

Integrating ArcGIS and SciPy

By Kevin Butler, Product Engineer, Esri Geoprocessing and Analysis Team

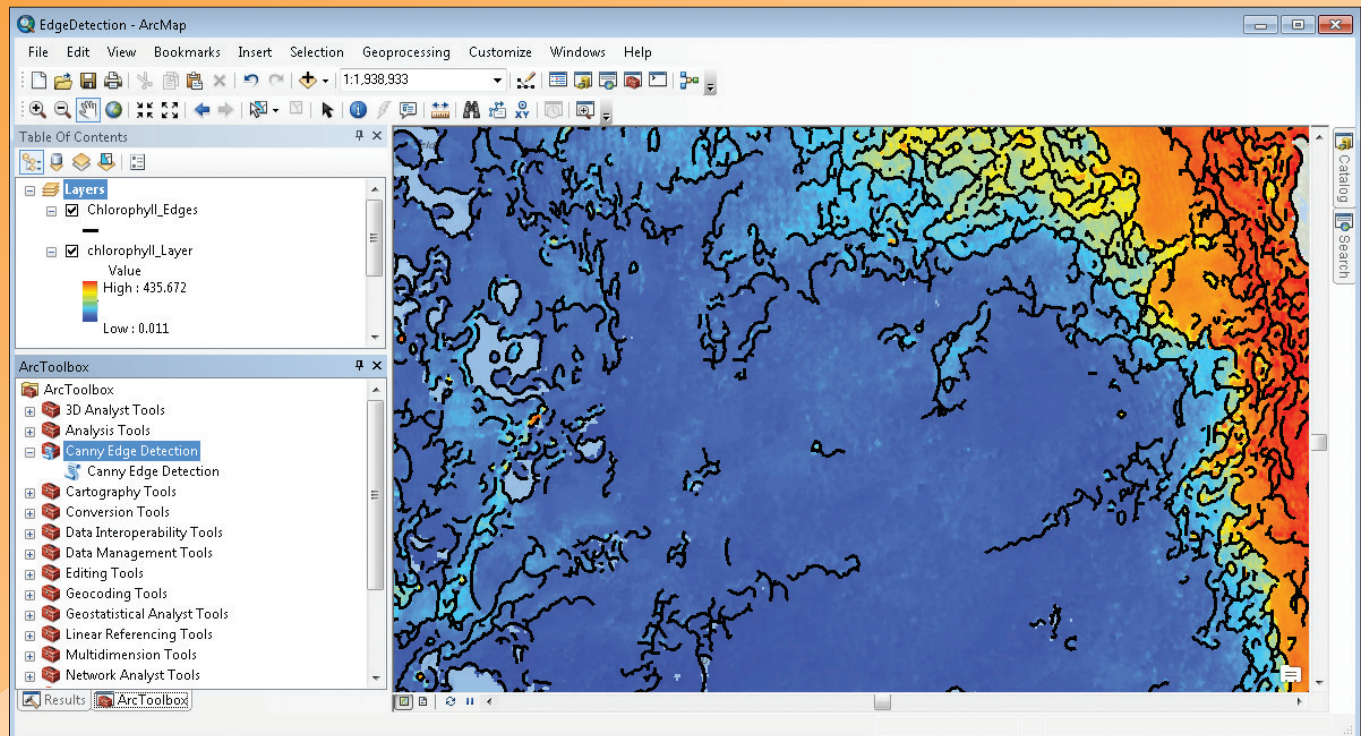
At the Esri Ocean GIS Forum in November 2014, Esri announced the integration of ArcGIS with SciPy, a software library that helps scientists, engineers, and GIS analysts perform custom scientific and technical computing.

SciPy is an open-source library built using Python, the easy-to-learn, highly scalable, stable scripting language of choice for ArcGIS. Python is both a programming language and a collection of modules (small sections of code that perform some useful operations). For example, a Python module called math performs trigonometric functions and angular conversions. Python ships with a predefined set of modules that meet most general scripting needs. SciPy extends the basic functionality of Python by adding modules that perform functions useful to the scientific and engineering communities.

SciPy is not just one Python module but a collection of modules and is sometimes referred to as the SciPy Stack, which is composed of

- NumPy for numerical computation using arrays
- SciPy, a collection of numerical algorithms
- Matplotlib for 2D and 3D plotting
- Pandas for high-performance data structures
- SymPy for symbolic mathematics and computer algebra
- IPython, an interactive interface for quickly testing scripts
- Nose for testing Python code

NumPy and Matplotlib have shipped with the ArcGIS platform for the last several releases. Adding the remaining core packages fully integrates the SciPy Stack with the ArcGIS platform to make scientific and technical computing easier for users.



↑ Chlorophyll edges detected using filters provided by the SciPy multidimensional image processing module

While the SciPy Stack is open-source software and can be freely downloaded, it can sometimes be difficult to integrate with ArcGIS. The strength of SciPy lies in its integration of many software modules. Installing these modules can be time-consuming and error prone. For example, Pandas and Matplotlib may depend on a particular version of NumPy. If the correct version of NumPy is not loaded on your computer, Pandas and Matplotlib may not function as expected. Getting the correct versions of all the components of the SciPy Stack can be challenging. With SciPy integrated into the ArcGIS platform, all these interdependencies are taken care of for you.

