Alessio Arnese

Applying Ecological Niche Factor Analysis for Predictive Modelling in the Kaulonia Field Survey

Abstract: During the archaeological survey in Kaulonia (Calabria, Southern Italy), we observed a pattern in the distribution of Bronze Age finds. Due to that observation, we tried to define both a model of these findings and a map of their probability of presence. We have chosen a method applied to model animal distribution, the Ecological Niche Factor Analysis (ENFA) that works without absence data. ENFA computes the suitability function by comparing the species distribution in the eco-geographical variables (EGV) space with that of the whole set of cells. The model describes the Bronze Age findings as more influenced first by high elevations, then by the distance from sea. In the next survey campaigns, we have tested the Habitat Suitability map, surveying new areas and we found new concentrations of Bronze Age finds where there is a high potential on the Habitat Suitability map.

Introduction: the Kaulonia Survey Project

From 2001 to 2006, the Laboratorio di Storia Archeologia e Topografia del Mondo Antico undertook eight survey campaigns in Calabria, near the Greek city of Kaulonia. It was a systematic intensive research on the ground on a territory divided in two different areas *(Fig. 1)*. The principal area is the northern zone surrounding Kaulonia, characterised by two small rivers (Assi and Stilaro from north to south) with an extension of about 58 km². The second area is a strip of 5 km² parallel to the coast, walked in one campaign. We have excluded from the survey the area inside the walls of Kaulonia, the zones with modern cities and others impracticable areas (like rivers, scattered rural houses etc.). Ian Hodder and Christine Malone had already studied this territory with systematic surveys in the late seventies (Hodder / Malone 1984). We classified the archaeological presence as topographical units (zones with high concentration of finds, often in the same place of old sites) and zones of sporadic finds, where there are few finds, moved by time from their original places. At the end of our survey campaigns, we have walked 35% of the whole area and have found 174 topographical units and 230 areas of sporadic findings (for more information about the survey see Facella / Arnese 2003), with a chronology from Prehistory to the Middle Ages. From the second campaign, every topographical unit has been georeferenced using a GPS receiver (two Garmin GPS Map 76S), with an error of approximately three meters. With a Geographical Information System we managed all data: topographi-

cal units and areas of sporadic findings, a map of the walked zone with their cultivations and visibility, a DEM of the whole area, rivers, new cities and roads. We also georeferenced raster and vector maps at a scale of 1 : 25 000 and raster maps at a scale of 1 : 5000, the same as that used in the field during the survey.

Protohistoric Sites: from Survey to Model

The work that I illustrate dates back to the fourth campaign, when we registered 81 topographical units, excluding the southern area from the survey. During the first half of the survey, the most relevant result was the finding of protohistoric sites, never found before (the Hodder / Malone surveys have recorded many prehistoric traces). The topographical units of this period show similar characteristics with others protohistoric sites found in Calabria. These settlements are located on spurs not near the sea. Among the others protohistoric topographical units, we found two necropoles in a good state of preservation, dating back to the Final Bronze Age and First Iron Age (see Facella / Arnese 2003). The aim of this work is to describe the ecological conditions of every protohistoric presence and to make a prediction of the most probable places where we would find new protohistoric topographical units, without any interest in reconstructing the old landscape. The first question we asked was about the quality of the data collected during the survey. With an intensive survey like this, the archaeologists can register
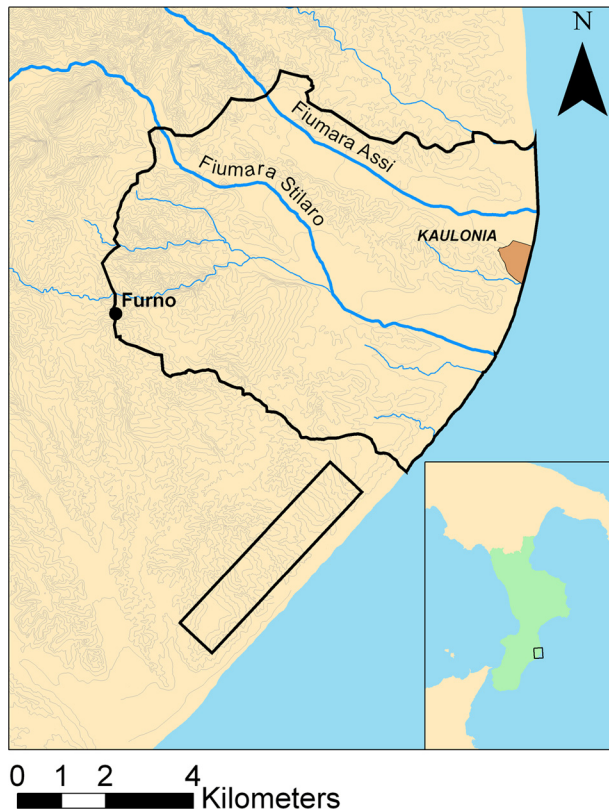
Fig. 1. Kaulonia Survey boundaries.

