\mathbf{i=1}}^{\mathbf{ND}} \boldsymbol{\omega}_{\mathbf{i}} \left[\mathbf{y}_{\mathbf{i}} - \mathbf{y}_{\mathbf{i}}^{'}(\underline{\mathbf{b}})\right]^{\mathbf{2}}$$

in matrix form

$$\mathbf{S(b)} = \left[\underline{\mathbf{y}} - \underline{\mathbf{y}}'(\underline{\mathbf{b}})\right]^{\mathbf{T}} \underline{\underline{\boldsymbol{\omega}}} \left[\underline{\mathbf{y}} - \underline{\mathbf{y}}'(\underline{\mathbf{b}})\right]$$

T indicates rows and columns of matrix are transposed

$$S(b) = \left[\underline{y} - \underline{y}'(\underline{b})\right]^T \underline{\underline{\omega}} \left[\underline{y} - \underline{y}'(\underline{b})\right]$$

$$\underline{e} = \begin{bmatrix} y_1 - y_1'(\underline{b}) \\ y_2 - y_2'(\underline{b}) \\ \bullet \\ \bullet \\ y_{ND} - y_{ND}'(\underline{b}) \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ \bullet \\ \bullet \\ e_{ND} \end{bmatrix} \quad \underline{\underline{\omega}} = \begin{bmatrix} \omega_{1,1} & \omega_{1,2} & \cdots & \cdots & \omega_{1,ND} \\ \omega_{2,1} & \omega_{2,2} & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \omega_{ND,1} & \cdots & \cdots & \cdots & \omega_{ND,ND} \end{bmatrix}$$

$$S(b) = [\underline{e}]^T \underline{\underline{\omega}} [\underline{e}]$$

Recall matrix multiplication (review from Wikipedia):

The product of an $m \times p$ matrix A with an $p \times n$ matrix B is an $m \times n$ matrix denoted AB whose entries are

$$(AB)_{i,j} = \sum_{k=1}^{p} A_{ik} B_{kj}$$

where $1 \le i \le m$ is the row index and $1 \le j \le n$ is the column index.

Example from Wikipedia:

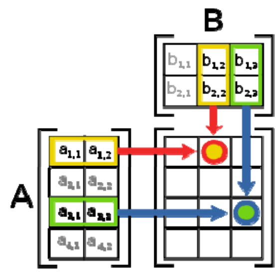
$$A = \begin{bmatrix} 14 & 9 & 3 \\ 2 & 11 & 15 \\ 0 & 12 & 17 \\ 5 & 2 & 3 \end{bmatrix}$$

$$B = \begin{bmatrix} 12 & 25 \\ 9 & 10 \\ 8 & 5 \end{bmatrix}$$

$14 \times 12 + 9 \times 9 + 3 \times 8$ AB11=273

$14 \times 25 + 9 \times 10 + 3 \times 5$ AB21=455

$$AB = \begin{bmatrix} 14 & 9 & 3 \\ 2 & 11 & 15 \\ 0 & 12 & 17 \\ 5 & 2 & 3 \end{bmatrix} \begin{bmatrix} 12 & 25 \\ 9 & 10 \\ 8 & 5 \end{bmatrix} = \begin{bmatrix} 273 & 455 \\ 243 & 235 \\ 244 & 205 \\ 102 & 160 \end{bmatrix}$$



A useful rule to remember when evaluating matrix multiplication is that the adjacent dimensions must match and the final matrix dimensions will be the # of rows of the first matrix and the number of columns of the second.

So for us there are N observations

1x1 results from 1xN NxN Nx1

$$S(b) = \left[\underline{y} - \underline{y}'(\underline{b})\right]^T \underline{\omega} \left[\underline{y} - \underline{y}'(\underline{b})\right]$$

Sometimes it is assumed the off diagonal terms of the weight matrix are zero and it is presented as follows

$$S(b) = \left[\underline{y} - \underline{y}'(\underline{b})\right]^T \underline{\omega} \left[\underline{y} - \underline{y}'(\underline{b})\right]$$

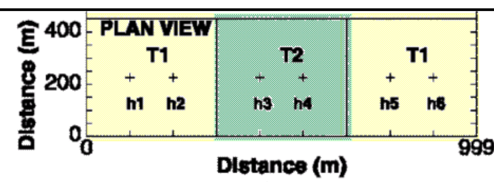
$$\underline{e} = \begin{bmatrix} y_1 - y_1'(\underline{b}) \\ y_2 - y_2'(\underline{b}) \\ \bullet \\ \bullet \\ y_{ND} - y_{ND}'(\underline{b}) \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ \bullet \\ \bullet \\ e_{ND} \end{bmatrix} \qquad \underline{\omega} = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \bullet \\ \bullet \\ \omega_{ND} \end{bmatrix}$$

$$S(b) = [\underline{e}]^T \underline{\omega} [\underline{e}]$$

but the w is an Nx1 that is the diagonal of the NxN which is how the math is conducted and a
1x1 results from 1xN NxN Nx1

