

# Geographic profiling of invading species

Alessio Papini and Ugo Santosuosso, University of Florence, Italy, [alpapini@unifi.it](mailto:alpapini@unifi.it)

## Abstract

Use of geographic profiling for finding first introduction point of a biological invasion

**Citation:** Alessio Papini and Ugo Santosuosso, University of Florence, Italy, [alpapini@unifi.it](mailto:alpapini@unifi.it) Geographic profiling of invading species. **protocols.io**

[dx.doi.org/10.17504/protocols.io.kytcxwn](https://doi.org/10.17504/protocols.io.kytcxwn)

**Published:** 24 Nov 2017

## Protocol

### Step 1.

### Step 2.

Silhouette with `kmeans_sil_0_0_2.py` python `kmeans_sil_0_0_2.py caulerpaprol.csv 4 G` The meaning of the command is: python coordinates.csv Number\_of\_cluster Graph "If instead of Graph You insert G on the command line You will get the graphic silhouette. If You insert N, no graphic will be shown" Graph = G/N "print" G = present a Graphical representation of clusters and centroids "the best number of clusters is the input for the next step" It is possible to save the Silhouette profiles as png images. Once You have chosen the right number of clusters with Silhouette, go on with the script `scriptgeokmeans_1.0.py`

### Step 3.

Use the script `scriptgeokmeans_1.0.py` python `scriptgeokmeans_1.0.py` You need to modify at the beginning of the script the names of the files regarding the observations in csv format (currently `caulerpaprol.csv`); the name of the map (currently `mediterraneolarge.bmp`); the number of clusters (evaluated with Silhouette, currently 4) it will run the following procedures (all python programs contained in the folder): (the default map name is "`mediterraneolarge.bmp`", while the data set is `caulerpaprol.csv` (csv file with the sites of *Caulerpa prolifera* in the Mediterranean)) 2a) python `kmeans0_57.py caulerpaprol.csv 4 N` [N means that no graph is shown; G would mean that a graph is shown, showing the clusters of points on the map, a map of the same dimension of the input map, but white] 2b) Geoprofiling on all the available data python `Geoprof2_0_5.csv.py mediterraneolarge.bmp caulerpaprol.csv results B E 10 0.4 0.4` [results is the prefix of the files arising as results; B is the type of output graph as default (only bulleye); E is the type of distance used for the geoprofiling (here euclidean); 10, 0.4 and 0.4 are the values of variables, respectively, B, F and G, in the Geoprofiling formula] 2c) The results are: file `risuts.png` with the geoprofiling e `vorondiagblend_5_clust.png` with the space divided in Voronoi tessels.