10th ACES International Workshop, Awaji Island, September 25-28, 2018

Clustering GPS Velocities

Granat, R.[1] and Donnellan, A. [1]

1 *Jet Propulsion Laboratory, California Institute of Technology, Pasadena CA, USA*

We present a data driven approach to clustering (grouping) GPS stations according to their velocities and other selected characteristics. Clustering GPS stations not only has the potential for identifying useful scientific information (e.g., separating regions of post-seismic motion), but also is a necessary initial step in other GPS analysis methods, such as those used to detect aseismic transient signals (Granat et. al. [2013]). Using this approach, the scientist selects the desired features of interest, which could include some subset of the three velocity components, uncertainty estimates, the station location, and any other relevant information present in the data set. Based on those selections, the method proceeds to autonomously group the GPS stations according to the desired method; some methods require that the number of groups be specified in advance, while others estimate the number of groups from the data. We have implemented this approach as a Python application, allowing us to draw up on the full range of open source clustering methods available in Python's scikit-learn package (Buitinck et. al. [2013]). The application returns the GPS stations labeled by group in both tabular form and as a color coded KML file for overlay in Google Earth with other sources of information as in Figure 1. Our implementation is designed to work with the GPS velocity information available from GeoGateway (Donnellan et. al. [2016]), a map-based science gateway supported by NASA, but is easily extendable to other data sources or output formats.

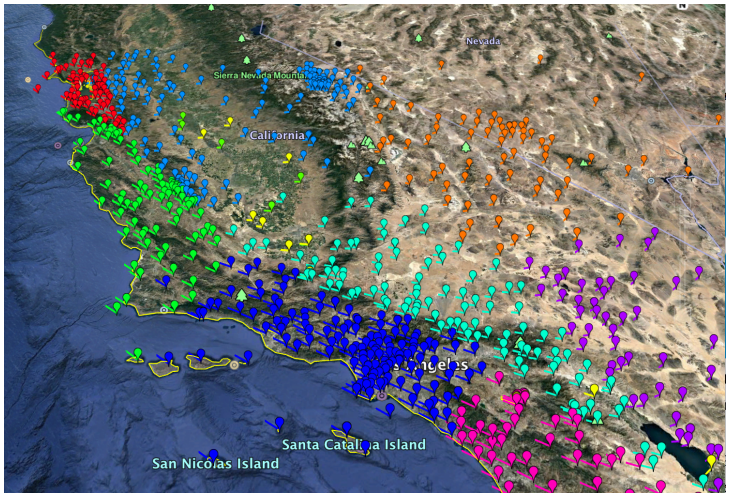

Figure 1. An example of GPS clustering results for stations in California, as displayed in Google Earth. Stations are color coded according to group label.