REFLECT DELIVERABLE D4.3B

UNCERTAINTY QUANTIFICATION GITHUB EXAMPLE GUIDE

Summary:

This document contains a brief guide to using the uncertainty quantification example provided through https://github.com/poortjp/REFLECT_D4-3_example

Authors:

Jonah Poort, TNO, Research Scientist

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement nº 850626.

This document reflects only the author's view and the European Commission is not responsible for any use that may be made of the information it contains.

1 INTRODUCTION

This example has been prepared as part of the dissemination of task 4.2 of work package 4 of the REFLECT project. The goal of this task was to investigate the impact of uncertainties in geothermal brine composition (arising for instance from measurement errors) on the amount of mineral scaling within (parts of) the geothermal system. This was done through the creation of a MATLAB script that automatically creates input files for the geochemical speciation software PHREEQC[1] containing different brine compositions.

This example focuses on uncertainty quantification (UQ) of barite for an example brine composition. This analysis is done by specifying which brine components in the PHREEQC model have to be varied, what their nominal concentrations are, and what the maximum percentage deviation from this nominal value is. Once defined, the user can specify a number of samples to create, and the pressure and temperature at which the scaling potential should be calculated.

When the script is executed, it generates the specified amount of brine samples by selecting a random concentration for each component from a uniform distribution within the bounds of the maximum and minimum deviation from its nominal value. Next, for each brine sample, the script reads the default PHREEQC input file, overwrites the concentration, pressure, and temperature parameters with the sample values, saves the file, and executes it in the command line. After the PHREEQC file has been executed, the script reads the output file generated by PHREEQC and saves the results of the current sample to a larger file that stores the results for all generated samples together. In addition to the random samples generated, the script also calculates the scaling potential for the nominal brine composition. This nominal result is always stored as the last row in the result file.

Once the scaling amount for all samples have been calculated, a different script can be used to generate a histogram that shows the distribution of the barite scaling amount including the mean, standard deviation, and nominal value.

For more information on the details of this workflow and extensions to it not included in this example, please refer to the publication[2] on the website of the REFLECT website: https://www.reflect-h2020.eu/.

[1] Parkhurst, D. & Appelo, C., 2013. Description of input and examples for PHREEQC version 3 - A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. *U.S. Geological Survey Techniques and Methods.*

[2] Poort, J., de Zwart, H., Wasch, L., Shoeibi Omrani, P.: *The H2020 REFLECT project: Deliverable 4.3 Impact of geochemical uncertainties on fluid production and scaling prediction*, GFZ German Research Centre for Geosciences, DOI: 10.48440/gfz.4.8.2022.001

2 FOLDER STRUCTURE

The main folder of this example contains a number of individual files, these include:

- "example_guide.pdf": this document.
- "main_UQ.m": MATLAB script in which the uncertainty quantification settings can be defined, and the PHREEQC input files are created and executed. Outputs a file containing calculated barite scaling amounts for different geothermal brine compositions.
- "process_results_UQ.m": MATLAB script that processes the result file generated through the "main_UQ.m" script and creates a plot of the barite scaling amount disctribution.
- "phreeqc.log": file generated by PHREEQC after executing an input file. Can be ignored.

In addition, the main folder also includes multiple subfolders, including:

- "dbase": subfolder containing the thermodynamic database used by PHREEQC to perform the speciation calculations.
    - Contains the file "PITZER.DAT", which is the database used for the UQ analysis within the REFLECT project.
- "inputs": subfolder containing the PHREEQC input files.
    - Contains an additional subfolder called "source", which contains a pre-prepared PHREEQC example input file. When the UQ script is executed, this file is read, its contents changed, and the updated contents are saved in the "inputs" folder.
    - Once the UQ script is executed, this folder will also contain a ".txt" and ".phr.out" file with the outputs of the PHREEQC calculations. During execution of UQ script, the ".txt" file is read to get the outputs of interest for each brine composition that is included in the UQ analysis.
- "outputs": subfolder in which the output files of the UQ analysis are saved. If the "main_UQ.m" script has not been executed yet, the only file in this folder will be an example csv.
- "phreeqc": subfolder containing "phreeqc.exe", the executable which is called in the "main_UQ.m" script in order to perform the PHREEQC calculations.
- "plots": folder in which plots of the results are saved. If the "process_results_UQ.m" script has not been executed yet, the only file in the folder will be an example plot.
- "utils": subfolder containing a number of MATLAB functions that are used by the "main_UQ.m" script.

3 EXECUTING THE MAIN SCRIPT

To run the UQ analysis, the "main_UQ.m" script has to be executed. This is best done by opening the file with MATLAB and running it from the editor menu. Before running, there are a number of settings or options that can be defined by the user, these are divided into three sections:

- The "Set-up" section, in which the user can define general settings defining filenames, and save options. Here, the user can change the following variables:
  - 'projectname': a name that will be used to save the results.
  - 'save_format': the format in which the results are saved. This can be either csv or xlsx. Please note that saving as xlsx is much slower, and can run into permission errors in certain cases.
  - 'PHREEQC_filename': the filename (without extension) of the source PHREEQC input file. Should only be changed if using an input file that is different from the default example provided. The script always assumes that this file is located in the "source" folder within the "inputs" subfolder.
- The "UQ inputs" section, in which the user can specify the details of the elements to include in the UQ analysis, and their concentration variations, etc. Includes:
  - 'elements': a cell array of the elements to be included in the UQ analysis. Elements should be given the same name (abbreviations) as used in the PHREEQC input file.
  - 'nominal_vals': the nominal concentration of each of the elements. Should be given in the same order as the array of elements and using the same units as the PHREEQC input file: mg/L.
  - 'var_percentages': the percentage by which the nominal values is expected to vary. A single number per element should be given which determines both the maximum and minimum percentual deviation from the nominal value. Brine concentrations will be sampled between these maximum and minimum values. Should be given in the same order as the array of elements.
  - 'num_samples': the number of brine samples to generate and calculate barite scaling amounts for. In addition to the number of samples specified, the workflow also calculates the scaling for the nominal values, so in total, the script will execute 'num_samples' + 1 samples.
- The "Pressure and temperature" section in which the pressure and temperature at which the scaling calculations should be done are specified. Includes:
  - 'p': the pressure value in bar.
  - 'T': the temperature value in °C.

Once all inputs are given, the script can be ran. After each sample, the current progress and estimated time remaining will be printed to the command window. The final file will be saved in the "outputs" folder under a name that includes the specified 'projectname', the pressure, temperature and number of samples, and the current timestamp (in seconds).

4 EXECUTING THE DATA PROCESSING SCRIPT

In order to not have to re-run the main sampling script every time the UQ results have to be analyzed or plotted, a separate script was written in which a result file can be loaded and processed. The main output of this script is a histogram plot of the distribution of the calculated barite scaling amounts.

The only user input that has to be specified in this script is the variable 'result_file', which gives the filename (including extension) of the file in which the results to be plotted are saved. The script assumes that this file is located in the "outputs" folder.