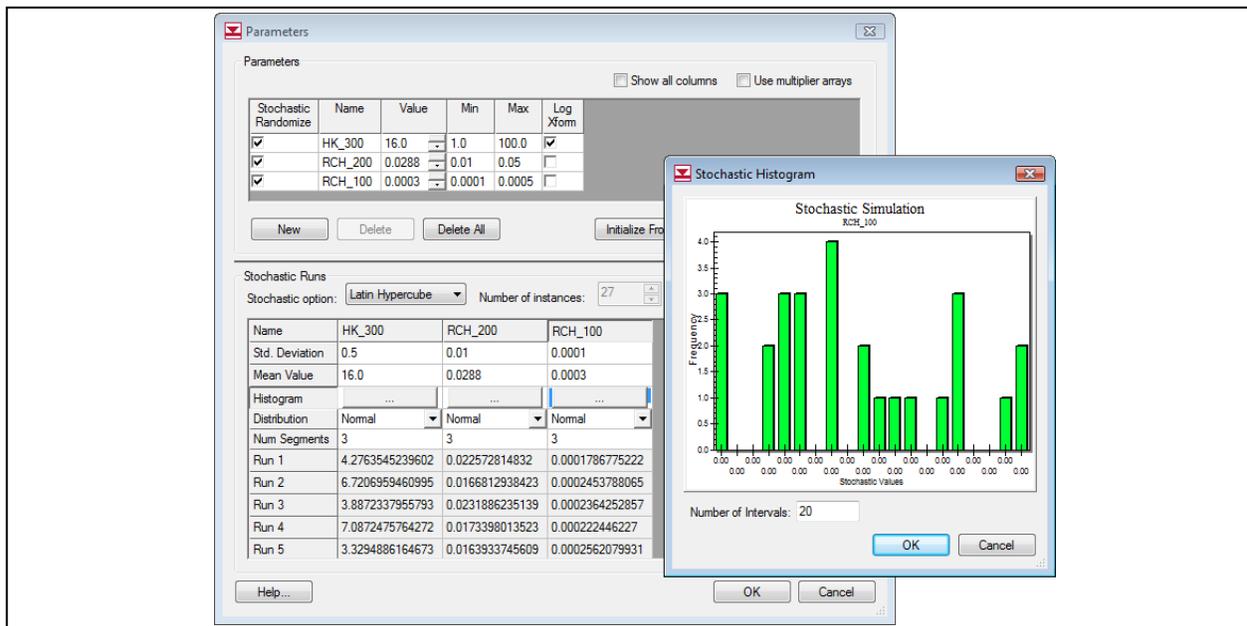


GMS 10.5 Tutorial

MODFLOW – Stochastic Modeling, Parameter Randomization

Run MODFLOW in Stochastic (Monte Carlo) Mode by Randomly Varying Parameters



Objectives

This tutorial shows how to develop a stochastic (Monte Carlo) simulation using parameter randomization. It explains Random Sampling versus Latin Hypercube. It will go over defining parameters and parameter zones, as well as, running MT3DMS on all the MODFLOW solutions and performing a threshold analysis.

Prerequisite Tutorials

- MODFLOW – Automated Parameter Estimation

Required Components

- Grid Module
- Map Module
- MODFLOW
- Stochastic Modeling

Time

- 40–60 minutes



1	Introduction	2
1.1	Random Sampling vs. Latin Hypercube	3
2	Getting Started	4
2.1	Importing the Project	5
3	Model Parameterization	5
3.1	Defining the Parameter Zones	6
3.2	Mapping the Key Values to the Grid Cells	7
3.3	Selecting the Stochastic Option	7
3.4	Editing the Parameter Data	7
3.5	Saving the Project and Running MODFLOW	9
3.6	Importing and Viewing the MODFLOW Solutions	9
4	MT3DMS	10
4.1	Importing the MT3DMS Project	10
4.2	MT3DMS Model	10
4.3	Selecting the MODFLOW Stochastic Simulation	11
4.4	Saving and Running MT3DMS in Stochastic Mode	11
4.5	Importing and Viewing the MT3DMS Solutions	11
4.6	Threshold Analysis	12
4.7	Statistical Analysis	13
5	Conclusion	14

1 Introduction

A significant amount of uncertainty is always 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 somewhat reduce this uncertainty. 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 a 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, it is recommended that that tutorial be completed prior to beginning this one.

