

The role of research software at the Council for Geoscience

P. Cole



Council for Geoscience

Disclaimer

The information contained in these documents is the sole property of the Council for Geoscience and all rights are reserved. The information is intended solely for the recipient and is confidential. Any use of this information by any person, including the recipient, and any further dissemination or disclosure of the information, either in whole or in part, is prohibited except with the prior written consent of the Council for Geoscience. The opinions expressed are in good faith and while every care has been taken in preparing these documents, the Council for Geoscience makes no representations and gives no warranties of whatever nature in respect of these documents, including, but not limited to, the accuracy or completeness of any information, facts and/or opinions contained therein. The Council for Geoscience, its Board, employees and agents cannot be held liable for the use of and reliance on the opinions, estimates, forecasts and findings in these documents.

The Council for Geoscience

- The Council for Geoscience is the successor of the Geological Survey of South Africa
- The Geoscience Act (no. 100 of 1993) established the Council for Geoscience in its present form.



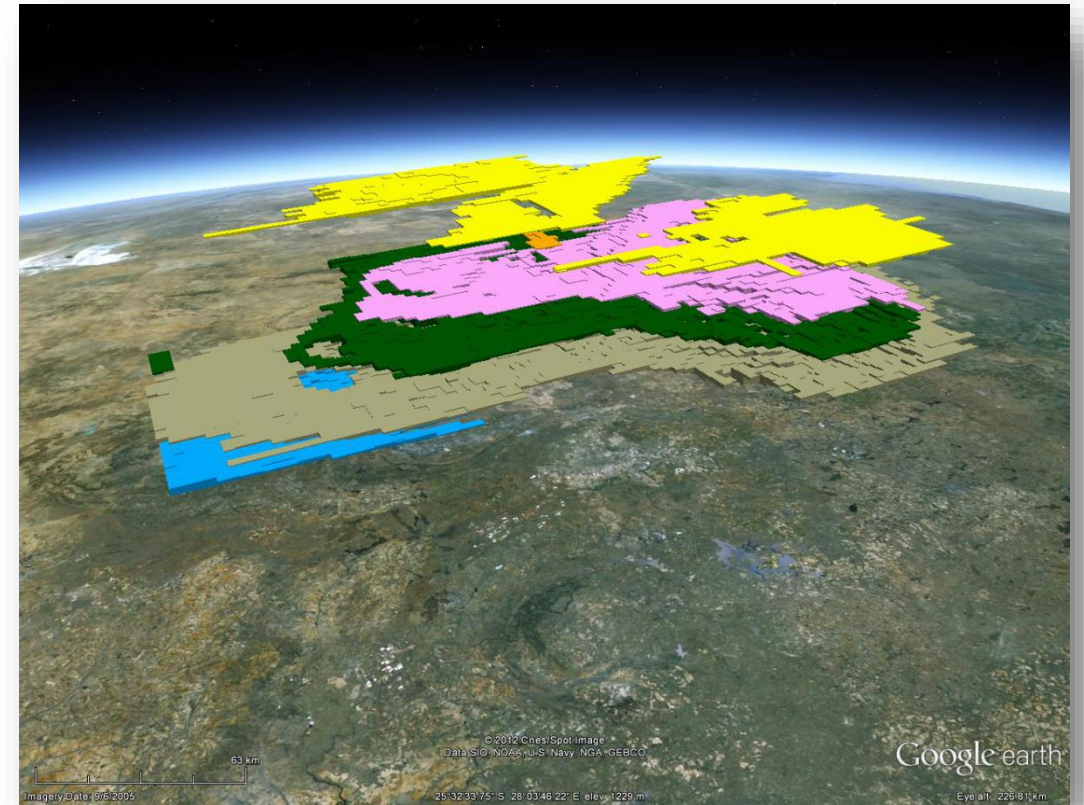
The Council for Geoscience

- Providing integrated, systematic and thematic maps and conducting research on the onshore and offshore geology of South Africa, as mandated.
- National advisory authority on geo-environmental pollution.
- Provides engineering geology studies – for example sinkholes.
- Manage the national seismic network.



Why Research Software?

- Vast quantities of data need a way to speed up interpretation.
- Includes:
 - Geology maps
 - Geochemical data
 - Magnetic and Gravity Data
 - Seismic Data
 - Laboratory analysis
 - Marine data
 - Remote sensing data
- Machine learning and AI present opportunities to do this.



Research Software Opportunities

- Geophysics
- Geochemistry
- Groundwater modelling
- Engineering Geology
- Seismology
- Geology
- Mineral systems for target identification
- Environmental studies



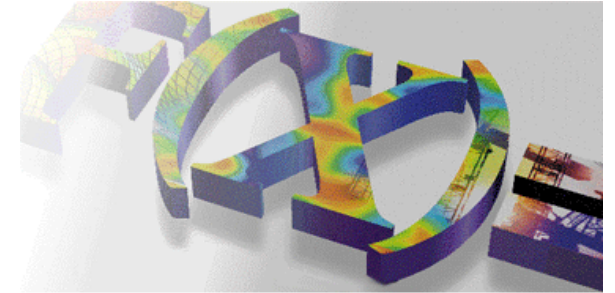
Geoscience Research Gaps

- Machine learning
 - Random forests
 - Neural Networks
- 3D Modelling
 - Geophysical
 - Geological
 - Groundwater
- Automation
 - Remote sensing



- Statutory projects
 - National priorities
- Academic studies
 - Fresh exchange of ideas