every significant trace of the past and such traces are often "ex silentium", but as the presence is a simplified representation of old settlements system, also absence reveals, or hides in some cases, more information about the old landscape and its relation with humans. In our study area, a great quantity of modern villages and scattered houses as well as intensive agricultural use of land amplifies the background noise. An analysis of our method and data reveals that the absence data can have three different origins. We record an absence every time we cannot see archaeological material, but we are not able to exclude its presence (false absence) or when a field is not accessible (false absence) or, in the last case, when walked zones have no archaeological material (true absence). As one can see, only one of these three causes of absence is a real absence. For that reason, we preferred a statistical method based only upon presence data. These kinds of problems are common in the zoological domain, where the absence data can be false. Like zoological models, for the archaeological ones also there is a strict relation between species/humans and ecology. It was not without reason that in 1957 the limnologist Hutchinson said: "only when anthropology and archaeology enter the field of human demography does some-

thing comparable to animal demography" (Hutchinson 1957). With this work we have applied to humans the concept of the ecological niche, as elaborated by Hutchinson: a hyper-volume in the multidimensional space of ecological variables within which a species can maintain a population.

Ecological Niche Factor Analysis

The ecological niche is the subset of cells in the ecogeographical space where the focal species has a reasonable probability to occur. This multivariate niche can be quantified on any of its axes by an index of marginality and specialization. Since the ecological variables are not independent, Ecological Niche Factor Analysis, like the PCA, determines relationships between variables and finds combinations of these variables to produce uncorrelated indices or components. In ENFA, however, unlike PCA, the components have direct ecological meaning.

The first factor is the marginality: it is the direction on which the species niche differs at most from the available conditions in the global area. The higher the absolute value of marginality, the more species habitat differs from study area. A positive marginality means that the species prefers higher-than-mean values on the ecological variable. Specialisation factors indicate how restricted the species' niche is in relation to the study area. It is extracted by computing the direction that maximises the ratio of the variance of the global distribution to that of the species distribution. The higher the absolute value, the more restricted is the range of the focal species on the corresponding variable. Like others statistical methods, ENFA requires that the study area is represented by raster maps. Each cell of a map contains the value of one variable. Eco-geographical maps contain continuous values, measured for each of the V descriptive variables. Species maps contain boolean values (0 or 1), a value of 1 meaning that the presence of the focal species was proved on this cell. A value of zero simply means absence of proof (not the absence of the species). Each cell is thus associated to a vector whose components are the values of the EGV in the underlying area, and can be represented by a point in the multidimensional space of the EGVs. If distributions are multinormal, the scatterplot will have the shape of a hyper-ellipsoid. The cells where the focal species was observed constitute a subset of the global distribution and are plotted as a smaller hy-
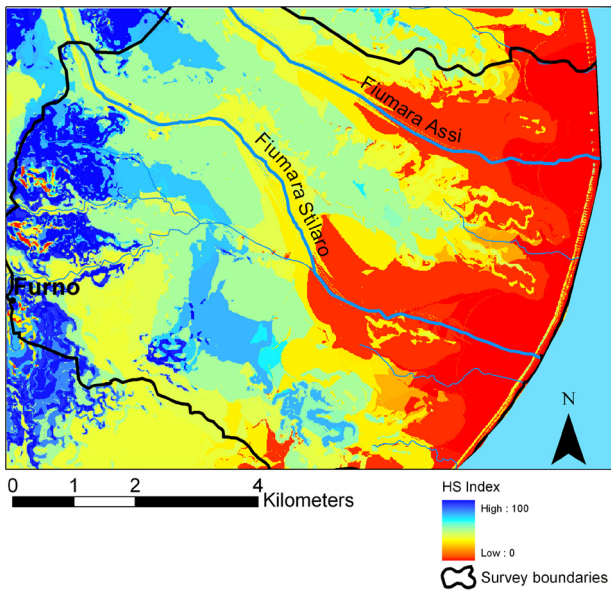
Fig. 2. The Habitat Suitability map.

