

Open Data Science Toolkits & Repositories for Geoscientists

Mustafa Kaouri

 mkaouri@gmail.com

 [@mustafakaouri](https://www.linkedin.com/in/mustafakaouri)

DISCLAIMER

THE MATERIAL OR INFORMATION CONTAINED ON THIS DOCUMENT IS FOR GENERAL INFORMATION PURPOSES ONLY. YOU SHOULD NOT RELY UPON THE MATERIAL OR INFORMATION AS A BASIS FOR MAKING ANY BUSINESS, LEGAL OR ANY OTHER DECISIONS.

MUSTAFA KAOURI MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND, EXPRESS OR IMPLIED ABOUT THE COMPLETENESS, ACCURACY, RELIABILITY, SUITABILITY OR AVAILABILITY WITH RESPECT TO THE INFORMATION, PRODUCTS, SERVICES OR RELATED GRAPHICS CONTAINED ON THE DOCUMENT FOR ANY PURPOSE. ANY RELIANCE YOU PLACE ON SUCH MATERIAL IS THEREFORE STRICTLY AT YOUR OWN RISK.

MUSTAFA KAOURI WILL NOT BE LIABLE FOR ANY FALSE, INACCURATE, INAPPROPRIATE OR INCOMPLETE INFORMATION PRESENTED ON THIS DOCUMENT.

MUSTAFA KAOURI WHO ASSOCIATED A WORK WITH THIS DEED HAS DEDICATED THE WORK TO THE PUBLIC DOMAIN BY WAIVING ALL OF HIS OR HER RIGHTS TO THE WORK WORLDWIDE UNDER COPYRIGHT LAW, INCLUDING ALL RELATED AND NEIGHBORING RIGHTS, TO THE EXTENT ALLOWED BY LAW.

YOU CAN COPY, MODIFY, DISTRIBUTE AND PERFORM THE WORK, EVEN FOR COMMERCIAL PURPOSES, ALL WITHOUT ASKING PERMISSION. SEE OTHER INFORMATION IN LINK BELOW.

<https://creativecommons.org/publicdomain/zero/1.0/>



Contents

Seismic and Seismology

Ground-penetrating Radar

Well Log

Simulation and Modelling

Reservoir Engineering

Geostatistics

Geospatial

Geochemistry

Geophysics

Structural Geology

Visualization

Data Repositories

Seismic and Seismology



SeisComp

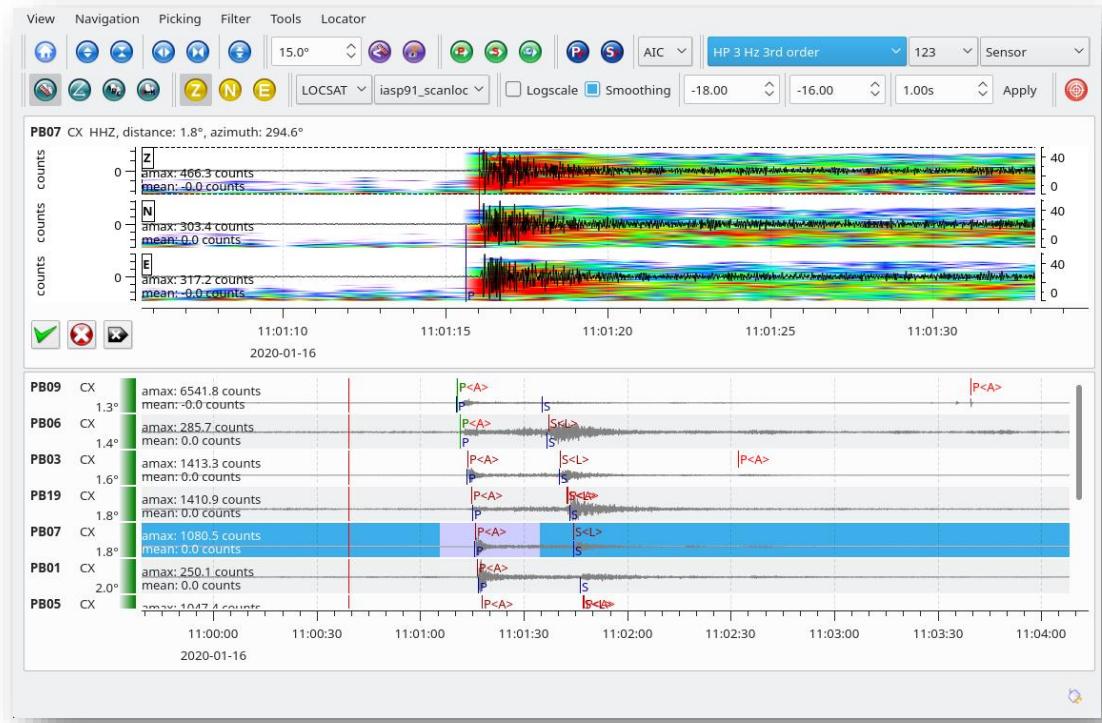
Seismic observatory automation toolkit.
Autodetection, storage, sharing, processing data and more

Language:

Python & C++

Link:

<https://github.com/SeisComP/seiscomp>



RedPy

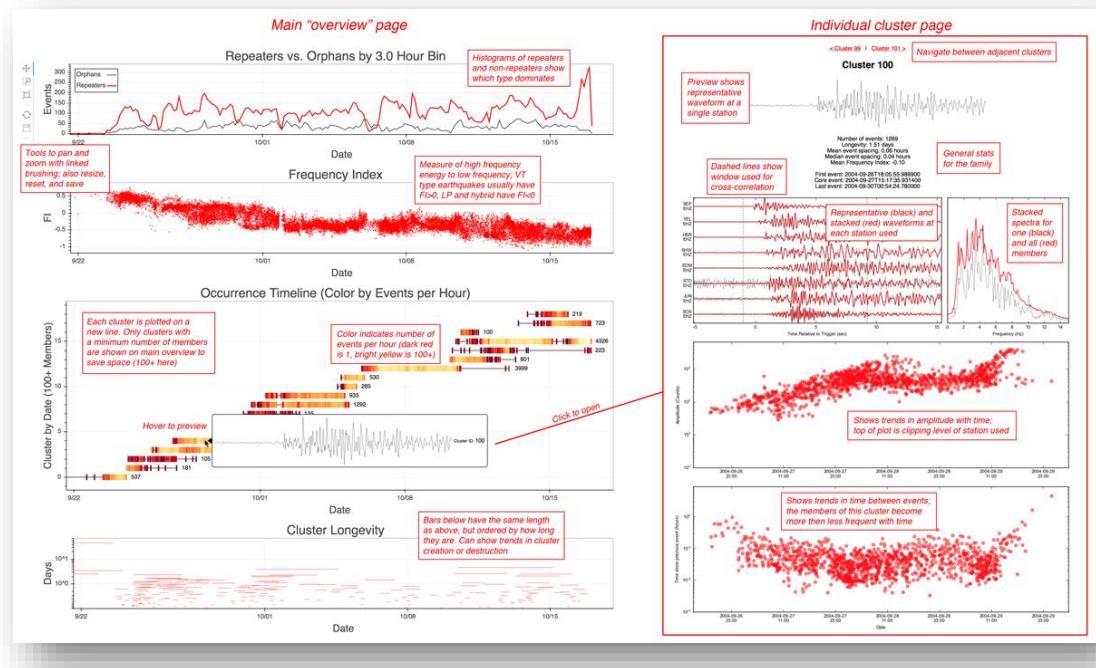
Auto-clustering for seismic events.
Automated detection and analysis of repeating earthquakes in continuous data

Language:

Python

Link:

<https://github.com/ahotovec/REDPy>



SeisUnix

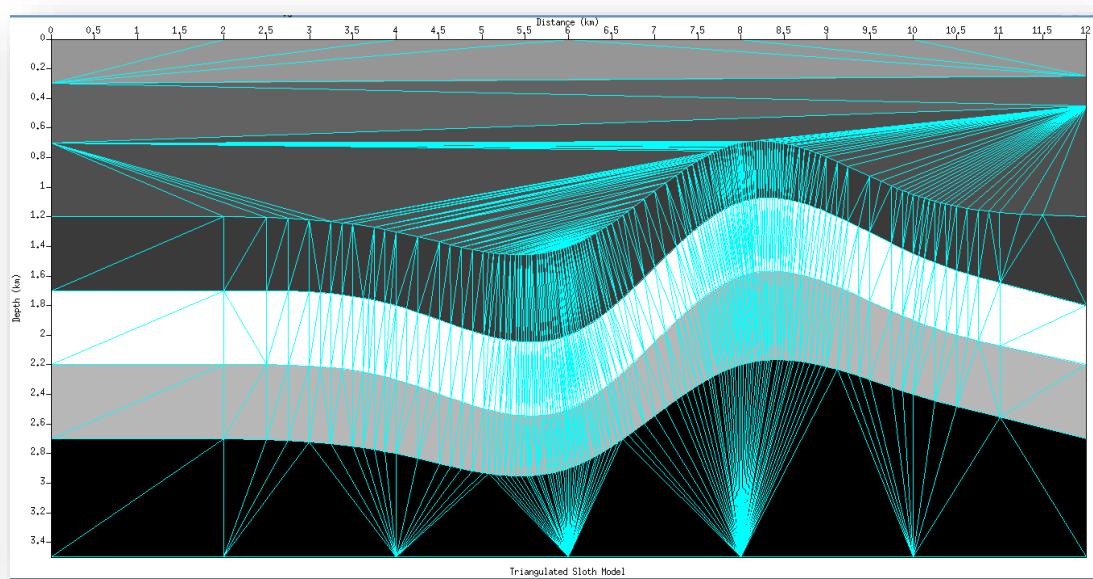
Seismic data processing
suite

Language:

C & other

Link:

[https://github.com/Joh
nWStockwellJr/SeisUnix](https://github.com/JohnWStockwellJr/SeisUnix)



OpenSeaSeis

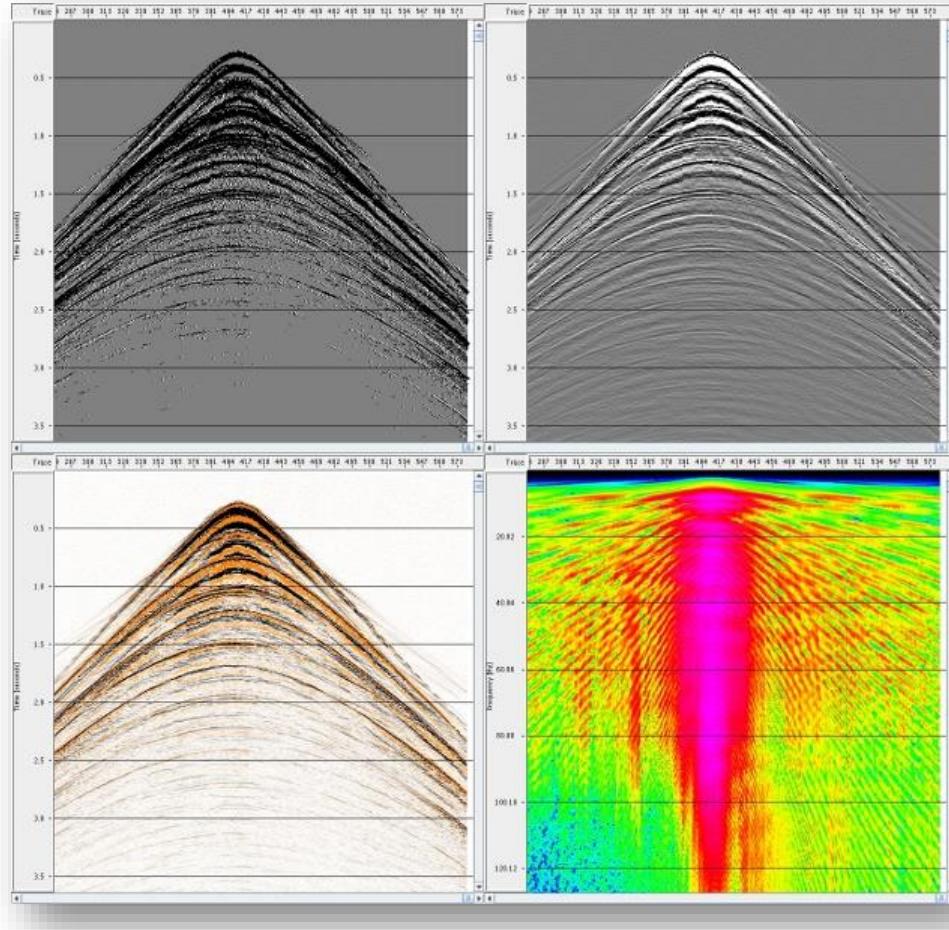
Seismic workflow
generator and seismic
viewer

Language:

C++ & Java

Link:

<https://github.com/JohNWStockwellJr/OpenSeaSeis>



Bruges

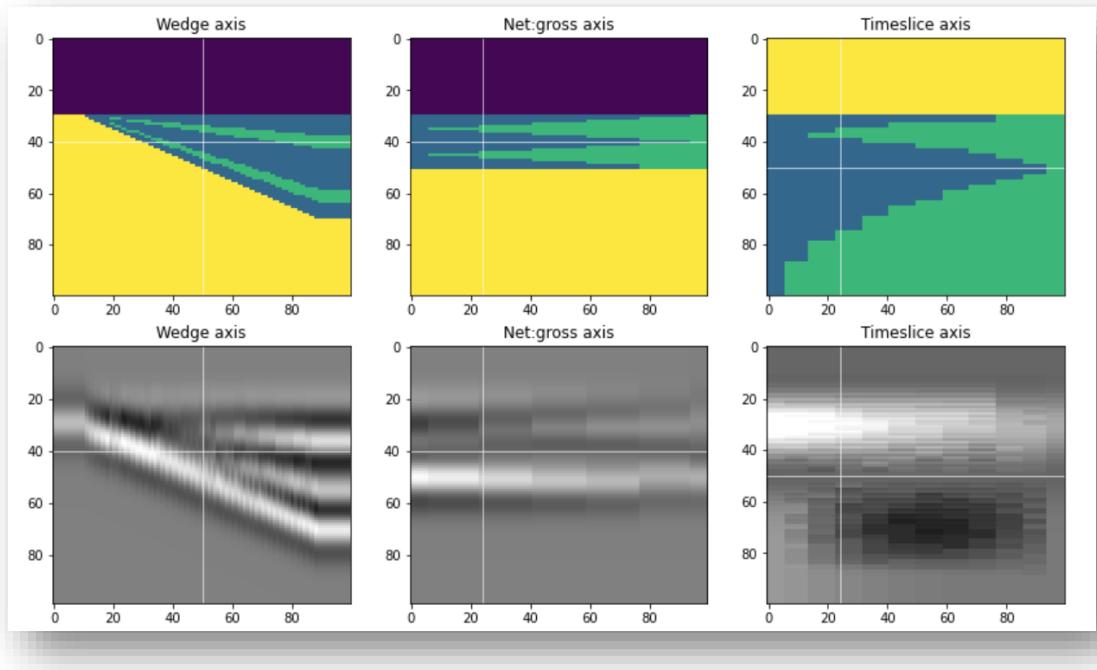
Various geophysical
equations and tools

Language:

Python

Link:

<https://github.com/agilegeoscience/bruges/tree/master/bruges>



ObsPy

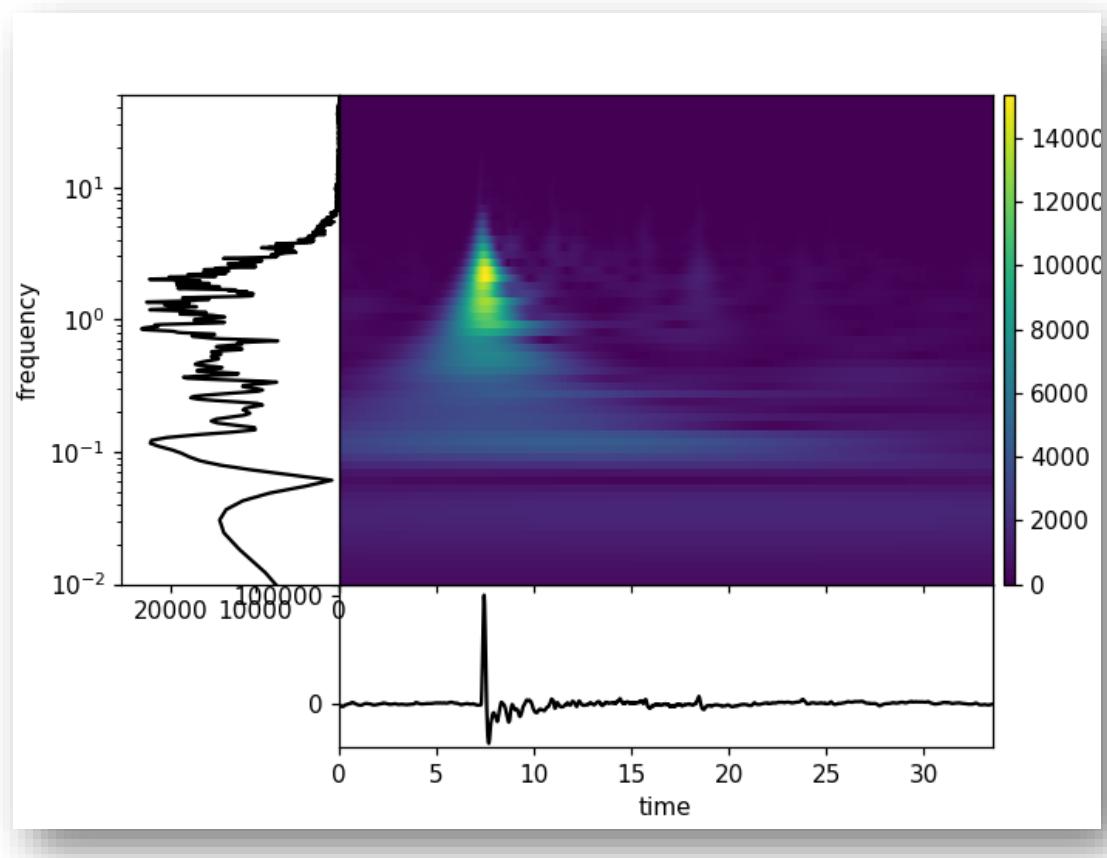
Framework for reading,
writing and processing
seismic and
seismological data

Language:

Python

Link:

<https://github.com/obspy/obspy/wiki>



Madagascar

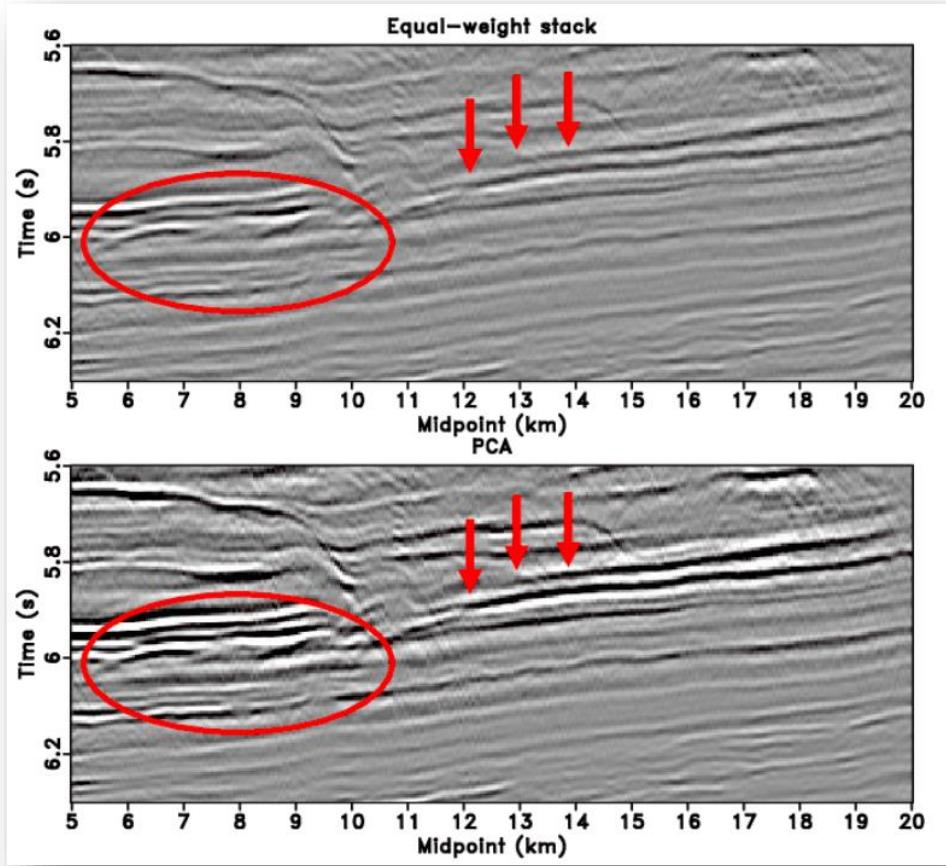
Multidimensional data analysis and reproducible computational experiments

Language:

Python & C

Link:

[https://reproducibility.org
/wiki/Download](https://reproducibility.org/wiki/Download)



Segyio

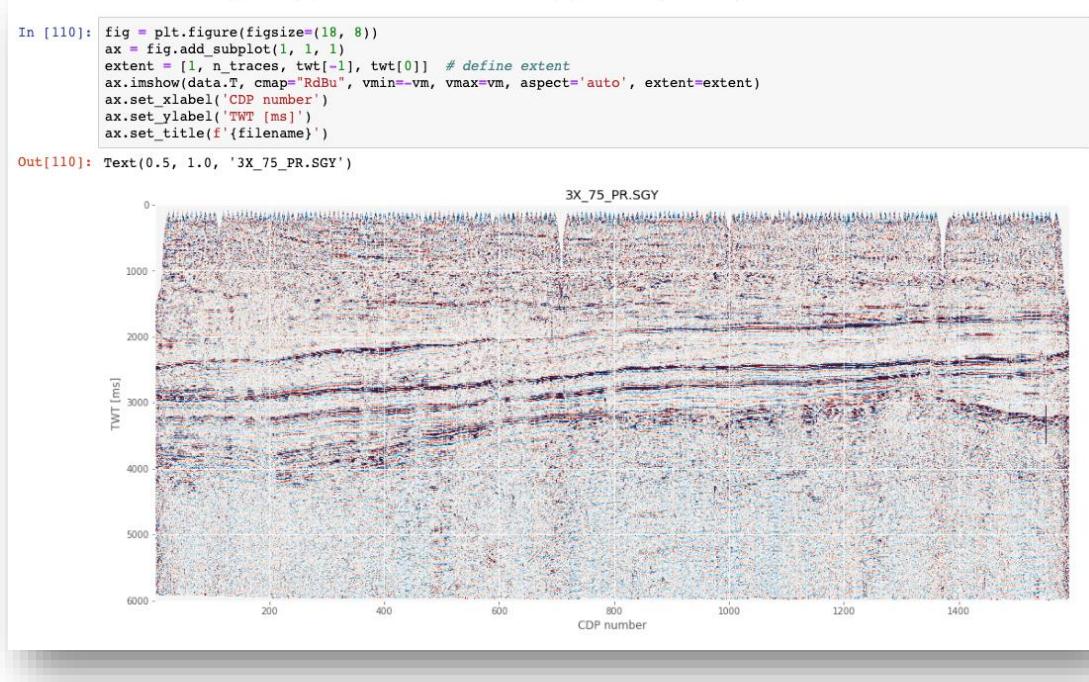
Library for easy interaction with SEG-Y seismic data

Language:

Python, MATLAB, C++, & C

Link:

<https://github.com/equinor/segyio>



Pyrocko

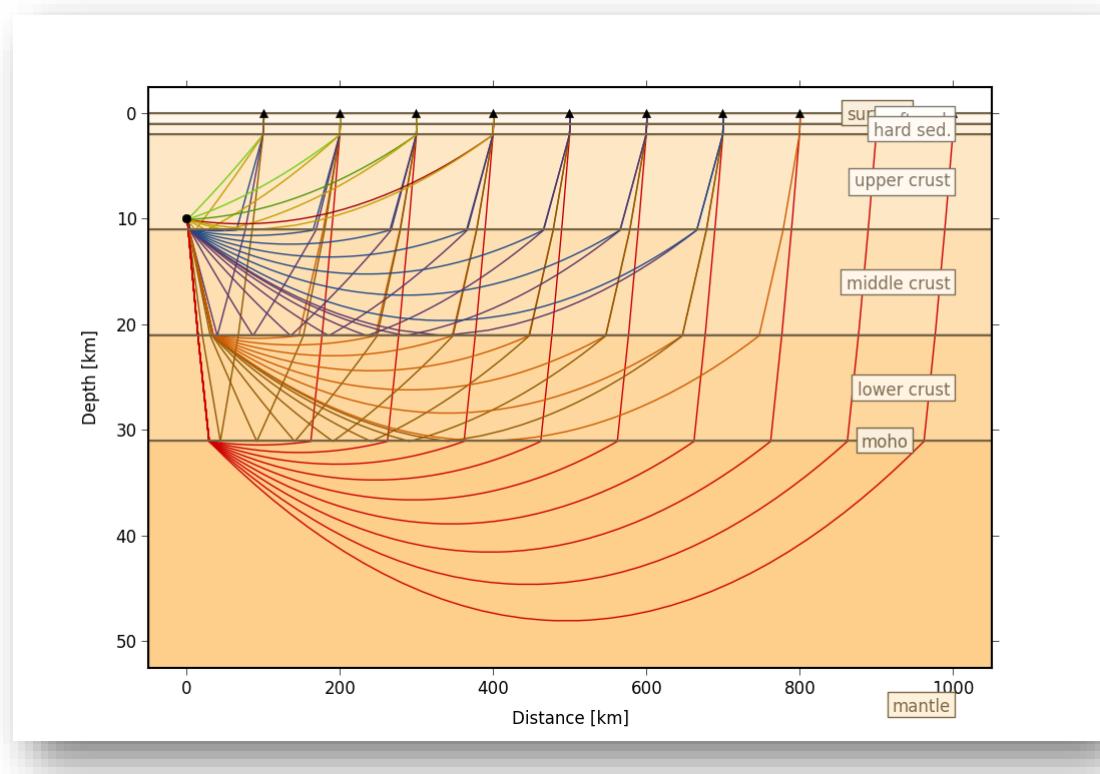
A seismology toolkit

Language:

Python

Link:

[https://git.pyrocko.org/
pyrocko/pyrocko](https://git.pyrocko.org/pyrocko/pyrocko)



Auralib

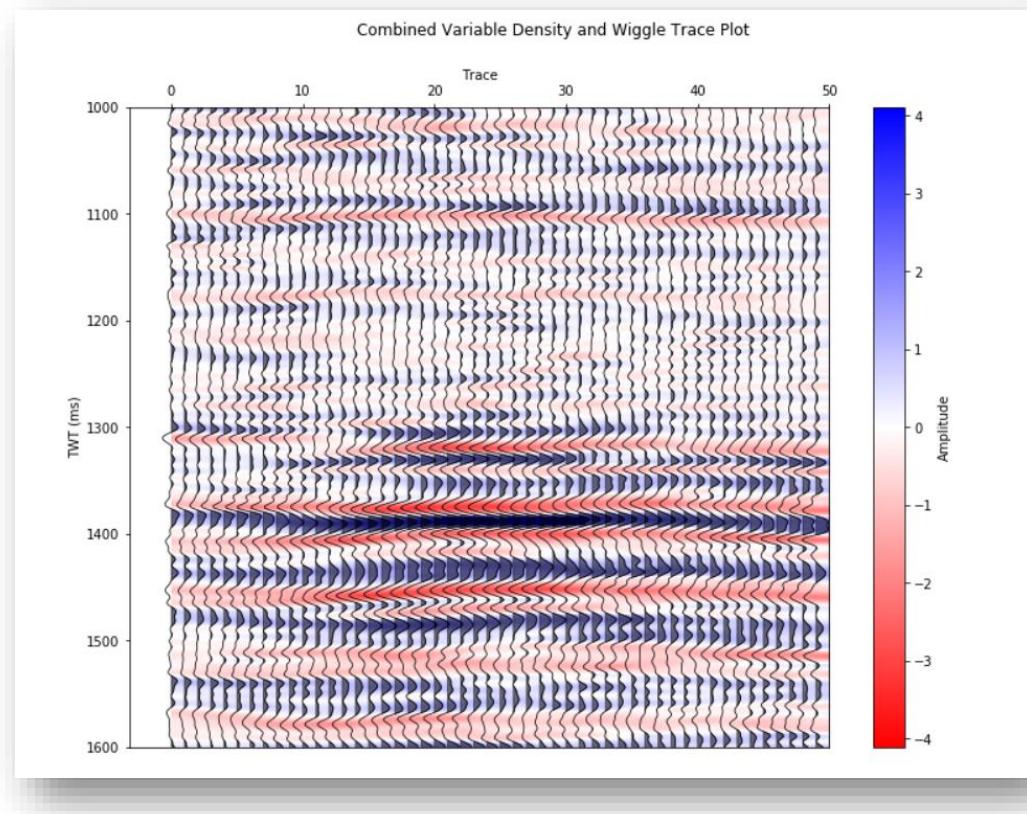
Support investigation of geoscience problems including geophysics, rock physics, petrophysics

Language:

Python

Link:

<https://github.com/wamllyn/auralib>



OpendTect Plugins

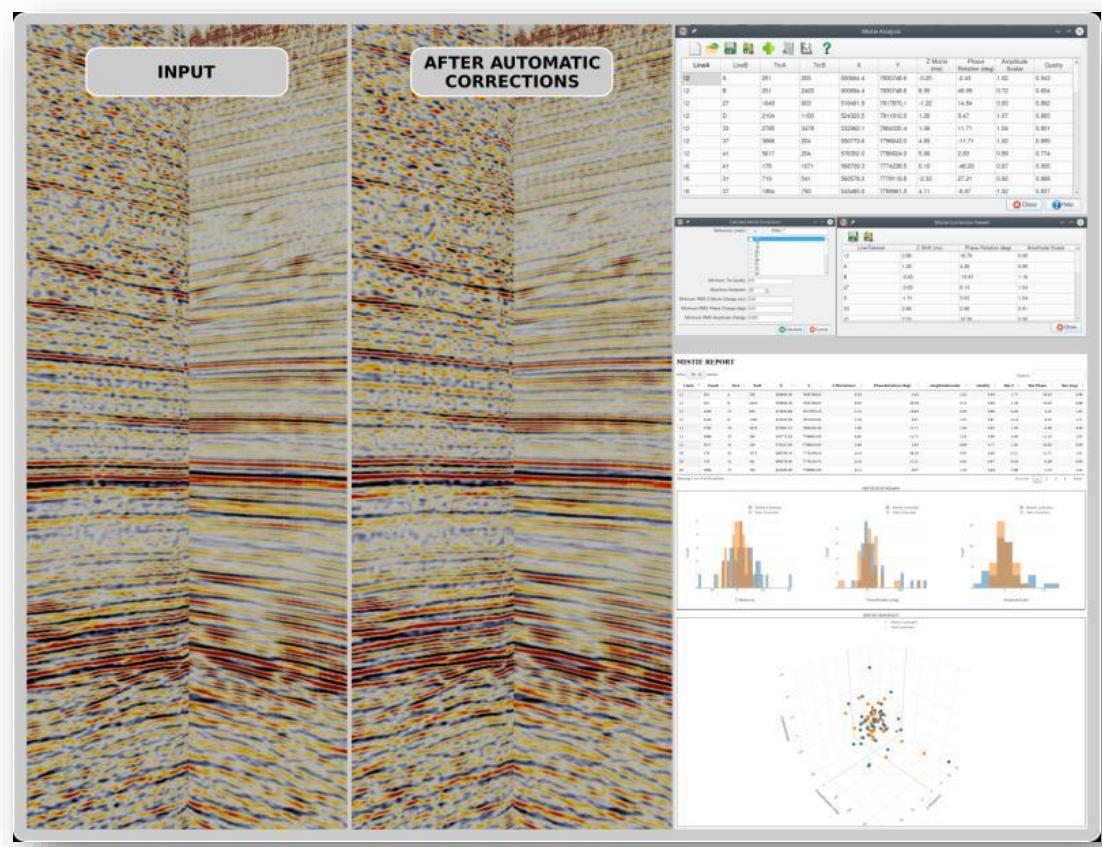
Open-source plugins for the OpendTect seismic interpretation platform

Language:

Python & C++

Link:

<https://github.com/waynegm/OpendTect-Plugins>



rsudp

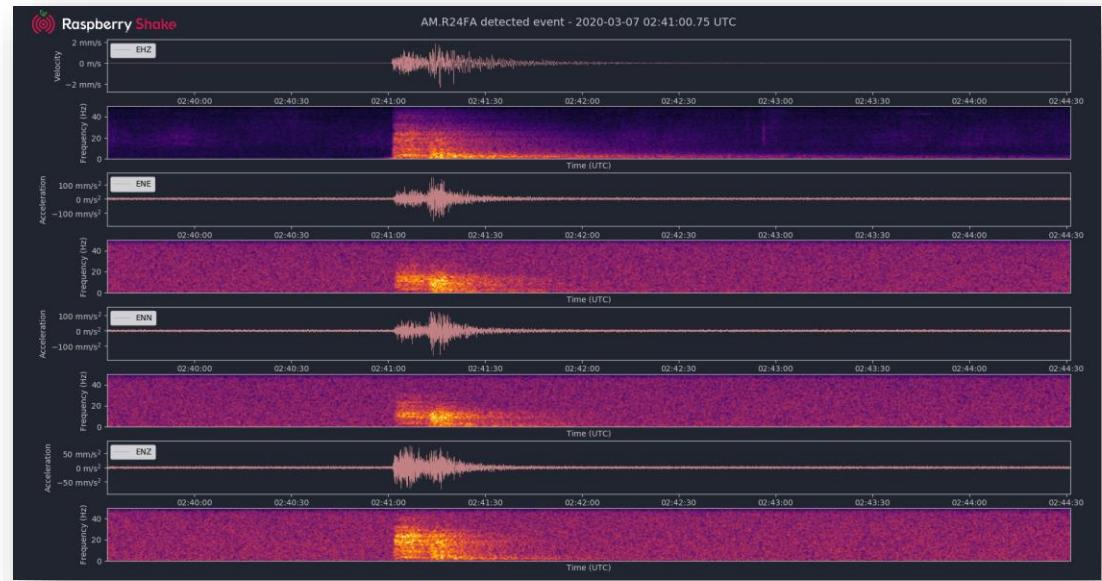
Tool for receiving and interacting with data casts from Raspberry Shake personal seismographs and Raspberry Boom pressure transducer instruments

Language:

Python

Link:

<https://github.com/raspisha/ke/rsudp>



SeismicZFP

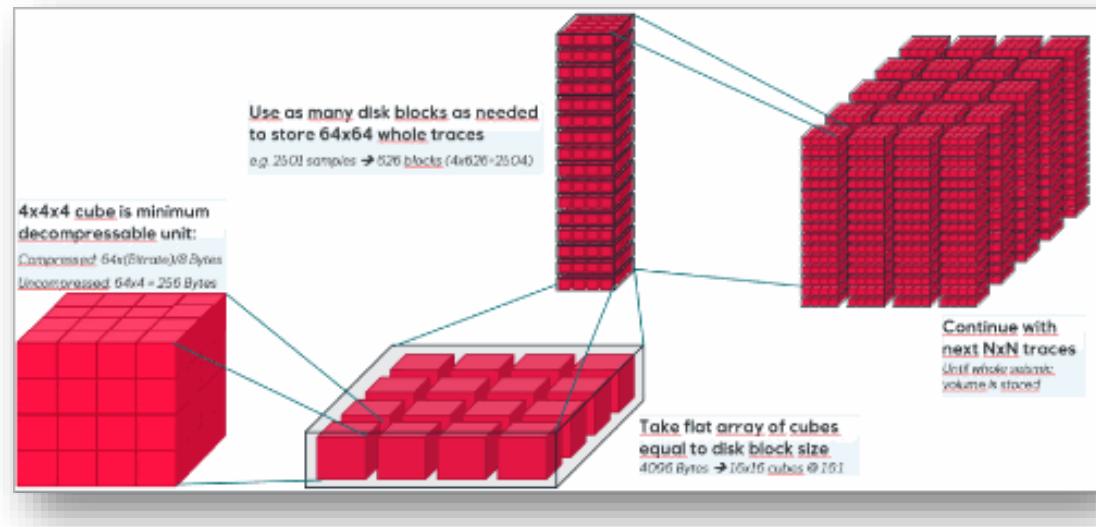
Convert SEG-Y files to compressed cubes and retrieve arbitrary sub-volumes from these, fast

Language:

Python

Link:

<https://github.com/equinor/seismic-zfp>



Ground-penetrating Radar



gprMax

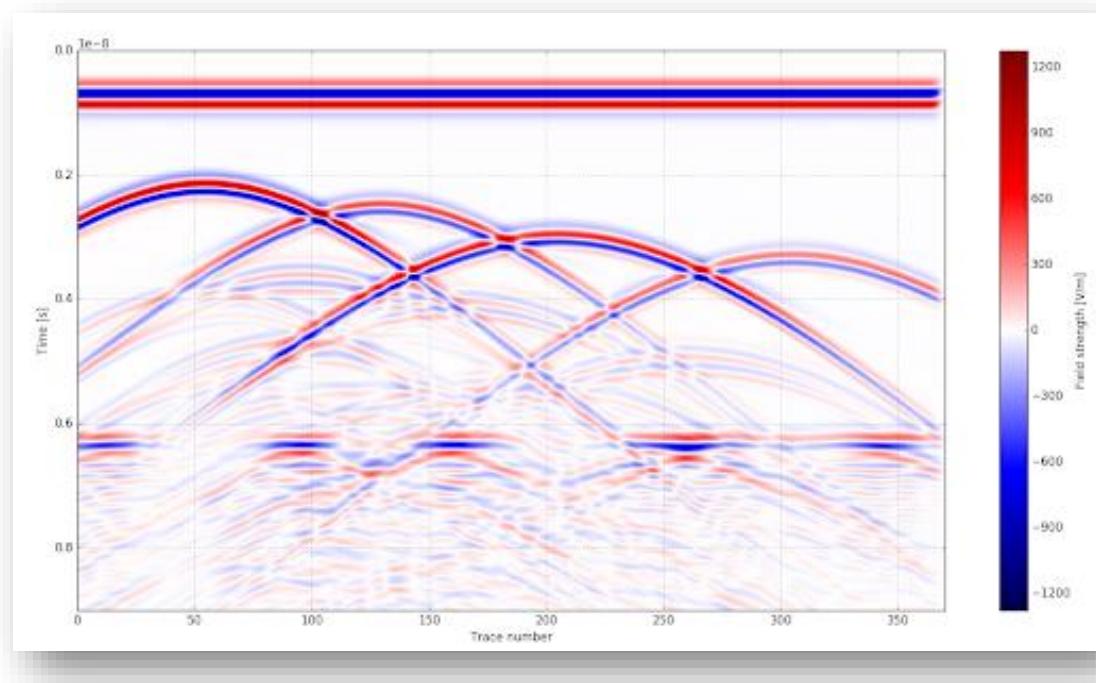
Finite-difference time-domain electromagnetic wave propagation simulator

Language:

Python & Cuda

Link:

<http://www.gprmax.com>



RGPR

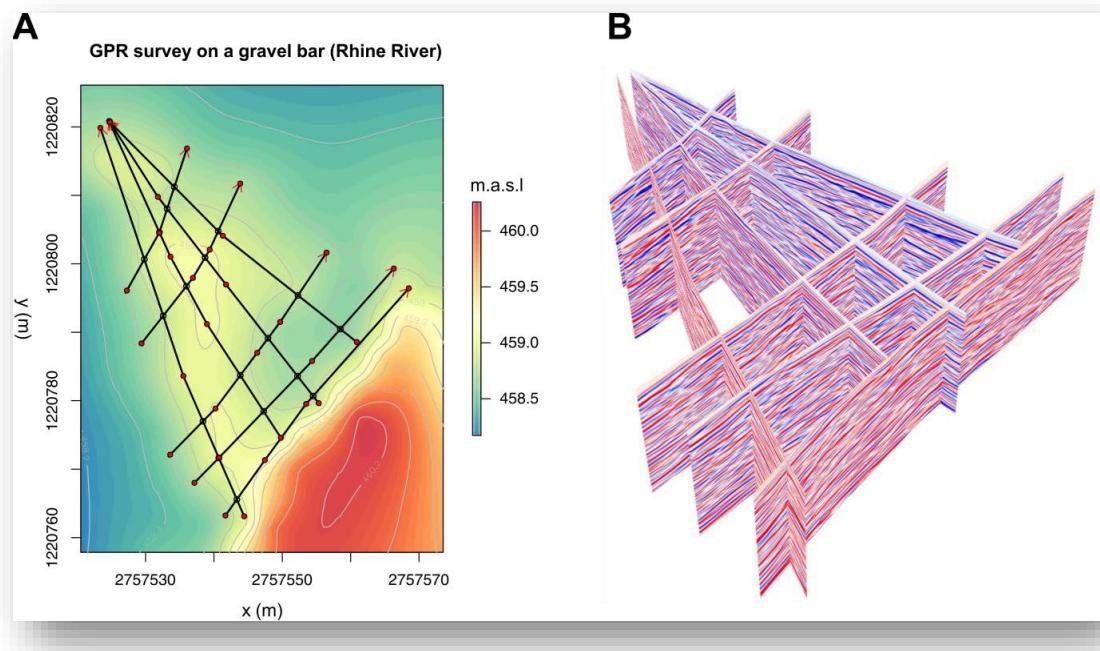
Reads, exports,
processes, and plots
ground-penetrating
radar data

Language:

R

Link:

<https://github.com/manuelhuber/RGPR>



readgssi

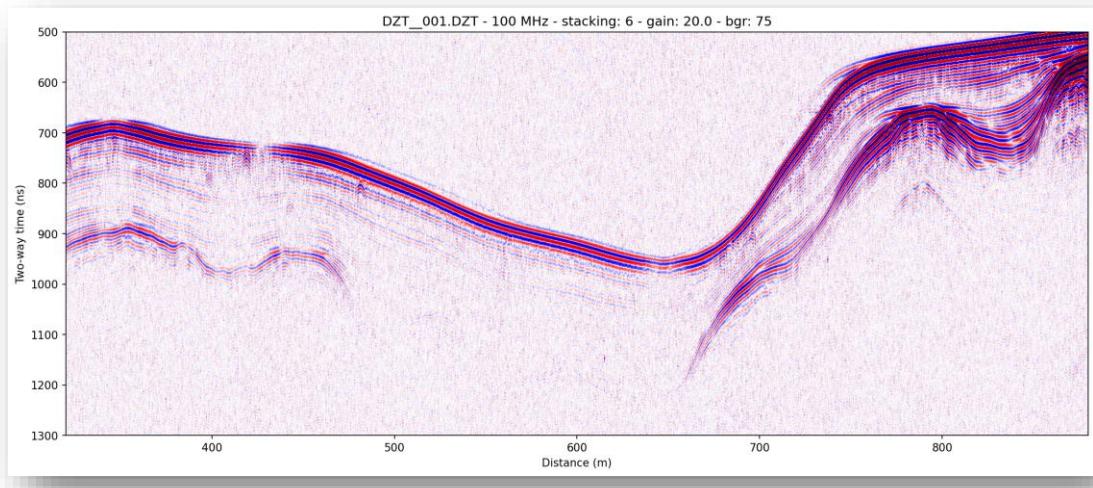
Fast command line or console-based visualization, filtering, and translation of GSSI radar data

Language:

Python

Link:

<https://github.com/iannesbitt/readgssi>



GPRPy

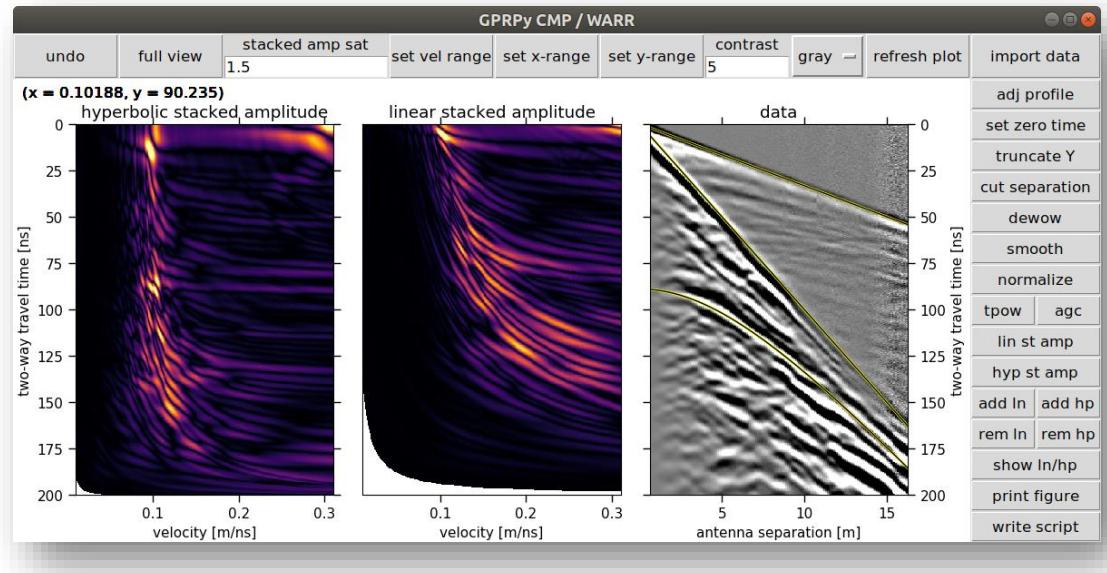
Multi-format, GUI-based GPR processing and visualization

Language:

Python

Link:

<https://github.com/NSG-eophysics/GPRPy>



RAGU

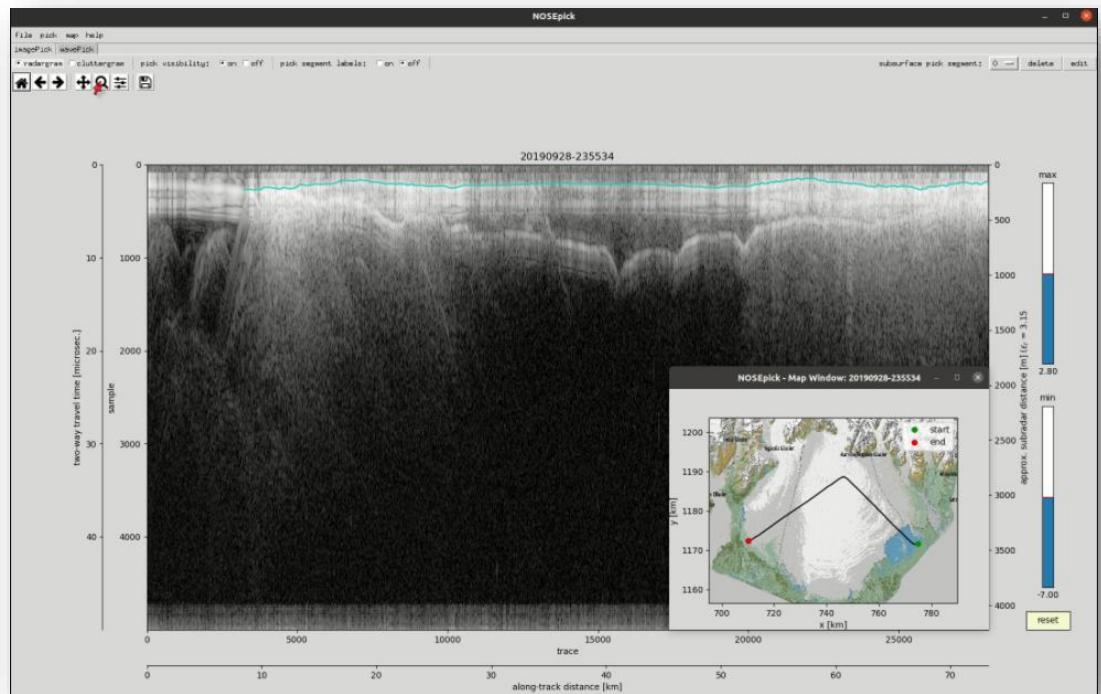
Radar interpretation
GUI compatible with
multiple radar datasets

Language:

Python

Link:

<https://github.com/btobers/RAGU>



Well Log



lasio

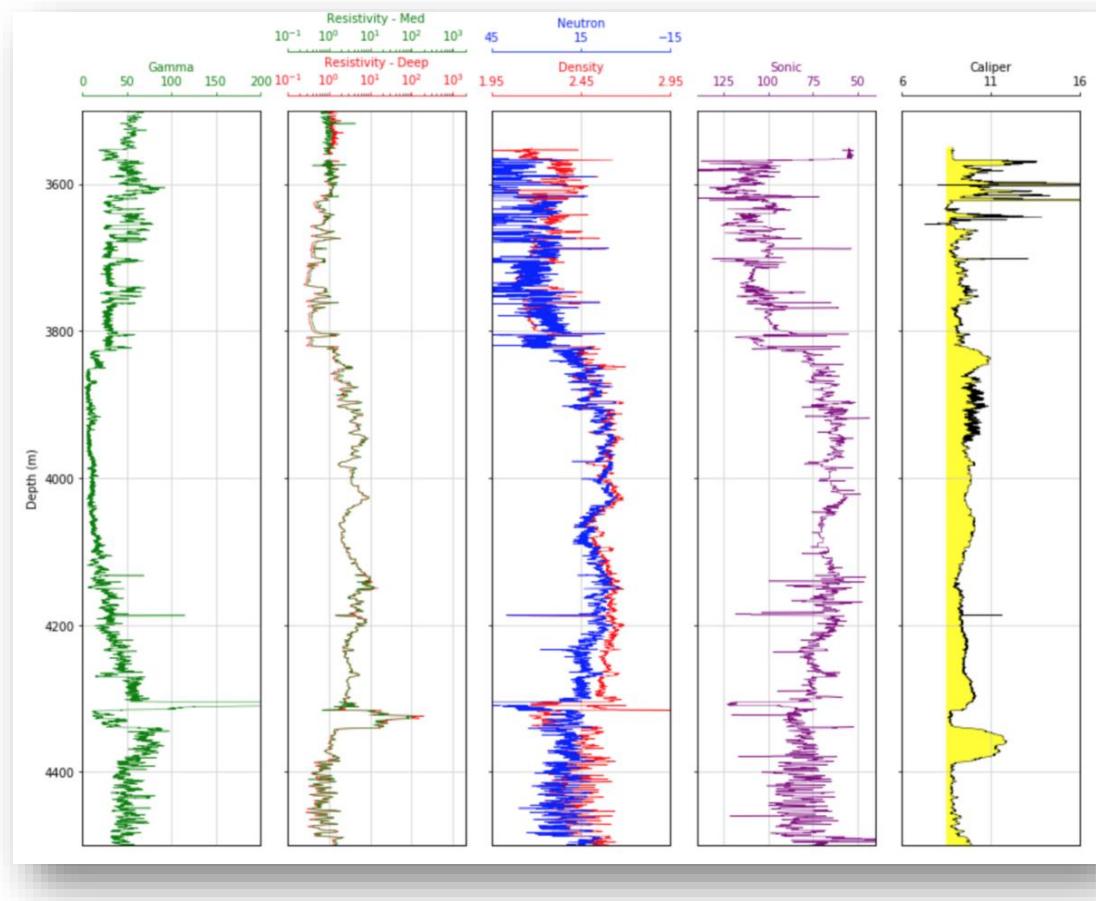
Reading and writing
well data using Log
ASCII Standard (LAS)
files

