
DESIGNSAFE-CI

A NATURAL HAZARDS ENGINEERING COMMUNITY

# Use Case Teams: Jupyter Workflows

Scott Brandenberg

# Jupyter Workflows

- Jupyter provides a flexible framework in which to integrate a range of research activities that enable a research outcome

- Members of this Use Case team will develop domain-specific Jupyter workflows that take advantage of existing (and new) DesignSafe capabilities in terms of data analytics, computational simulation, visualization

# Use Case Teams

| Researcher | Use Case Themes | Researcher | Use Case Themes |
|---|---|---|---|
| Arduino (Washington) | HPC, Workflows | Kumar (UT) | Viz, HPC |
| Brandenberg (UCLA) | Workflows, Data | Lowes (Washington) | Data, HPC |
| Dawson (UT) | HPC, Workflows | Mosqueda (UCSD) | Workflows, Data |
| Haan (Calvin) | Viz, Data | Padgett (Rice) | Viz, Workflows |
| Kareem (Notre Dame) | HPC, Viz | Pinelli (FIT) | Data, Viz |

# Use Case 1: Jupyter Workflows and HPC

**PI: Pedro Arduino**

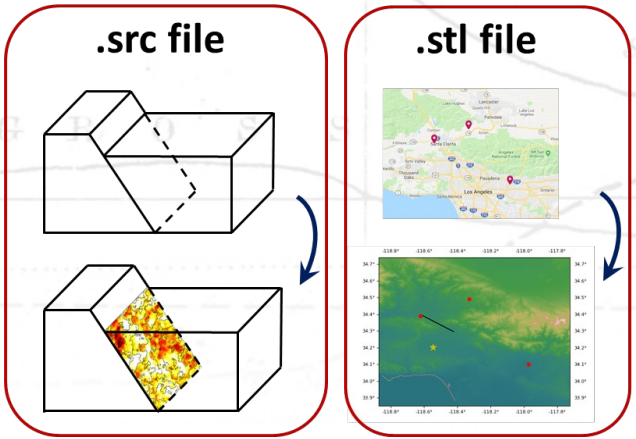
- In this project, we propose to develop cloud computing tools for performing ground motion simulations using existing broad band platforms (BBP) coupled with 1 dimensional soil columns that accept excitations in 3 directions (a.k.a 1D-3C columns). Several aspects will be considered:

1. Existing BBP (Broad Band Platform) will be used to generate rock outcrop motions.
2. 1D-3C columns will be used to resolve the wave equation in one dimension and account for soil nonlinearity. This will include the used of advanced 3D constitutive models.
3. The tool will leverage the Tapis platform provided by DesignSafe-CI to allow running BBP and 1D-3C site response simulations using web-based interfaces.
4. Jupyter notebooks with access to HPC resources will be developed to facilitate running simulations and visualize results directly from the DesignSafe portal.

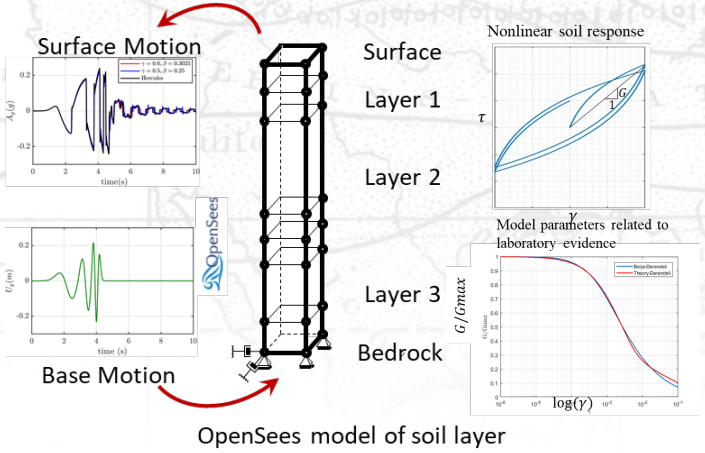
This is a multiyear project. In year 1 we will work on the following tasks:

- Enabling SCEC BBP platform in DesignSafe in TACC,
- Implementing, validating and verifying a 1D-3C wave propagation tool, and
- Implementing a nonlinear elasto-plastic constitutive model in the tool

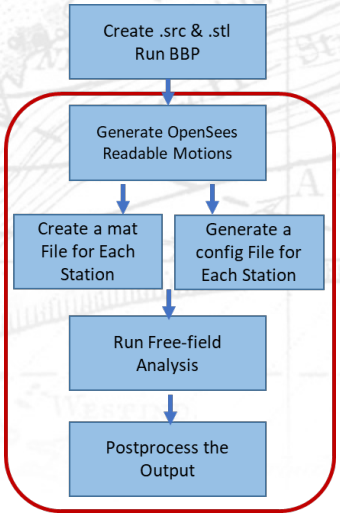
## Broad Band Platform (BBP)




