

GMS TUTORIALS

MODFLOW – Stochastic Modeling – Parameter Randomization

There is always a significant amount of uncertainty associated with a groundwater model. This uncertainty can be associated with the conceptual model or with the data and parameters associated with the various components of the model. Some model parameters such as hydraulic conductivity and recharge are particularly prone to uncertainty. Calibrating a model to a rich set of observation data (monitoring wells, stream flows, etc.) may reduce this uncertainty somewhat. However, calibration data are often scarce and even well-calibrated models have a high level of uncertainty.

One method for dealing with uncertainty is to utilize a stochastic modeling approach. With a non-stochastic approach, a single model is developed that represents the best estimate of the real system being simulated. This model is used to make predictions. With a stochastic approach, a set of models is constructed where each model in the set is thought to be equally probable. Each model is then used to make the prediction or simulate a given scenario and the results are used to estimate a probability or risk that a certain outcome will occur. While this approach still relies on underlying model assumptions to generate initial parameter estimates, it more honestly reflects the uncertainty associated with modeling.

GMS includes two basic methods for generating stochastic simulations: parameter randomization and indicator simulations. With the parameter randomization method, selected model parameters are randomized using either a Random Sampling or Latin Hypercube approach. Each combination of input parameters defines a model instance. With the indicator simulation approach, multiple equally probable realizations of the aquifer heterogeneity are generated and each realization is used to define a model instance.

This tutorial illustrates how to develop a stochastic simulation using parameter randomization. Parameter randomization utilizes many of the tools described in the *MODFLOW – Automated Parameter Estimation* tutorial. Therefore, we recommend that you complete that tutorial prior to beginning this one.

1.1 Outline

This is what you will do:

1. Open a MODFLOW model and solution.
2. Set up parameters in coverages and map to MODFLOW.
3. Run MODFLOW in stochastic mode.
4. Open a MT3DMS project.
5. Run MT3DMS in Stochastic mode.
6. Run a risk analysis and a statistical analysis.

1.2 Required Modules/Interfaces

You will need the following components enabled to complete this tutorial:

- Grid
- Map
- MODFLOW
- Stochastic Modeling

You can see if these components are enabled by selecting the *File | Register* command.

2 Description of Problem

The model we will be using in this tutorial is the same model featured in the *MODFLOW - Conceptual Model Approach* tutorial. This is a two layer model representing an aquifer in East Texas. The model is bounded on the bottom and right sides by rivers represented with specified head boundaries and, on the north, by a no-flow boundary corresponding to a bedrock outcropping. The model includes two extraction wells and three drains. The model was developed to analyze the long-term consequences of a proposed landfill.

For this tutorial, we will randomize the recharge, the leakage from the landfill, and the hydraulic conductivity associated with the top layer. After developing multiple MODFLOW simulations using the Latin Hypercube sampling method, we will then simulate the contaminant transport resulting from each flow model using MT3DMS. Finally, we will process the results using the threshold analysis option in the *GMS Risk Analysis Wizard*.

3 Random Sampling vs. Latin Hypercube

GMS provides two methods for performing parameter randomization: Random Sampling and Latin Hypercube. With the Random Sampling method, the user specifies a mean, a standard deviation, a minimum value, and a maximum value for each parameter. In addition, the parameter can be specified as log transformed, which is typically the case for hydraulic conductivity. The user also specifies the number of simulations. For each simulation, a random number is generated for each parameter according to the specified distribution using the mean, standard deviation, maximum and minimum. GMS supports both normal and uniform distributions. The more simulations generated, the greater the confidence that all options have been explored.

The Latin Hypercube method is an attractive alternative to the Monte Carlo method since it allows for a greater degree of confidence with fewer model runs. This can be especially useful for complex models that require large amounts of computational time. As with the Random Sampling method, the user specifies the mean, standard deviation, minimum, and maximum for each parameter. For each parameter, the user also specifies a number of segments. The probability distribution curve for each parameter is then divided up into n segments of equal probability. Figure 3-1 shows a normal distribution for a parameter that has six segments. Note that each segment has an equal area, not an equal distance between values.