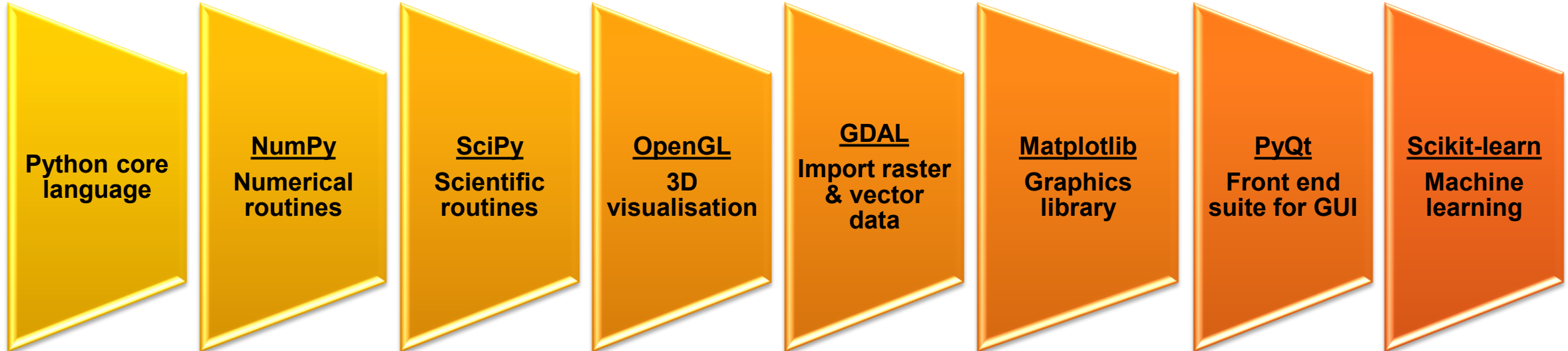
Programming Languages



Software development

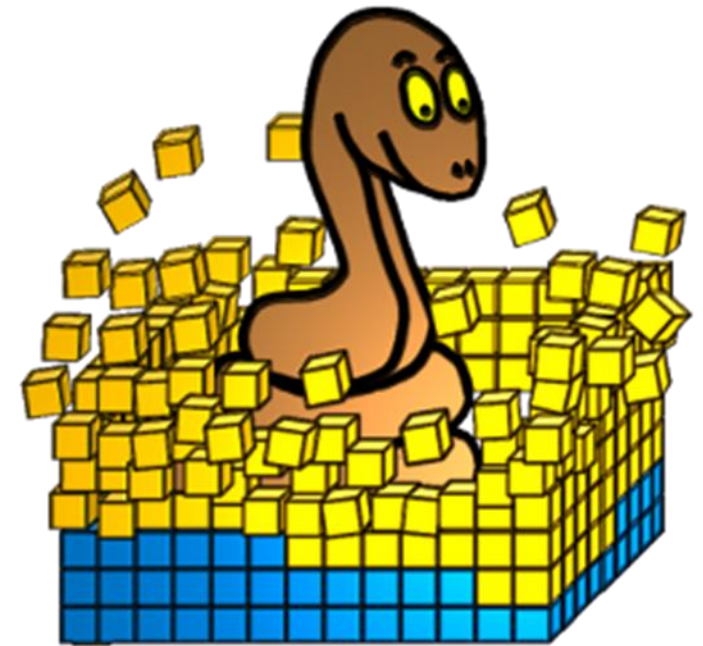


- Python is a popular language which allows easy implementation of research software.



CGS Research Software - PyGMI Project

- PyGMI stands for Python Geoscience Modelling and Interpretation. It is a 3D modelling and interpretation suite aimed at magnetic, gravity and other datasets.
- Cluster Analysis, including use of scikit-learn libraries.
- Seismological functions for SEISAN data.
- Remote sensing ratios and improved imports.
- It is a Qt based application.
- It is released under the Gnu General Public License version 3.0