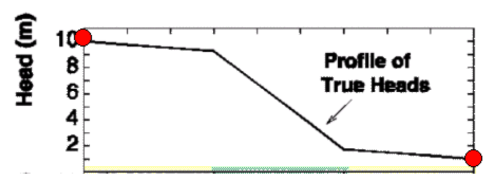
But in this case it is easy to sum the weighted squared residuals by hand to confirm the matrix algebra

Let's look at some simple examples of the sum of squared weighted residuals



T1 = 1.0 m^2/s T2 = 0.1 m^2/s

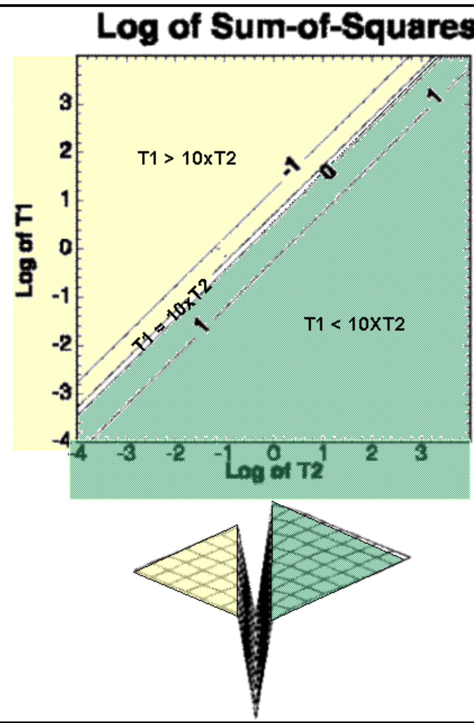
The heads are fixed at each end, on the left at 10m and on the right at 1m, for a gradient of 0.009.



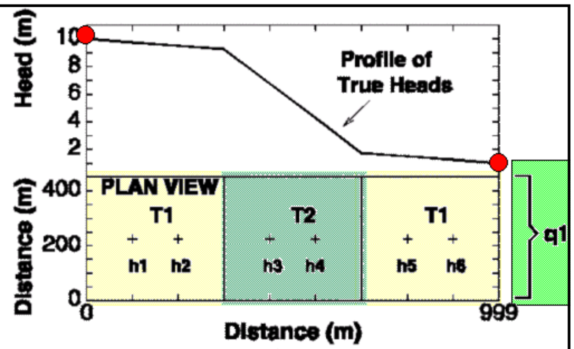
This results in the following accurate head observations from the system:

Observation Type and #	Value (m)
h1	9.75
h2	9.50
h3	5.75
h4	4.25
h5	1.50
h6	1.25

With the sum of squared weighted residuals being one value, for a simple 2 parameter estimation problem we can plot it on a graph and look at the surface



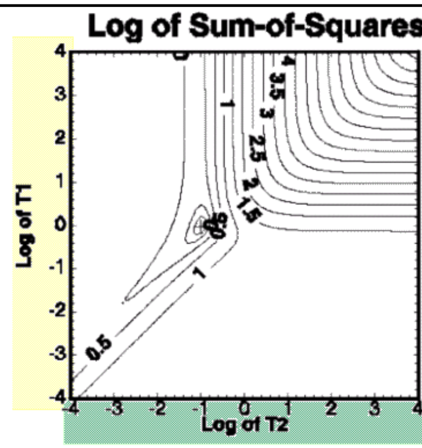
Here is the same example but now we add a flow observation and note the affect on the sum of squared weighted residuals

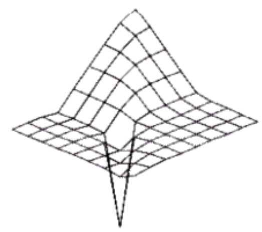


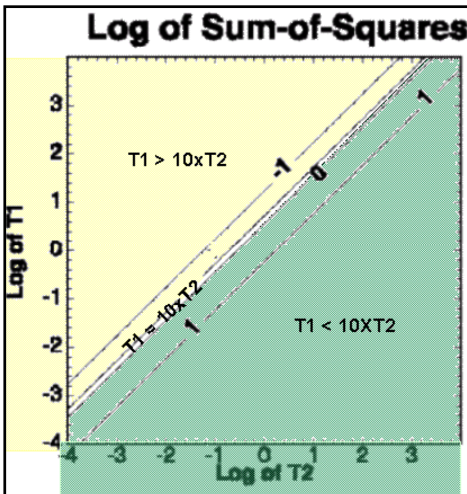
T1 = 1.0 m^2/s T2 = 0.1 m^2/s

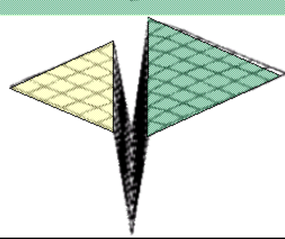
Observation Type and #	Value (m)
h1	9.75
h2	9.50
h3	5.75
h4	4.25
h5	1.50
h6	1.25
q1	-1.0135