## Nonlinear Soil Column (1D-3C)




## DesignSafe WorkFlow

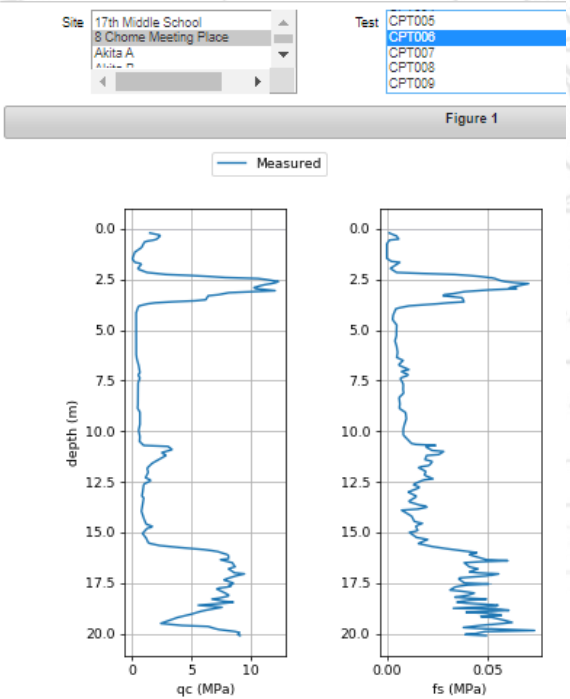


# DesignSafe Enabling Features

- The SCEC Broad Band Platform is already running on DesignSafe HPC resources. These are computationally demanding algorithms that require HPC.
- Ground response analysis platforms such as DeepSoil and OpenSees are already implemented on DesginSafe HPC resources.
- Jupyter notebooks will be used to integrate these tools into a holistic workflow.

# Use Case 2: Jupyter Workflows and Data

**PI: Scott Brandenberg**




- In this project, we propose to use Jupyter notebooks to run SQL queries on datasets in the Next Generation Liquefaction database, and use an unstructured machine learning clustering algorithm to identify soil layers and quantify their properties.
- The defined properties will be useful for systematically including layering effects in liquefaction triggering and consequence models.

# Motivation

- Liquefaction triggering evaluation has largely been handled by small research groups working with small datasets.
- There has been a proliferation of liquefaction observations and we now have more data than we can handle using traditional interpretation methods.
- Soil layering is an important part of liquefaction assessment.
- We need structured liquefaction data, and tools to analyze the data to automatically identify soil layers and quantify their properties.

