

1 Downloading example data and configuration files

Under Handouts on the course web page, there are three different sets of files for use with GISolve. We will have different groups of students run GISolve on different files. Please listen for which set your group should download.

You will need to use Firefox (not Internet Explorer) for running GISolve during this lab, so you might as well use Firefox for these beginning steps, too.

2 Checking the resource reservation status

The following web site shows which TeraGrid resources are available to GISolve users at what times.

<https://www.cigi.uiuc.edu/doku.php/projects/gisolve/tg-resources>

Note which sites are available and how many CPUs at each. With multiple groups about to run jobs, we want to avoid starting jobs that require the same CPUs as this slows everything down tremendously.

3 Logging into GISolve

Go to www.gisolve.org.

Log in using your email address and password. Select the Bayesian Geostatistical Modeling tab. Examine the screen, but don't enter anything yet.

We will refer to the plot of speed ups (slide 37 on the TeraGrid talk handout) to determine how many CPUs should be used for each of the three dataset sizes. We will then decide how to allocate the available resources among the groups.

- number of CPUs is total number to be divided among all the chains you are running at the site
- make number of CPUs *per chain* a perfect square to use PLAPACK efficiently
 - how big a perfect square determined by size of dataset (see graph of speedups in next slide)
- run more than one chain on a site if enough CPUs are available
 - helps in assessing convergence

- generates more samples per unit time if CPUs are available
- samples from different chains are independent

Please fill in below what is agreed on for your group.

Resource _____ No of CPUs _____ No of chains _____

Resource _____ No of CPUs _____ No of chains _____

4 Uploading data files and configuration files

Edit your configuration file to make sure that it has the right number of lines at the end for the number of chains that you are going to run. If you are using more than one resource and want a different number of chains on different resources, you need to prepare more than one configuration file.

Upload your data file and configuration file, using the "Browse" capability provided.

5 Running the job

- specify maximum wall clock time
 - must be long enough for the number of requested iterations to finish
 - must not run past the end of the reserved time on resource
- submit job
- click "Visualize output" periodically to view plots of accumulating samples
- download zip files of plots and numeric output