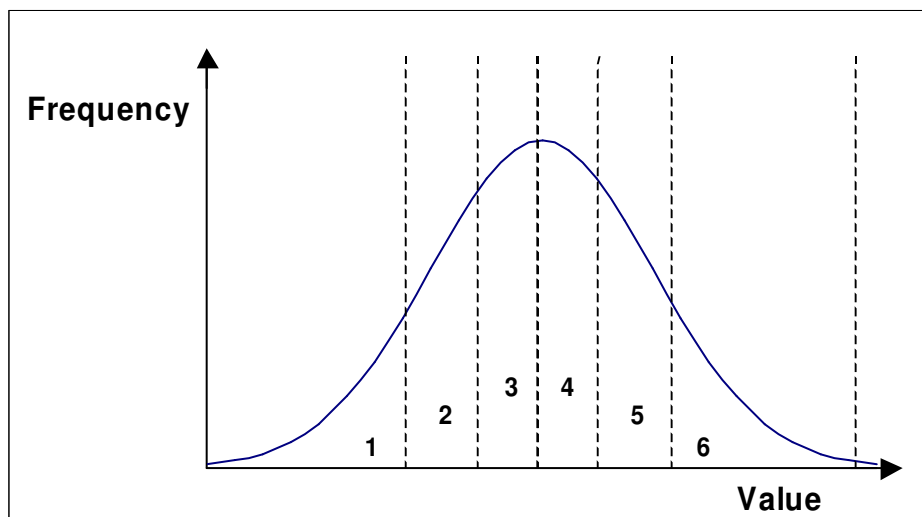


Figure 3-1 Latin Hypercube Segmentation for a Parameter with a Normal Distribution and Six Segments.

The idea behind the Latin Hypercube approach is that the parameter space (all possible combinations of parameter values) should be sampled as completely as possible with a limited number of model runs. Once the segments are defined, each parameter is then randomized until a value is found that lies within each probability segment. The random numbers for each parameter are combined with the random numbers from the other parameters such that all possible combinations of segments are sampled. The total number of model runs is the product of the number of segments for each parameter. For example, if you have three parameters with four segments each and one with five, the

total number of simulations would be $4 \times 4 \times 4 \times 5 = 320$. In GMS, the total area under the probability curve is further limited by the specified maximum and minimum to give the maximum parameter range while still maintaining the best chance for model stability. The greater certainty from a smaller number of runs comes from guaranteeing that a more complete set of parameter combinations is tested. We will be using the Latin Hypercube approach in this tutorial.

4 Getting Started

Let's get started.

1. If necessary, launch GMS. If GMS is already running, select the *File | New* command to ensure that the program settings are restored to their default state.

5 Reading in the Project

First, we will read in a modeling project representing a completed MODFLOW model for the East Texas site:

1. Select the *Open* button .
2. Locate and open the **tutfiles\MODFLOW\stochastic1** directory.
3. Open the file entitled **tex.gpr**.

You should see a MODFLOW model with a solution and a set of GIS coverages. The four coverages include a source/sink, a recharge, and two layer attribute coverages.

6 Model Parameterization

The first steps in setting up a stochastic model are similar to those used to start an inverse model – you need to “parameterize” the input. This involves identifying which parts of the model input we wish to randomize. The parameters with the highest uncertainty are obvious candidates for parameterization. When parameterizing a model, care should be taken to keep the number of selected parameters small. If too many parameters are chosen, unreasonably large numbers of model runs must be completed in order to adequately explore a sufficient combination of parameters. For this model, we will only use three parameters to ensure that the model run times will be fifteen minutes or less (depending on the speed of your computer).

7 Defining the Parameter Zones

The conceptual model approach used in GMS allows us to quickly define our parameter zones because the conceptual model consists of zones of recharge and hydraulic conductivity defined by polygons. We will mark the polygons as parameter zones by

assigning a “key value” to each polygon. The key value should be a value that is not expected to occur normally in the MODFLOW input. A negative value typically works well.

7.1 Setting up the Recharge Zones

First we will set up the recharge zones. The recharge polygons are shown in Figure 7-1. We will define a parameter for both the small polygon defining the landfill and the larger polygon encompassing the rest of the site. The key values associated with the two zones are shown on the polygons in the figure. The -200 zone represents recharge to the aquifer from rainfall. The -100 zone represents leakage from the landfill to the aquifer.