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

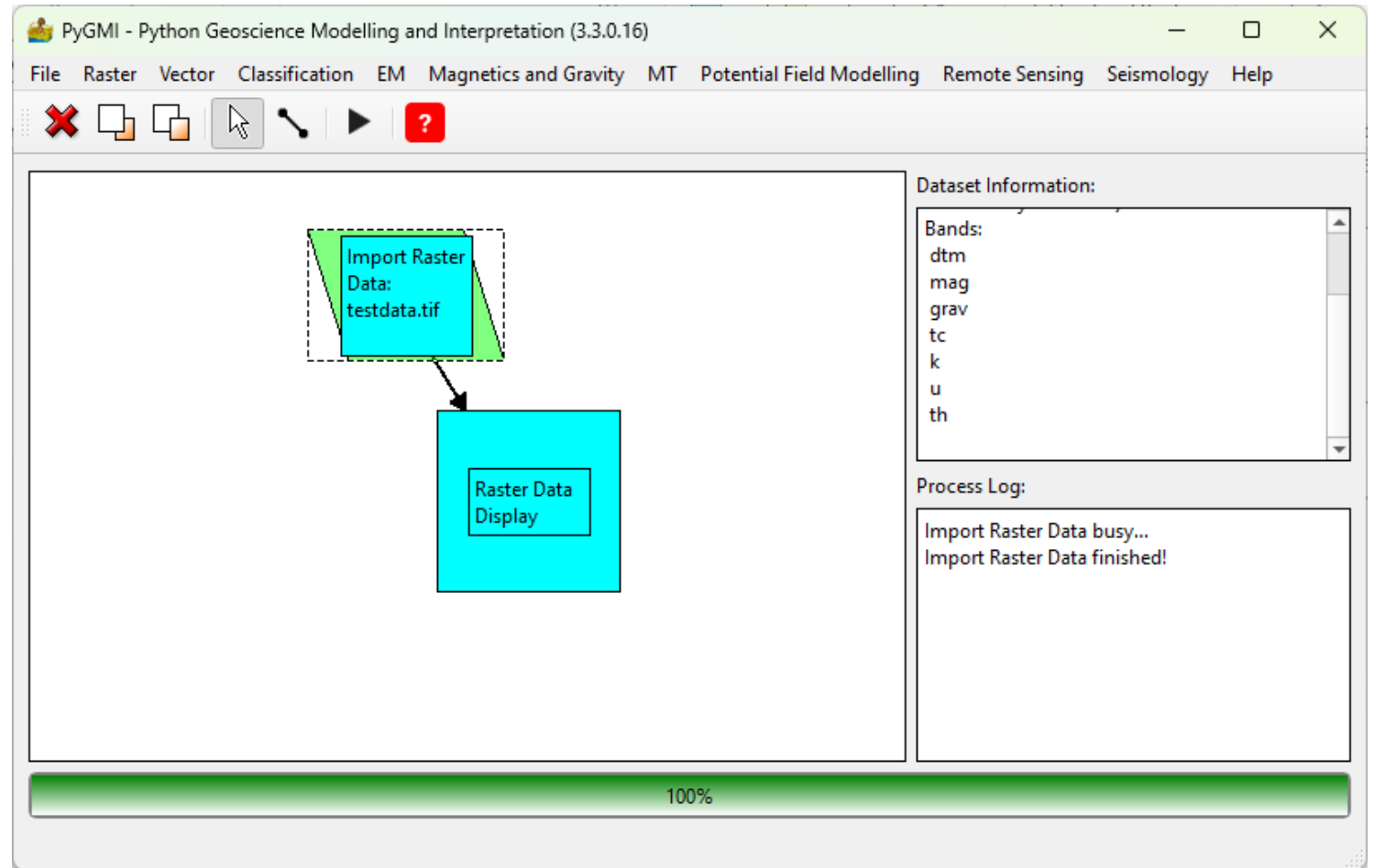
PyGMI Open Source Software

- Provides best opportunities for research software support.
- Gnu General Public License Version 3.
 - Allow free use while giving credit to developers.
- GitHub hosting of project.
 - Version tracking
 - Online documentation via Sphinx
- Windows installer
 - Inno Setup Compiler
 - WinPython distribution



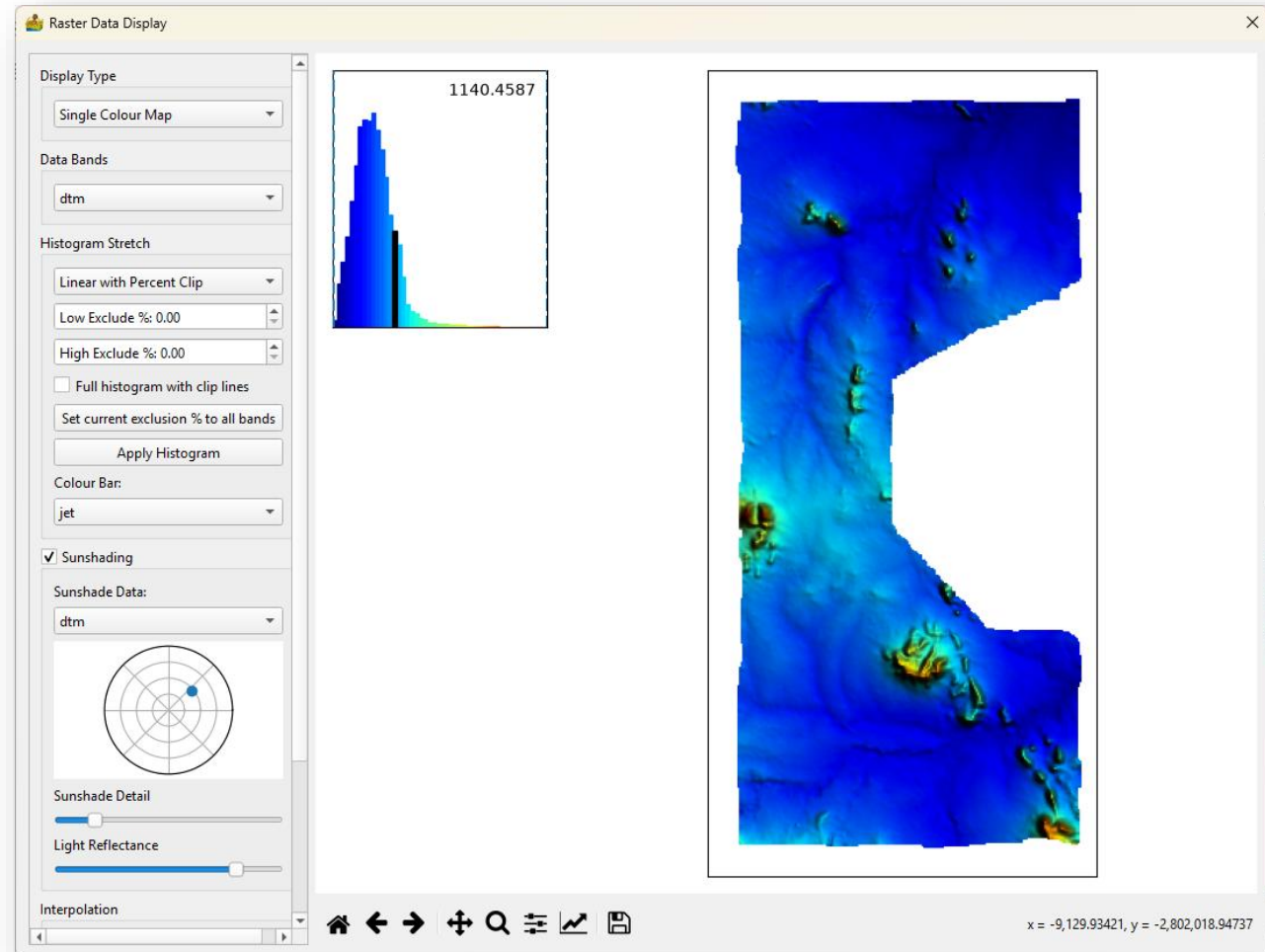
PyGMI Functionality

- Raster
- Vector
- Classification
- EM
- Magnetism and Gravity
- MT
- Potential Field Modelling
- Remote Sensing
- Seismology