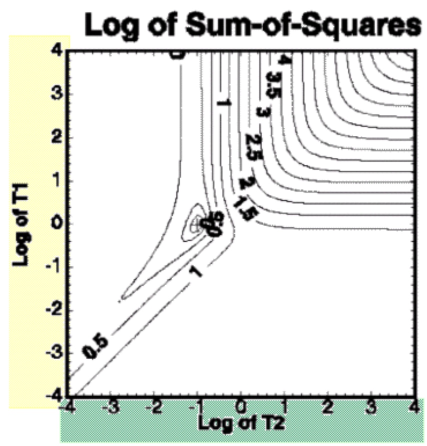
Here is the sum of squared weighted residuals surface with the flow observation included. Note that now it is possible to find a unique solution

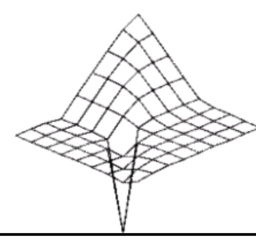












Sometimes we include prior knowledge of the parameter values from independent tests as observations (for diagonal weights):

$$S(b) = \sum_{i=1}^{ND} \omega_i \left[y_i - y_i'(\underline{b}) \right]^2 + \sum_{p=1}^{NPR} \omega_p \left[P_p - P_p'(\underline{b}) \right]^2$$

NPR number of *prior* information values

P_p pth prior estimate

$P_p'(\underline{b})$ pth modeled equivalent of prior estimate

ω_p weight on pth modeled equivalent of prior estimate

Considering a full weight matrix

$$\underline{e} = \begin{bmatrix} y_1 - y_1'(\underline{b}) \\ y_2 - y_2'(\underline{b}) \\ \bullet \\ \bullet \\ y_{ND} - y_{ND}'(\underline{b}) \\ p_1 - p_1' \\ p_2 - p_2' \\ \bullet \\ p_{N\,\mathrm{Pri}} - p_{N\,\mathrm{Pri}}' \end{bmatrix} = \begin{bmatrix} e_1 \\ e_2 \\ \bullet \\ \bullet \\ e_{ND} \\ e_{1\mathrm{Pri}} \\ e_{2\mathrm{Pri}} \\ \bullet \\ e_{N\,\mathrm{Pri}} \end{bmatrix} \qquad \underline{\omega} = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & \cdots & \cdots & 0 & 0 & 0 & 0 \\ w_{2,1} & \cdots & \cdots & \cdots & \cdots & 0 & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & w_{ND,ND} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & w_{pri1,pri1} & w_{pri1,pri2} & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & w_{pri2,pri1} & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & \cdots & \cdots & w_{Npri,Npri} \end{bmatrix}$$

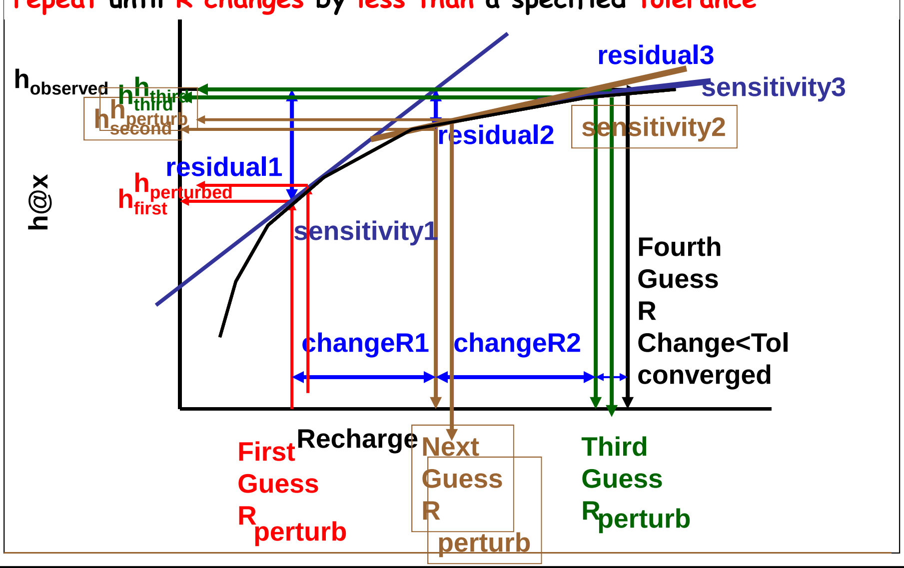
MODFLOW

Observation / Sensitivity / Parameter Estimation
User's manual is a separate physical document but is integrated into the on-line guide for MODFLOW

All 3 processes in MODFLOW2000
But this is being discontinued

MODFLOW2005 some OBSERVATION Packages
These will be enhanced
Sensitivity will be added
Parameter estimation will be replaced by UCODE

Consider how we could go about estimating parameter values for the following nonlinear model. We guess a recharge R, calculate h, determine residual, use that and the slope (sensitivity) to make a linear estimate of the best R, and because it is nonlinear, we repeat until R changes by less than a specified tolerance