The model in this tutorial is a two-layer model representing an aquifer in East Texas. This is the same model featured in the “MODFLOW – Conceptual Model Approach 1” tutorial. The model is bounded on the bottom and right sides by rivers represented with specified head boundaries and bounded 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.

This tutorial randomizes 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, the contaminant transport resulting from each flow model will be simulated using MT3DMS. Finally, the results will be processed using the threshold analysis option in the *GMS Risk Analysis Wizard*.

This will be accomplished by opening a MODFLOW model and solution, setting up parameters in coverages, mapping them to MODFLOW, and then running MODFLOW in stochastic mode. An MT3DMS project will then be opened, MT3DMS will be run in stochastic mode, and both a risk and a statistical analysis will be run.

1.1 Random Sampling vs. Latin Hypercube

GMS provides two methods for performing parameter randomization: Random Sampling and Latin Hypercube.

With the Random Sampling method, a mean, a standard deviation, a minimum value, and a maximum value are specified for each parameter. In addition, the parameter can be specified as log transformed, which is typically the case for hydraulic conductivity. The number of simulations is also specified. 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 Random Sampling 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 mean, standard deviation, minimum, maximum, and the number of segments for each parameter, are specified for each parameter. The probability distribution curve for each parameter is then divided up into n segments of equal probability. Figure 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.

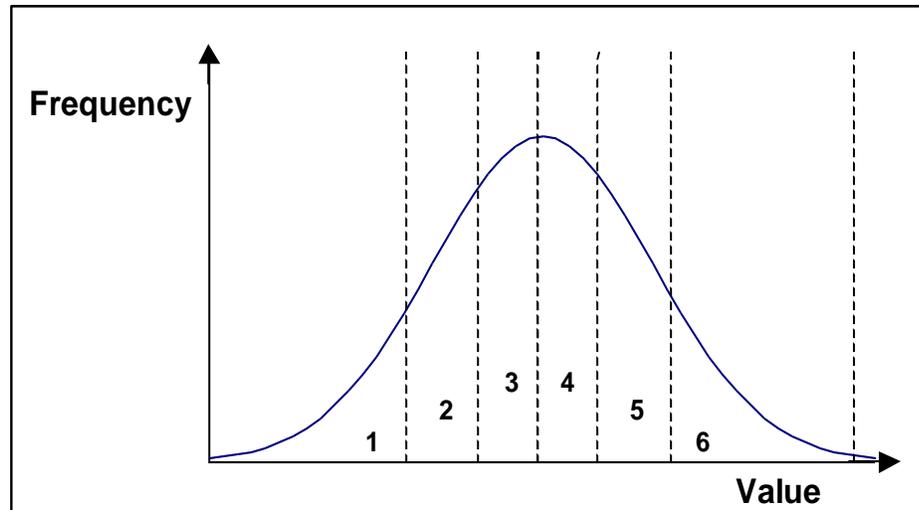


Figure 1 Latin hypercube segmentation for a parameter with a normal distribution and six segments

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 there are three parameters with four segments each and one with five, the total number of simulations would be $4 * 4 * 4 * 5 = 320$. In GMS, the total area under the probability curve is further limited by the specified maximum and minimum parameter range. This gives 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.

This tutorial uses the Latin Hypercube approach.

2 Getting Started

Do the following to get started:

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

2.1 Importing the Project

The first step is to import a modeling project representing a completed MODFLOW model for the East Texas site.