per-ellipsoid within the global one. The first factor, or marginality factor, is the straight line passing through the centroids of the two ellipsoids. The species marginality is the distance between these centroids. To obtain the specialization factors, the reference system is changed in order to transform the species ellipsoid into a sphere, the variance of which equals unity in each direction. In this new metrics, the first specialization factor is the one that maximizes the variance of the global distribution (while orthogonal to the marginality factor). The other specialization factors are then extracted in turn, each step removing one dimension from the space, until all V factors are extracted. The University of Lausanne has designed a software called Biomapper, which builds Habitat Suitability (HS) models and maps for any kind of animal. With a few simple steps, it prepares the eco-geographical maps (maps that represent the ecological variables), computes the Ecological Niche Factor Analysis and then a Habitat Suitability map. The eco-geographical variables maps need to be able to be perfectly overlaid, with the same raster resolution and quantitative. After the calculation of Ecological Niche Factor Analysis, Biomapper produces a HS map, where suitability values range from 0 to 1 *(Fig. 2)*.

Results

Our presence data consist of 23 topographical units with protohistoric pottery, without any considera-

tion about relevant concentrations or ratio between these materials and other pottery classes. This is a small sample, but sufficient to obtain a good model. The choice of eco-geographical variables is based on our surveying experience, on the study of others protohistoric sites from Calabria and on data availability. The ecological variables are elevation, slope, aspect and distance from coastline and from principal rivers (Assi / Stilaro). Every variable is a raster surface with a resolution of 10 m, derived from vector maps at a scale of 1 : 25 000. The ENFA method provides for our presence set an overall marginality of M = 1.067, showing that protohistoric topographical units habitat is very different from the mean habitat of the study area (protohistoric presence needs specific ecological conditions) and an overall specialization of S = 2.765, which means that we found protohistoric materials in a small range of conditions. If we observe the marginality coefficients of every eco-geographical variable *(Tab. 1)*, we can see that the protohistoric findings are influenced primarily by elevation, then by distance from coastline and slope. Positive values mean that the values of variables inside site areas are greater than values in the whole study area (i.e. the mean elevation of protohistoric presence is 292 m, versus the mean elevation of 121 m for the study area). The values of specialization factors indicate that protohistoric sites have a small range of elevations, but the range of distances from the coast is more restricted. The marginality and the first specialization factor together explain the 91% of information. For that reason, these factors alone are sufficient to produce the Habitat Suitability map, where the suitability values range from 0 to 1. In order to calculate this map we used the median algorithm, while making the assumption that the best habitat is at the median of the species distribution on each factor. Overlaying the HS map with that of the presence set, we can identify two different groups of topographical units: the first group has high elevations and it is at great distance from the coastline; the second group is nearest to the sea and at lower elevation *(Fig. 3)*. Moreover, on the Habitat Suitability map the first group of sites has values greater than 0.5, while the second group has values lower than 0.5. By observing the archaeological data from these areas, we find that the first group is more homogeneous than the second group: each topographical unit has a certain chronology and, in most cases, only protohistoric material.
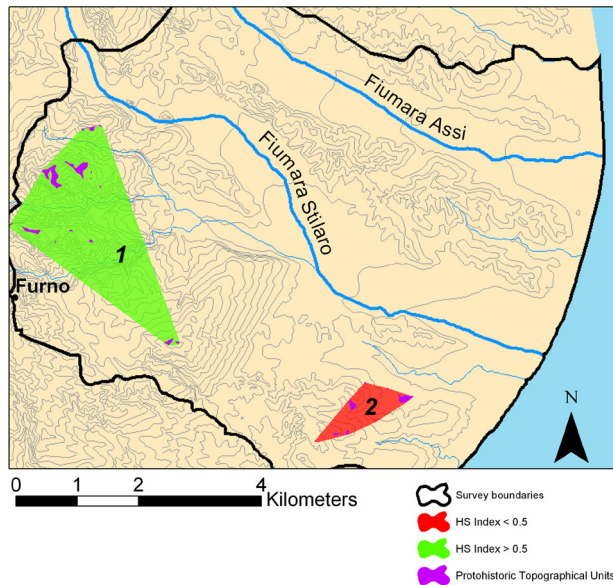
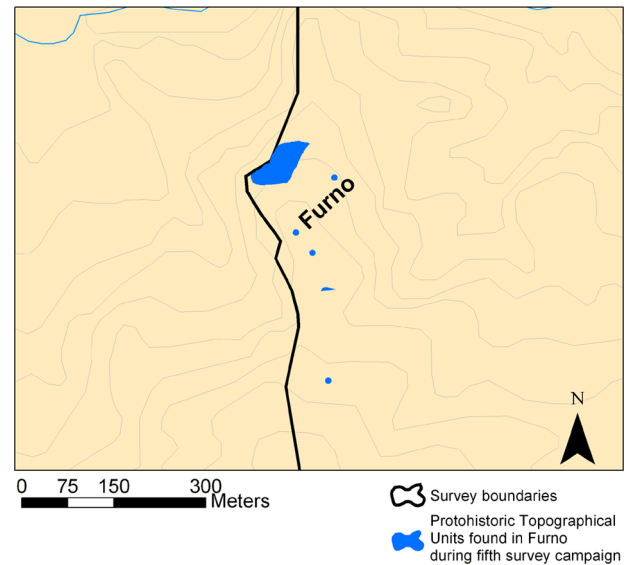Fig. 3. Two different groups of protostoric topographical units.