To examine how we find the parameters that produce the minimum sum-of-squared residuals, reconsider the simplest form of the objective function:

$$S(b) = \left[\underline{y} - \underline{y}'(\underline{b})\right]^T \underline{\underline{\omega}} \left[\underline{y} - \underline{y}'(\underline{b})\right]$$

take the derivative of the objective function with respect to the parameters and set it to zero:

$$\frac{\partial}{\partial \underline{b}} \left[\underline{y} - \underline{y}'(\underline{b})\right]^T \underline{\underline{\omega}} \left[\underline{y} - \underline{y}'(\underline{b})\right] = \underline{0}$$

$\underline{0}$ has NP elements, all zero

linearize the objective function with

the first 2 terms of the Taylor Series Expansion

$$S(b) = \left[\underline{y} - \underline{y}'(\underline{b})\right]^T \underline{\underline{\omega}} \left[\underline{y} - \underline{y}'(\underline{b})\right]$$

$$\underline{\mathrm{y}}'(\underline{\mathrm{b}})\ is\ \underline{\mathrm{y}}^{\mathrm{linearized}}(\underline{\mathrm{b}})$$

The residuals are determined as observed-simulated (at the current parameter values)

They form a 1D array (ND, # observations long)

$$\lfloor \underline{y} - \underline{y}'(\underline{b}) \rfloor =$$

$$\begin{bmatrix} y_1 \\ y_2 \\ . \\ y_{ND} \end{bmatrix} - \begin{bmatrix} y'_1(b_o) \\ y'_2(b_o) \\ . \\ y'_{ND}(b_o) \end{bmatrix} = \begin{bmatrix} residual_1 \\ residual_2 \\ . \\ residual_{ND} \end{bmatrix}$$

The value of y at b, where we anticipate the residuals will be minimal, is approximated linearly by the value at b_0 + the "slope" times the residual.

We have an array of b and b_0 and the slope is the sensitivity linearized $\underline{y}(\underline{b}) =$

$$y^{lin}(\underline{b}) \cong \underline{y}'(\underline{b}_0) + \left.\frac{\partial \underline{y}'(\underline{b})}{\partial \underline{b}}\right|_{\underline{b}=\underline{b}_0} (\underline{b} - \underline{b}_0)$$

we can write the linearized form in terms of the sensitivity matrix X evaluated at b_0

$$\underline{y}^{lin}(\underline{b}) \cong \underline{y}'(\underline{b}_0) + \underline{\underline{X}}\Big|_{\underline{b}=\underline{b}_0} (\underline{b} - \underline{b}_0)$$

$\underline{\underline{X}}$ = sensitivity matrix (Jacobian)

elements are $\frac{\partial y_i'}{\partial b_i}$

CONSIDER the SENSITIVITY MATRIX

y'(b) has ND + NPR elements

ND = # observations

NPR = # prior observations of parameters

b has NP elements

NP = # parameters

So the sensitivity matrix X has ND+NPR rows & NP columns

$$\begin{array}{cc} & \xrightarrow{NP} \\ \begin{array}{c} ND \\ + \\ NPR \end{array}\Bigg\downarrow & \underline{\underline{X}} = \begin{bmatrix} \frac{\partial y_1'}{\partial b_1} & \frac{\partial y_1'}{\partial b_2} & \cdots & \frac{\partial y_1'}{\partial b_{NP}} \\ \frac{\partial y_2'}{\partial b_1} & \frac{\partial y_2'}{\partial b_2} & \cdots & \frac{\partial y_2'}{\partial b_{NP}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial y_{ND}'}{\partial b_1} & \frac{\partial y_{ND}'}{\partial b_2} & \cdots & \frac{\partial y_{ND}'}{\partial b_{NP}} \end{bmatrix} \end{array}$$

Each sensitivity is determined as:
[simulated(current b values)- simulated(perturbed b values)] / [(current b)-(perturbed b)]

i.e.

$$\frac{\text{simulated}(b_o) - \text{simulated}(b')}{b_o - b'}$$

Estimating Parameter Values that Minimize the Sum of Weighted Squared Residuals via Nonlinear Regression using the Modified Gauss-Newton Gradient Method (also called Marquardt-Levenberg)

An iterative form of linear regression (i.e. solves normal equations like you do to fit a straight line to data, but repeatedly with updated parameter values)