1. Click **Open**  to bring up the *Open* dialog.
2. Select “Project Files (*.gpr)” from the *Files of type* drop-down.
3. Browse to the `sto_param\sto_param\` directory and select “tex.gpr”.
4. Click **Open** to import the project and exit the *Open* dialog.
5. In the Project Explorer, fully expand “ Map Data” to make the coverages visible.

A MODFLOW model with a solution and a set of GIS coverages should now be visible (Figure 2). The four coverages include a source/sink, a recharge, and two layer attribute coverages.

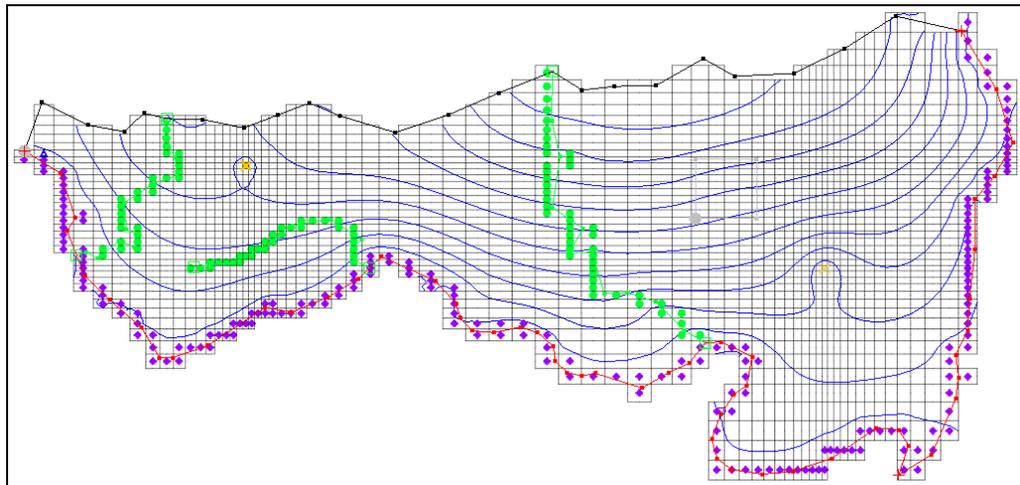


Figure 2 The initial MODFLOW model

3 Model Parameterization

The first steps in setting up a stochastic model are similar to those used to start an inverse model because the input needs to be “parameterized”. This involves identifying which parts of the model input need to be randomized. 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, only three parameters will be used to ensure that the model run times will be as short as possible.

3.1 Defining the Parameter Zones

The conceptual model approach used in GMS allows for quickly defining the parameter zones because the conceptual model consists of zones of recharge and hydraulic conductivity defined by polygons. The polygons will be marked 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.

Setting up the Recharge Zones

First, set up the recharge zones. The recharge polygons are shown in Figure 3. 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.

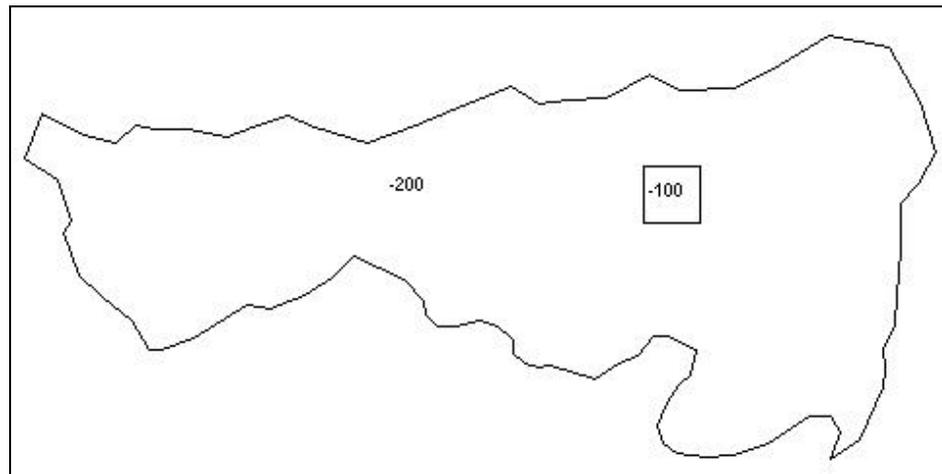


Figure 3 Recharge zones representing aquifer recharge (-200) and landfill leakage (-100)