Integrating SciPy with ArcGIS makes developing scientific and technical geoprocessing tools and scripts easier and more efficient. You won't have to start from scratch each time you begin solving a scientific or technical problem. The SciPy Stack represents tens of thousands of lines of useful code that have already been thoroughly tested and documented. The work of hundreds of other programmers can be leveraged simply by reusing code from a SciPy module.

There is another important reason that SciPy is being integrated with ArcGIS. This will ensure that any custom scientific or technical geoprocessing tools you create are shareable. One of the fundamental design goals of Esri's geoprocessing framework is the ability to distribute your tools among the ArcGIS community of users. If you develop a custom geoprocessing tool that depends on SciPy, you can share that tool as a geoprocessing package or a geoprocessing service and not have to worry whether the person you're sharing with has the correct version of SciPy loaded.

The possibilities for integrating SciPy with ArcGIS are really only limited by the imagination of the GIS community. Python, and by extension SciPy, is embedded in the ArcGIS platform to make automating repetitive geoprocessing tasks and extending the geoprocessing framework easier. For the scientific and engineering

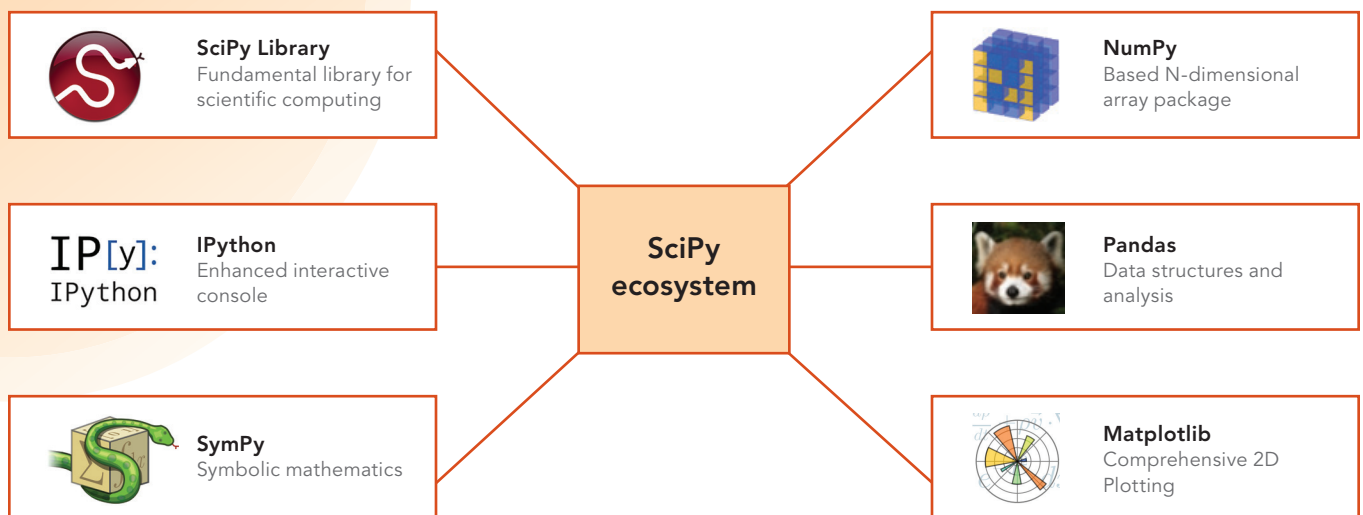
communities, extending the geoprocessing framework will be the more important part.

The SciPy Stack has so many rich functions that it is hard to imagine all the creative spatial problems that GIS users with scientific or engineering domain expertise will be able to tackle—but there are a few that come to mind. An atmospheric scientist may use the image filtering modules to delineate zones of horizontal transport of water vapor. A transportation geographer may use the Markov chain modules to simulate traffic flow. A geoscientist may use the symbolic mathematics routines to trace faults and model crustal movement. An ocean scientist may use the calculus module to calculate ocean dynamics. A fisheries scientist or resource manager may use the linear algebra modules to set a harvest quota for a fish stock.

SciPy will be integrated with ArcGIS through a staged release. It will be available for ArcGIS Pro with the 10.3 release as an optional install. (ArcGIS Pro is the premier application for visualizing, editing, and performing analysis that comes with ArcGIS 10.3 for Desktop.) At the 10.3.1 release, SciPy will be automatically installed with ArcGIS for Desktop.

SciPy is being integrated with ArcGIS Pro first because ArcGIS Pro has a 64-bit architecture and leverages threading to keep the user interface responsive and utilize additional CPU cores on the local machine, which is a very important feature when running computationally intensive scientific algorithms.

Future plans for SciPy and ArcGIS will depend on the scientific and engineering practitioners in the GIS community. SciPy is open-source software, so it grows and improves as users contribute new and/or improved code or modules. Newer versions of the SciPy Stack will be integrated and supported in future releases of ArcGIS and documented in the familiar "What Is New in ArcGIS" document that details each release.



↑ The SciPy Stack