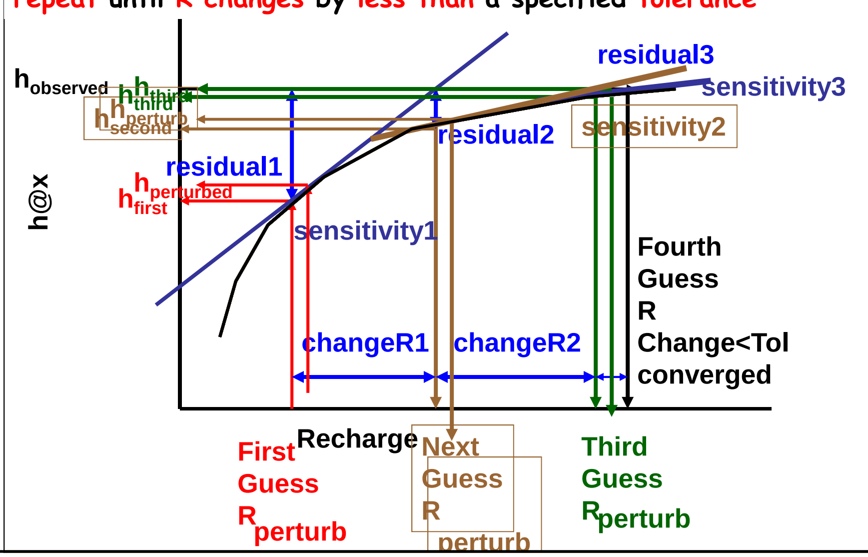
To do this we minimize the objective function (i.e. we obtain the normal equations by assuming linearity and taking the derivative with respect to the parameters, then set the derivative equal to zero to find the parameter values that would minimize the function)

The ground water flow equations are not linear with respect to the parameters, so we repeat the process using the new parameter values and continue until there is little change in the parameter values

This only works well for non-linear problems IF MODIFIED to include:

* scaling
* adjusting to gradient correction
* damping

Consider how we could go about estimating parameter values for the following nonlinear model. We guess a recharge R, calculate h, determine residual, use that and the slope (sensitivity) to make a linear estimate of the best R, and because it is nonlinear, we repeat until R changes by less than a specified tolerance



Recall to find the parameters that produce the minimum sum-of-squared residuals, we set the derivative of the objective function to zero. This produces the normal equations.

We'll do this using the simplest form of the objective function:

$$S(b) = \left[\underline{y} - \underline{y}'(\underline{b})\right]^T \underline{\underline{\omega}} \left[\underline{y} - \underline{y}'(\underline{b})\right]$$

$$\frac{\partial}{\partial \underline{b}} \left[\underline{y} - \underline{y}'(\underline{b})\right]^T \underline{\underline{\omega}} \left[\underline{y} - \underline{y}'(\underline{b})\right] = 0$$

Substitute the linear expression for $\underline{y}'(\underline{b})$

$$\underline{y}^{lin}(\underline{b}) \cong \underline{y}'(\underline{b}_0) + \underline{\underline{X}}\Big|_{\underline{b}=\underline{b}_0} (\underline{b} - \underline{b}_0)$$

$$\frac{\partial}{\partial \underline{b}} \left[\underline{y} - \left(\underline{y}'(\underline{b}_0) + \underline{\underline{X}}(\underline{b}_0) * (\underline{b} - \underline{b}_0)\right)\right]^T$$

$$\underline{\underline{\omega}} \quad \left[\underline{y} - \left(\underline{y}'(\underline{b}_0) + \underline{\underline{X}(\underline{b}_0)} * (\underline{b} - \underline{b}_0)\right)\right] = 0$$

abbreviate :

$$\frac{\partial}{\partial \underline{b}} \left[\underline{y} - \left(\underline{y}' + \underline{\underline{X}} \Delta \underline{b}\right)\right]^T \underline{\underline{\omega}} \left[\underline{y} - \left(\underline{y}' + \underline{\underline{X}} \Delta \underline{b}\right)\right] = 0$$

After some mathematical considerations that we will not take time for here, we calculate the change in the parameters that is required (assuming a linear model) to minimize the residuals for 1 iteration:

Vector $(\underline{d}_i)$ defines the amount each parameter needs to change

$$[\underline{\underline{X}}]_{iter}^{T}\, \underline{\underline{\omega}}\, [\underline{\underline{X}}]_{iter}\, \underline{d}_{iter} = \underline{\underline{X}}_{iter}^{T}\, \underline{\underline{\omega}} \left(\underline{y} - \underline{y}'(\underline{b_{iter}})\right)$$

updated parameters are $\underline{b}_{iter+1} = \underline{b}_{iter} + \underline{d}_{iter}$

but not best fit for a nonlinear model,

so repeat at new $(\underline{b})$

Conceptually:

$$\left(\underline{\underline{X}}_r^{T}\, \underline{\underline{\omega}}\, \underline{\underline{X}}_r\right) \underline{d}_r = \underline{\underline{X}}_r^{T}\, \underline{\underline{\omega}} \left(\underline{y} - \underline{y}'(\underline{b}_r)\right)$$