To assign the key values to the polygons:

1. Select the “ Recharge” coverage to make it active.
2. Using the **Select Polygon**  tool, double-click on the landfill polygon (see Figure 3) to bring up the *Attribute Table* dialog.
3. Enter “-100” in the *Recharge rate (m/d)* column in the spreadsheet.
4. Click **OK** to close the *Attribute Table* dialog.
5. Double-click on the aquifer recharge polygon (anywhere outside the landfill polygon) to bring up the *Attribute Table* dialog.
6. Repeat steps 3–4, entering “-200” for the *Recharge rate (m/d)*.

Setting up the Hydraulic Conductivity Zone

The next step is to parameterize the hydraulic conductivity for the top layer. For simplicity, assume that the hydraulic conductivity is constant for the entire layer.

1. Select “ Layer 1” in the Project Explorer to make it active.
2. Using the **Select Polygon**  tool, double-click in the aquifer recharge polygon to bring up the *Attribute Table* dialog.
3. Enter “-300” in the *Horizontal K (m/d)* column in the spreadsheet.
4. Click **OK** to close the *Attribute Table* dialog.

3.2 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 *Feature Objects* / **Map** → **MODFLOW** to bring up the *Map* → *Model* dialog.
2. Click **OK** to accept the defaults and close the *Map* → *Model* dialog.

3.3 Selecting the Stochastic Option

Before editing the parameter data, it is necessary to turn on the stochastic option. This option is located in the *MODFLOW Global/Basic Package* dialog.

1. Select *MODFLOW* | **Global Options...** to open the *MODFLOW Global/Basic Package* dialog.
2. In the *Run options* section, select *Stochastic*.
3. Click **OK** to exit the *MODFLOW Global/Basic Package* dialog.

3.4 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. This tutorial will use the Latin Hypercube method and three parameters, giving each parameter three segments for a total of 27 (3*3*3) MODFLOW runs.

The next step is to 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. It is also necessary to define bounds for the parameters to keep the values within an acceptable range.

1. Select *MODFLOW* | **Parameters...** to open the *Parameters* dialog.

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

2. Click **Initialize From Model** to automatically set all three parameters.

Notice that the parameters have been given default names. The next step is to enter a starting (mean), minimum, and maximum value for each parameter.

3. Check the box in the *Stochastic Randomize* column for each of the parameters.
4. Enter the following data into the parameters spreadsheet:

Name	Value	Min	Max
HK_300	16.0	1.0	100.0
RCH_200	0.0288	0.01	0.05
RCH_100	0.0003	0.0001	0.0005

5. Check the box in *Log Xform* column on the hydraulic conductivity parameter row (“HK_300”)
6. Uncheck the box in the *Log Xform* column for each of the two recharge parameters (“RCH_200” and “RCH_100”).
7. Click **Yes** if a dialog appears asking to set the standard deviation between a max and min value.

The spreadsheet in the *Stochastic Runs* section displays the parameter values for each stochastic run.

8. In the *Stochastic Runs* section, select “Latin Hypercube” from the *Stochastic option* drop-down.

In the *Stochastic option* drop-down, three different options are available: “Random Sampling”, “Latin Hypercube”, and “User-Defined”. The “User-Defined” option lets the user set the number of runs and enter in the parameter values for each run. This is a convenient way to set up customized model sensitivity runs or stochastic simulations.

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

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

10. Click **Repopulate Runs**, then scroll down in the spreadsheet and notice that the parameter values for each run are displayed.
11. Click **OK** to exit the *Parameters* dialog.