Sites
Field Performance
Field Investigation
Earthquake

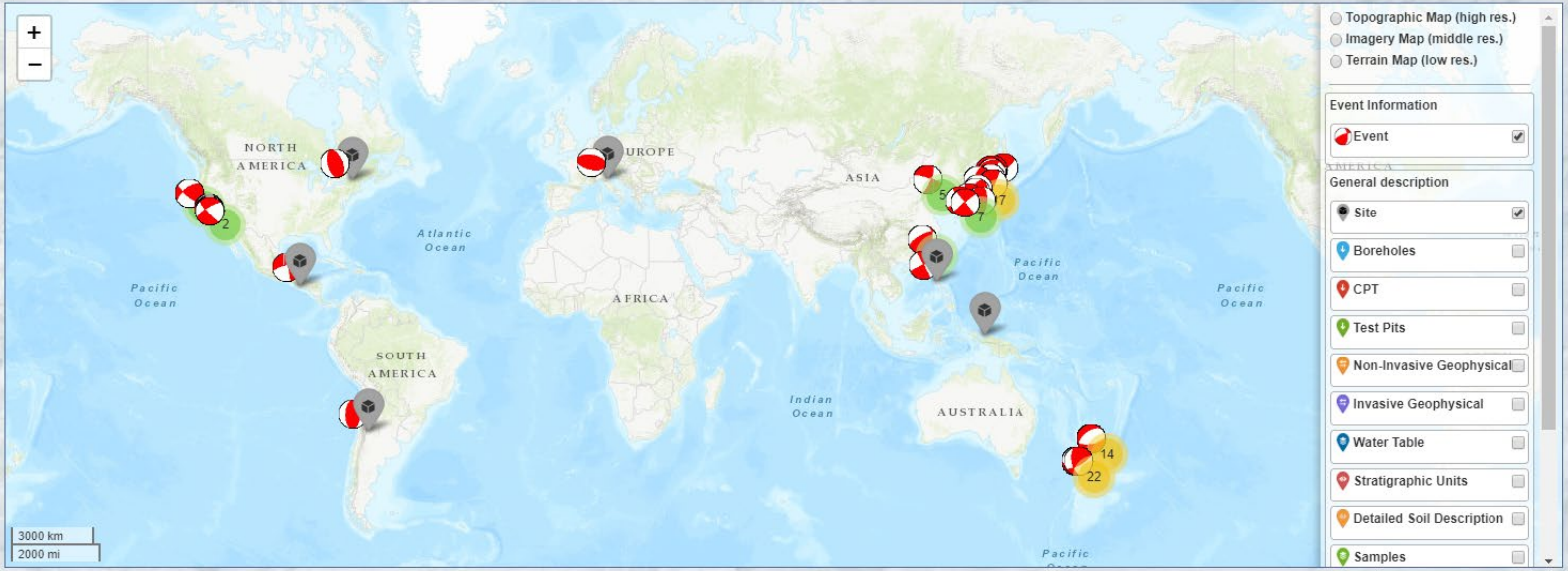
Type event name

Magnitude
min max

M6.9 Kobe, Japan
M6.5 Imperial Valley-06
M7.2 El Mayor-Cucapah
M7.7 Nihonkai-Chubu - near the v
M6.2 Hokkaido
M8.3 Tokachi
M6.9 Obihiro - Hokkaido

