

Modelling the center of origin and the spreading pattern of *Caulerpa* invasion in the mediterranean

Alessio Papini¹ & Ugo Santosuosso²

¹ Department of Biology, University of Florence, Via Micheli 3, Florence, 50121, Italy. E-mail: alessio.papini@unifi.it

² Department of Clinical and experimental Medicine, University of Florence, Largo Brambilla 3, 50134 Florence, Italy. E-mail: ugo@unifi.it

Two species of genus *Caulerpa*, *Caulerpa taxifolia* (M. Vahl) C. Agardh and *C. racemosa* (Forskål) J. Agardh are among the most relevant invasive species in the Mediterranean, possibly also in relation to climate change. Their presence is affecting the presence of native species of benthonic algae and the prairies of *Posidonia oceanica*.

The origin of the invasions was previously investigated and, while the center of origin of the invasion of *C. taxifolia* is known (accidental release from an aquarium), it is still a matter of debate for *C. racemosa*.

One of the method employed for this type of analysis, is the Geographic Profiling, a method that was first used to simplify the localization of serial killers home on the basis of the site of the offenses. In the case of invasive species the employed data are the observations in the field reported on a map. The map presents the area of most probable point of origin of the invasion with a red color and the areas with lower probability with other colours (normally yellow and green). Since the method implies the presence of a buffer zone around the origin, where the spreading would not occur (hence dispersal possible only at a given distance, variable depending on the species), ad hoc assumptions based on the biology of the investigated invasive species should be done case by case.

The method can be further refined by dividing the data set in more homogenous subsets of observations with K-means algorithm; increasing the robustness of the observations with a jackknife technique and comparing the results with other methods of analysis of the invasion pattern, such as DB-scan and Isolation Forest.

The use of these methods allowed for *C. racemosa* to exclude a lessepsian origin of the invasion and rather to point at an unvolunteer transport by ship, with the possible center of the origin of the invasion localized most probably in South Italy.

All these methods, implemented in Python programming language by the authors, are open source and available on the internet. These methods can be useful for analyzing also other pattern of distribution of invasive species or, in general the biologicla data set characterized by a starting point on a map.