3.5 Saving the Project and Running MODFLOW

Now to save the project and run MODFLOW in Stochastic mode:

1. Select *File* | **Save As...** to bring up the *Save As* dialog.
2. Select “Project Files (*.gpr)” from the *Save as type* drop-down.
3. Enter “mfsto.gpr” as the *File name*.
4. Click **Save** to save the project under the new name and close the *Save As* dialog.
5. Select *MODFLOW* | **Run MODFLOW** to bring up the *Stochastic MODFLOW* model wrapper dialog.

MODFLOW is now running in Stochastic mode. The spreadsheet at the top shows the set of parameter values associated with each model run. The lower window shows the text output from MODFLOW.

3.6 Importing and Viewing the MODFLOW Solutions

Once all the MODFLOW runs are completed, the next step is to import the solutions.

1. Turn on *Read solution on exit* and *Turn on contours (if not on already)*.
2. Click **Close** to close the *Stochastic MODFLOW* dialog, and open the *Reading Stochastic Solutions* dialog.

The first column indicates whether or not each model run has converged. Some combinations of parameter values result in unstable models.

Now, choose which of the solutions to import; all converged model solutions are checked by default.

3. Click **OK** to import the solutions and exit the *Reading Stochastic Solutions* dialog.
4. Fully expand the “ 3D Grid Data” folder in the Project Explorer.
5. Select the “ mfsto001 (MODFLOW)” solution to make it active.

This folder contains the results of the first stochastic simulation (Figure 4). Feel free to review all 27 of the solutions to see 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, it is possible to use the up and down arrows on the keyboard to cycle through the solutions.

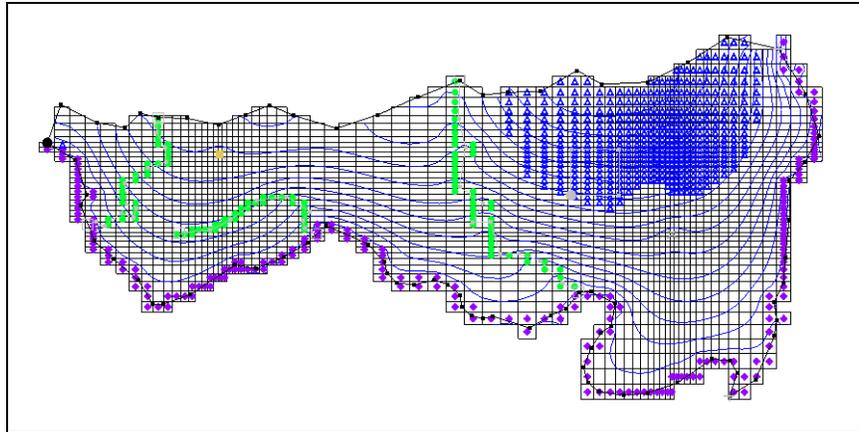


Figure 4 Results of the "mfsto001" stochastic MODFLOW simulation

4 MT3DMS

The next step is to 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. A concentration will be assigned 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. Then the computed transport solutions will be imported and probabilistic threshold concentration analysis will be performed.

4.1 Importing the MT3DMS Project

First, read in the MT3DMS project:

1. **Save**  the project in its current state, including the stochastic solutions.
2. Click **Open**  to bring up the *Open* dialog.
3. Select "Model Super Files" from the *Files of type* drop-down.
4. Browse to the `sto_param\sto_param\leachate_MT3DMS` directory and select "leachate.mts".
5. Click **Open** to import the file into the current project and exit the *Open* dialog.

4.2 MT3DMS Model

The MT3DMS model that was 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, the river, or both.

4.3 Selecting the MODFLOW Stochastic Simulation

The MT3DMS model is already set up, but it is still necessary to specify that MT3DMS should run in batch mode with the stochastic MODFLOW solution.

1. Select *MT3DMS* | **Run Options...** to open the *Run Options* dialog.
2. Select *Batch run with selected MODFLOW solution set*.