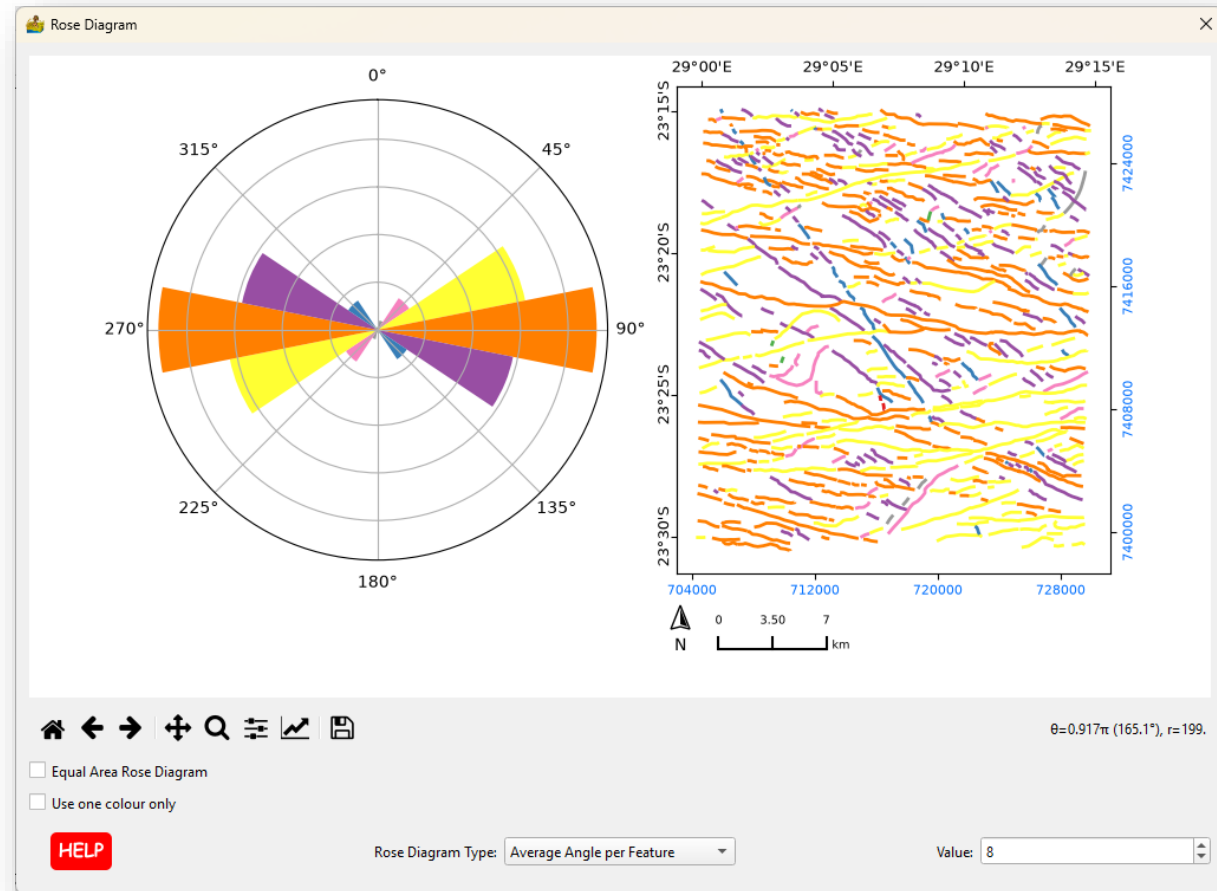
Raster Function Examples

- Imports and exports
- Filters
- Reprojection
- Statistics
- Equations
- Data display
 - Anaglyphs
 - Single layer and ternary
 - Sun shading