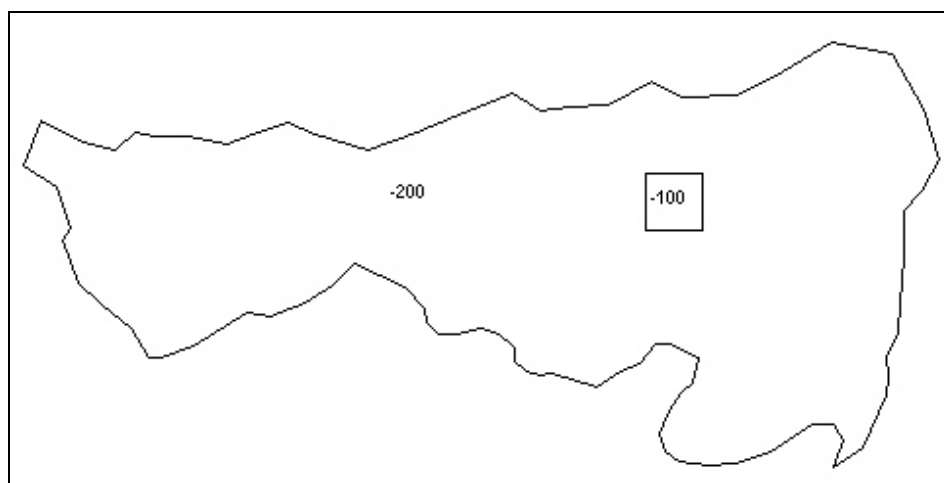





Figure 7-1 Recharge Zones Representing Aquifer Recharge (-200) and Landfill Leakage (-100).

To assign the key values to the polygons:

1. Expand the *East Texas* conceptual model  by clicking on the plus symbol next to it in the *Project Explorer*.
2. Switch to the *Recharge* coverage  by selecting it from the *Project Explorer*.
3. Select the *Select Polygon* tool .
4. Double click on each of the polygons show in Figure 7-1 and assign the appropriate key value to the *Recharge rate* input field.

7.2 Setting up the Hydraulic Conductivity Zone

Next, we will parameterize the hydraulic conductivity for the top layer. For simplicity, we will assume that the hydraulic conductivity is constant for the entire layer.

1. Switch to the *Layer 1* coverage  by selecting it from the *Project Explorer*.

2. Double-click on the polygon and assign a key value of **-300** to the *Horizontal K* input field.


7.3 Mapping the Key Values to the Grid Cells

Once the key values are assigned to the polygons, they must be mapped to the cells in the MODFLOW grid.

1. Select the Feature Objects | Map → MODFLOW command.
2. Select OK at the prompt.

8 Selecting the Stochastic Option

Before we edit the parameter data, we will turn on the *Stochastic* option. This option is located in the *Global Options* dialog.

1. Select the *3D Grid Data* Folder  in the *Project Explorer*.
2. Select the *MODFLOW | Global Options* command.
3. In the *Run options* section of the dialog, select the *Stochastic Simulation* option.
4. Choose *OK* to exit the dialog.

9 Editing the Parameter Data

Two options are available when running a stochastic simulation using parameter zones – Random Sampling and Latin Hypercube. With Random Sampling, the total number of runs is specified directly in the *Parameters* dialog. With the Latin Hypercube method, the total number of MODFLOW runs is calculated from the number of segments specified for each parameter. We are going to use the Latin Hypercube method and we will use three parameters and give each parameter three segments for a total of 27 (3*3*3) MODFLOW runs.

Now we will create a list of parameters and enter a mean and standard deviation for each parameter that defines a normal probability distribution curve for the parameter value. We will also define bounds for the parameters to keep the values within an acceptable range.

1. Select the *MODFLOW | Parameters* command.

This dialog is used to manage the list of parameters and it is explained in detail in the *MODFLOW – Automated Parameter Estimation* tutorial.

2. Select the *Initialize From Model* button.

Note that all three parameters were automatically found. Also note that the parameters have been given a default name. The next step is to enter a starting (mean), minimum, and maximum values for each parameter. You should notice that the usage fields have been set to *stochastic*. GMS will only perturb the parameters that are labeled as *stochastic* regardless of the values in the other fields.

3. Enter the following data into the parameters spreadsheet

Name	Value	Min	Max
HK_300	16	1	100
RCH_200	0.0288	0.01	0.05
RCH_100	0.0003	0.0001	0.0005

4. Toggle on the check box in the *Stochastic Randomize* column for each of the parameters.
5. For the hydraulic conductivity parameter (**HK_300**), ensure that the *Log Xform* option is on.

The spreadsheet in the lower half of the parameters dialog displays the parameter values for each stochastic run.