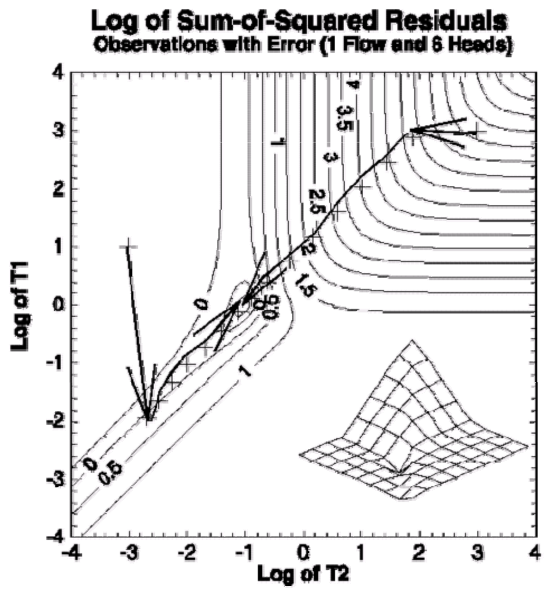
right hand side = steepest descent

left coefficient = modifies direction

for a better route to the minimum

Some modifications are needed to put this to work in a practical sense:
SCALING
DAMPING
ADJUSTING to GRADIENT DIRECTION

First, just an image of “the route” to the minimum:



Scale to account for large differences in parameter values and sensitivities for improved accuracy of d

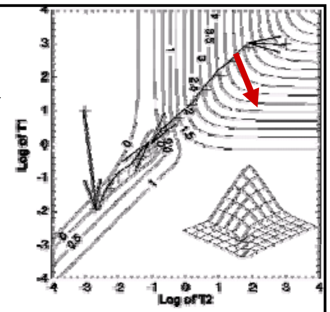
$$\left(\underline{\underline{C}}^T \underline{\underline{X}}_r^T \underline{\underline{\omega}} \underline{\underline{X}}_r \underline{\underline{C}}\right)\underline{\underline{C}}^{-1}\underline{d}_r = \underline{\underline{C}}^T \underline{\underline{X}}_r^T \underline{\underline{\omega}}\left(\underline{y} - \underline{y}'(\underline{b}_r)\right)$$

$\underline{\underline{C}}$ = diagonal scaling matrix with element

c_{jj} equal to $\left(\underline{\underline{X}}^T \underline{\underline{\omega}}\underline{\underline{X}}_{jj}\right)^{-0.5}$

producing a scaled matrix with the smallest condition number

If the direction vector nearly parallels the contours of the objective function



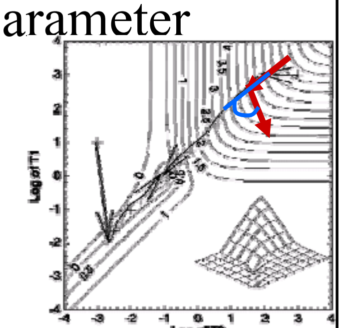
Marquardt parameter changes direction to be more down gradient

$$\left(\underline{\underline{C}}^T \underline{\underline{X}}_r^T \underline{\underline{\omega}}\,\underline{\underline{X}}_r \underline{\underline{C}} + \underline{\underline{I}}\, m_r\right)\underline{\underline{C}}^{-1}\underline{d}_r = \underline{\underline{C}}^T \underline{\underline{X}}_r^T \underline{\underline{\omega}}\left(\underline{y} - \underline{y}'(\underline{b}_r)\right)$$

$\underline{\underline{I}} = NPxNP$ identity matrix

$m_r = Marquardt\ \ parameter$ this iteration

if the down gradient vector of this iteration and the last is more than some angle, commonly $\approx 87^{\circ}$ apart, then the Marquardt parameter is included in calculating the parameter change vector



$$m_r^{new} = 1.5 m_r^{old} + 0.001$$

with each iteration

to avoid repeated overshoot of the minimum

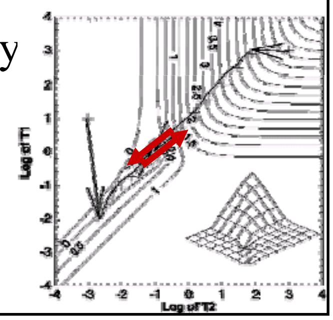
a damping parameter is used, ρ_r from 0 to 1.0

changes magnitude, but not direction of new parameter estimates

$\underline{b}_{r+1} = \underline{b}_r + \rho_r \underline{d}_r$

one criterion is to keep change less than a fractional maximum (specified by user) for any parameter

$$e.g. \frac{\underline{b}_i^{r+1} - \underline{b}_i^r}{\underline{b}_i^r} < 2$$



Gauss-Newton approach:

We solve iteratively for d:

$$\underline{d}_r = \left(\underline{\underline{X}}_r^T \underline{\underline{\omega}} \underline{\underline{X}}_r\right)^{-1} \underline{\underline{X}}_r^T \underline{\underline{\omega}} \left(\underline{y} - \underline{y}'(\underline{b}_r)\right)$$

Modified Gauss-Newton approach

scale(C) adjust direction(m) damp(ρ)

$$\underline{d}_r = \left(\underline{\underline{C}}^T \underline{\underline{X}}_r^T \underline{\underline{\omega}} \underline{\underline{X}}_r \underline{\underline{C}} + \underline{\underline{I}} m_r\right)^{-1} \underline{\underline{C}} \underline{\underline{C}}^T \underline{\underline{X}}_r^T \underline{\underline{\omega}} \left(\underline{y} - \underline{y}'(\underline{b}_r)\right)$$

