

# The Appennino Tool: a GIS-Tool for the analysis of wildlife habitat selection

## INTRODUCTION

Wildlife management studies identify the resources (e.g. food items or habitats) used by animals and document their availability. Resources availability is defined as the quantity accessible to the animal or populations of animals and is distinguished from abundance, which is defined as the resources in the environment (Allredge et al., 1998). Resource usage is the quantity taken by an animal or population of animals. Used resources may be consumed, in the case of food items, or simply visited, in the case of habitats (Manly et al., 1993). A wide variety of methods can be adopted to studies the animal resource selection (Manly et al., 2002). One of this methods is the Compositional analysis.

Compositional analysis is an application of multivariate analysis of variance (Aebischer et al., 1993; Pendleton et al., 1998; Manly et al., 2002). The technique can be applied to the analysis of data considered for selection among categories of a simple categorical variable. Resource use is defined in terms of the proportion of different types of resources within estimated home range boundary or other use area. A compositional analysis proceeds as follows. It assumes there are  $D$  types of available resource units, and that the single animal's proportional resource usage is described by the composition  $o_{u1}, o_{u2}, \dots, o_{uD}$ , where  $o_{ui}$  is the estimated proportion of the  $i$  type resource used by the individual. The proportions sum to one. Similarly, the analysis sets the available proportions for the same animals to be  $\pi_{a1}, \pi_{a2}, \dots, \pi_{aD}$ . Compositional analysis cannot be applied when resources units are described by continuous variables, unless those continuous variables are expressed as discrete classes. It can be shown that for any component  $o_j$  of a composition, the log-ratio transformation  $y_i = \log_e(o_{ui}/o_{uj})$  produces new variables that are linearly independent, with a specific choice for  $j$ . Based on this result, the differences

$$d_i = \log_e(o_{ui}/o_{uj}) - \log_e(\pi_{ai}/\pi_{aj})$$

are calculated for  $i$ th animal to represent the difference between the relative use and availability of resources  $i$  and  $j$ . With no selection, the mean value of  $d_i$  is expected to be zero over all

animals, for all  $i$ . If necessary, zero values have to be replaced by small positive number when calculating  $d_i$  values. Using the animal as the unit of observation may avoid problems related to the sampling level (Kenward, 1992):

- the unit-sum constraint whereby the avoidance of one resource type leads to some selection for alternatives (Aebischer et al., 1993);
- the differential use of resources by different groups of animal (Aebischer et al., 1993), and pseudoreplications (Hurlbert, 1984) provided animals behave independently.

Despite these advantages, there are important assumptions underlying compositional-based analyses. These include the data independence from different animals and multivariate normality. The need to substitute, under some circumstances (Pendleton et al., 1998), zero values with an arbitrary constant has also caused the method to be criticized.

The method is often used for the analysis of habitat preference (Fearer & Stauffer, 2004; M. Ferretti et al., 2010). It allows to study the animal's preference at both home range or single point (fix) levels. At the moment, it can be performed using several free software like the R Project for Statistical Computing as well as some commercial solutions (Compos Analysis v6.3 – Smith ecology). Microcal Origin, GraphPad Prism, SPSS and SigmaPlot are examples of specialized computer programs commonly adopted to carry out statistical analysis. These programs tend to be expensive and do not include just statistical analysis features but also curve fitting, modeling and graphing ones. Cheaper Microsoft© Excel add-ins, such as JMP-IN and Analyze- It, are available for much lower fees. If the goal of the user is to simply carry out the statistical analysis described in this paper, a lot of resources are wasted on a vast excess of unnecessary features.

Worldwide, Geographical Information Systems (GIS) are being used to solve problems associated with wildlife management. From calculating existing wildlife densities, assessing sustainability of habitat for wildlife populations, to evaluating human impact on the wildlife and

land, GIS can prove to be an invaluable tool. A geographic information system integrates hardware, software, and data for capturing, managing, analyzing, and displaying all forms of geographically referenced information. GIS allows to view, understand, query, interpret, and visualize data in many ways that reveal relationships, patterns, and trends in the form of maps, globes, reports, and charts. A GIS helps answering questions and solve problems by looking at data in a way that is quickly understood and easily shared. GIS technology can be integrated into any enterprise information system framework ([www.gis.com/content/what-gis](http://www.gis.com/content/what-gis)). The ArcGIS© software is a suite consisting of a group of geographic information system (GIS) tools produced by the US-based ESRI company. At the desktop level, ArcGIS© includes ArcView which allows viewing spatial data, creating layered maps, and performing basic spatial analysis. Over the years, the use of the GIS software in the study of wildlife has become essential. In particular, utilities have been developed by experts in the fauna field for the most popular frameworks. Such extensions are commonly used for applied research. Among these we can surely remind Animal Movement (Hooge & Eichenlaub, 1997) for ESRI ArcView 3.x, and Hawts Tools (<http://www.spataleecology.com/htools/tooldesc.php>), for ESRI ArcGIS 9.x. Both of them allow to perform home ranges and animal movements analysis. Other software for the study of habitat preferences, are available on the web (for example <http://www.resourceselectionbyanimals.com/rsba/ProgramListing.aspx>).

The development of radio-collars and their use to track animal movements, led to the need of storing and transpose on digital maps thousands of positions (GPS fixes) for each studied individual. The management of such dataset has become complex and is liable of great simplification.