6. Enter the following values for the standard deviation and number of segments in the lower spreadsheet:

Name	Std. Deviation	Num Segments
HK_300	0.5	3
RCH_200	0.01	3
RCH_100	0.0001	3

7. Click on the *Repopulate Runs* button. Scroll down in the spread sheet and notice that the parameter values for each run are displayed.

Notice that there is a drop down selection box just above the lower spreadsheet that is labeled *Stochastic option*. There are three different options listed: Random Sampling, Latin Hypercube, and User-Defined. By selecting the User-Defined option you can set the number of runs and enter in the parameter values for each run. This is a convenient way to set up your own model sensitivity runs or your own stochastic simulation.

8. Select *OK* to exit the dialog.

10 Saving the Project and Running MODFLOW

We are now ready to save the project and run MODFLOW in Stochastic mode.

1. Select the *File | Save As* command.
2. Save the project with the name **mfsto.gpr**.
3. Select the *MODFLOW | Run MODFLOW* command.

MODFLOW is now running in Stochastic mode. The spreadsheet at the top shows the set of parameter values associated with each model run. The first column indicates whether or not each model run has converged; some combinations of parameter values result in unstable models. The lower window shows the text output from MODFLOW.

Note: If you have room on your screen, you may wish to resize the output window by dragging the handle in the lower right corner of the window.

11 Reading in and Viewing the MODFLOW Solutions

Once all the MODFLOW runs are completed, you can read in the solutions.

1. Make sure the *Read solution on exit* toggle is checked and select the *Close* button.

When the MODFLOW dialog closes, another dialog appears that lists each MODFLOW solution and whether it converged. You now have the option of choosing which of the solutions you want to read in, but all converged model solutions are checked by default.

2. Select *OK* to exit the dialog.


You should see a new folder named **mfsto (MODFLOW)(STO)** appear in the *Project Explorer*. You can expand this folder to see the results of the stochastic simulation. Clicking on individual solutions within the folder updates the contours on the MODFLOW grid to reflect the solution results. You may wish to click on several of the solutions to view the variety in the range of answers. Some of the solutions may indicate flooded cells (blue triangle symbols). This means that the computed water table elevation is above the ground surface. Note that once an individual solution folder is selected in the *Project Explorer*, you can use the up and down arrows on your keyboard to cycle through the solutions.

12 MT3DMS

Now we will run an MT3DMS model using the results from each MODFLOW solution generated by the stochastic flow simulation. The MT3DMS model consists of one species, leachate from the landfill, entering the model from mass flux in the recharge and traveling toward the river and nearby well. We will assign a concentration to the recharge coming from the landfill polygon. Because the recharge for the landfill area in the MODFLOW model was varied as a parameter, the mass flux (recharge rate X concentration) of contaminant leaving the landfill will vary for each model run. We will then read in the computed transport solutions and perform a probabilistic threshold concentration analysis.

13 Reading in the MT3DMS Project

First, we will read in the MT3DMS project:

1. Select the *File | Save* command. This will save the project in its current state, including the stochastic solutions.
2. Select the *Open* button .
3. Locate and open the **tutfiles\MODFLOW\stochastic1\leachate_MT3DMS** directory.
4. Change the *Files of type* to **Model Super Files**.
5. Open the file entitled **leachate.mts**.

14 MT3DMS Model

The MT3DMS model we have imported is a simple transport model with a concentration assigned to the recharge at the cells in the location of the landfill. The initial concentration for the entire model is set to zero. All of the other sources/sinks have a zero concentration. The simulation is set to run for 3000 days with output every 300 days. The leachate plume should migrate to the south and be captured by the well or the river or both.

15 Selecting the MODFLOW Stochastic Simulation

The MT3DMS model is already set up for us, but we still need to specify that we want to run MT3DMS in batch mode with our stochastic MODFLOW solution.

1. Select the *MT3D | Run Options* command.
2. Select the *Batch run with selected MODFLOW solution set* option. We currently have only one MODFLOW solution set in GMS so *mfsto (MODFLOW)(STO)* should appear in the combo box below the selected radio button.
3. Select the *OK* button to exit the dialog.

16 Saving and Running MT3DMS in Stochastic Mode

We are now ready to save the project and run MT3DMS using the MODFLOW stochastic results.

1. Select the *File | Save As* command.

Note that we can specify different names for the MODFLOW and MT3DMS model files using the controls at the bottom of the dialog. However, in this case, we will use the default names.