There is currently only one MODFLOW solution set in GMS so “mfsto (MODFLOW)(STO)” should appear in the combo box below the selected radio button.

3. Click **OK** to exit the *Run Options* dialog.

4.4 Saving and Running MT3DMS in Stochastic Mode

Now to save the project and run MT3DMS using the MODFLOW stochastic results:

1. **Save**  the project.
2. Select *MT3DMS* / **Run MT3DMS...** to bring up the *MT3DMS Using Stochastic MODFLOW* dialog.

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, depending on the speed of the computer being used.

4.5 Importing and Viewing the MT3DMS Solutions

Once all the MT3DMS runs are completed, import the solutions as follows:

1. Turn on *Read solution on exit* and *Turn on contours (if not on already)*.
2. Click **Close** to exit the *MT3DMS Using Stochastic MODFLOW* dialog and bring up the *Reading Stochastic Solutions* dialog.

This dialog appears after the MT3DMS run dialog closes, listing each MT3DMS solution and whether it converged (just as with the MODFLOW stochastic run).

3. Click **OK** to import the solutions and exit the *Reading Stochastic Solutions* dialog.
4. Select the “ mfsto001 (MT3DMS)” solution under the “ mfsto (MT3DMS)(STO)” folder.

Contours should be visible around the landfill (Figure 5). Feel free to explore the remaining solutions under the “mfsto (MT3DMS)(STO)” folder to see the results of the stochastic transport simulation.

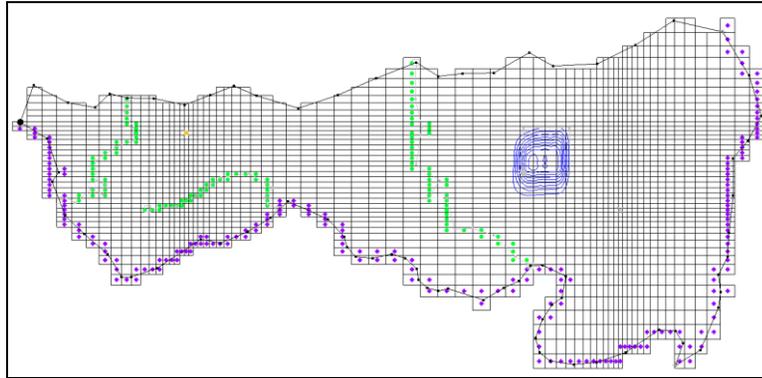


Figure 5 Contours visible around the landfill

4.6 Threshold Analysis

Now that the MT3DMS solution set has been imported, a threshold analysis can be completed on the computed leachate concentrations. A threshold analysis can be used to generate a dataset representing the probability that a set of conditions is satisfied. For this case, it is necessary to generate a plot indicating the probability that the leachate concentration exceeds “10.0”.

To do this, a rule will be set up to mark leachate concentrations above “10.0”. GMS then searches through all the selected solutions and counts how many times the leachate concentration is exceeded for each cell. This number is then divided by the total number of solutions (27, in this case). These results are then contoured as a probability threshold dataset.

1. Right-click on “mfsto (MT3DMS)(STO)” in the Project Explorer (the folder containing the individual MT3DMS stochastic solutions) and select **Risk Analysis...** to bring up the *Risk Analysis Wizard* dialog.
2. Select “MT3DMS” in the box at the top.
3. Click **Next** to go to the *Threshold Analysis* dialog.

This dialog allows the threshold analysis rules to be defined. Because there is only one contaminant—leachate—only one rule is needed.

4. On row 1, select “>” from the drop-down in the *Operator* column.
5. In the *Value* column, enter “10.0” as the concentration.
6. Below the spreadsheet, enter “above10” as the *Analysis title*.
7. Click **Finish** to close the *Threshold Analysis* dialog.

- When the analysis finishes, select the new “ above10” dataset in “ mfsto (MT3DMS)(STO)” in the Project Explorer.