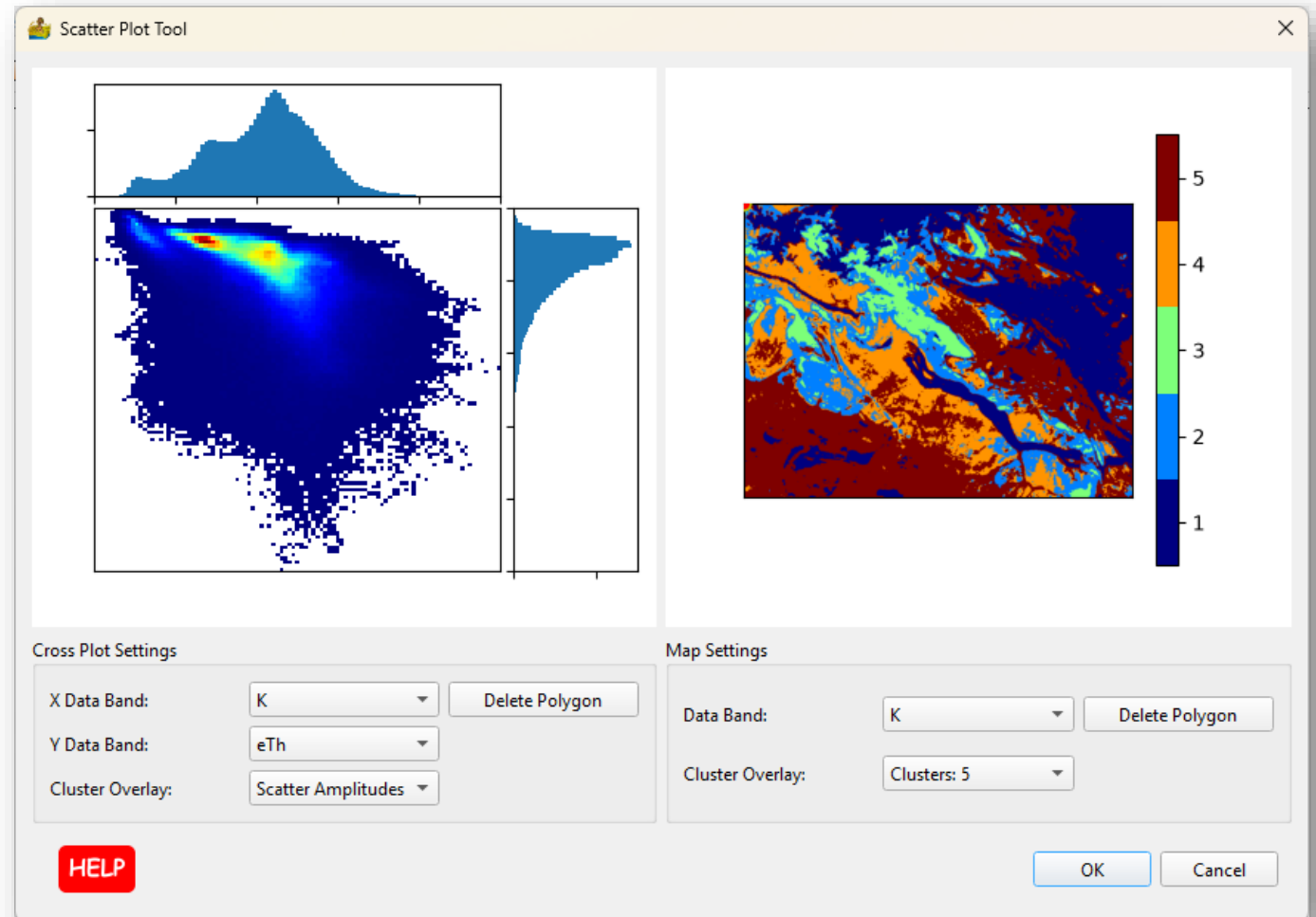
Vector Function Examples

- Imports and exports
- Gridding
- Reprojection
- Statistics
- Structural Complexity
- Data display
 - Rose diagrams
 - Vector plots
 - Histograms
 - Correlation coefficients



Machine Learning

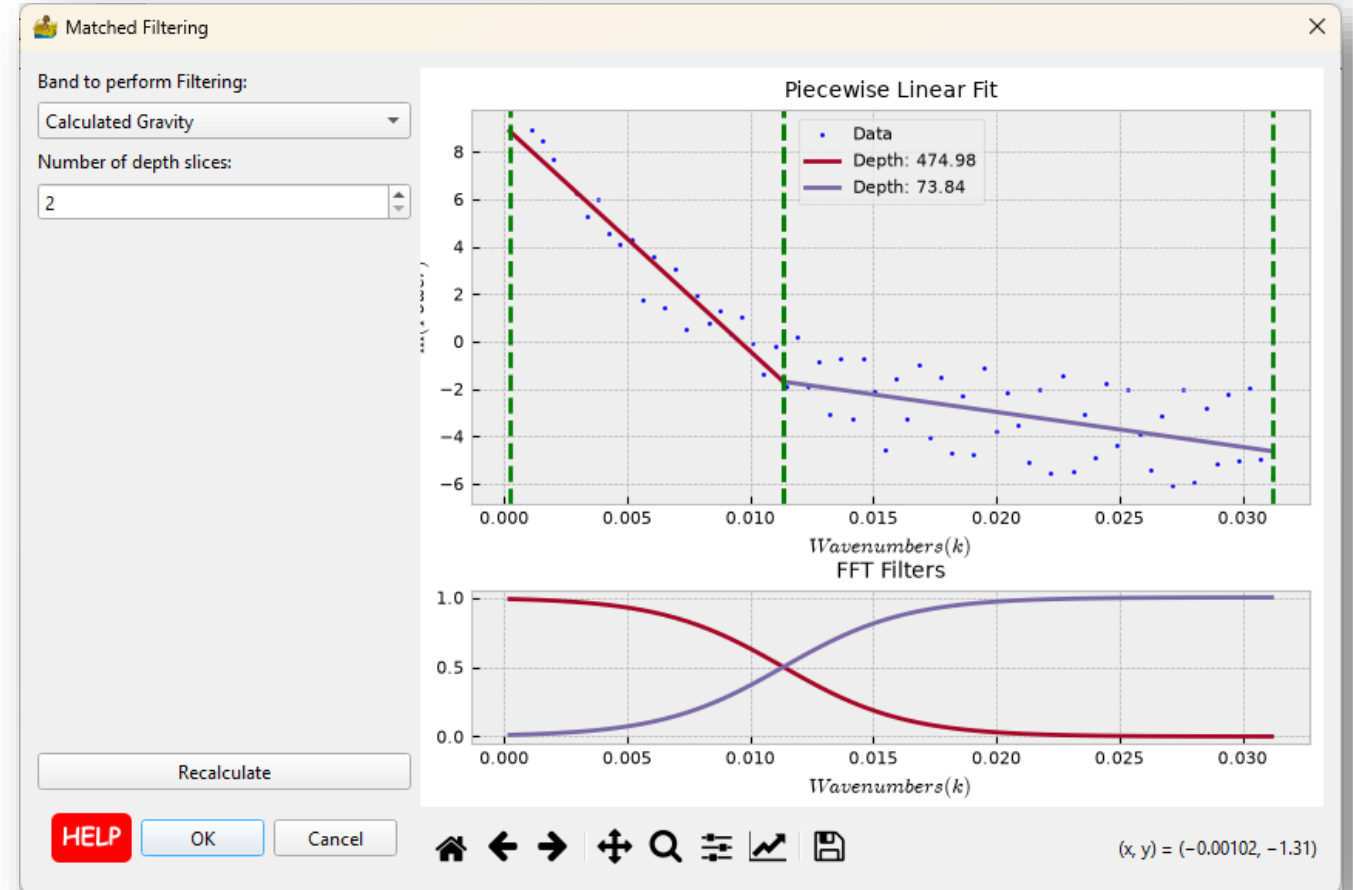
- Cluster analysis
 - Includes crisp and fuzzy
- Image segmentation
- Supervised classification
- Random forests and neural networks
- Data display
 - Scatter plot tool
 - Class results
 - Metrics