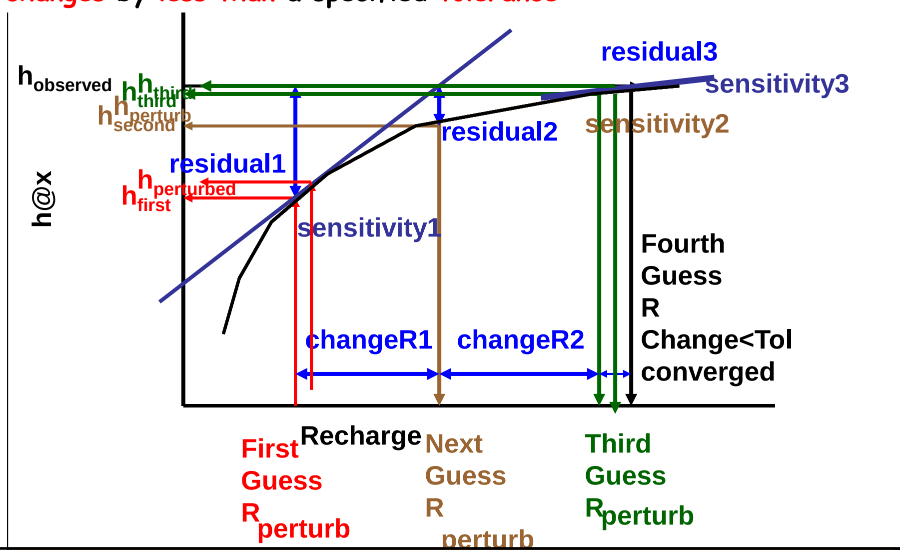
And update b:

$$\underline{b}_{r+1} = \underline{b}_r + \rho_r \underline{d}_r$$

REPEAT UNTIL THE DISPLACEMENT VECTOR d
is LESS THAN TOLERANCE
Typically 1% change in parameters

Once optimal parameters are found, evaluate:
PARAMETER STATISTICS
RESIDUAL STATISTICS
To assess quality of the model

RECALL: When the situation is nonlinear we assume linearity and keep trying until parameter values do not change much. We guess R, calculate h, determine residual, use that and the slope (sensitivity) to linearly estimate the "best" R, and because it is nonlinear, repeat until R changes by less than a specified tolerance



groundwater flow equations are not a linear function of the parameters,
even though confined groundwater flow equations are a linear function of space and time

$$Q = -KA\frac{\partial h}{\partial x} = -KA\frac{\Delta h}{\Delta x} = -KA\frac{h_x - h_1}{x}$$

Expand gradient in Darcy's Law

$$\frac{Q}{-KA}x + h_1 = h_x$$

Rearrange for h of x

$$h_x = h_1 - \frac{Q}{KA}x$$ *derivative* h *with respect to x is*

$$\frac{\partial h}{\partial x} = \frac{Q}{KA}$$

linear because independent of x

$$h_x = h_1 - \frac{Q}{KA}x$$ derivative of h with respect to Q is

$$\frac{\partial h}{\partial Q} = -\frac{1}{KA}x$$

linear, because independent of Q, but is dependent on K which can have a nonlinear impact on h

derivative with respect to K

$$\frac{\partial h}{\partial K} = \frac{Q}{K^2 A}x$$ is nonlinear

SUM OF WEIGHTED SQUARED RESIDUALS

$$S(b) = \sum \omega \left(s_{RESIDUAL}\right)^2$$

CALCULATED ERROR VARIANCE

$$cev = s^2 = \frac{S(b)}{ND - NP}$$

STANDARD ERROR

$$\mathbf{s = \sqrt{s^2}}$$

VARIANCE/COVARIANCE MATRIX

$$COV = cev\left(\underline{\underline{X}}^T \underline{\underline{\omega}}\,\underline{\underline{X}}\right)^{-1}$$

$$\begin{array}{c|cccc|}
 & j=1 & \bullet & \bullet & j=NP \\
i=1 & 1,1 & 1,2 & \bullet & 1,NP \\
\bullet & 2,1 & 2,2 & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet \\
i=NP & NP,1 & NP,2 & NP,3 & NP,NP
\end{array}$$

If 2 parameters were estimated:

$$\begin{array}{c|cc|}
 & b1 & b2 \\
b1 & Var_1 & Cov_{1,2} \\
b2 & Cov_{2,1} & Var_2
\end{array}$$

