

GM Simulation Investigation Tool

Functions:

- Detect anomalies in the IMs produced by a GM simulation with respect a benchmark (either empirical or a previous CS)
- Investigation and Visualization of GM simulations to allow identifying interesting science (and improvements made wrt. earlier versions)

Components:

1. Generation of Comparison IM CSVs (fault based)
 - a. From empirical GMMs
 - b. From previous GM simulation runs
2. Implementation of residual metrics (and other investigation metrics?)
 - a. At individual simulation level
 - b. At fault level
 - c. Per Site
 - i. At Run/CS level
 - ii. Fault Level
3. Scripts for result checking directly after completion (i.e. integrated into workflow as a job?)
 - a. At Rel Level
 - b. At Fault Level
4. Visualization
 - a. Spatial Residuals maps
 - i. At Rel Level
 - ii. At Fault Level
 - b. Reference IM Plots
 - i. At Rel Level
 - ii. At Fault Level
 - c. Summary Spatial Residual Plots (i.e. histogram)
 - i. At Rel level
 - ii. At Fault level
 - d. Single Site
 - i. Residual distribution (i.e. histogram)
 1. At Fault Level
 - ii. Response spectra comparison
 1. At Rel Level
 2. At Fault Level
 - e. Response Spectra Means
 - i. Across all Sites
 1. At Cybershake Level (one line per event)
 - ii. Across all Cybershake Events
 1. At each Site (one line per site)
 - f. Spatial Mean Log Ratio across Events run for each Site
 - i. At Cybershake Level

Structure / Overview

https://github.com/ucgmsim/cybershake_investigation

Base directory (**cybershake_investigation**)

1. **cybershake_investigation**
 - a. **configs** (Holds config files for acceptable station difference values)
 - b. **scripts** (Contains scripts to call functionality)
 - i. **compare_csvs** (Script to compare IMs from empirical models to cybershake results)
 - ii. **gen_empirical** (Script to generate empirical data for a cybershake_root structure, has different args to archived version)
 - iii. **gen_empirical_archive** (Script to generate empirical data for archived versions of cybershake)
 - iv. **plot_avg_event_residual_spectra** (Script to produce a single plot across all events at every site across the response spectrum)
 - v. **plot_avg_site_residual** (Script to produce a single plot across all sites for every event across the response spectrum)
 - vi. **plot_avg_site_residual_spectra** (Script to produce a single plot that shows the log ratio residuals values across the NZ map for every event / site)
 - vii. **plot_hist_residual** (Script that shows the ratio value data in a histogram instead of a map view for each fault and or realisation)
 - viii. **plot_single_site_residual** (Script to plot a histogram of a single site / fault / IM showing each realisation ratio value)
 - ix. **plot_single_site_spectrum** (Script to plot pSA values for averaged simulation / empirical values and shows each realisation in grey lines for a single site / fault)
 - x. **plot_spatial_im** (Script that plots the IM values directly for each IM / fault and or realisation on a map)
 - xi. **plot_spatial_residual** (Script that plots the ratio values between empirical and simulation for each fault and or realisation)
 - c. **empirical** (Contains the functions to calculate empirical values for archived or cybershake run results)
 - d. **_spectra** (Contains the functions to plot response spectra with all events and sites)
 - e. **hist** (Contains the functions to plot all histogram plots)

- f. **ratios** (Contains the function to compute the ratios between simulation and empirical results and test against acceptable station difference values)
 - g. **single_site** (Contains the functions to plot single site residual histograms and pSA values)
 - h. **spatial** (Contains the function to plot spatially the IM, ratio values and also a histogram of those spatial values)
 - i. **utils** (Contains common functions used by multiple different scripts)
2. **Readme, setup, requirements, jenkinsfile** etc.