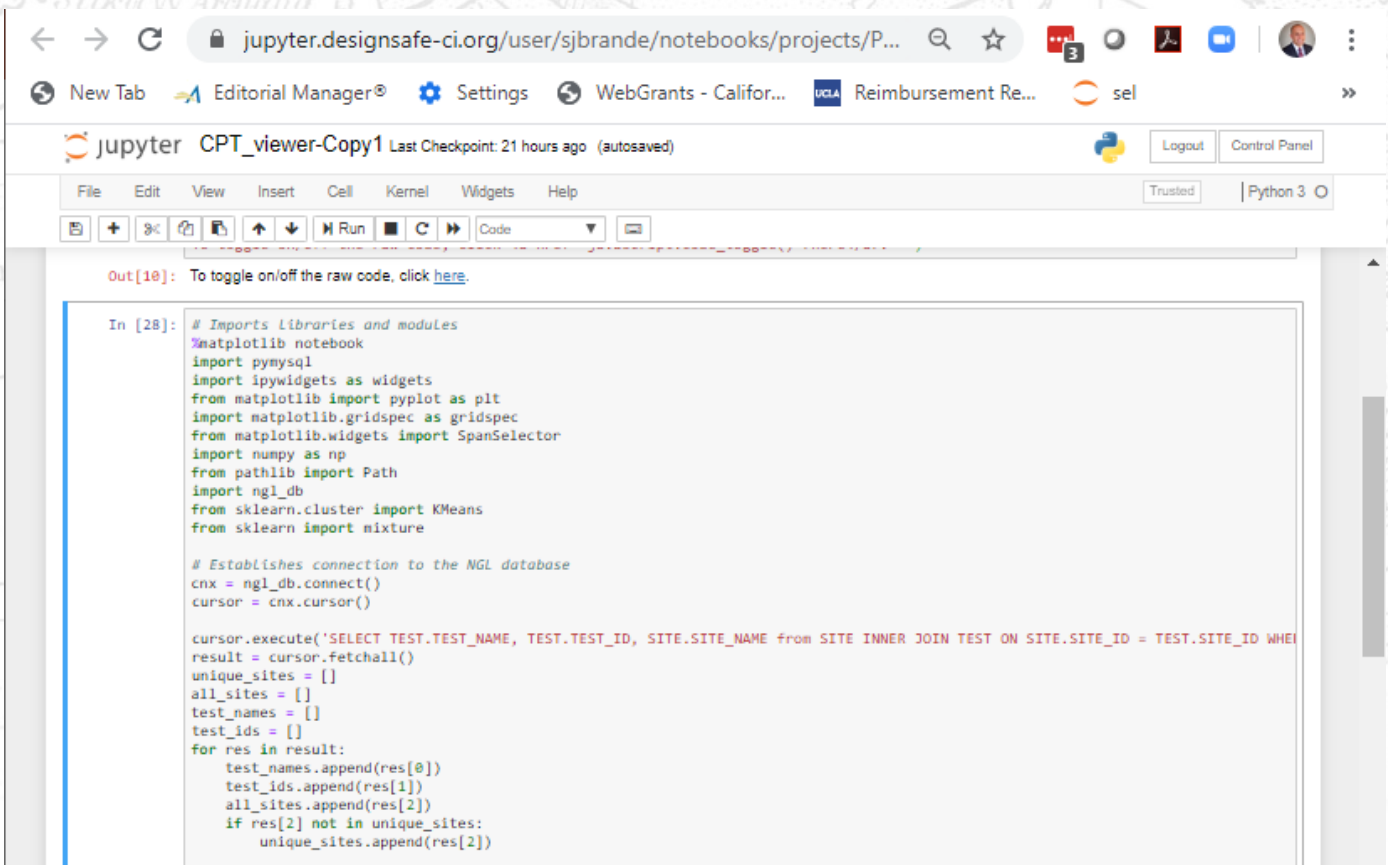
Reset Submit

Statistics

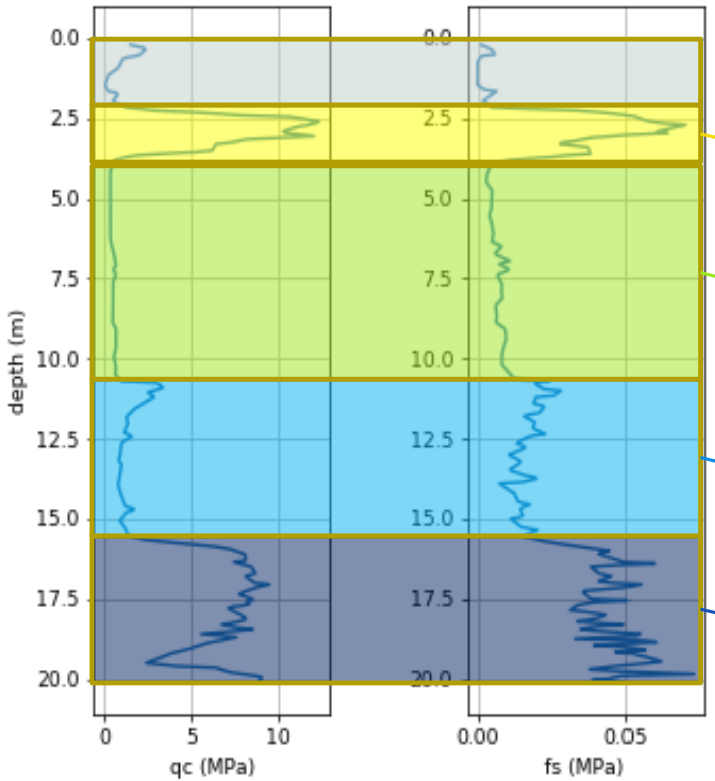



SwRI PEER Caltrans U.S.NRC MPC Mountain-Plains Consortium UDOT Keeping Utah Moving

NSF DESIGNSAFE-CI NHERI: NATURAL HAZARDS ENGINEERING RESEARCH INFRASTRUCTURE
TEXAS The University of Texas at Austin UCLA TACC RICE Florida Tech


jupyter.designsafe-ci.org/user/sjbrande/notebooks/projects/P...
New Tab
Editorial Manager®
Settings
WebGrants - Califor...
Reimbursement Re...
sel
jupyter CPT_viewer-Copy1 Last Checkpoint: 21 hours ago (autosaved)
Logout
Control Panel
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3
Run
Code
Out[10]: To toggle on/off the raw code, click here.
In [28]:
# Imports libraries and modules
%matplotlib notebook
import pymysql
import ipywidgets as widgets
from matplotlib import pyplot as plt
import matplotlib.gridspec as gridspec
from matplotlib.widgets import SpanSelector
import numpy as np
from pathlib import Path
import ngl_db
from sklearn.cluster import KMeans
from sklearn import mixture
# Establishes connection to the NGL database
cnx = ngl_db.connect()
cursor = cnx.cursor()
cursor.execute('SELECT TEST.TEST_NAME, TEST.TEST_ID, SITE.SITE_NAME from SITE INNER JOIN TEST ON SITE.SITE_ID = TEST.SITE_ID WHE
result = cursor.fetchall()
unique_sites = []
all_sites = []
test_names = []
test_ids = []
for res in result:
test_names.append(res[0])
test_ids.append(res[1])
all_sites.append(res[2])
if res[2] not in unique_sites:
unique_sites.append(res[2])

# Layer Observations




Small layer thickness, boundary effects at top and bottom

Thick layer without much variability

Layer with medium amount of variability

Layer with significant variability

# DesignSafe Enabling Features

- The NGL structured dataset is replicated in DesignSafe and can be queried via Jupyter notebooks. We anticipate tens of thousands of soil profiles will be uploaded. Each cone penetration test (CPT) profile may have thousands of data points.
- DesignSafe enables users to interact with data in the cloud without having to download and process all of it.
- The Sci Kit Learn Python library will be used in Jupyter to develop an unsupervised clustering algorithm to automate identification of soil layers from CPT profiles, and quantify soil layer properties.
- Identified soil properties will then be utilized in supervised machine learning algorithms to integrate observations of liquefaction (or lack thereof) with soil layer properties and shaking intensity.