1. Make sure that the Save simulations in project folder using project filename option is on.
2. Select the Save button.
3. Select Yes to overwrite the existing project.
4. Select the MT3D | Run MT3D command.

MT3DMS is now running in stochastic mode. The spreadsheet at the top shows each model, and as each model run is completed, the spreadsheet will update the status for the run as converged or not converged. It will take several minutes for all 27 model runs to be completed.

17 Reading in and Viewing the MT3DMS Solutions

Once all the MT3DMS runs are completed, you can read in the solutions as follows:

1. Make sure the *Read solution on exit* toggle is checked and select the *Close* button.


As was the case with the MODFLOW solution, when the MT3DMS dialog closes, another dialog appears that lists each MT3DMS solution and whether or not it converged.

2. Select *OK* to exit the dialog.

You should see a new folder name *mfsto (MT3DMS)(STO)* appear below the MODFLOW solutions in the *Project Explorer* window. Once again, you can expand this folder and click on individual solutions to see the results of the stochastic transport simulation.

18 Threshold Analysis

Now that we have imported the MT3DMS solution set, we can perform a threshold analysis on the computed leachate concentrations. A threshold analysis can be used to generate a data set representing the probability that a set of conditions is satisfied. For our case, we want to generate a plot indicating the probability that the leachate concentration exceeds 10.0. To do this, we will set up a rule to mark leachate concentrations above 10.0. GMS then searches through all the selected solutions and for each cell it counts how many times the leachate concentration is exceeded. This number is then divided by the total number of solutions (27 in our case). These results are then contoured as a probability threshold dataset.

1. Right-click on *mfsto (MT3DMS)(STO)* folder  in the *Project Explorer* window. This is the folder that contains the individual MT3DMS solutions.
2. Select the *Risk Analysis* command from the pop-up menu. This brings up the *Risk Analysis Wizard*.

3. Verify that **MT3DMS** is selected in the list box, and select the *Next* button.


This next step in the *Risk Analysis Wizard* allows you to set up rules. Because we only have one contaminant, leachate, we only need one rule.

4. In the *Value* field for the first rule, enter a concentration of **10**.
5. Make sure that the *Operator* is **>**.
6. Change the *Analysis title* to **above10**.
7. Choose the *Finish* button.

When the *Risk Analysis Wizard* finishes, a new data set, *above10*, will be added to the *mfsto (MT3DMS)(STO)* folder. This data set contains the probability that the leachate concentration will be above 10.0 for each of the time steps of our MT3DMS model. The best way to view this dataset is to turn on color filled contours.

1. Select the *Data | Contour Options* command.
2. Change the *Contour Method* to **Color Fill**.
3. Select the *Color Ramp* button to bring up the *Color Ramp Options* dialog.
4. Make sure the *Legend* box is checked.
5. Select *OK* in both dialogs to exit.

You should now see a probability plume extending from the landfill. You can cycle through the time steps in the *Time Step* window below the *Project Explorer* to see the probabilities at the different times.

6. Expand the *mfsto (MT3DMS)(STO)* item  in the *Project Explorer* and select the *above 10* data set. You can now select a different time step in the *Time Step* window.

19 Statistical Analysis

The mean, min, max, and standard deviation data sets can be generated from the output of a stochastic run.

1. Right-click on the *mfsto (MT3DMS)(STO)* folder  and select the *Statistical analysis* menu item.

It may take a short amount of time for the new datasets to be generated.

2. Once GMS is finished creating the data sets, select the *mean_leachate* data set in the *Project Explorer* and change the time step to 3000.0.

3. Now select the *max_leachate* data set. Notice the broad range of concentrations that are computed from the stochastic runs.

20 Conclusion

This concludes this tutorial. Here are the things that you should have learned in this tutorial:

- GMS supports two types of stochastic approaches: parameter randomization and indicator simulations
- With parameter randomization, you can do random sampling, latin hypercube sampling, or user defined sampling.
- You can run an MT3DMS model against all the stochastic MODFLOW solutions, but currently GMS does not support running MT3DMS stochastically itself (randomizing the MT3DMS parameters).
- You can run the *Risk Analysis Wizard* on any folder of solutions. One of the options in the *Risk Analysis Wizard* is a threshold analysis.
- You can generate the mean, min, max, and standard deviation of a stochastic simulation by selecting the *Statistical analysis* menu option when right-clicking on a folder of solutions.