VARIANCE (b1)

$$VAR(b1) = \left(\underline{\underline{X}}^T \underline{\underline{\omega}}\,\underline{\underline{X}}\right)^{-1}_{1,1}(EVAR)$$

$$Std\ Dev = \sqrt{VAR\ (b1)} \qquad 95\%\ Confid = b1 +/- 2 * StdDev$$

VARIANCE (b2)

$$VAR(b2) = \left(\underline{\underline{X}}^T \underline{\underline{\omega}}\,\underline{\underline{X}}\right)^{-1}_{2,2}(EVAR)$$

$$Std\ Dev = \sqrt{VAR\ (b2)} \qquad 95\%\ Confid = b2 +/- 2 * StdDev$$

Confidence interval on parameters

Later we use this for confidence interval on predictions

CORRELATION (normalized variance)

$$CORR(i,j) = \frac{COV(i,j)}{\sqrt{VAR(i)} * \sqrt{VAR(j)}}$$

$$\begin{array}{c c} & \begin{array}{cccc} j=1 & \bullet & \bullet & j=NP \end{array} \\ \begin{array}{c} i=1 \\ \bullet \\ \bullet \\ i=NP \end{array} & \begin{bmatrix} 1,1 & 1,2 & \bullet & 1,NP \\ 2,1 & 2,2 & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ NP,1 & NP,2 & NP,3 & NP,NP \end{bmatrix} \end{array}$$

If 2 parameters were estimated:

$$\begin{array}{c c} & \begin{array}{cc} b1 & b2 \end{array} \\ \begin{array}{c} b1 \\ b2 \end{array} & \begin{bmatrix} 1 & Cor_{b1,b2} \\ Cor_{b2,b1} & 1 \end{bmatrix} \end{array}$$

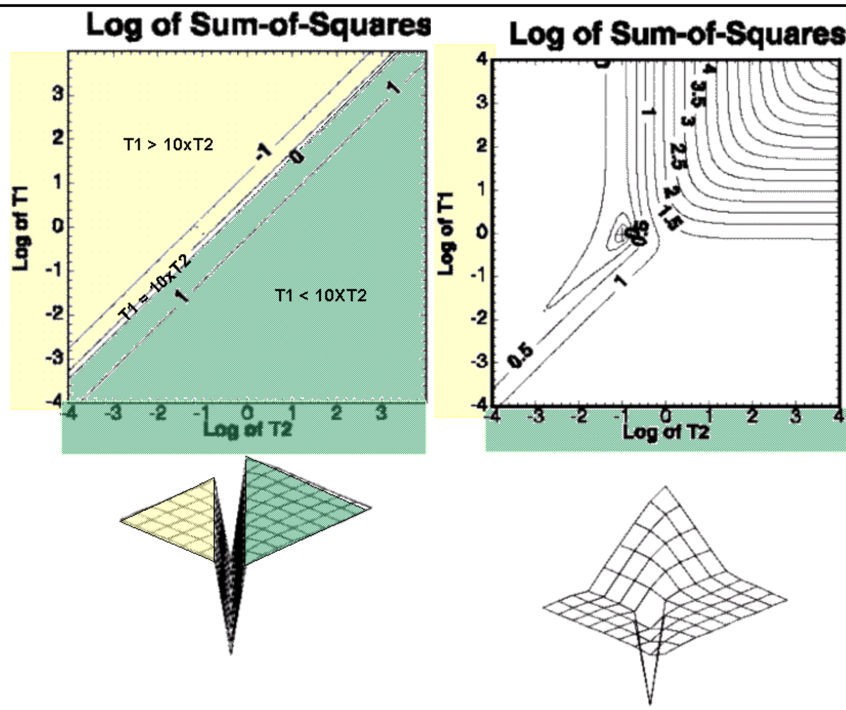
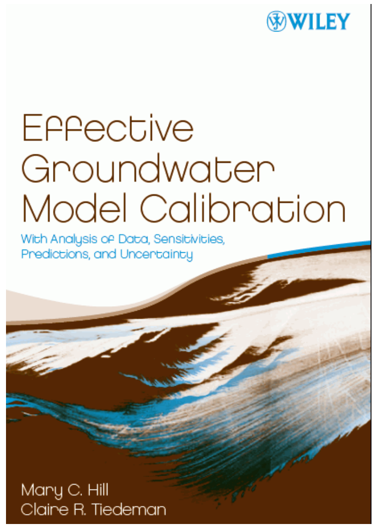


Table 1: Guidelines for effective model calibration (from Hill and Tiedeman, 2007 ; modified from Hill, 1998).

Model Development
1. Apply the principle of parsimony (start simple; build complexity slowly)
2. Use a broad range of information to constrain the problem
3. Maintain a well-posed, comprehensive regression problem
4. Include many types of observations in the regression
5. Use prior information carefully
6. Assign weights that reflect errors
7. Encourage convergence by improving the model and evaluating the observations
8. Consider alternative models
Test the Model
9. Evaluate model fit
10. Evaluate optimized parameters
Potential New Data
11. Identify new data to improve model parameter estimates and distribution
12. Identify new data to improve predictions
Prediction Accuracy and Uncertainty
13. Evaluate prediction uncertainty and accuracy using deterministic methods
14. Quantify prediction uncertainty using statistical methods

Learn much more about calibrating models via Hill and Tiedeman



DUE NEXT WEEK

SUBMIT THE OBSERVATION FILES ALONG WITH YOUR WORKING MODFLOW FILES FROM THIS WEEK

Include comments on the quality of fit

This submission will be considered as part of your assignment 6 grade