Fig. 4. The topographical units in Furno.

	Marginality (60%)	Spec. 1 (31%)	Spec. 2 (5%)	Spec. 3 (3%)	Spec. 4 (1%)
Elevation	0.84	0.20	0.41	0.24	0.40
Distance from sea	0.40	−0.37	−0.89	−0.38	−0.02
Slope	0.35	0.03	0.03	0.07	−0.88
Aspect	−0.08	0.04	−0.07	0.89	0.26
Distance from rivers	0.03	−0.91	0.18	0.01	−0.04

Tab. 1. Variance explained by the ecological factors. EGVs are sorted by decreasing absolute value of coefficients on the marginality factor. The amount of specialization accounted for is given in parentheses in each column heading.

Testing the Model in the Field

Following this study, during the fifth survey campaign, we tested obtained results in the field. We walked new areas and compared results from the field with the Habitat Suitability map. Where this map indicates a low Habitat Suitability value, we found no protohistoric pottery and in areas with high HS value we found new protohistoric topographical units. An interesting example is the place called Furno, located near the south-west boundaries of our study area: it is a hill covered by wild vegetation with a high Habitat Suitability value *(Fig. 4)*. We collected small fragments of protohistoric pottery already along the path leading to the plateau, where a big concentration of protohistoric ceramics was found. Another example that validates our model and Habitat Suitability map is Prano, an area covered by high vegetation and with bad visibility conditions. This zone has a high HS value and during the survey we had not found any protohistoric topographical units (the only presence data used to elaborate the model). By observing the map with all the findings, we can now see here an area of sporadic findings with protohistoric pottery. In the future we will apply this method to other topographical units dating back to other periods with other eco-geographical variables. We will also represent cultural elements as raster maps, decisive in choosing settlement location. This way we can describe a "Cultural Niche" by means of "geocultural" variables.

References

Facella / Arnese 2003
A. Facella / A. Arnese, Kaulonia: Ricognizioni Archeologiche nel territorio (2001–2005). Annali della Scuola Normale Superiore de Pisa, Classe di lettere e filosofia 8, 2003, 457–467.

Hirzel / Hausser / Perrin 2002
A. Hirzel / J. Hausser / N. Perrin, Biomapper 2.0 (Lausanne 2002). http://www.unil.ch/biomapper [31 Dec 2007].

Hirzel / Helfer / Métral 2001
A. Hirzel / V. Helfer / F. Metral, Assessing habitat-suitability models with a virtual species. Ecological Modelling 145, 2002, 111–121.

Hirzel et al. 2002
A. Hizel / J. Hausser / D. Chessl / N. Perrin, Ecological-niche factor analysis: How to compute habitat-suitability maps without absence data? Ecology 83, 2002, 2027–2036.

Hodder / Malone 1984
I. Hodder / C. Malone, Intensive survey of prehistoric sites in the Stilo region, Calabria. Proceedings of the Prehistoric Society 50, 1984, 121–150.

Hutchinson 1957
G. Hutchinson, Concluding remarks. Cold Spring Harbour Symposium on Quantitative Biology 22, 1957, 415–427.

Alessio Arnese

Scuola Normale Superiore
Piazza dei Cavalieri, 7
56126 Pisa
Italy
arnese@sns.it