This dataset 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.

- Click **Contour Options**  to open the *Dataset Contour Options – 3D Grid – above10* dialog.
- Turn on *Legend* in the lower left corner of the dialog.
- Select “Color fill” from the top drop-down in the *Contour method* section.
- Click **Color Ramp...** to open the *Color Options* dialog.
- In the *Palette method* section, select *Hue ramp*.
- Click **Reverse** button to make red the *Max* and blue the *Min*.
- Click **OK** to exit the *Color Options* dialog.
- Click **OK** to exit the *Dataset Contour Options – 3D Grid – above10* dialog.

A probability plume should now extend from the landfill (Figure 6). Feel free to cycle through the time steps to see the probabilities at the different times.

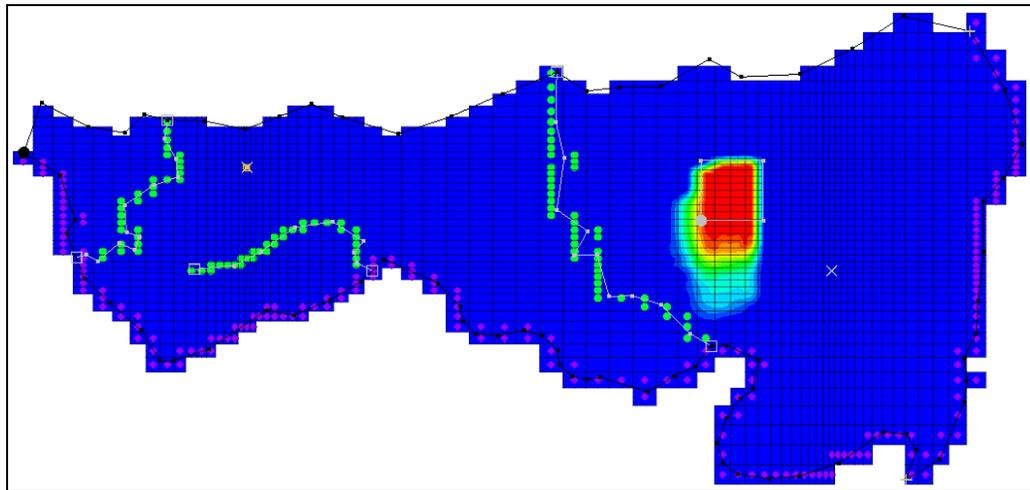


Figure 6 Leachate probability plume at time step 1 (300.0)

4.7 Statistical Analysis

The mean, min, max, and standard deviation datasets can be generated from the output of a stochastic run. Create these statistical datasets for the MT3DMS stochastic run using the steps below.

- Right-click on “ mfsto (MT3DMS)(STO)” and select **Statistical Analysis...** to generate several new statistical datasets from the MT3DMS stochastic simulation.

2. Once GMS is finished creating the datasets, select the “ mean_leachate” dataset in the Project Explorer.
3. Select time step 10 (it has a value of “3000.0”).
4. Select the “ max_leachate” dataset and select time step 10.

Notice the broad range of concentrations that are computed from the stochastic runs. Feel free to compare various time steps in the various datasets.

5 Conclusion

This concludes the “MODFLOW – Stochastic Modeling, Parameter Randomization” tutorial. The following key concepts were discussed and demonstrated in this tutorial:

- GMS supports two types of stochastic approaches: parameter randomization and indicator simulations.
- With parameter randomization, it is possible to do Random Sampling, Latin Hypercube sampling, or user-defined sampling.
- It is possible to run an MT3DMS model using a set of stochastic MODFLOW solutions, but currently GMS does not support running MT3DMS stochastically itself (randomizing the MT3DMS parameters).
- It is possible to run the *Risk Analysis Wizard* on any folder of solutions. One of the options in the *Risk Analysis Wizard* is a threshold analysis.
- It is possible to 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.