Language:

Python

Link:

<https://github.com/kinary1/lasio/>



welly

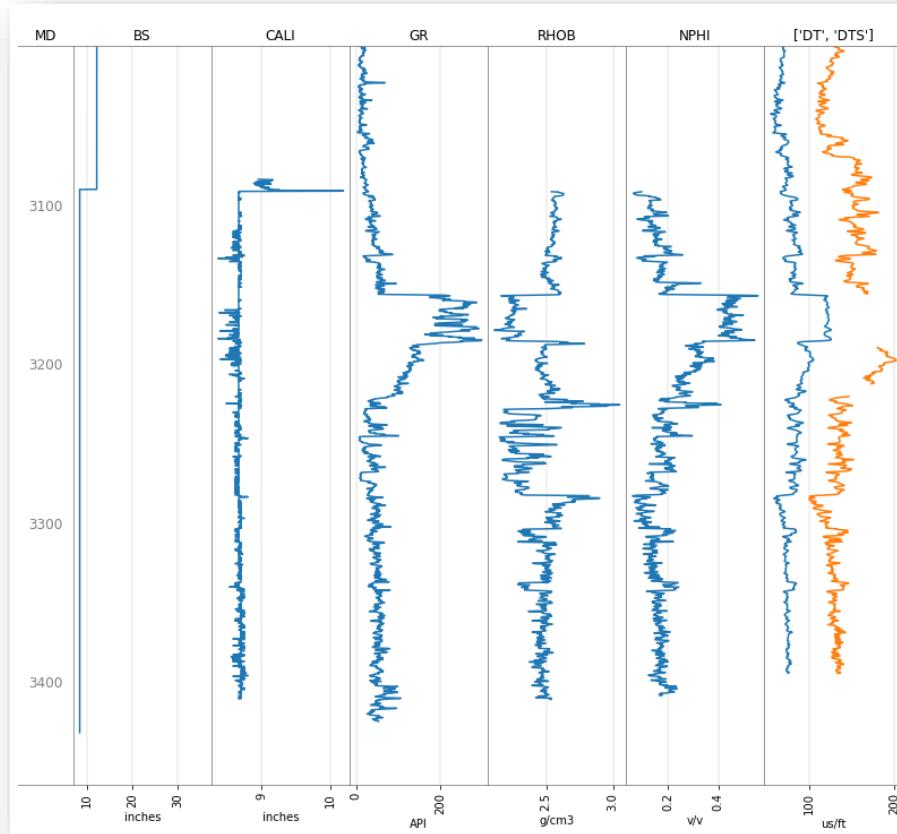
Analyzing and
processing well log data

Language:

Python

Link:

<https://github.com/agile-geoscience/welly>



dlisio



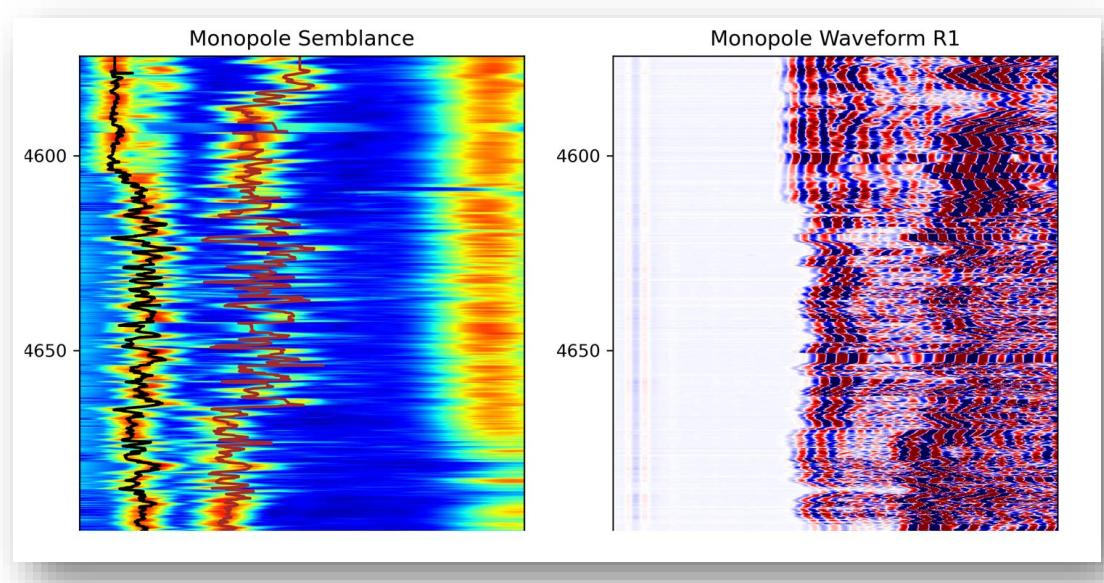
Parser for dlis well log files

Language:

Python & C++

Link:

<https://github.com/equinor/dlisio>



striplog



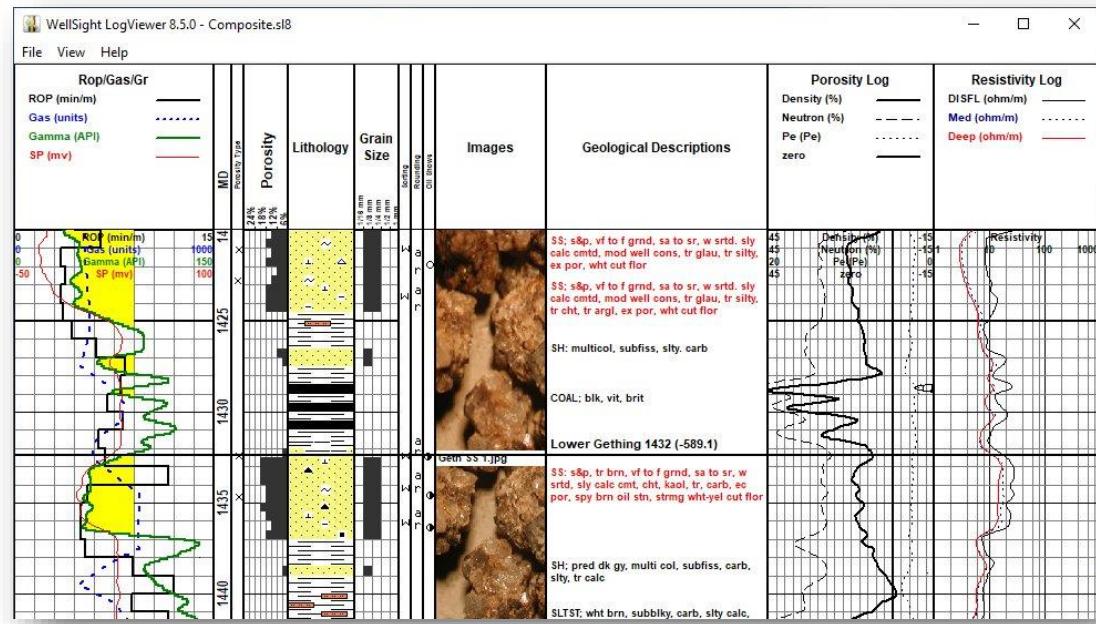
Display lithological and stratigraphic logs for wells and outcrop

Language:

Python

Link:

<https://github.com/agile-geoscience/striplog>



PetroPy

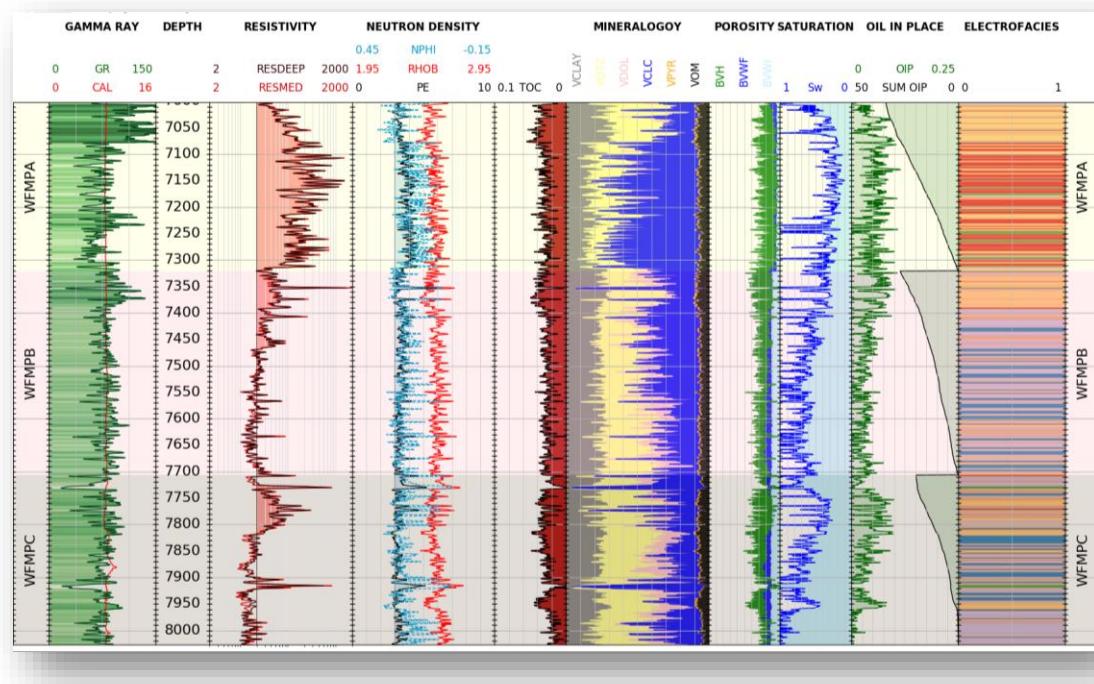
Petrophysics package
for conventional and
unconventional
formation evaluation

Language:

Python

Link:

<https://github.com/todheitmann/PetroPy>



wellpathpy

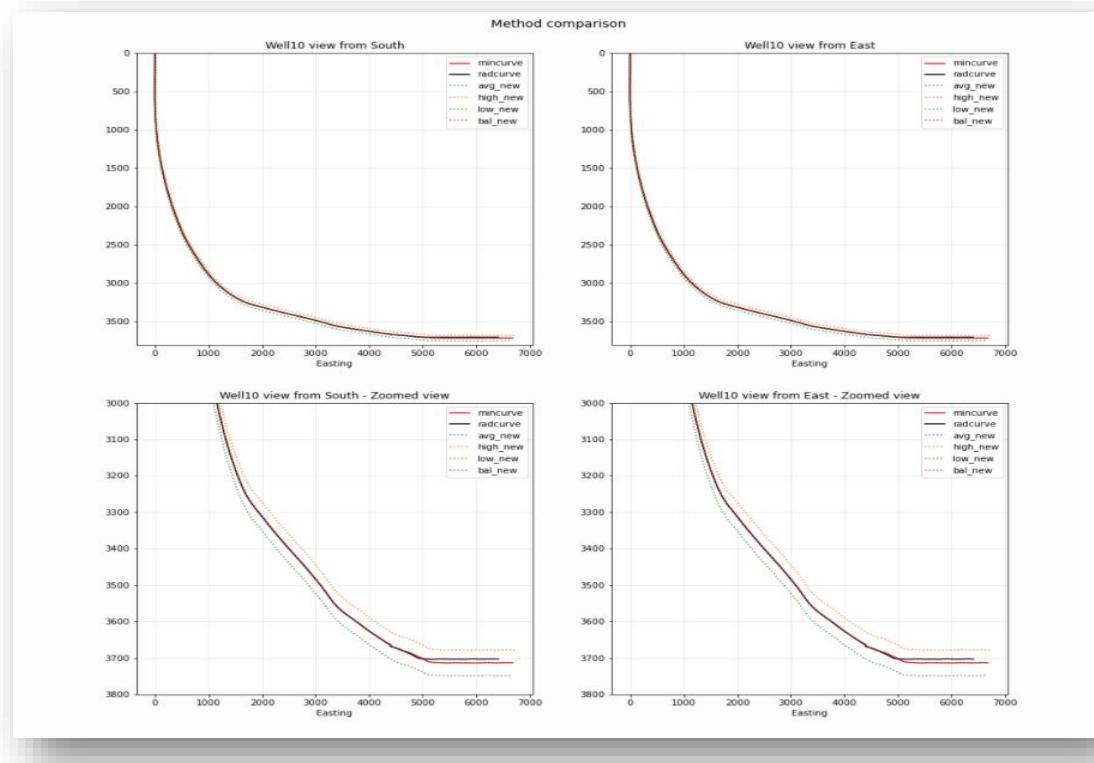
Library to import well deviations, calculate their TVD values using a choice of methods and return them as positional logs

Language:

Python

Link:

<https://github.com/Zabamund/wellpathpy>



Simulation and Modelling



SimPEG



Simulation and gradient based parameter estimation in the context of geophysical applications

Language:

Python

Link:

<https://github.com/simpeg/simpeg>

Open source software for simulations and inversions of airborne electromagnetic data

Lindsay J. Heagy
University of British Columbia,
Geophysical Inversion Facility
Vancouver, BC, Canada
lheagy@eos.ubc.ca

Seogi Kang
University of British Columbia,
Geophysical Inversion Facility
Vancouver, BC, Canada
skang@eos.ubc.ca

Rowan Cockett
University of British Columbia,
Geophysical Inversion Facility
Calgary, AB, Canada
rowanc@gmail.com

Douglas W. Oldenburg
University of British Columbia,
Geophysical Inversion Facility
Vancouver, BC, Canada
doug@eos.ubc.ca

Forward Simulation

Figure 1. Depth slice (left) and cross section (right) through the model of a conductive plate (8.1 S/m) in a resistive half-space (10^3 S/m).

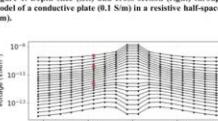


Figure 2. Vertical $\partial V / \partial t$ data over the conductive plate. The red dots correspond to the times and source locations shown in Fig 3.

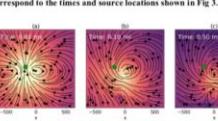


Figure 4. Magnetic flux density in the y=0m cross-section. The source and receiver location is shown by the green dot.

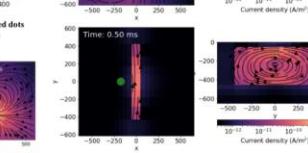
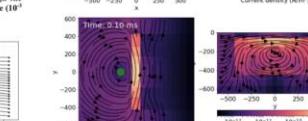
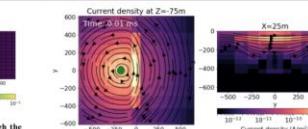
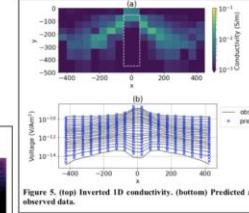


Figure 5. (top) Inverted 1D conductivity. (bottom) Predicted and observed data.

1D inversion



AEM2018/7th International Workshop on Airborne Electromagnetics

2D parametric inversion

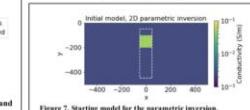


Figure 7. Starting model for the parametric inversion.

2D voxel inversion

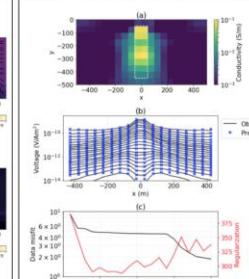


Figure 8. (a) Conductivity model recovered in a parametric inversion for a 2D plate. (b) residual predicted and observed data. (c) data misfit over the course of the inversion.

March 10, 2018

Devito

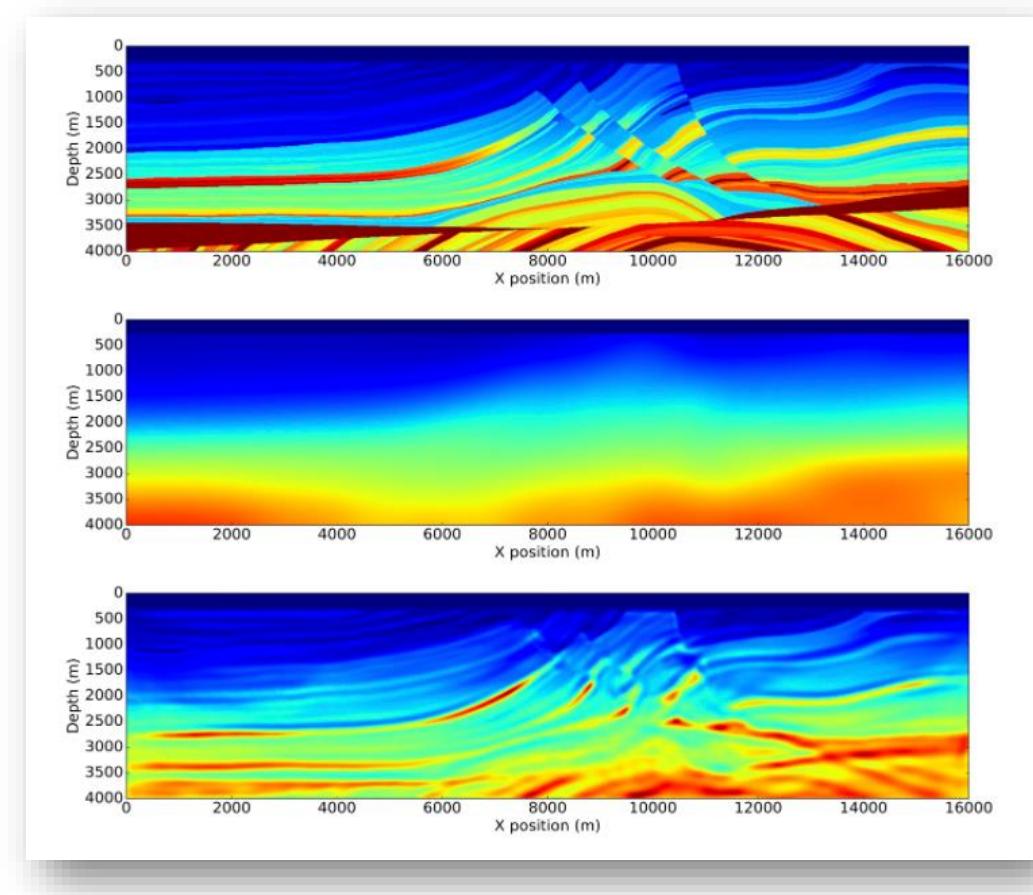
Finite-Difference
computation from high-
level symbolic problem
definitions

Language:

Python

Link:

<https://github.com/devitocodes/devito>



bh_tomo

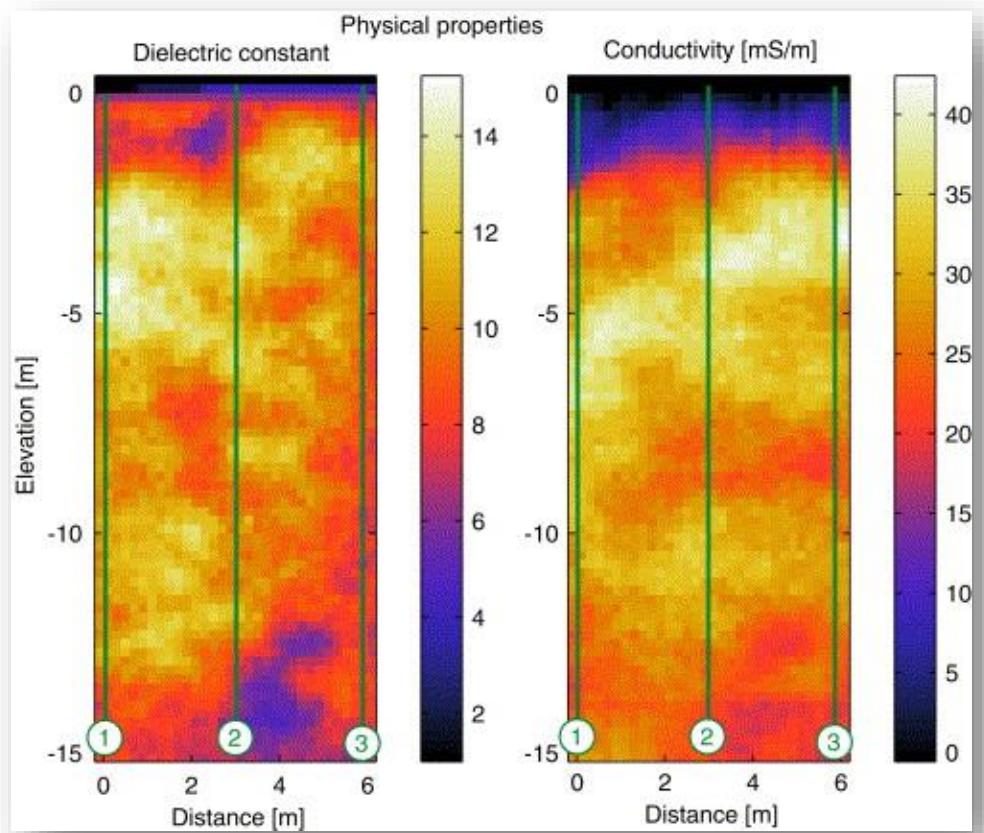
Borehole radar and seismic tomography package

Language:

MATLAB & C++

Link:

https://github.com/groupeLIAMG/bh_tomo



modelr

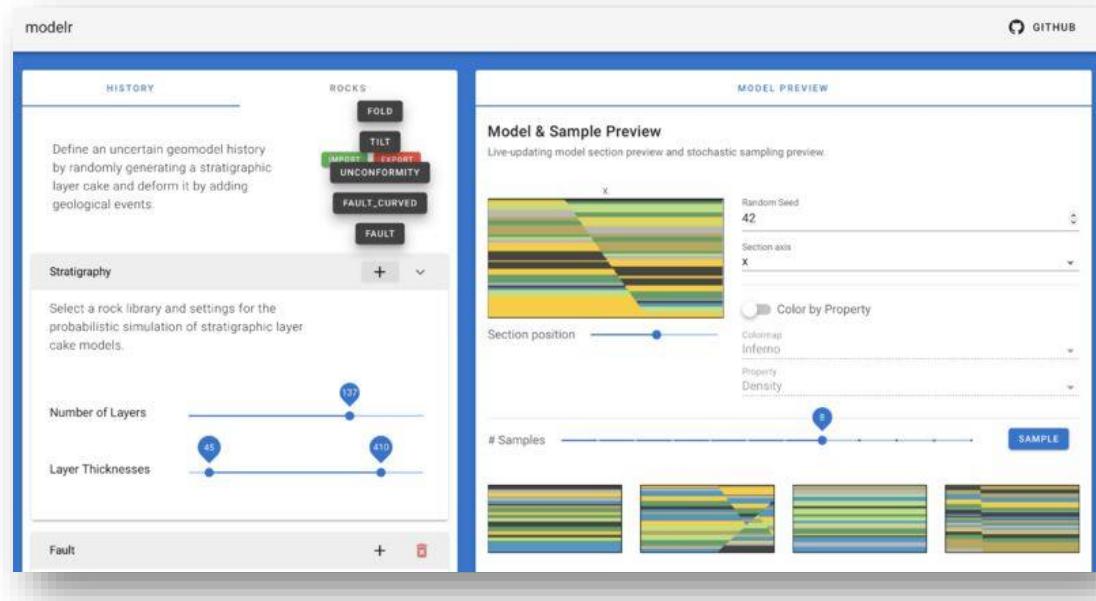
Web app for simple synthetic seismic forward modelling

Language:

Python, HTML &
Javascript

Link:

<https://github.com/agile-geoscience/modelr>



ModFlow

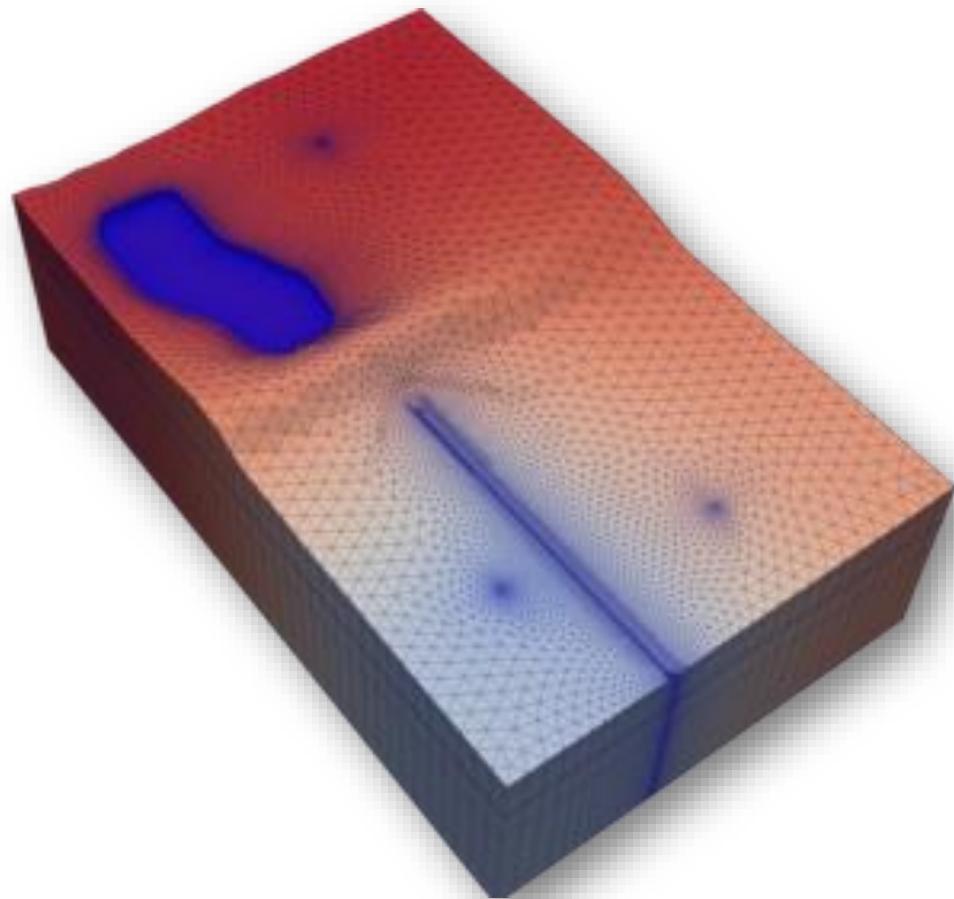
Flow modelling software
to simulate and predict
groundwater conditions
and groundwater
interactions

Language:

Python

Link:

<https://www.usgs.gov/software/modflow-6-usgs-modular-hydrologic-model>



pyGIMLi

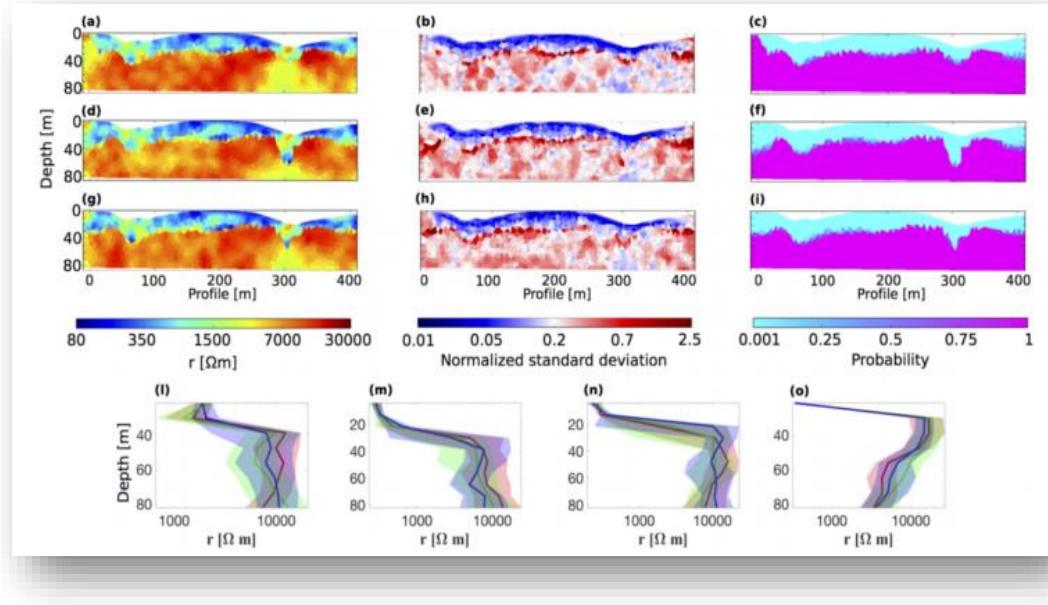
Multi-method library
for solving inverse and
forward tasks related to
geophysical problems

Language:

Python & C++

Link:

<https://www.pygimli.org/>



GemPy

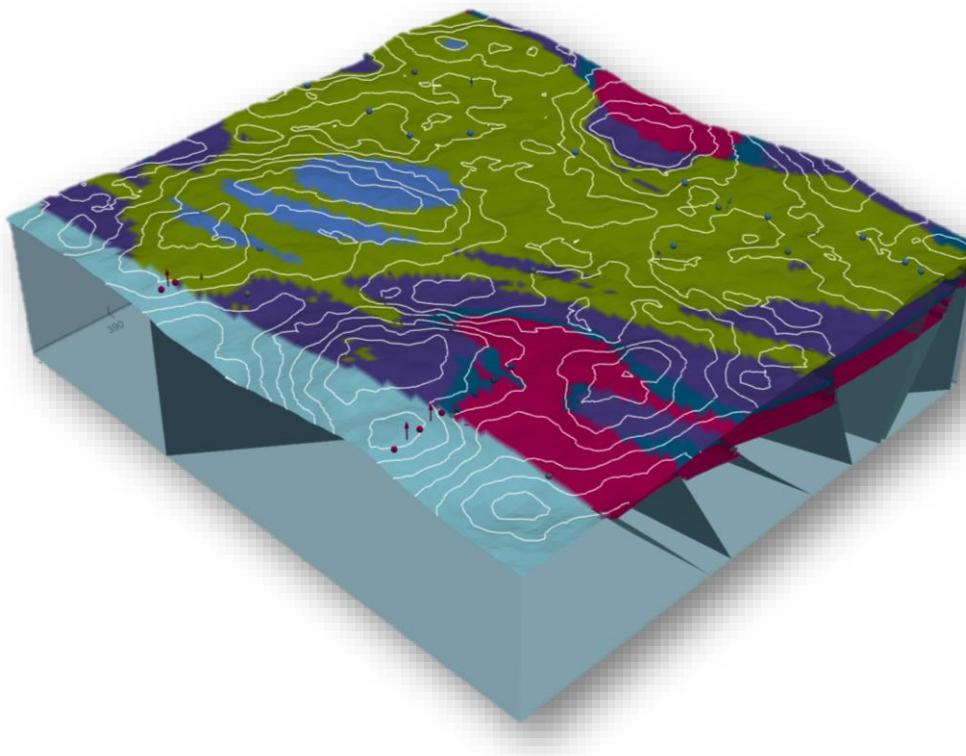
3-D structural geological modelling software with implicit modelling and support for stochastic modelling

Language:

Python

Link:

<https://github.com/cgreaachen/gempy>



HyVR



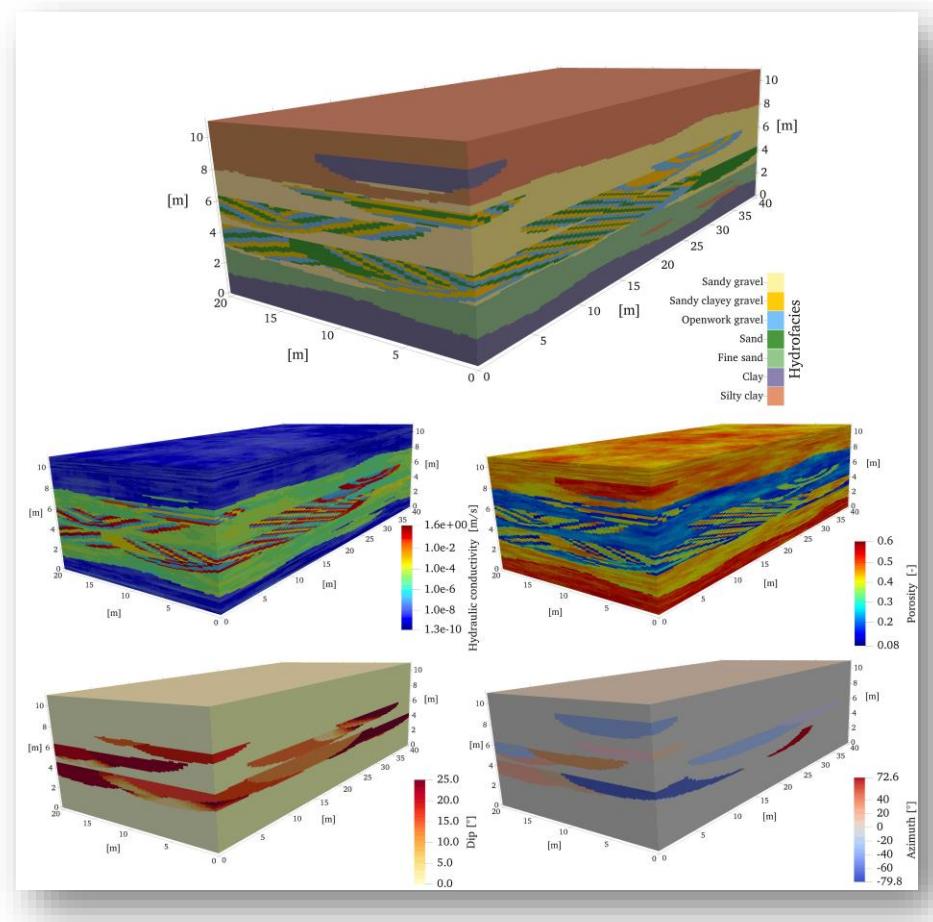
3-D anisotropic subsurface models based on geological concepts that can be used with groundwater flow simulators

Language:

Python

Link:

<https://github.com/driftingides/hyvr>



Landlab

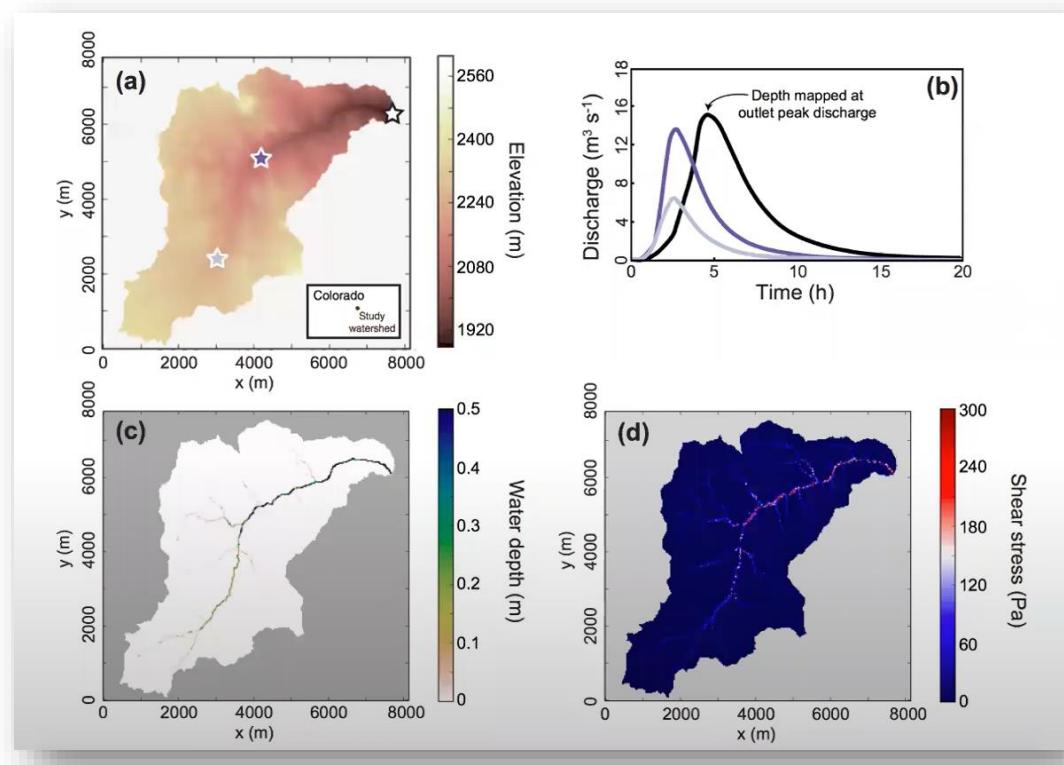
Simulate surface processes using a large suite of existing interoperable process components

Language:

Python

Link:

<https://github.com/landlab/landlab>



pyGeoPressure

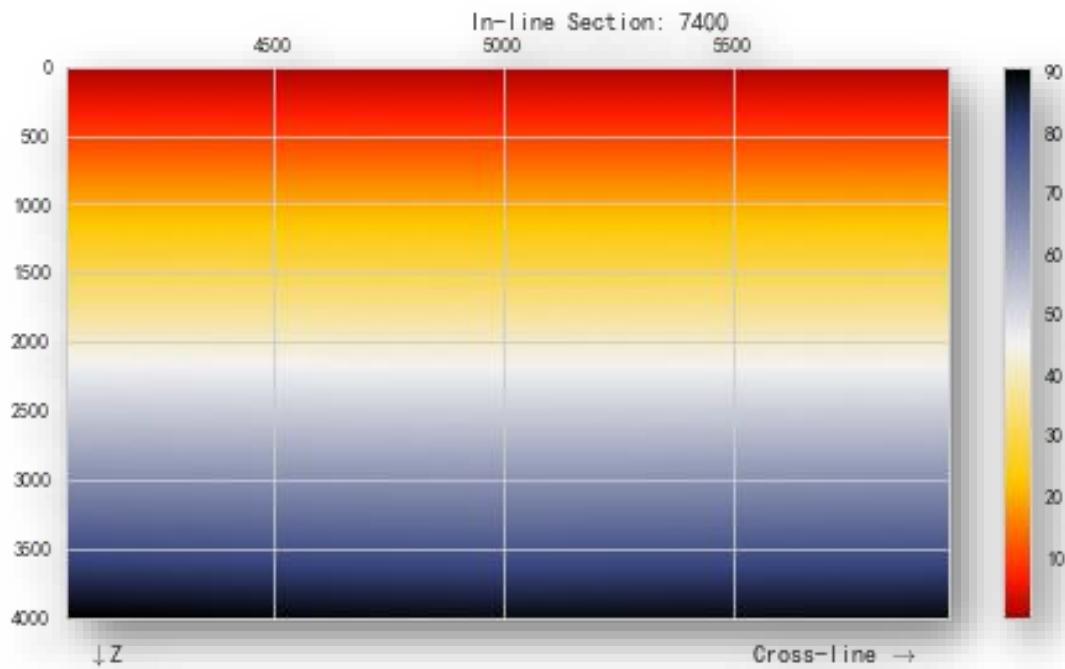
Pore pressure prediction using well log data and seismic velocity data

Language:

Python

Link:

<https://pygeopressure.readthedocs.io/en/latest>



emsig

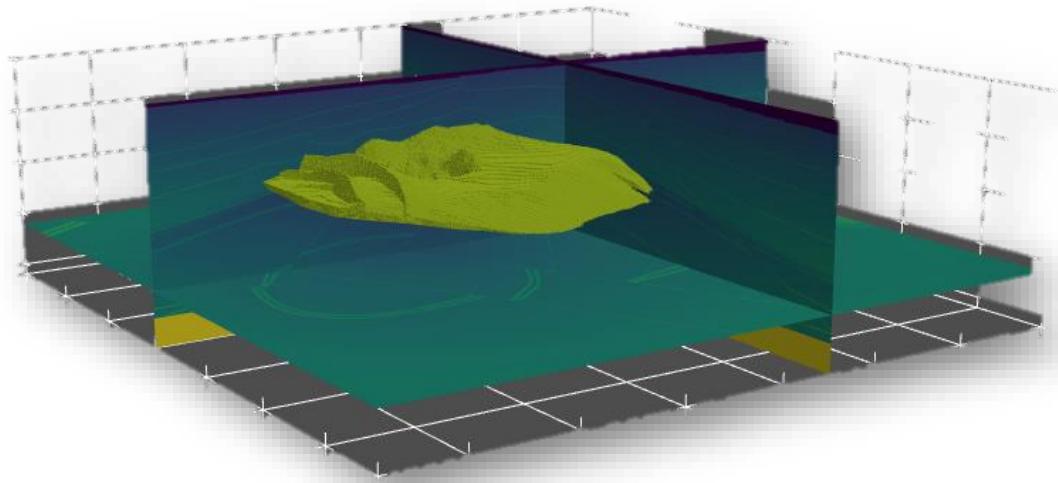
Controlled-source
electromagnetic
modelers for layered
and three-dimensional
anisotropic media

Language:

Python

Link:

<https://emsig.xyz>



PyLops

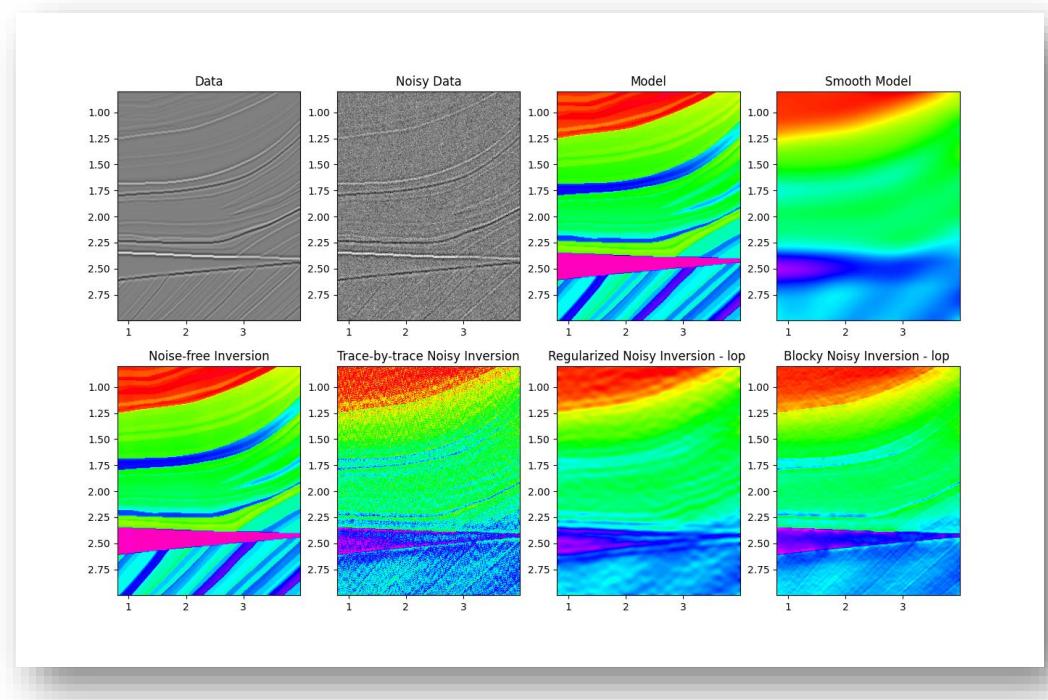
Linear Operators with
some geophysics /
seismic modules

Language:

Python

Link:

<https://pylops.readthedocs.io/en/latest>



ttcrpy

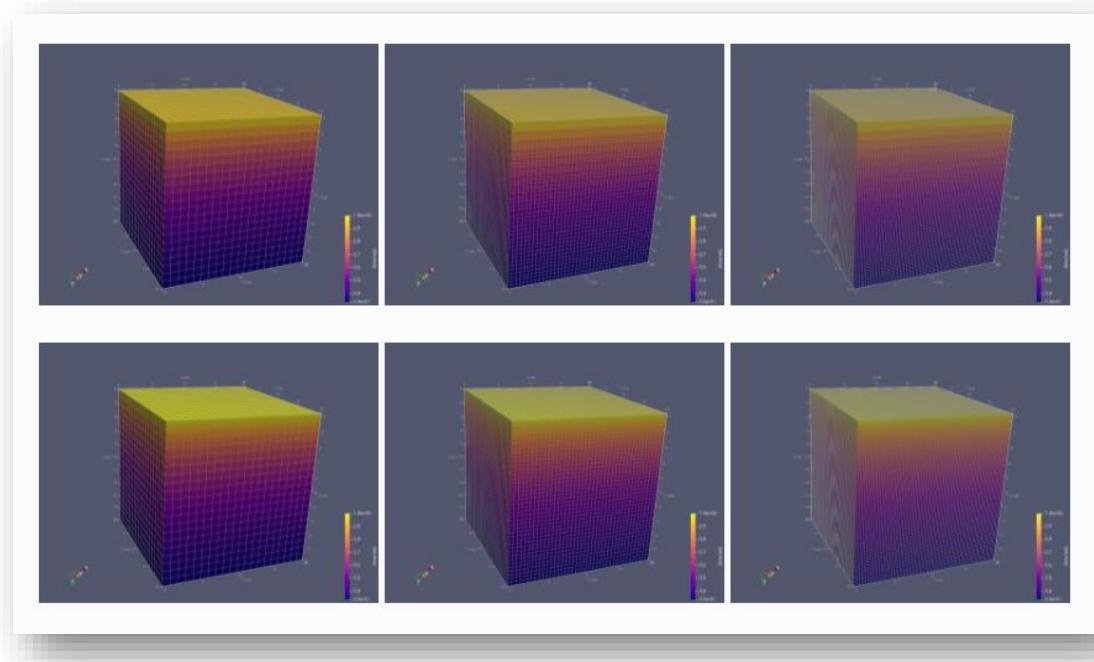
Travel-time computation and raytracing on 2D & 3D rectilinear grids and unstructured meshes

Language:

Python

Link:

<https://ttcrpy.readthedocs.io/en/latest/>



XTgeo

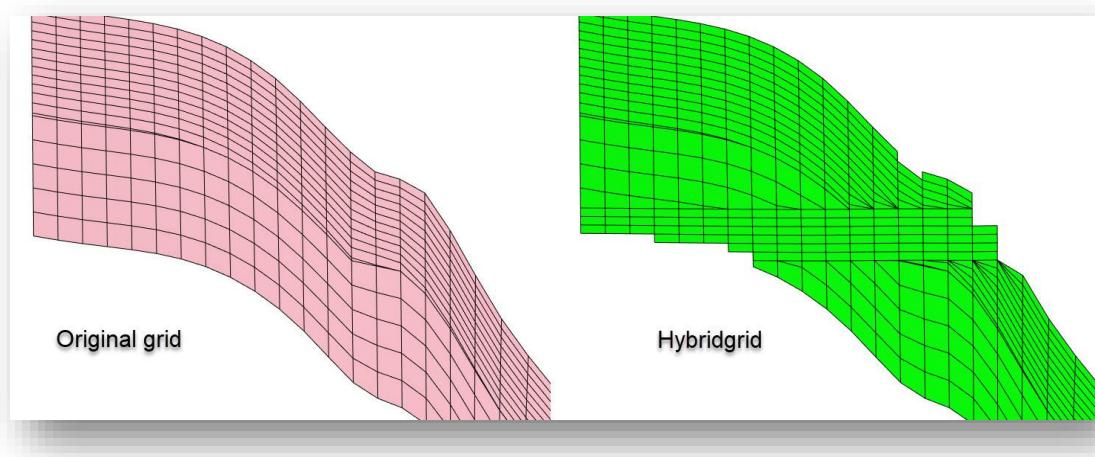
Support manipulation
of subsurface reservoir
modelling

Language:

Python

Link:

<https://xtgeo.readthedocs.io/en/latest/>



disba

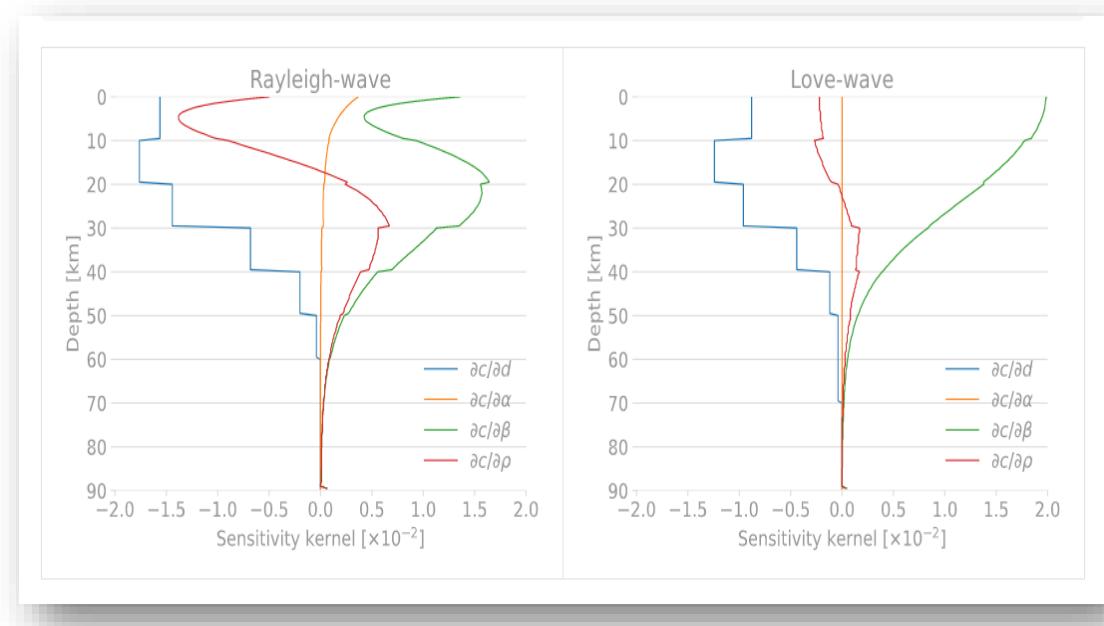
Numba-accelerated
computation of surface
wave dispersion

Language:

Python

Link:

<https://github.com/keurfonluu/disba>



PyGMI

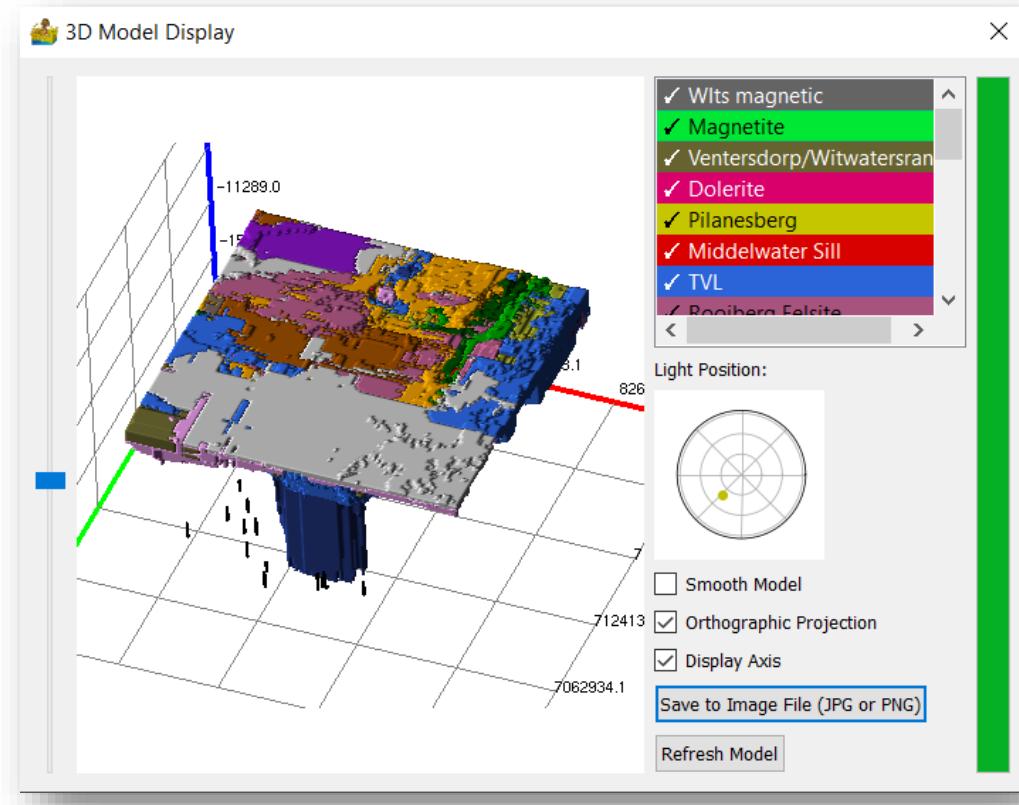
A modelling and interpretation suite aimed at magnetic, gravity and other datasets

Language:

Python

Link:

<https://patrick-cole.github.io/pygmi>



Reservoir Engineering



libres

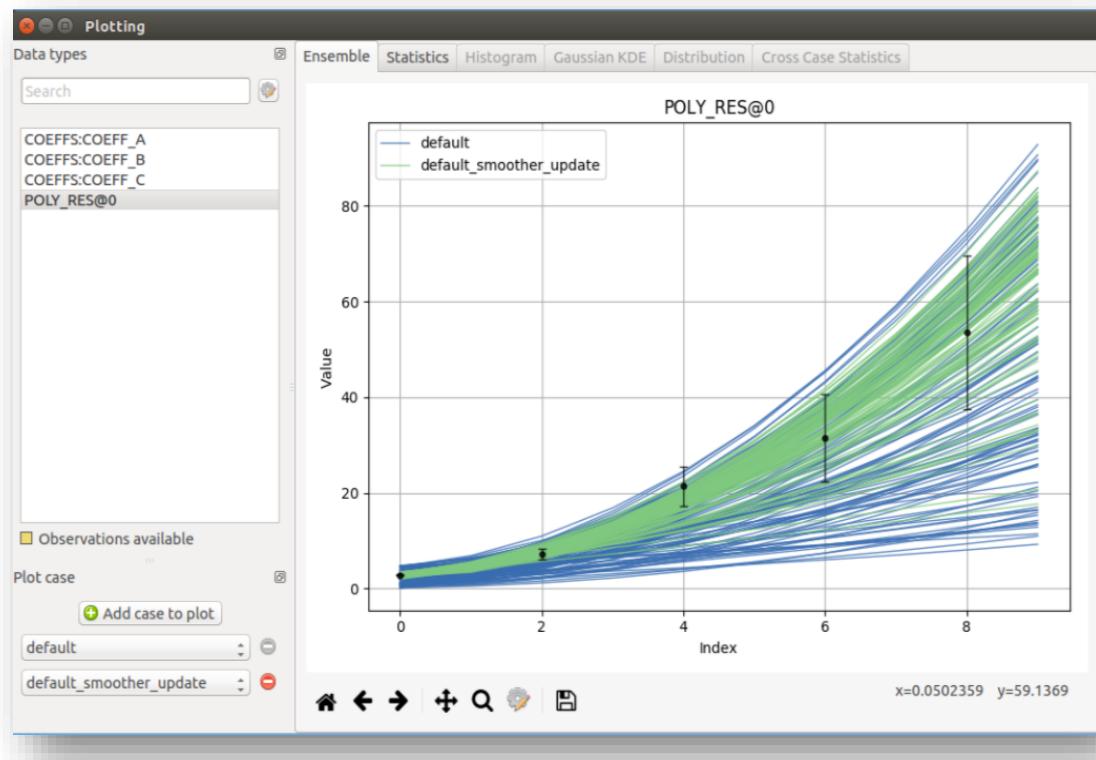
Tool for managing an ensemble of reservoir models

Language:

Python

Link:

<https://github.com/equinor/libres>



ecl

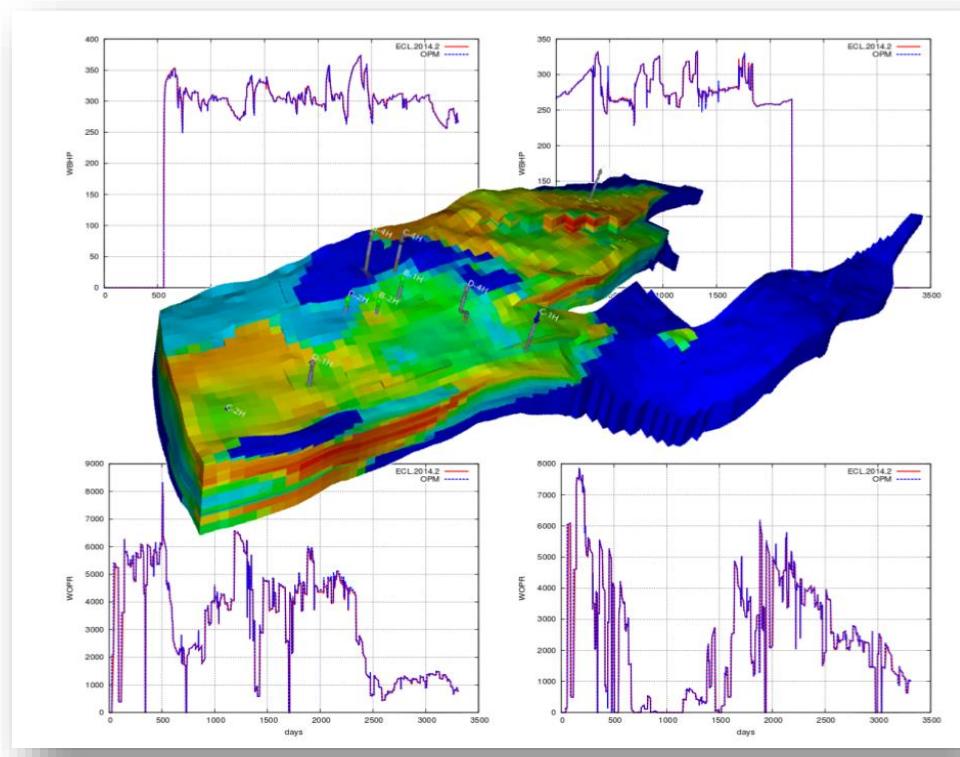
Reading and writing
Eclipse reservoir
simulator files

Language:

Python & C++

Link:

<https://github.com/equinor/ecl>



MRST

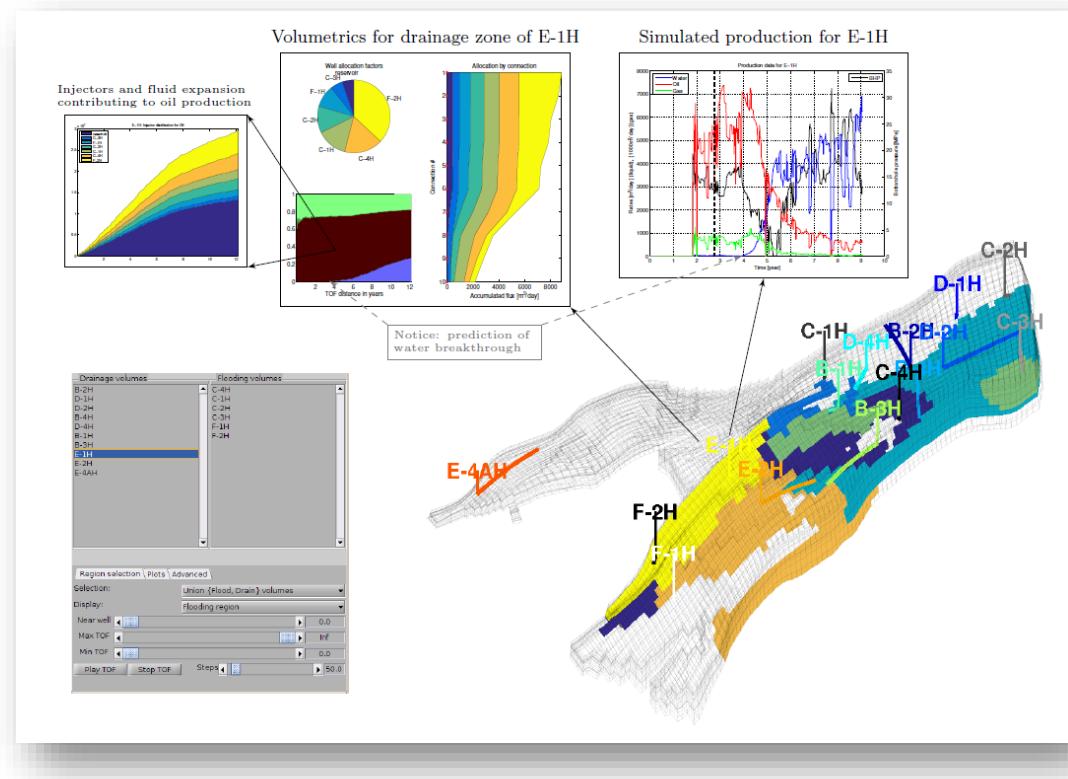
Rapid prototyping and demonstration of new simulation methods in reservoir modelling and simulation

Language:

MATLAB

Link:

<https://www.sintef.no/projectweb/mrst>



DuMux

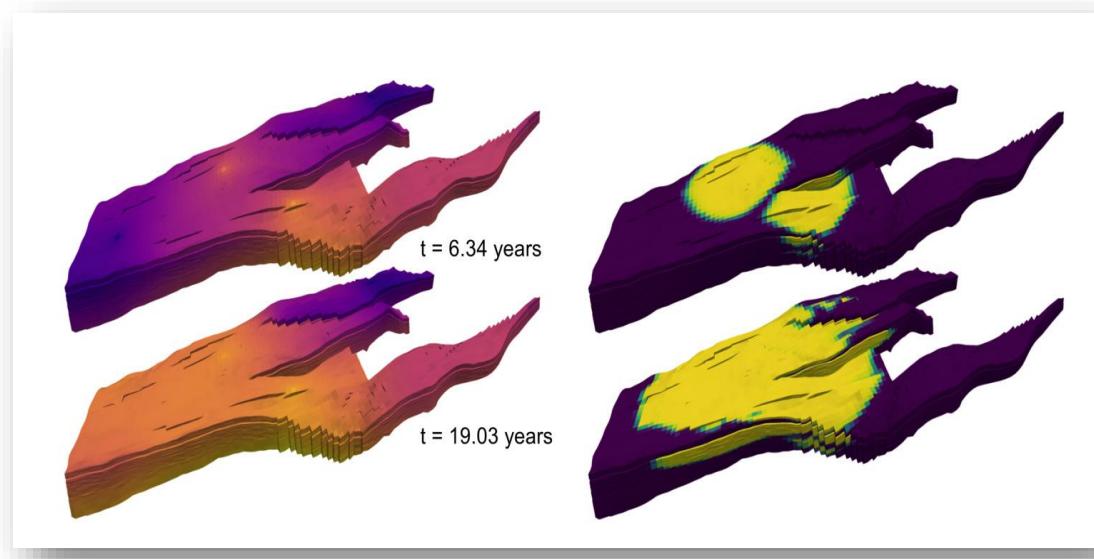
Simulator for flow and transport processes in porous media

Language:

C++

Link:

<https://dumux.org>



SHEMAT-Suite

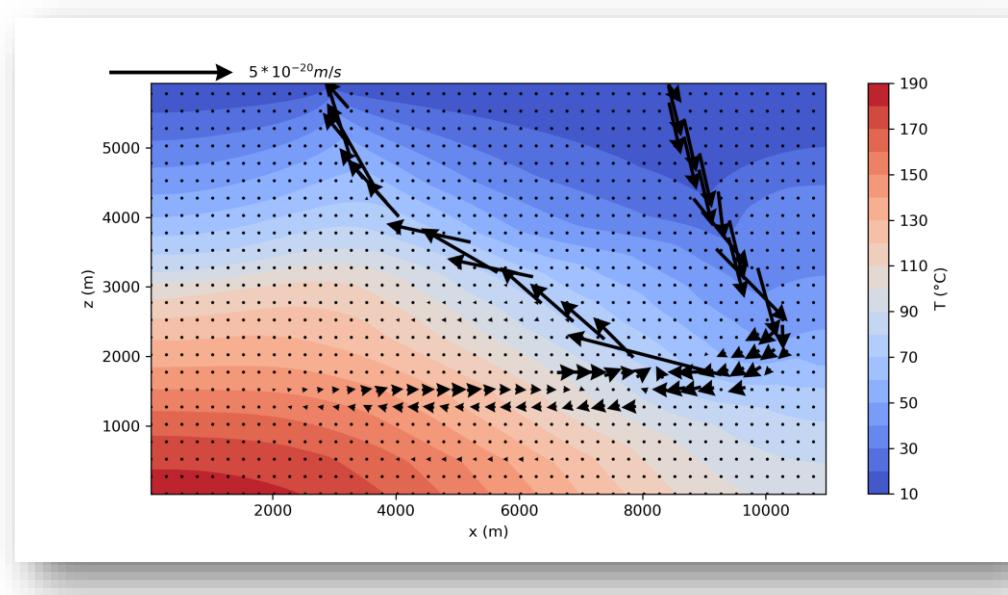
Simulator for flow, heat and species transport in porous media including stochastic and deterministic parameter estimation

Language:

Fortran-95

Link:

<https://git.rwth-aachen.de/SHEMAT-Suite/SHEMAT-Suite-open>



Geostatistics



pyKriging

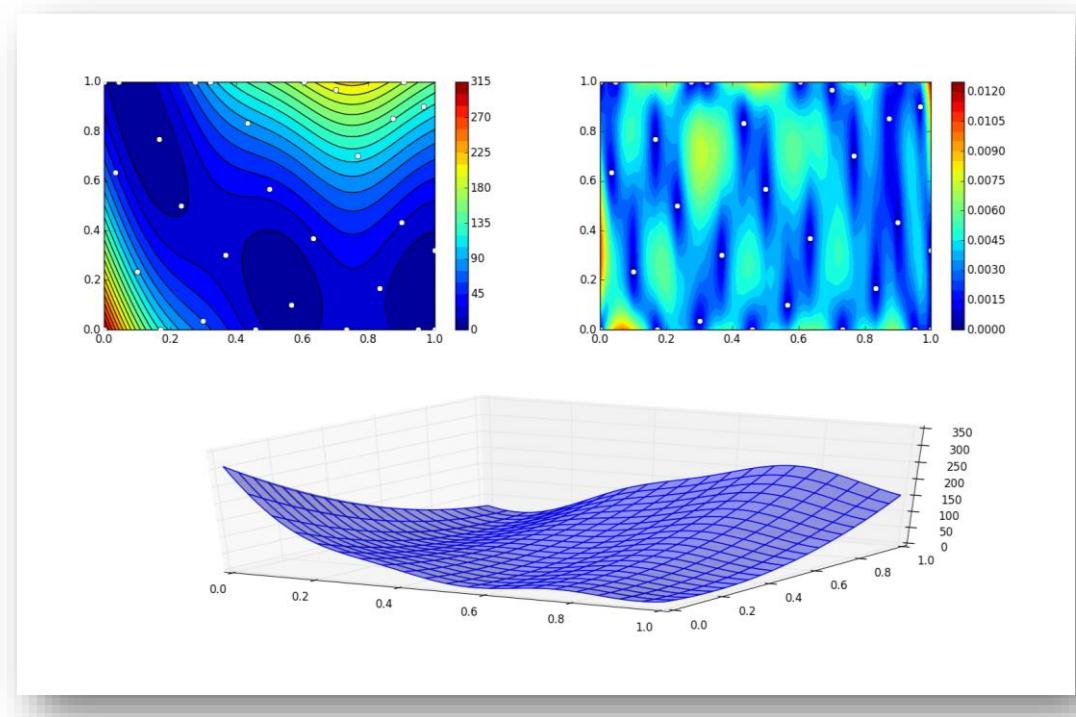
N-dimensional kriging

Language:

Python

Link:

<http://www.pykriging.com>



SGeMS

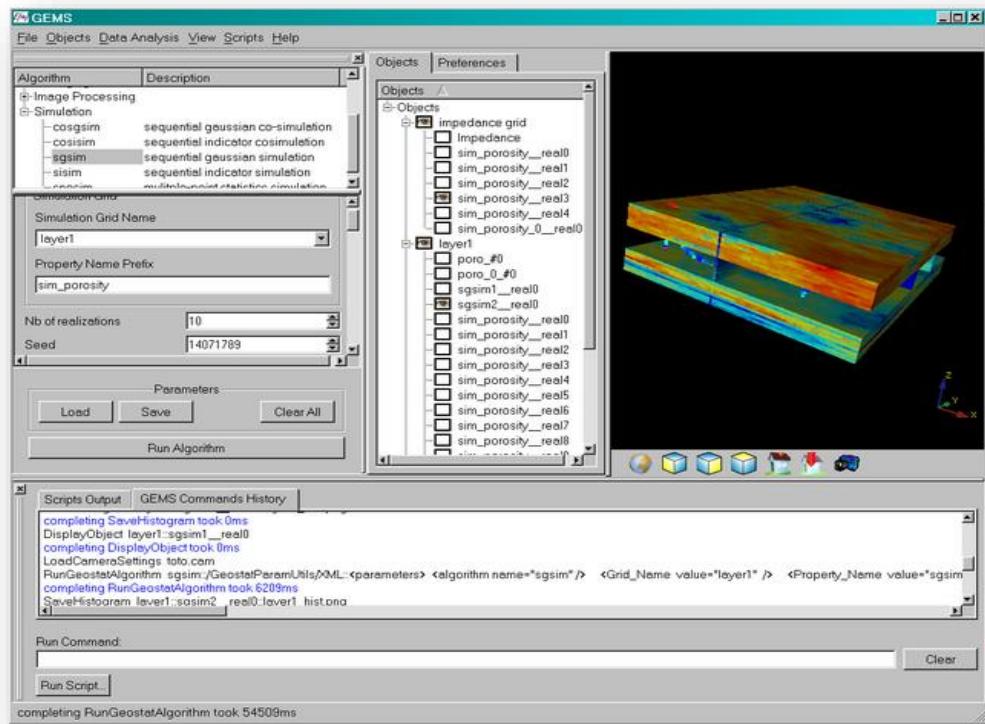
Stanford geostatistical
modelling software

Language:

QT

Link:

<http://sgems.sourceforge.net>



pysgems

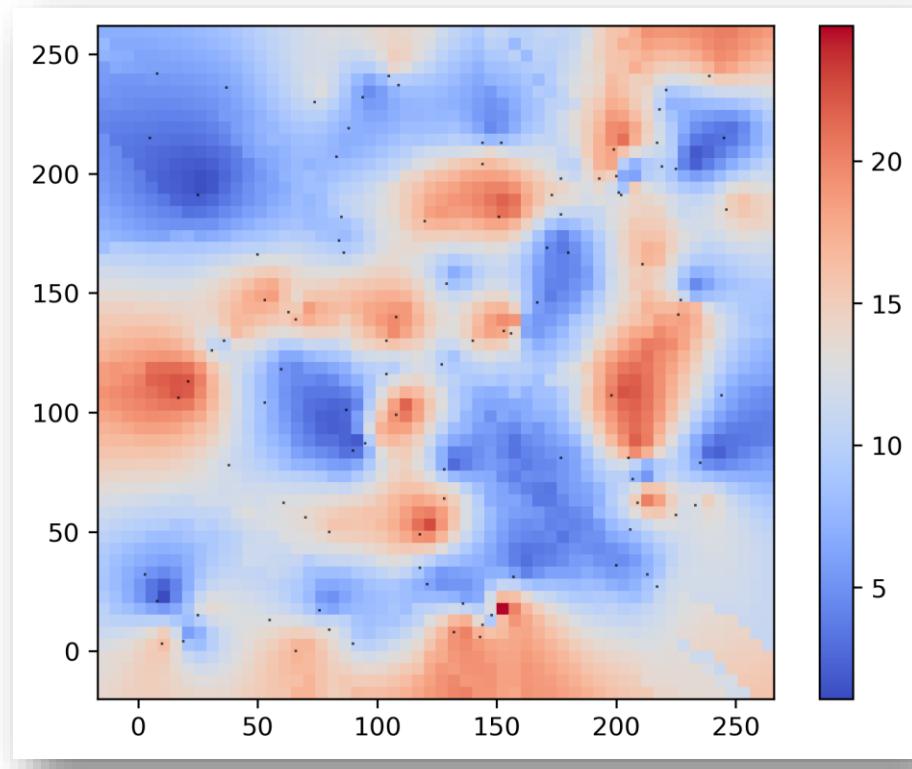
Use SGeMS (Stanford geostatistical modelling software) within Python

Language:

Python

Link:

<https://github.com/robinthibaut/pysgems>



HPGL

High performance
geostatistics library

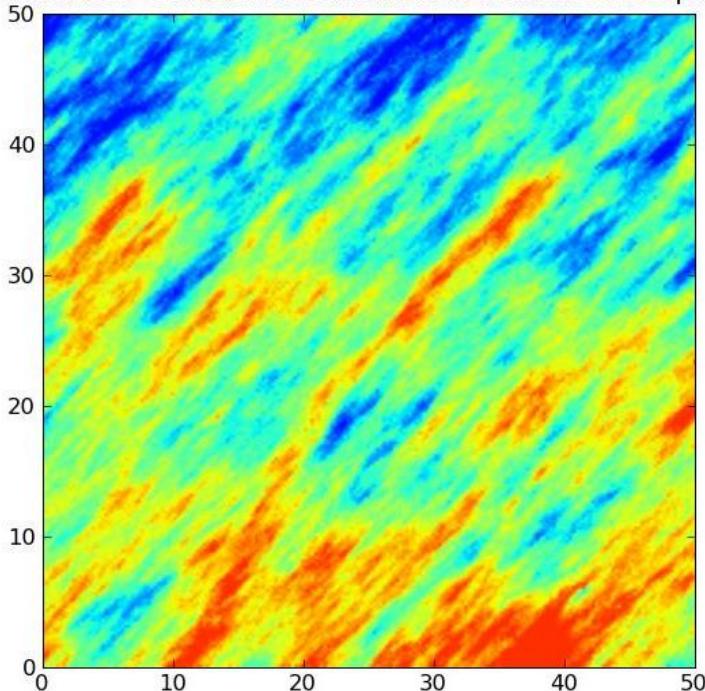
Language:

Python

Link:

[https://github.com/hpgl/
hpgl](https://github.com/hpgl/hpgl)

Simulated conditional realization in normal-score space



gstat

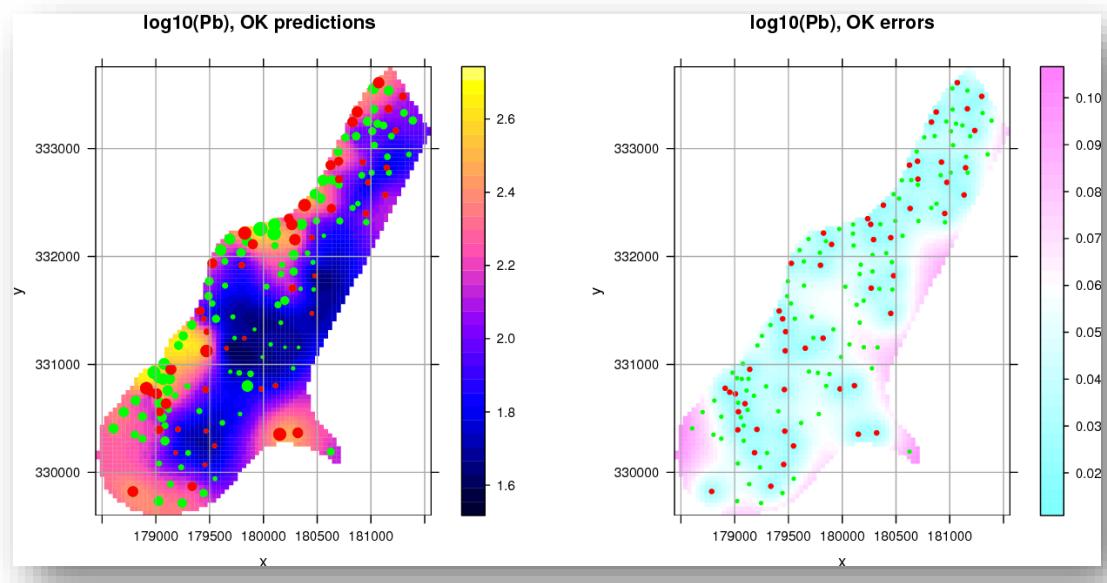
Spatial and spatio-temporal geostatistical modelling, prediction and simulation

Language:

R & C

Link:

<https://github.com/r-spatial/gstat/>



PyGSLIB

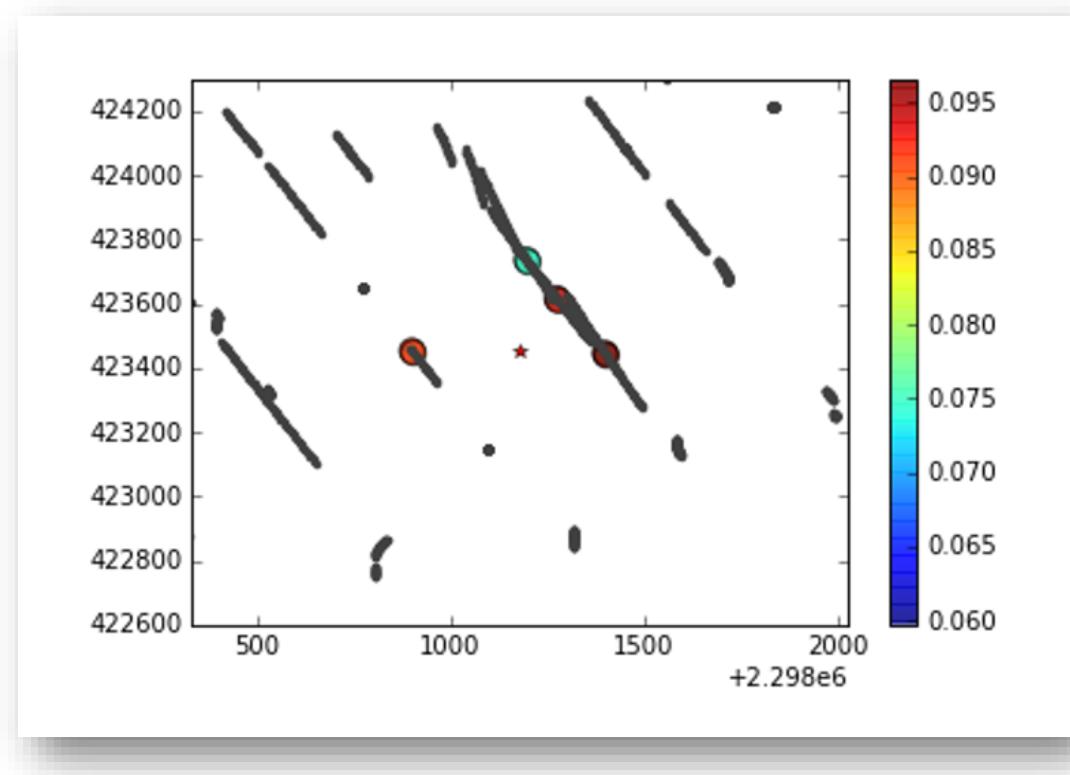
Mineral resource
estimations

Language:

Python

Link:

<https://opengeostat.git-hub.io/pygslib/>



GeoStats.jl



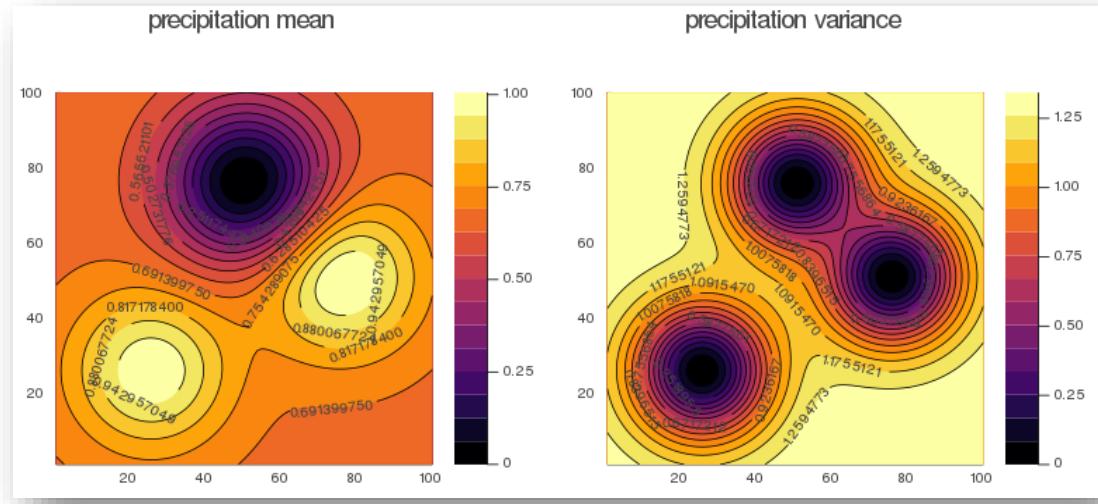
High-performance geostatistics

Language:

Julia

Link:

<https://github.com/JuliaEarth/GeoStats.jl>



GeostatsPy

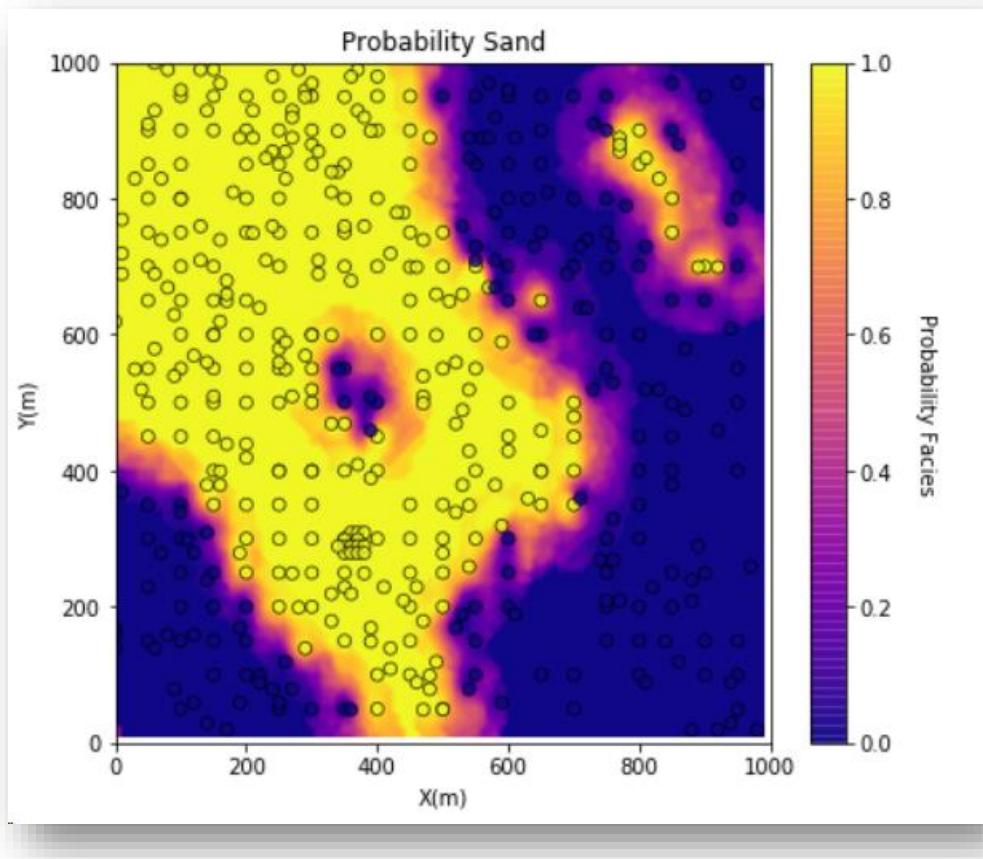
Brings Geostatistical Library (GSLIB) functions to Python.
GSLIB is a set of code for building spatial modeling workflows

Language:

Python

Link:

<https://github.com/GeostatsGuy/GeostatsPy>



GeoStat-Framework

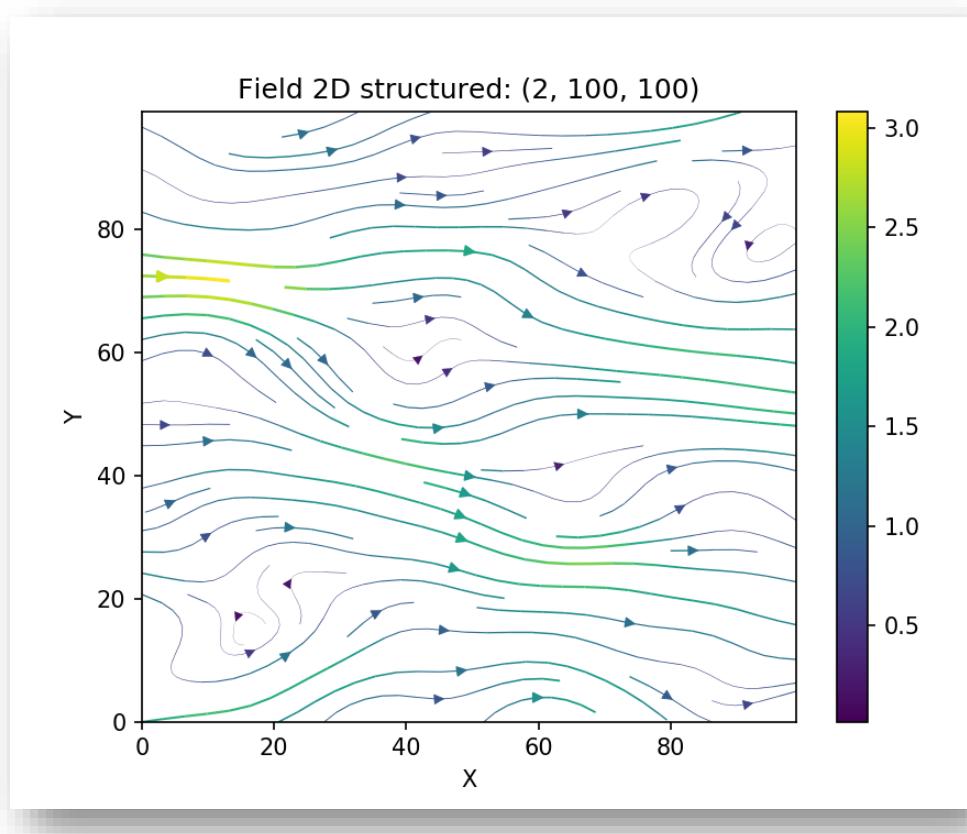
Framework for geostatistical simulations

Language:

Python

Link:

<https://github.com/GeoStat-Framework>



Geospatial



Generic Mapping Tools

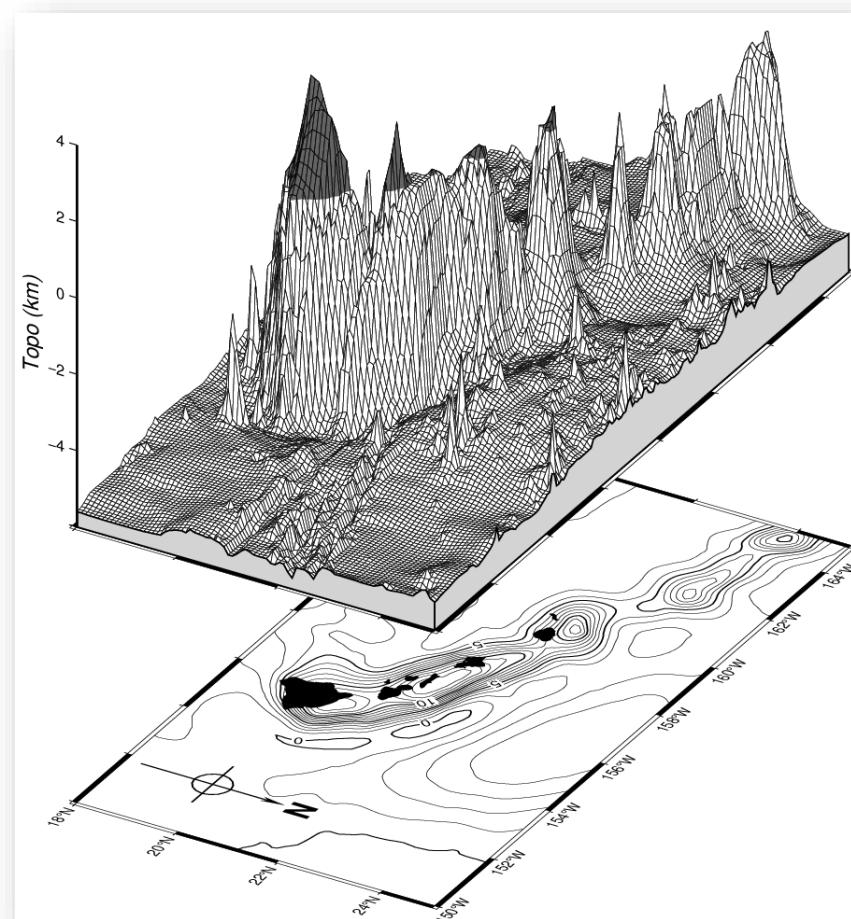
Command-line tools for manipulating geographic and Cartesian data sets

Language:

C, MATLAB, Julia, Python

Link:

<https://www.generic-mapping-tools.org>



geonotebook

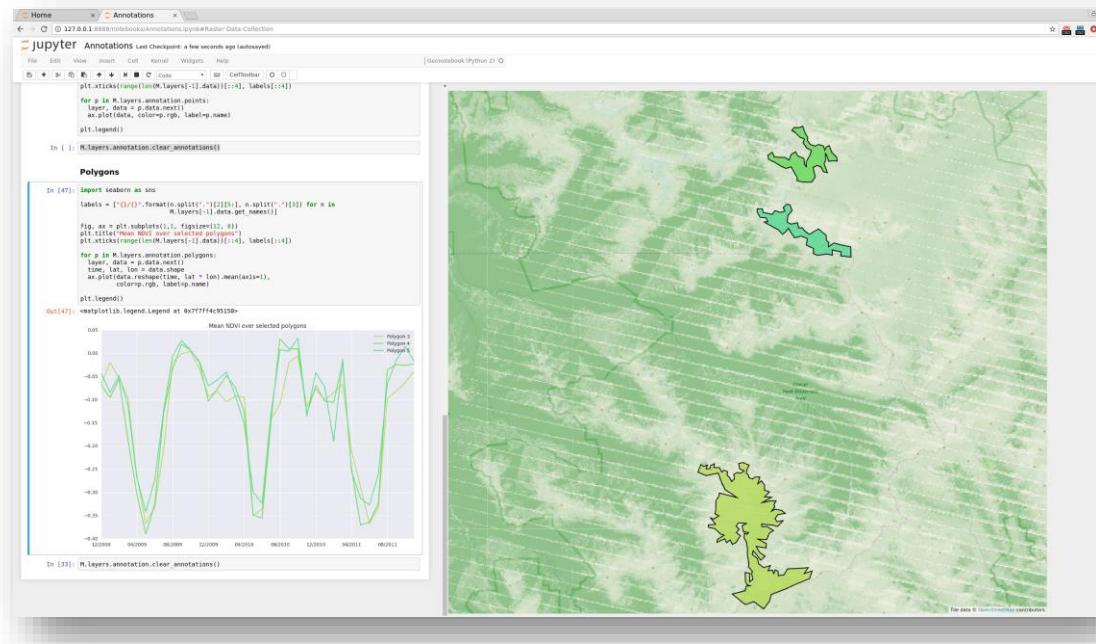
Geospatial visualization
and analysis developed
by NASA

Language:

Python

Link:

<https://github.com/OpenGeoscience/geonotebook>



Stress2Grid

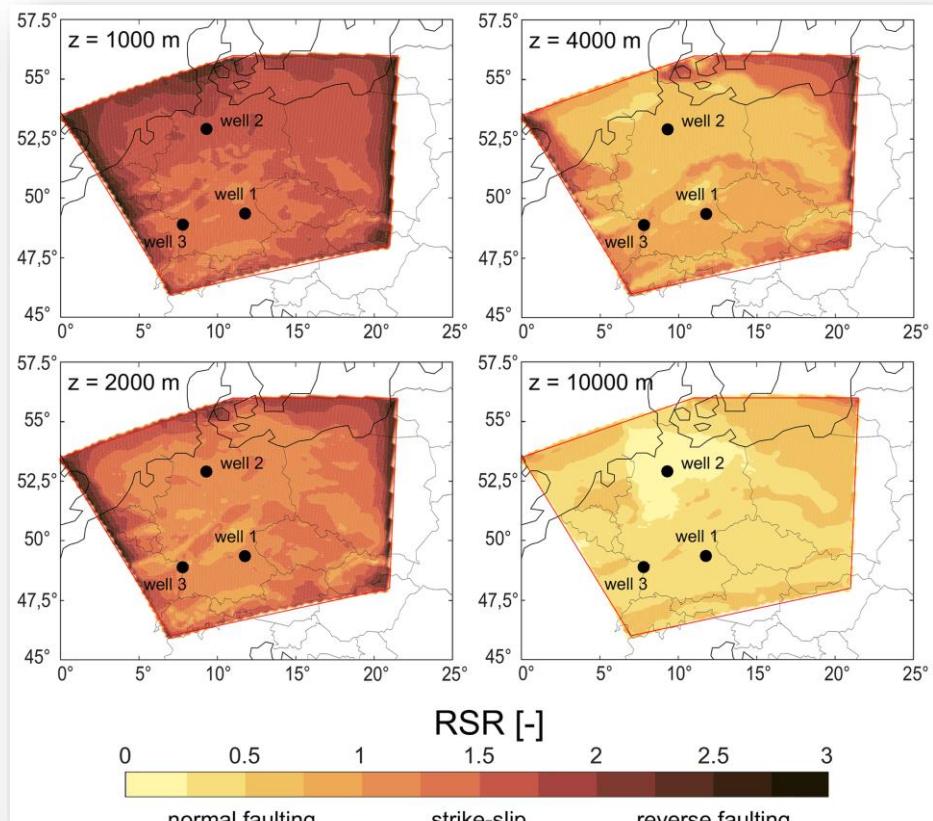
Stress orientation data analysis tool. Two concepts to calculate the mean SHmax orientation

Language:

MATLAB

Link:

<https://github.com/MorZieg/Stress2Grid>



Verde

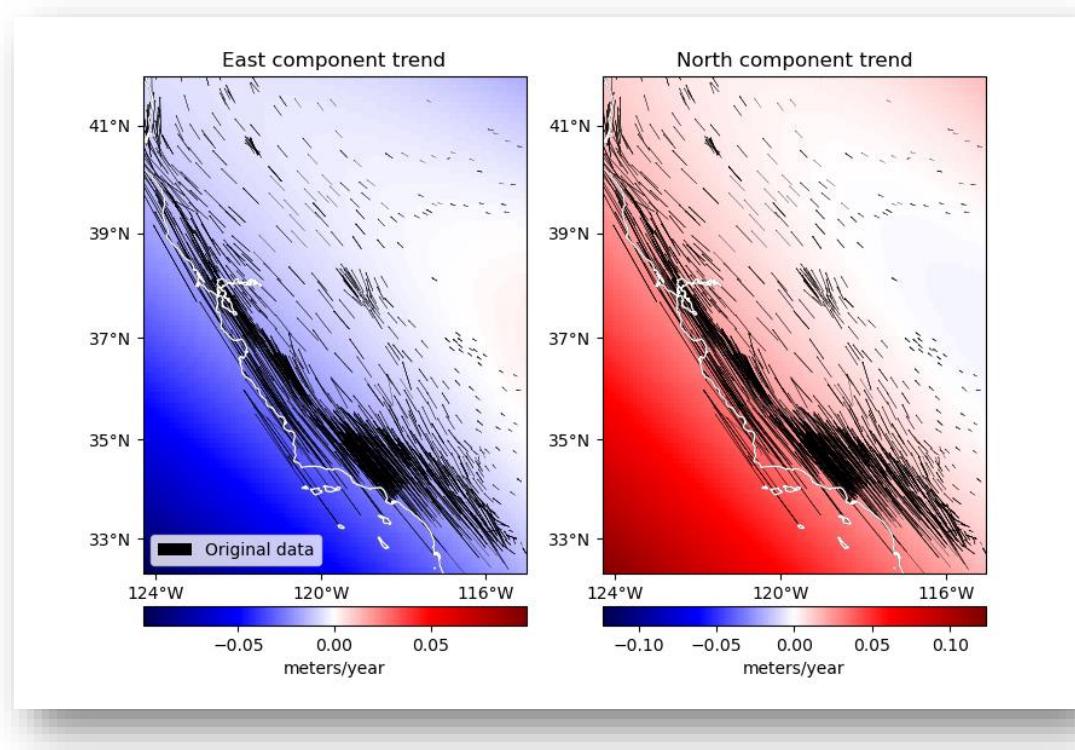
Processing spatial data
to regular grids

Language:

Python

Link:

<https://github.com/fatiando/verde>



GeoPHP

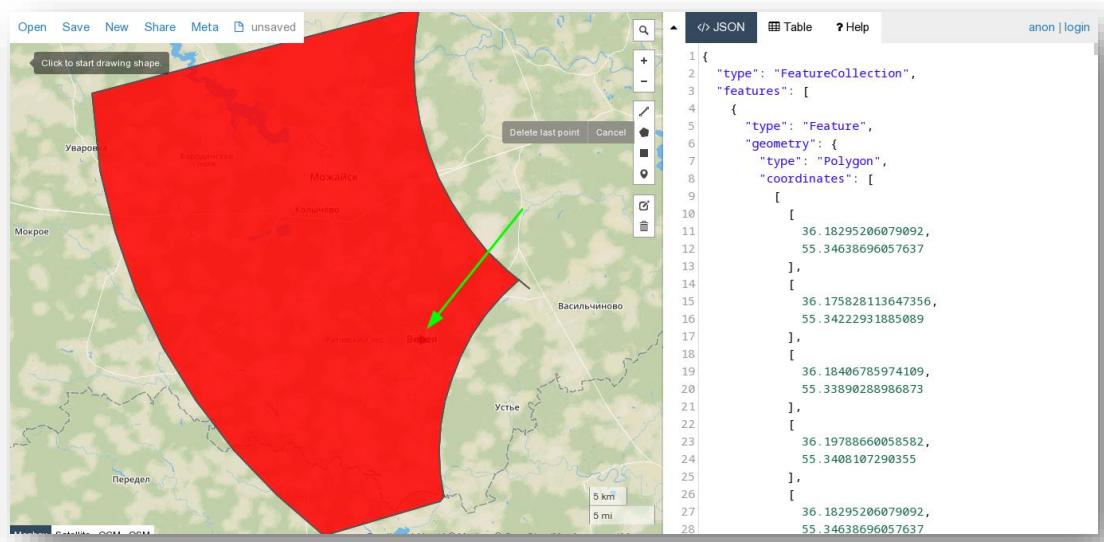
Geospatial library that works with many formats