Custom Functions

Magnetic and gravity:

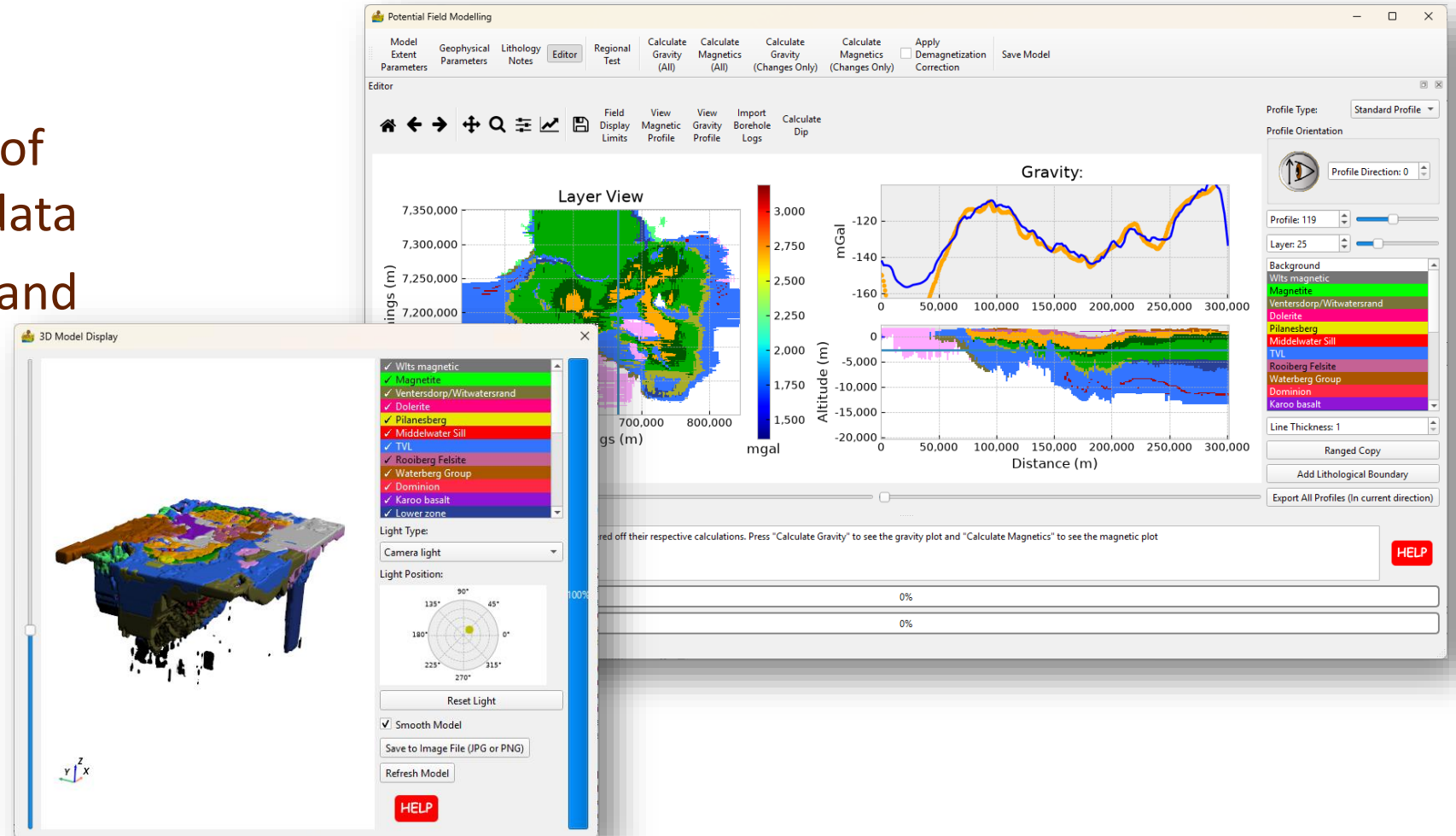
- Reduction to the Pole
- IGRF calculation
- Analytic signal
- Continuation
- Tilt angle and tilt depth
- Matched filtering
- Import and process gravimeter data.



3D Modelling Example

Potential field data:

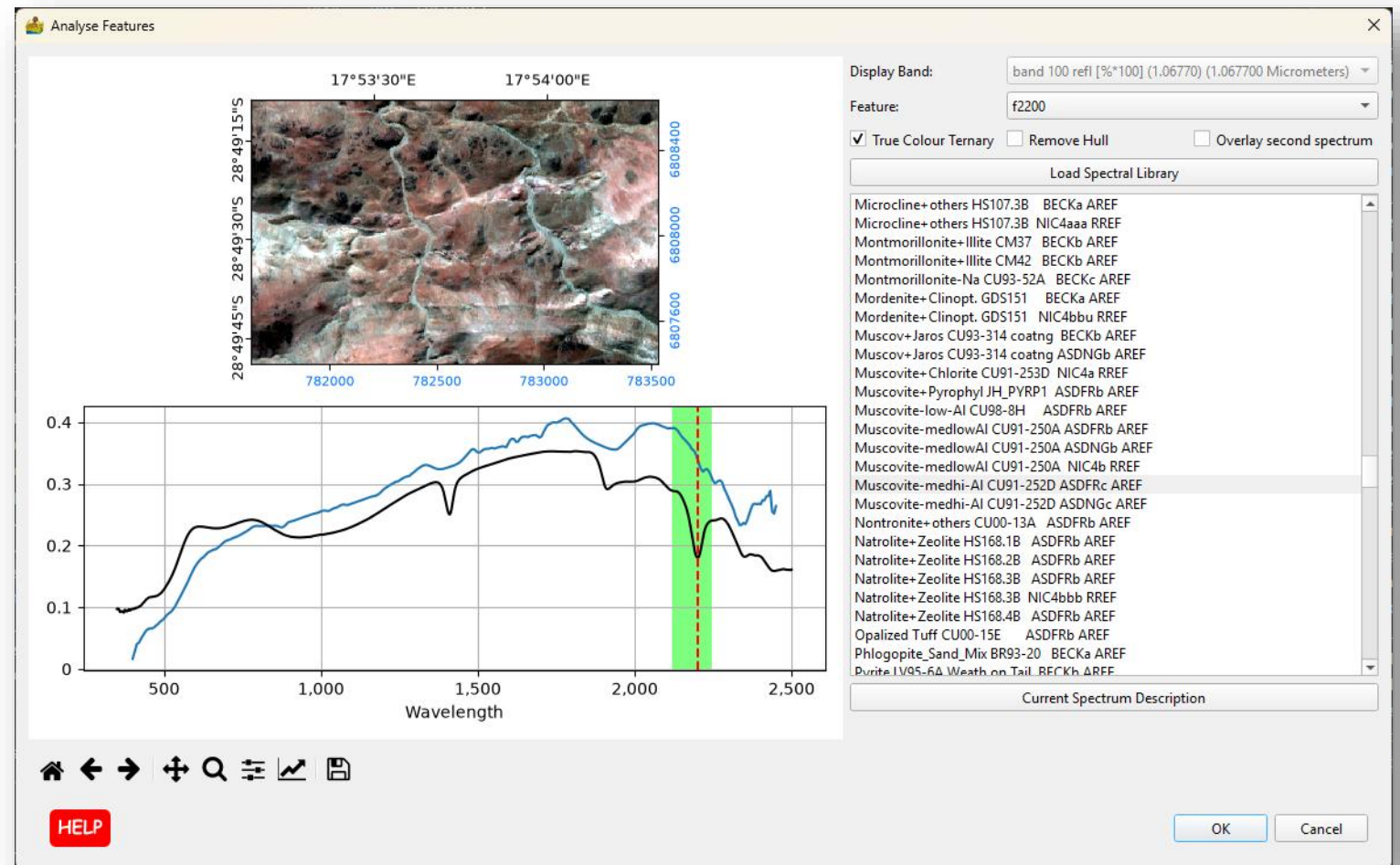
- 3D forward modelling of magnetic and gravity data
- Inversion of magnetic and gravity data.
- Statistics
- 3D display and export.



Spectral Analysis and Automation Example

Remote sensing:

- Imports
- Band ratios
- PCA and MNF
- Topographic correction
- Condition indices
- Hyperspectral analysis and feature detection
- Change detection