Language:

PHP

Link:

<https://geophp.net>



vtk_triangulate_points

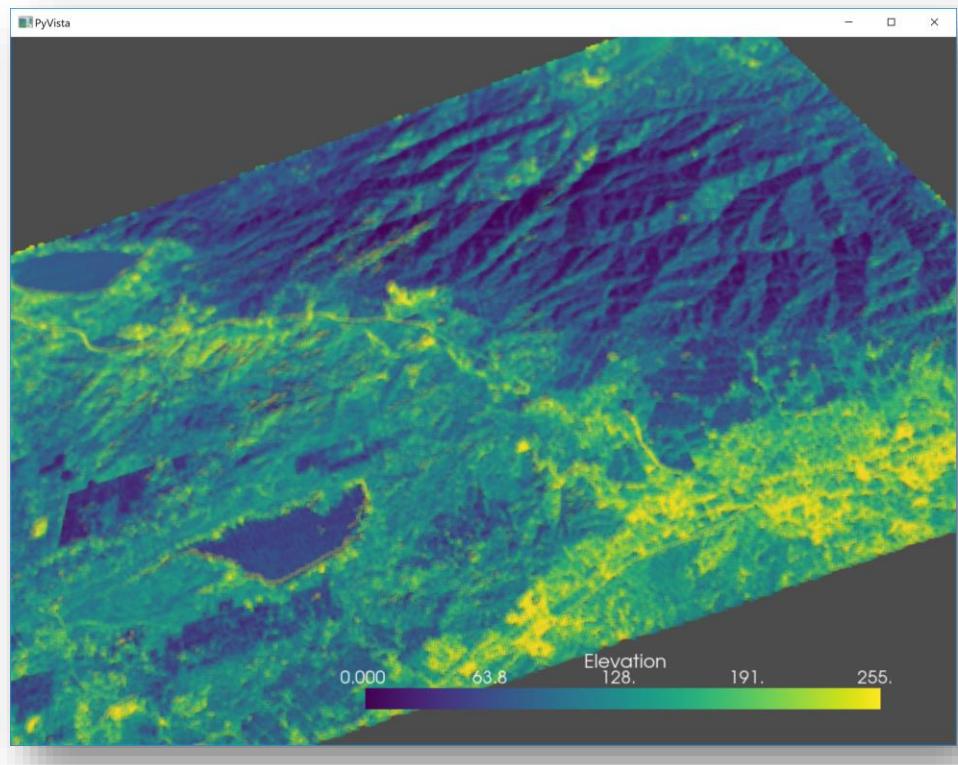
Generate simplified topography surfaces from dense point clouds (lidar/radar/drone)

Language:

Python

Link:

https://github.com/pemn/vtk_triangulate_points



Geochemistry



PhreeQC

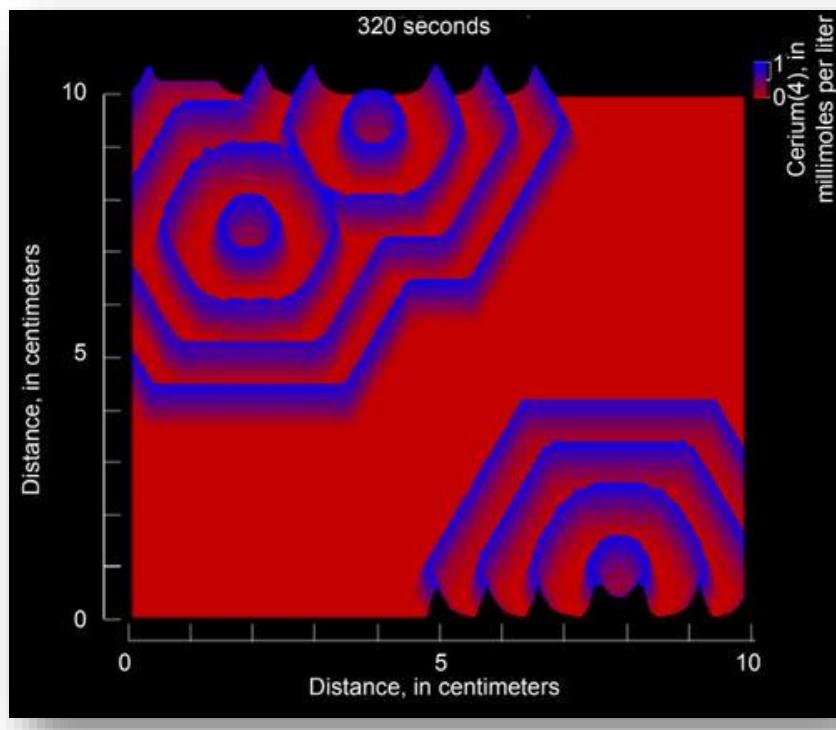
Reactions in water and between water and rocks and sediments (speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations)

Language:

C++

Link:

<https://www.usgs.gov/software/phreeqc-version-3>



Reaktoro

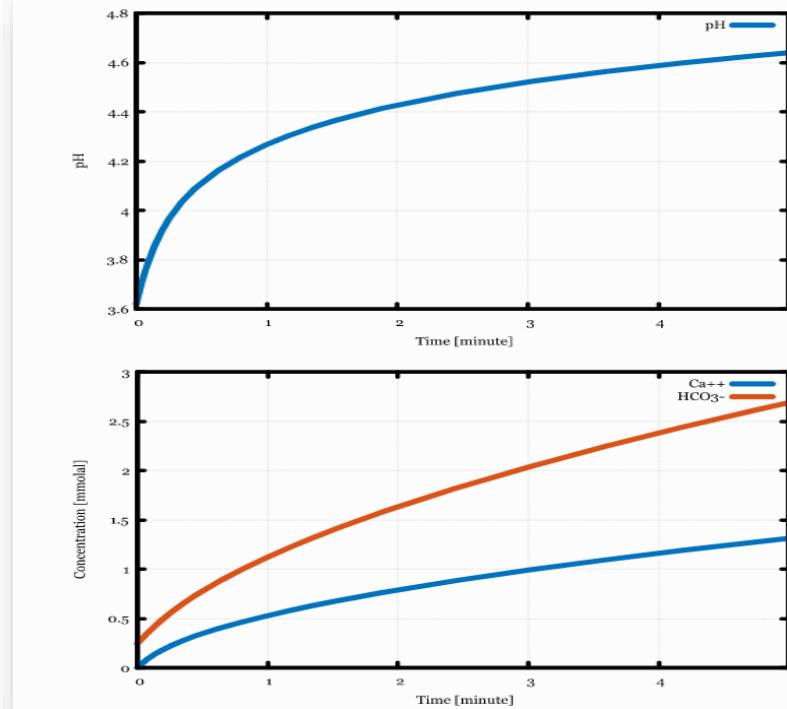
Unified framework for
modelling chemically
reactive systems

Language:

Python & C++

Link:

<https://reaktoro.org>



GeoPyTool

Application with
geochemical plotting
capabilities

Language:

Python

Link:

<https://github.com/GeoPyTool/GeoPyTool>

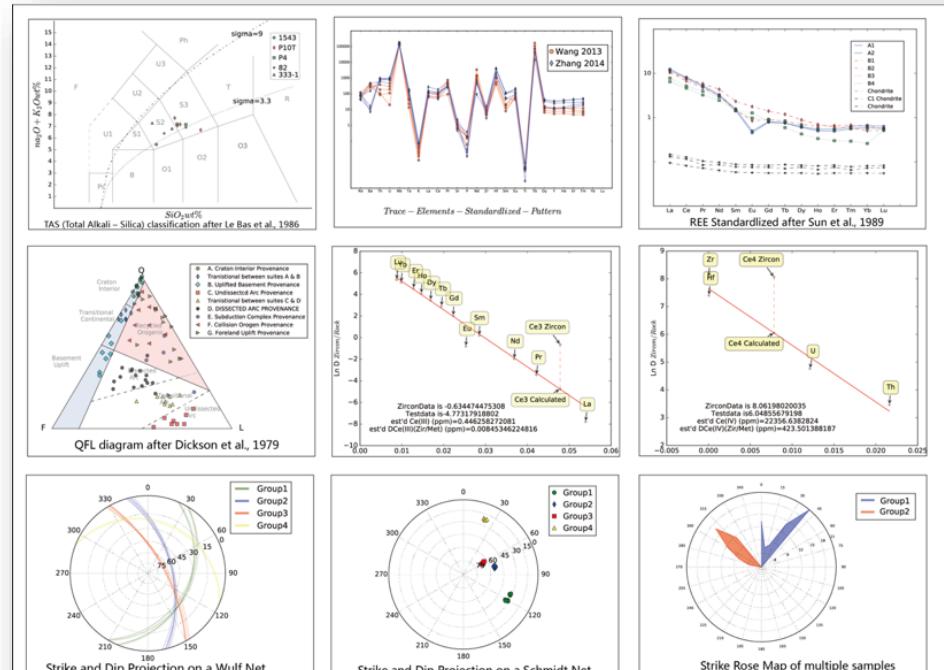


Figure 1. Diagrams generated by GeoPyTool

pyrolite

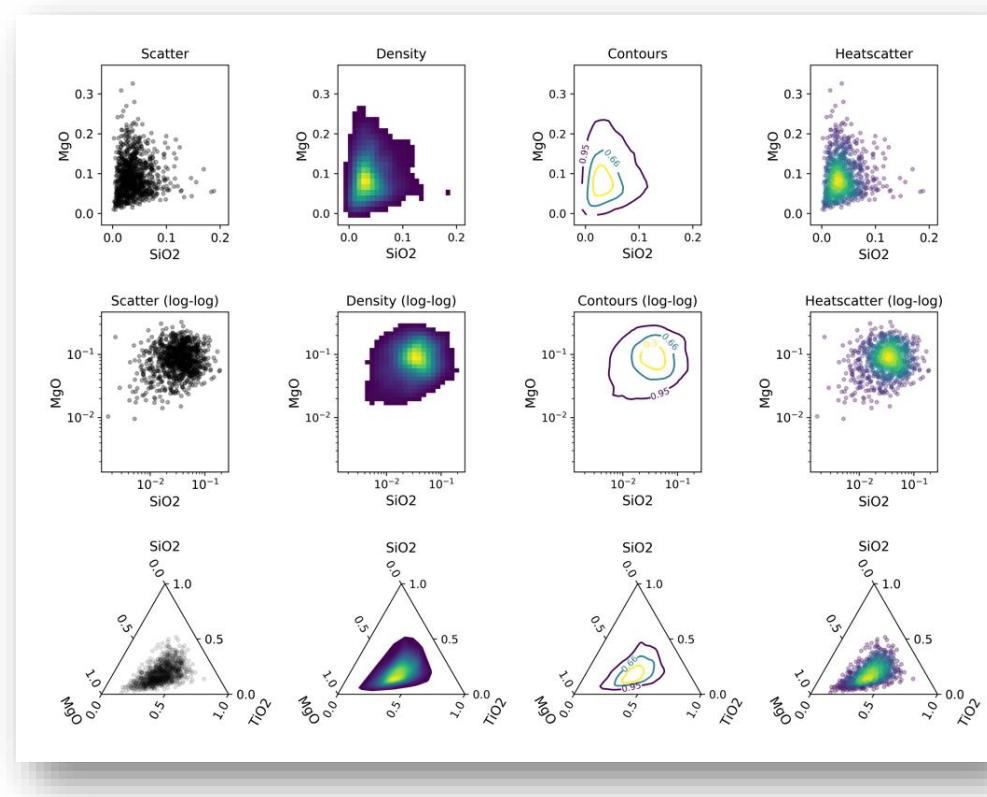
Geochemical
transformation and
visualization

Language:

Python

Link:

<https://github.com/morganjwilliams/pyrolite>



Geophysics



MTPy

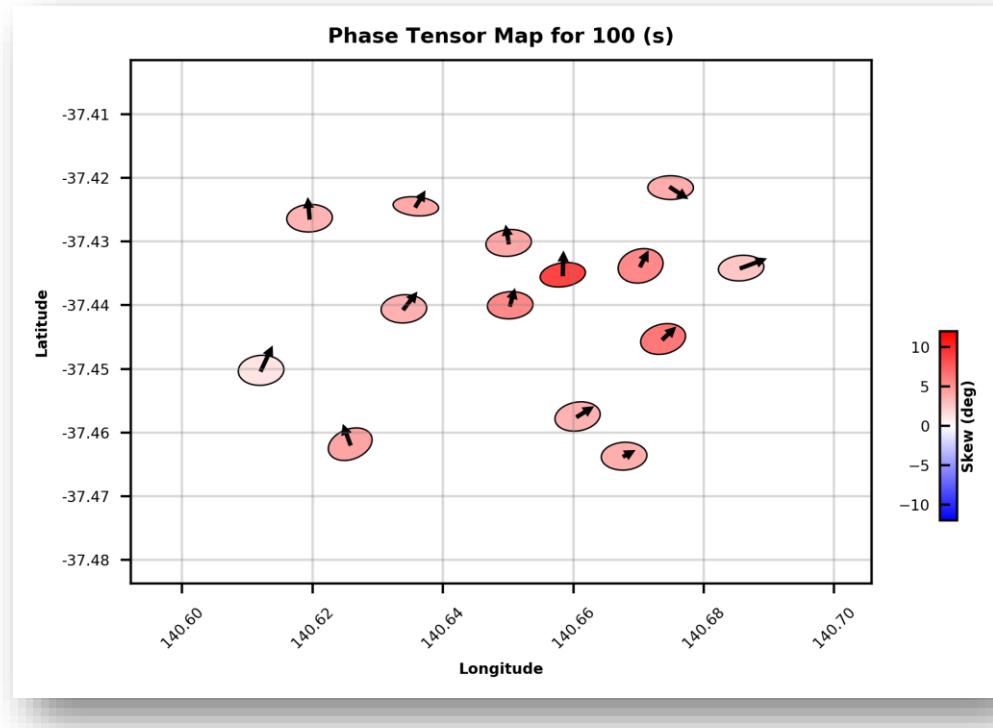
Toolbox for magnetotelluric
data processing, analysis,
modelling and visualization

Language:

Python

Link:

<https://github.com/MTgeophysics/mtpy>



MATE

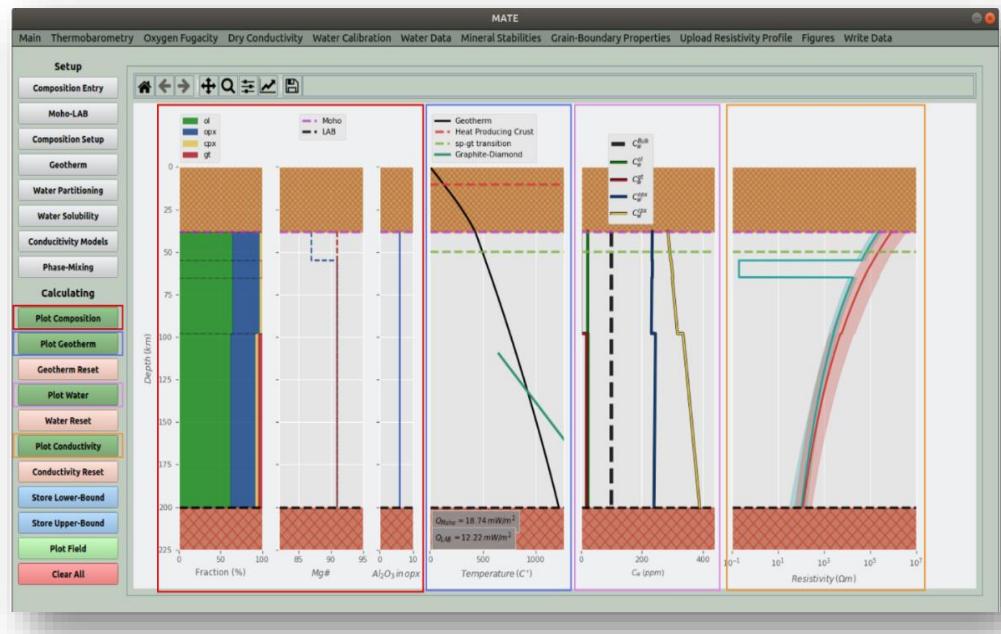
Program for Interpreting
magnetotelluric models of
the mantle

Language:

Python

Link:

<https://github.com/sinanozaydin/MATE>



Structural Geology



mplStereonet

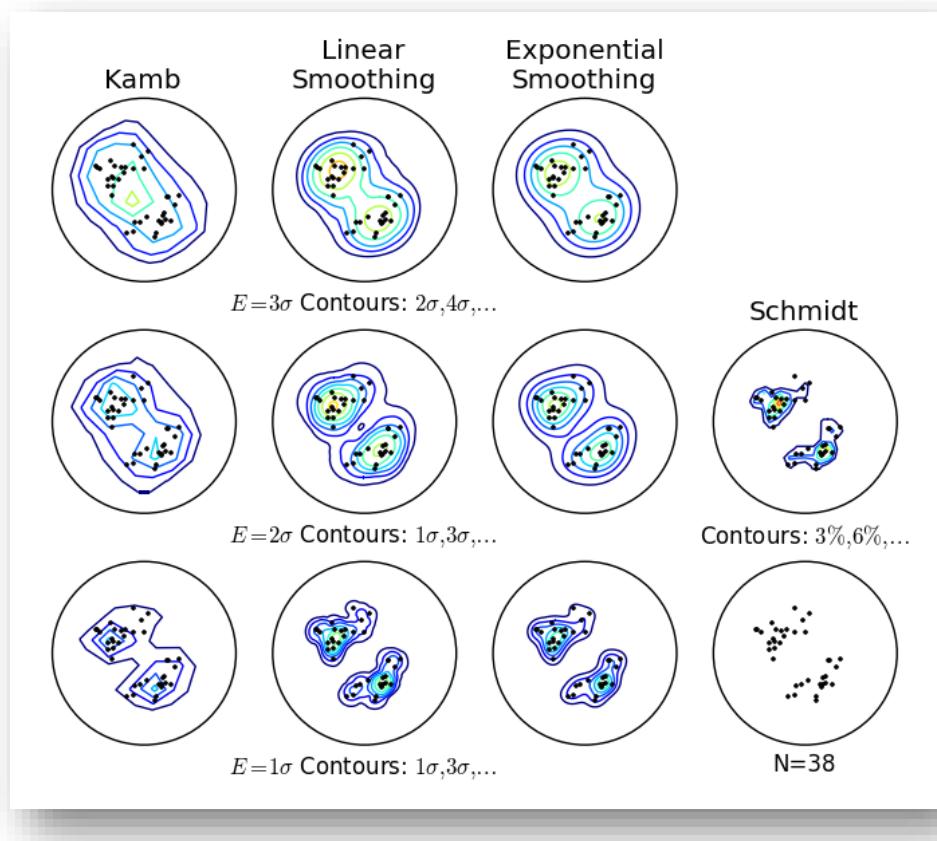
Provides lower-hemisphere equal-area and equal-angle stereonets for matplotlib

Language:

Python

Link:

<https://github.com/joferkington/mplstereonet>



apsg

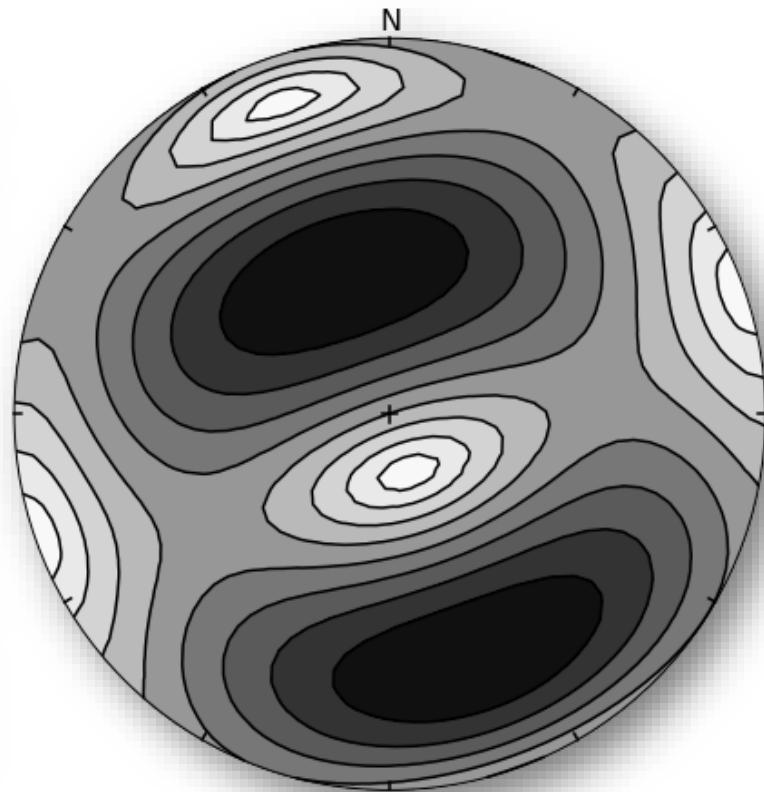
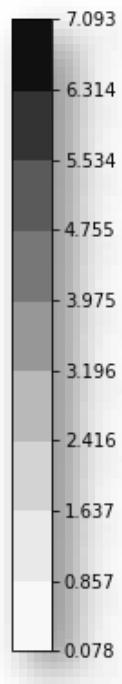
Advanced structural geology
analysis and visualisation
based on Matplotlib

Language:

Python

Link:

[https://github.com/ondrolexa/
apsg](https://github.com/ondrolexa/apsg)



Visualization



cmocean

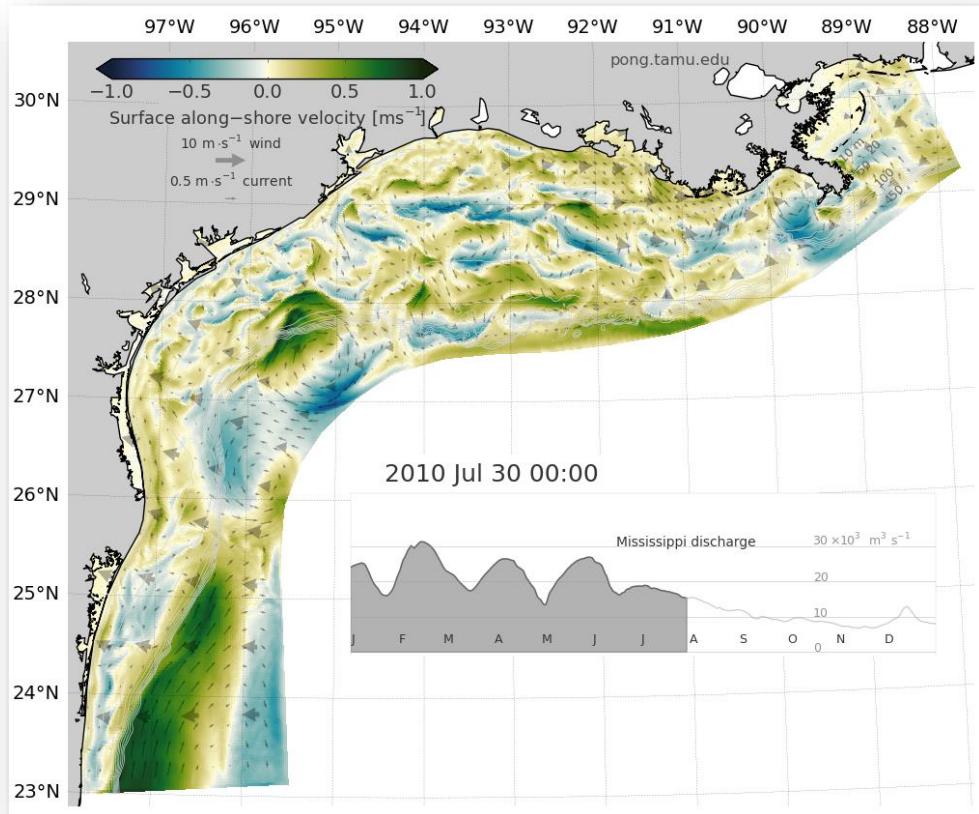
Collection of colormaps
for commonly-used
oceanographic variables

Language:

Python

Link:

<https://matplotlib.org/cmocean>



PyVista

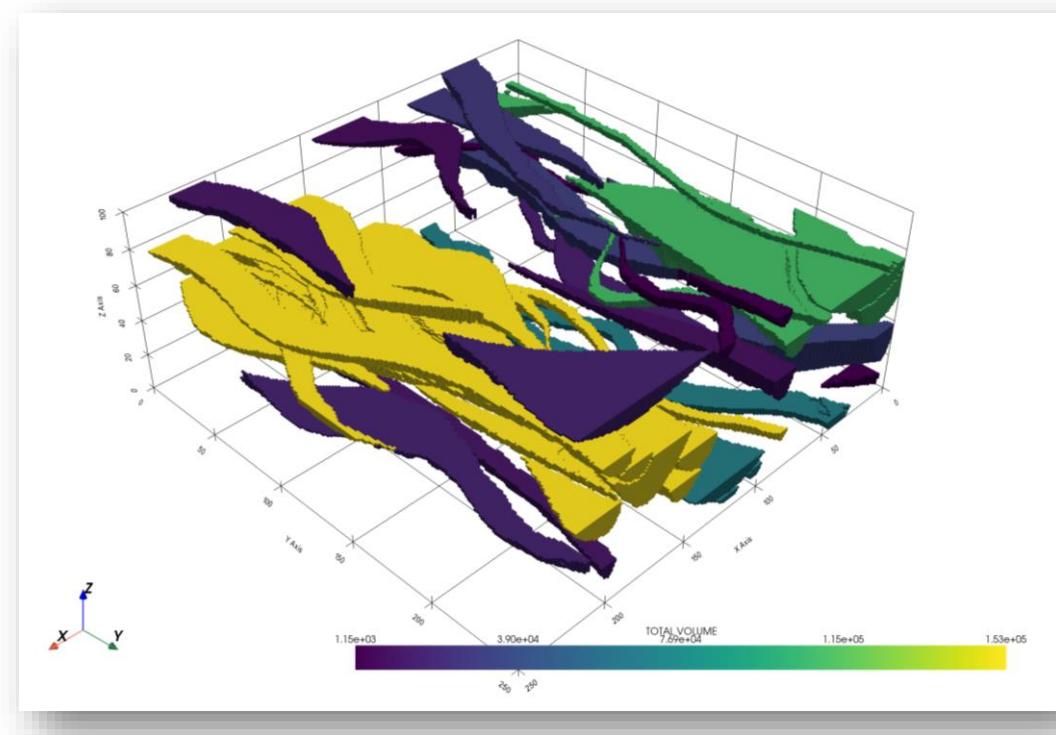
3D plotting and mesh analysis through a streamlined interface for the Visualization Toolkit (VTK)

Language:

Python

Link:

<https://github.com/pyvista/pyvista>



PVGeo

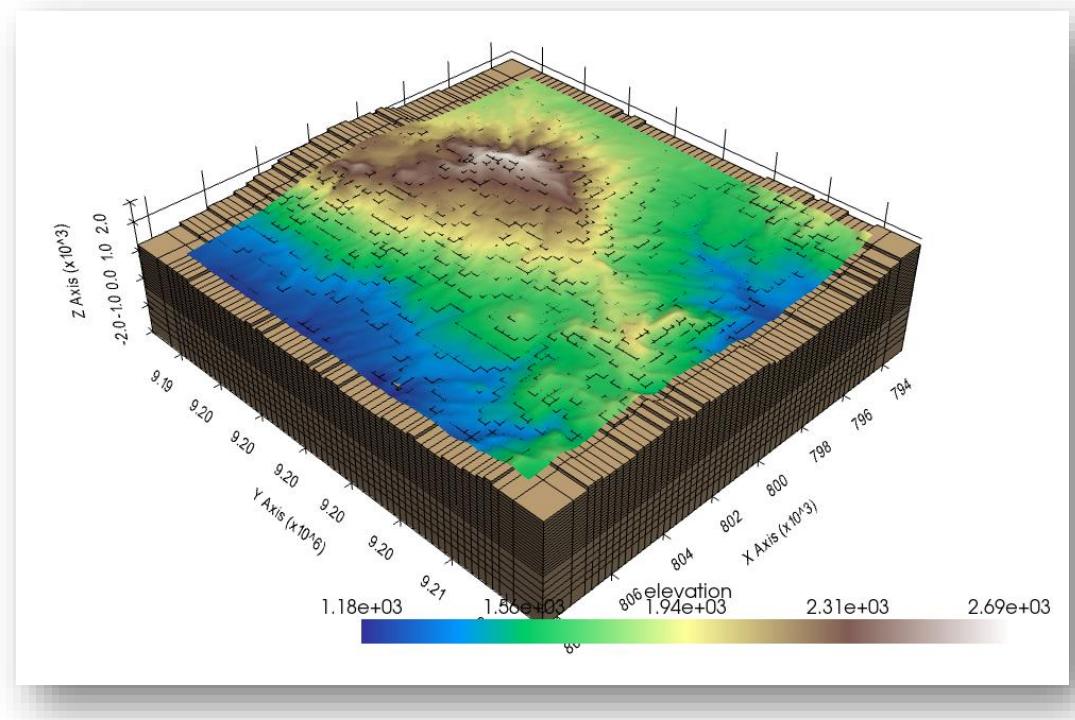
Data and model
visualization in ParaView
and Visualization Toolkit
(VTK)

Language:

Python

Link:

<https://github.com/OpenGeoVis/PVGeo>



omfvista

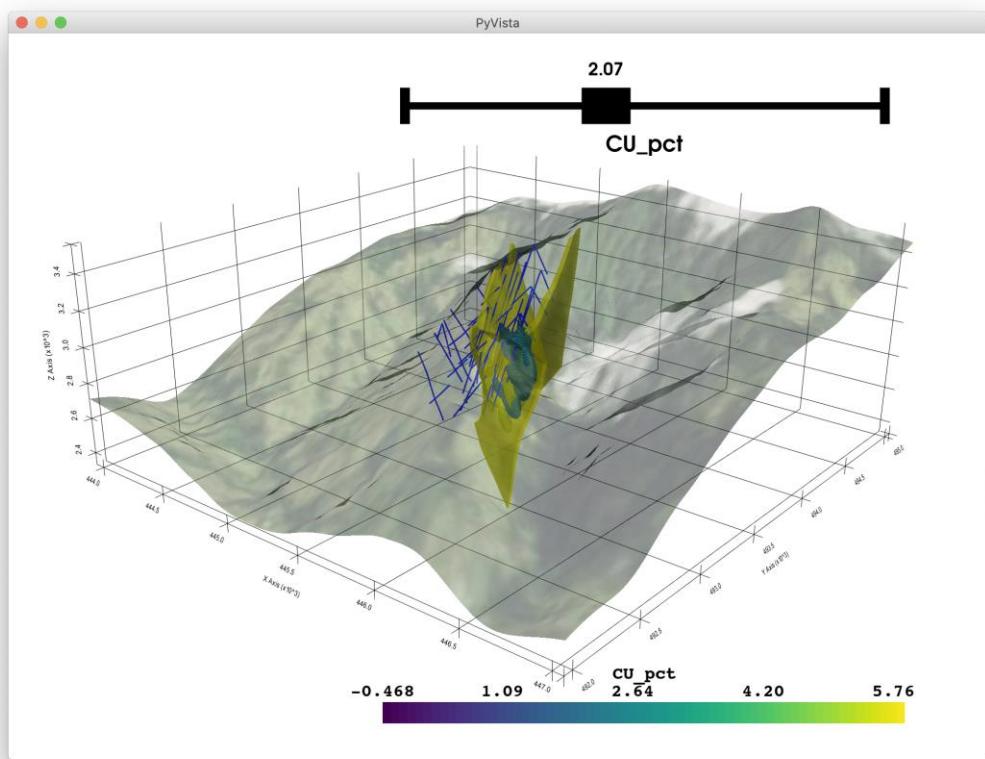
PyVista interface for the Open Mining Format package (omf) providing 3D visualization and useable mesh data structures for processing datasets

Language:

Python

Link:

<https://github.com/OpenGeoVis/omfvista>



ipyleaflet

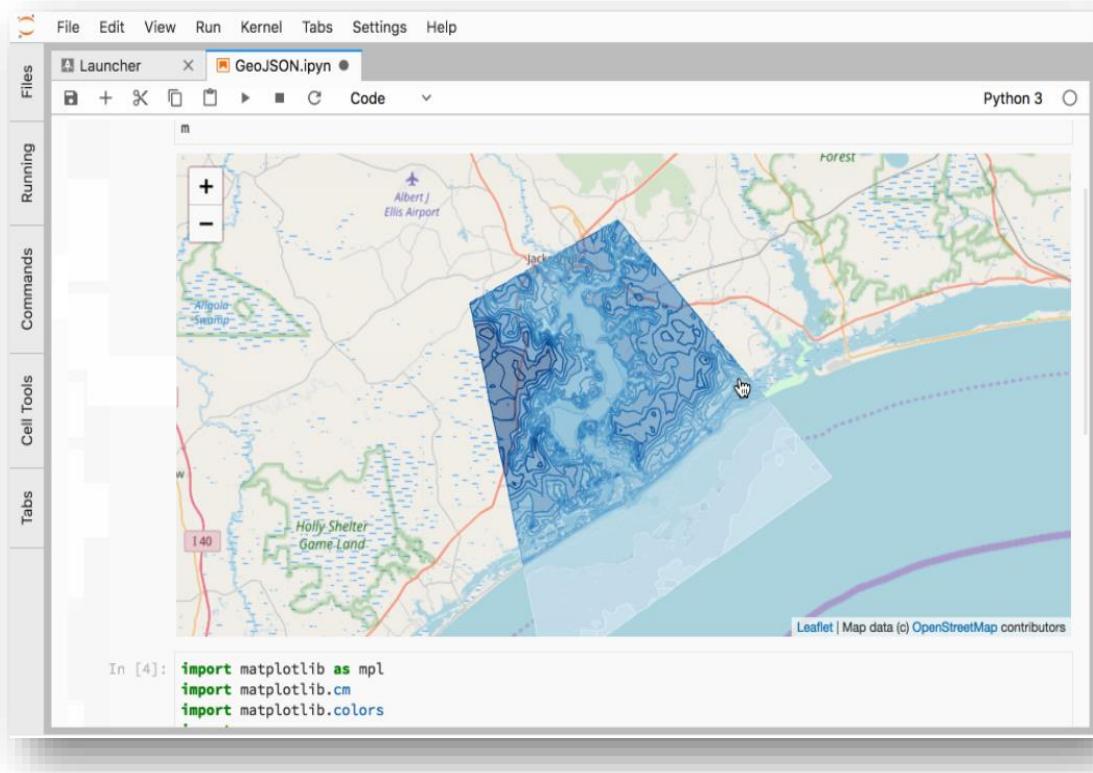
2D interactive maps and
GIS visualization in the
Jupyter Notebook

Language:

Python

Link:

<https://github.com/jupyter-widgets/ipyleaflet>



Colorcet



Collection of perceptually uniform colormaps for use with Python plotting programs

Language:

Python

Link:

<https://github.com/holoviz/colorcet>



Geologic Patterns

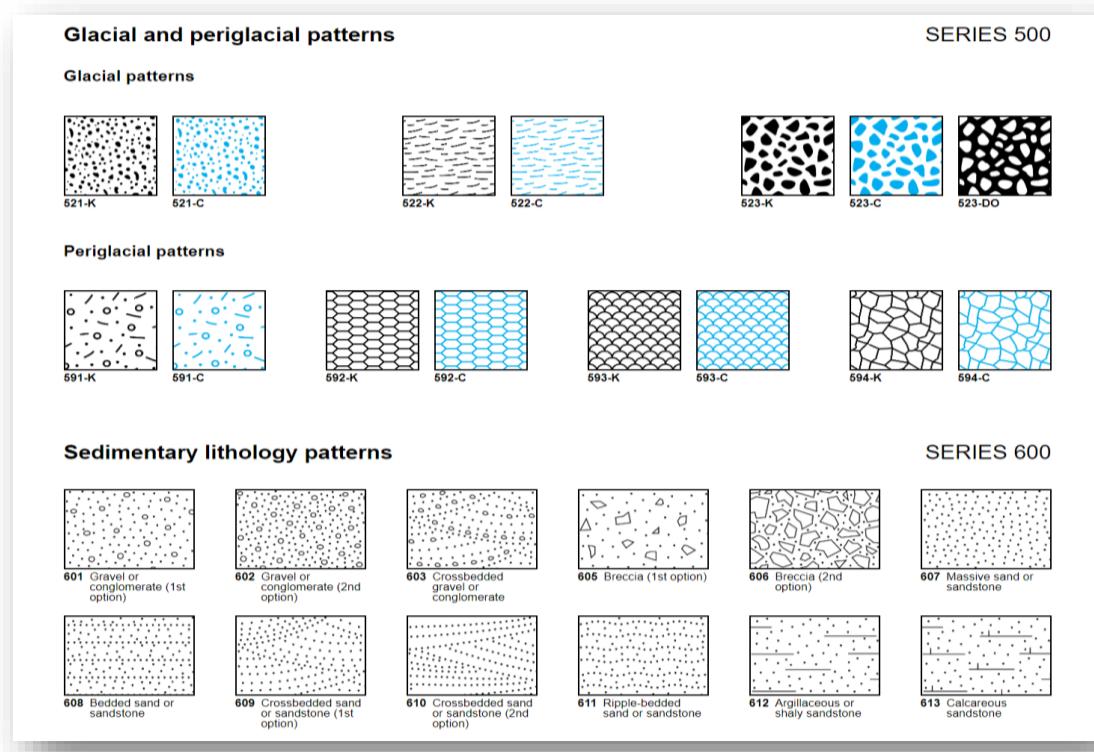
Entire FGDC pattern library extracted to SVG and PNG for use in geologic maps and stratigraphic columns

Language:

Python & Javascript

Link:

<https://davenquinn.com/projects/geologic-patterns>



Platforms



QGIS

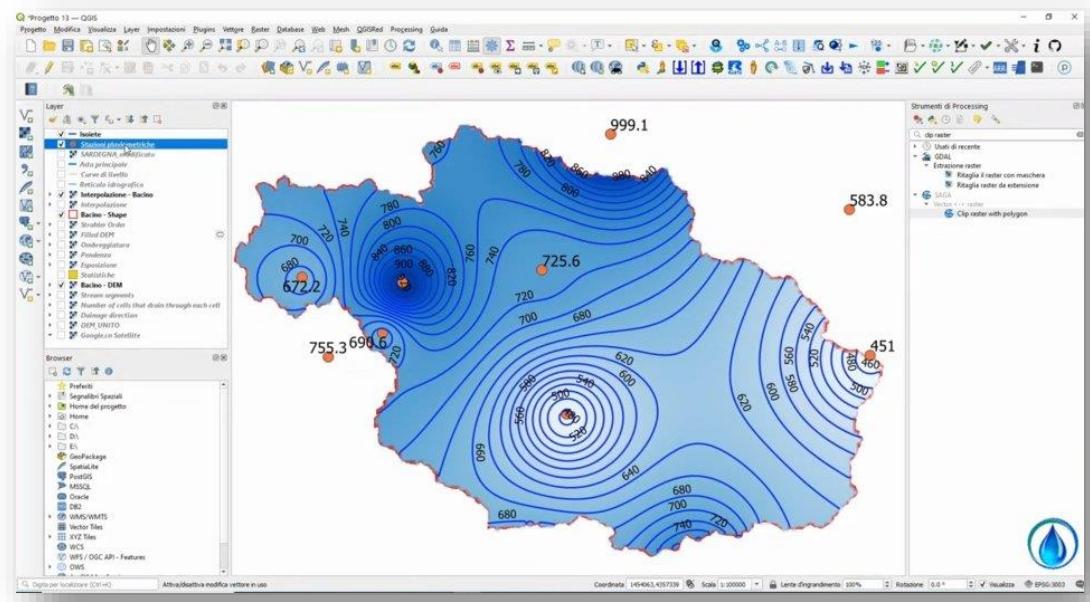
GIS platform to visualize,
manage, edit, analyze
data, and compose
printable maps

Language:

C++ & Python

Link:

<https://qgis.org/en/site/>



OpenTect

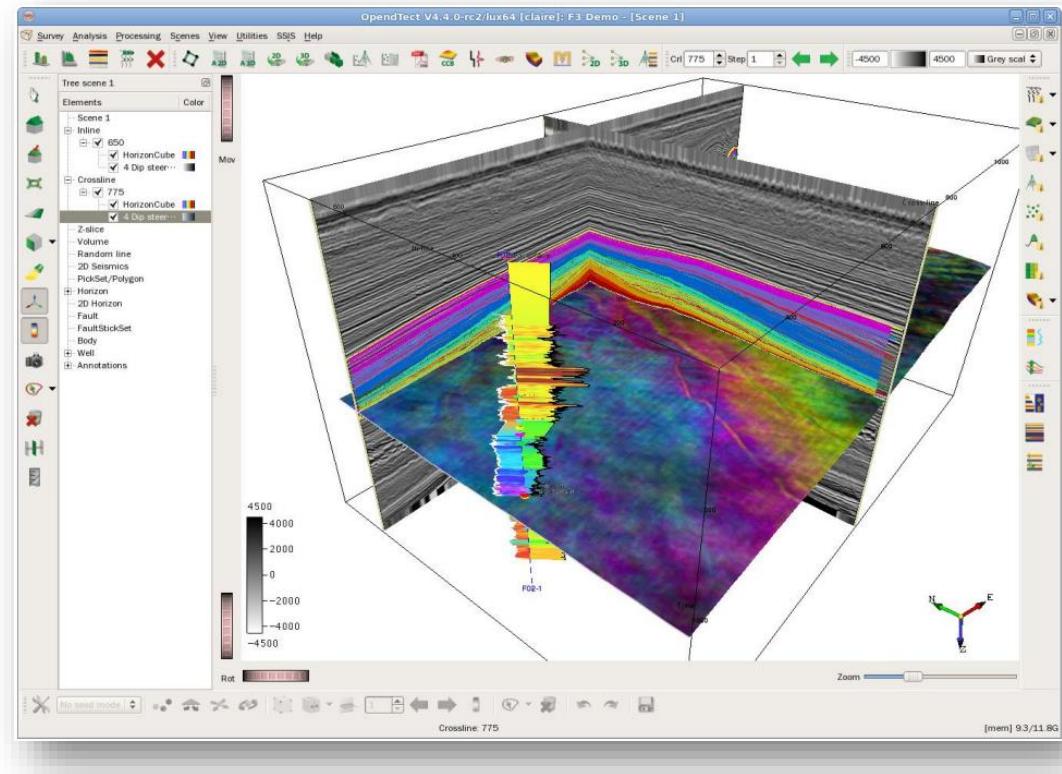
Seismic interpretation software system for 2D and 3D seismic data, and for fast-track development of innovative interpretation tools

Language:

C++

Link:

<https://dgbes.com/index.php/software#free>



Pangeo

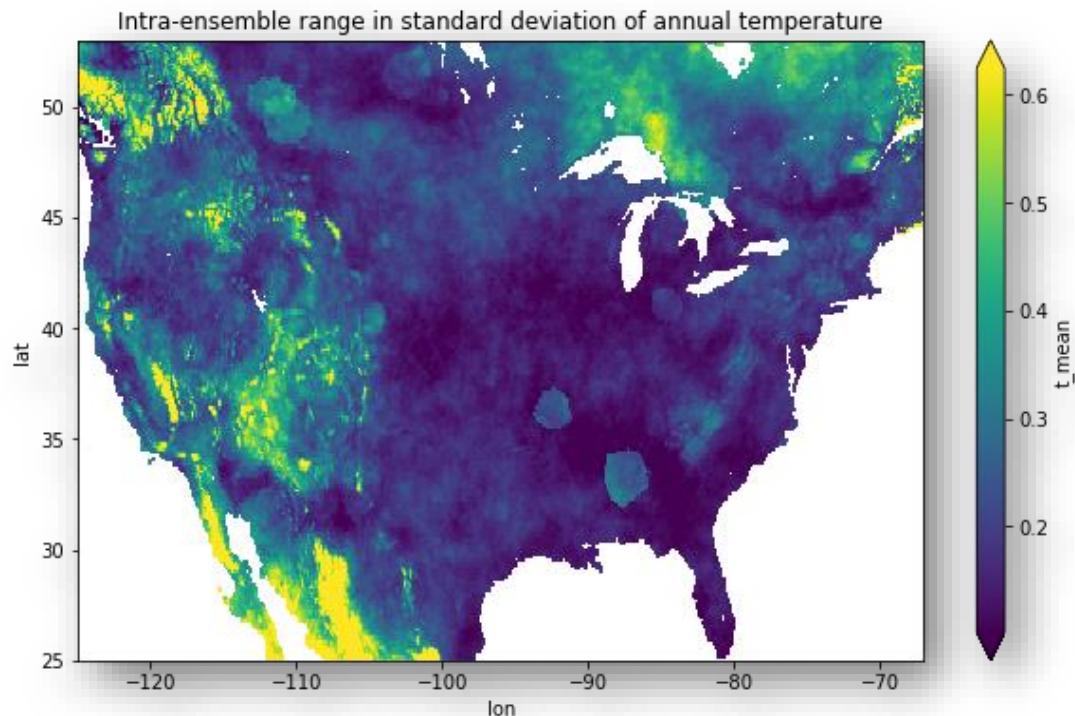
A community platform
for Big Data geoscience
built on top of the open
source scientific python
ecosystem

Language:

Python

Link:

<https://pangeo.io>



OpenGeode

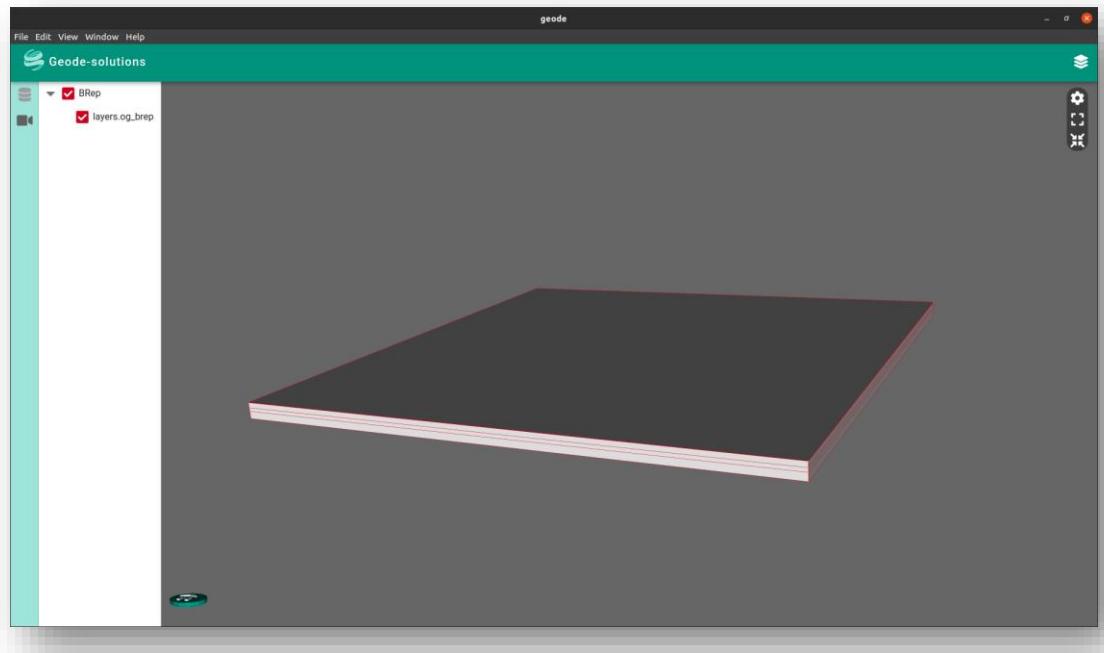
Open source framework for representing and manipulating geometric models, from simple meshes to complex boundary representations

Language:

Python & C++

Link:

<https://github.com/Geode-solutions/OpenGeode>



Data Repositories



- [Poseidon NW Australia](#) – Interpreted 3D seismic (32bit) including reports and well logs.
- [World Stress Map](#) – A global compilation of information on the crustal present-day stress field.
- [NOPIMS](#) – Open petroleum geoscience data from Western Australia made available by the Australian Government.
- [UK National Data Repository](#) – Open petroleum geoscience data from the UK Government (free registration required).
- [Athabasca Oil Sands Well Dataset McMurray/Wabiskaw](#) – Well logs and stratigraphic picks for 2193 wells, including 750 with lithofacies, from Alberta, Canada.
- [ICGEM](#) – Hosts gravity field spherical harmonic models and provides a webservice for generating grids of gravity functionals (geoid, gravity anomaly, vertical derivatives, etc).
- [TerraNubis](#) – The new *Open Seismic Repository*, includes the classic F3 and Penobscot seismic volumes (which both also have wells and other data assets).
- [Quantarctica](#) – User-configurable [QGIS](#) basemap for Antarctica with high-quality, peer-reviewed, free and open Antarctic scientific data.
- [Digital Rocks Portal](#) – Powerful data portal for images of varied porous micro-structures.
- [GSQ Open Data Portal](#) – Petroleum, coal, and mineral geoscience data from the Queensland resource industry and government, with supporting information from [GSQ GitHub Repository](#) for Data Models, RDF Vocabularies, and system design.
- [Geoscience Australia Portal](#) – Comprehensive map-based Australian data portal across multiple geoscience domains.
- [SARIG](#) – South Australian Resources and Information Gateway providing map-based statewide geoscientific and geospatial data with over 600 datasets.
- [SEG Open Data Catalog](#) – Catalog of "geophysical data that is readily available for download from the internet, via mail, or through special request", maintained by the Society of Exploration Geophysicists.
- [Volve data village](#) – A complete set of data from a North Sea oil field available for research, study and development purposes.

For more information about AI/ML & Data Science in the Upstream industry, please follow Mustafa Kaouri at LinkedIn
[@mustafakaouri](#)