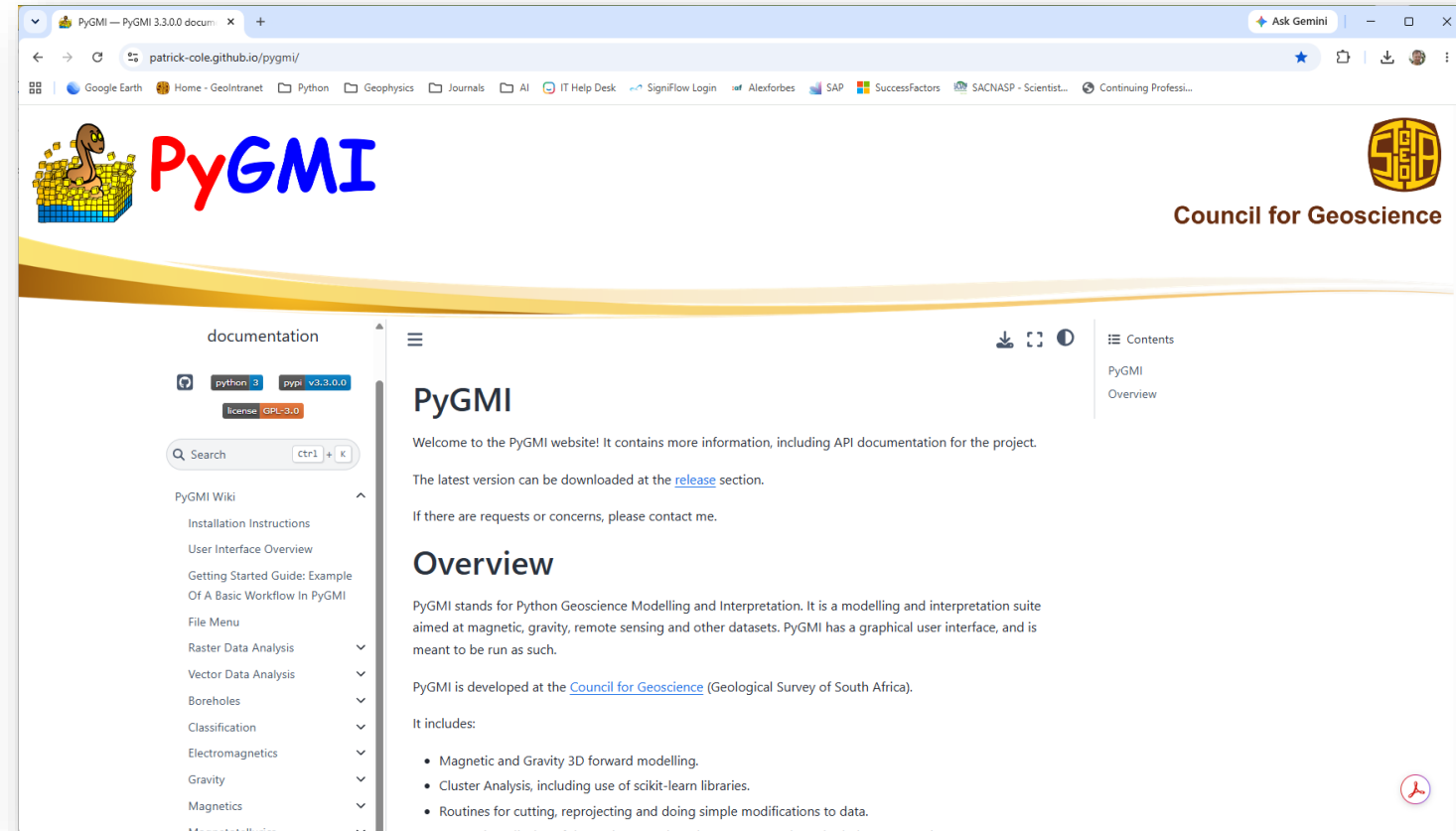
Feedback

- Important for software to be relevant
 - Evolution of software for new features
 - Removal of obsolete features
- Not only about finding bugs
 - But people do use things in (unpredictably) different ways



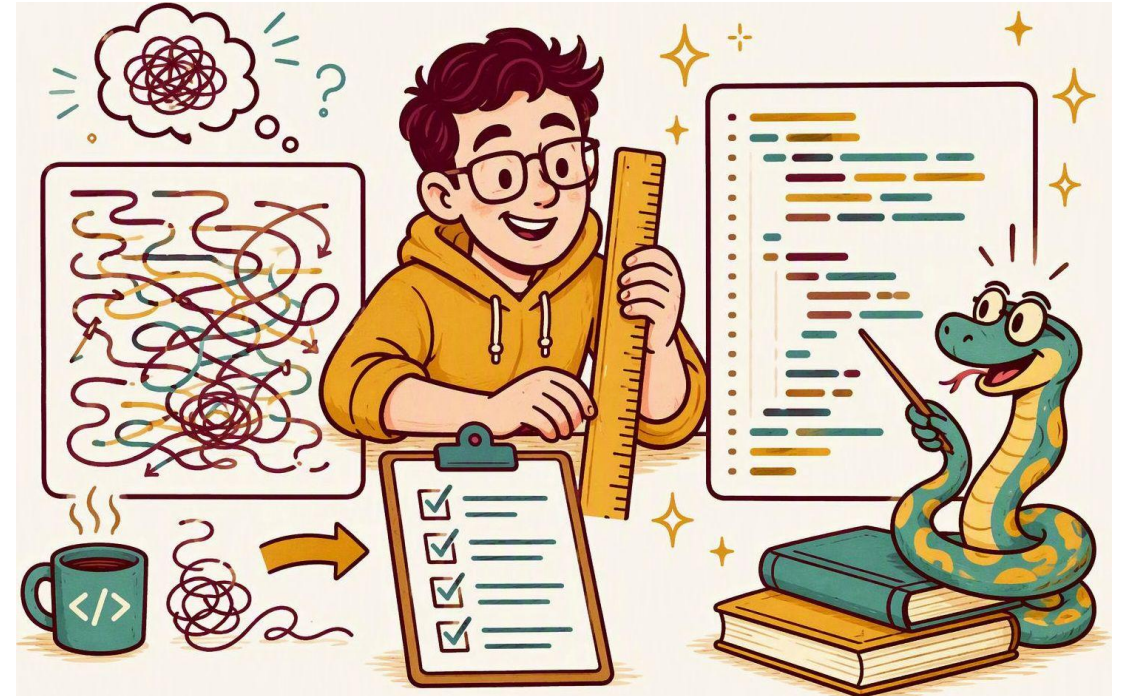
Documentation and Websites

- Website important for documentation and new releases
- Wiki and API
 - <https://patrick-cole.github.io/pygmi/>
- Github repository
 - <https://github.com/Patrick-Cole/pygmi>
- PyPI Repository
 - <https://pypi.org/project/pygmi/>



Coding Standards

- Important for reusability of code.
- For Python, this is the PEP (Python Enhancement Proposals) 8 standard:
 - Code layout and formatting
 - Naming conventions
 - Imports
 - Programming practices and idioms
- Automation tools available:
 - Flake8
 - Ruff
 - Black



Software challenges

- Need specific skillsets
 - Aptitude with programming
 - Applied Mathematics
 - Geology
- Needs organisational support in terms of time



From <https://www.binaryfolks.com/blog/software-development-problems>

The role of AI

- Very useful for coding examples, including what libraries are available for a specific problem.
- Sometimes prone to syntax or function parameter inaccuracies.
- Provides solutions based on phrased request, not necessarily best practice.
- Makes custom images for presentations easy!



Conclusions

- Open Source collaboration is possible.
- Frameworks – such as GitHub – allow for co-development and sharing of research software.
- Research software allows for cutting edge research and fresh ideas.
- Journal of Open Source Software (JOSS) is an excellent example of research software publication – encourages good coding standards as a part of the review process.

THANK YOU