This tool would allow to keep on working inside the GIS software and would save from the need of exporting the data to an external statistic software.

## PROGRAM DESCRIPTION

The presented tool was first coded in VBA (Visual Basic for Application) directly using the facilities provided by the ESRI ArcGIS (Burke 2003, Chang 2007) environment. Once it was completed and thoughtfully tested, a complete review of the modules was undertaken. Compositional Analysis is a very widespread study technique for the research of environmental preferences. A range of them is obtained from geographically processing environmental availability, home ranges and single animal locations (fix), elaborated to underline their relative weight only in presence of different types of environments. To apply this technique by traditional means, first of all, a spreadsheet should be used for data pre-processing. After that, it could be possible to use a statistical program to apply the ANOVA algorithm between environmental frequencies in the home ranges or for the single locations. Current workflows require to get out from the GIS environment to apply this type of analysis (as the Animal Movement Extension tool and the Hawts Tool documentations show). On the opposite, it would be advisable to have a utility that could let apply the Compositional Analysis within the GIS environment. This seems more desirable, since today the development of Gps radio-collars has arisen the need to store and transpose on digital maps thousands of positions for each animal. The management and the export of such dataset can become quickly complex and can be certainly simplified. The ArcGIS© tool described in this paper has been designed for ease of use and has been rigorously tested in conjunction with commercially available packages for accuracy.

Basic data needed to perform Compositional Analysis are:

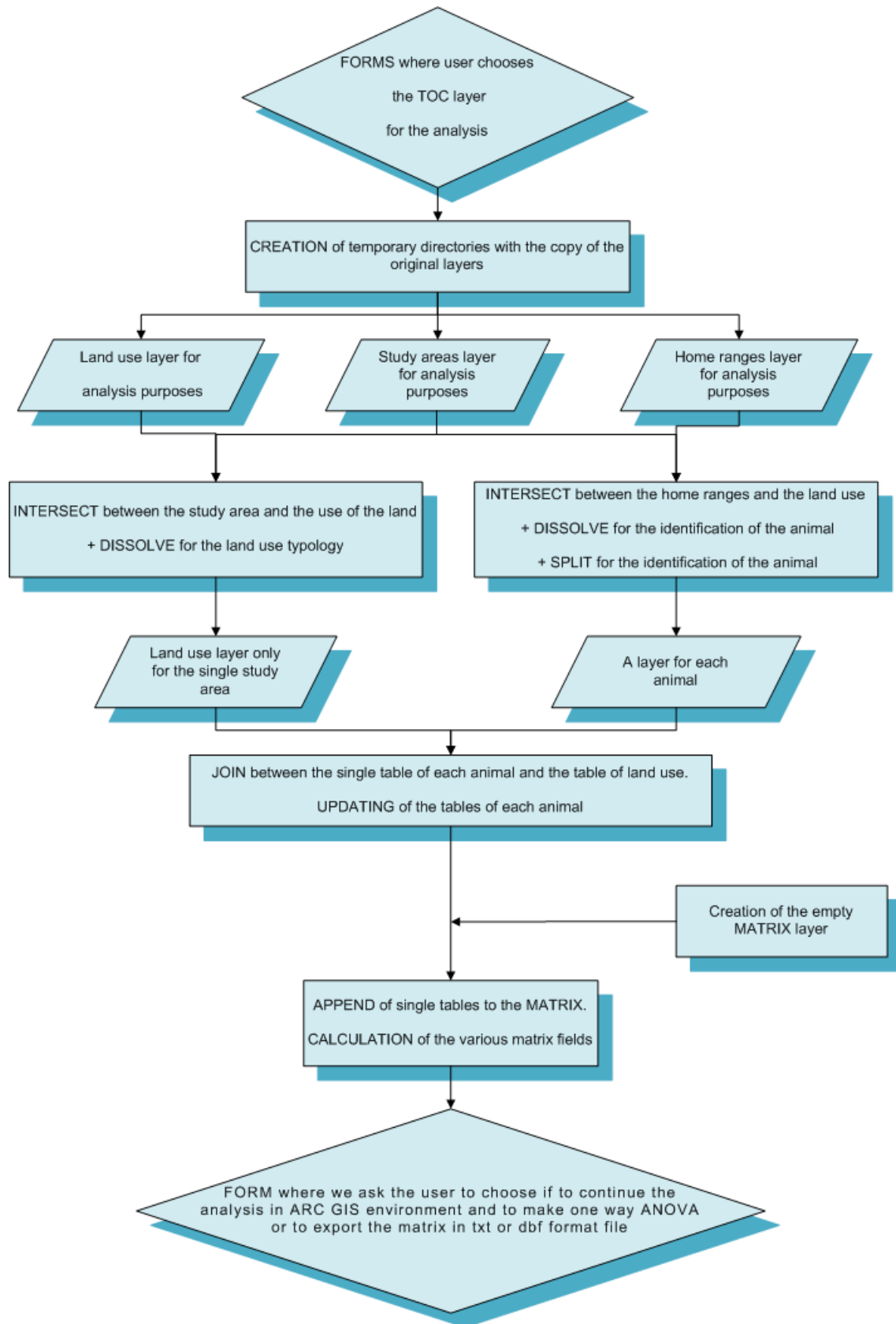
- use of the land
- home ranges of animals
- environmental availability of study area for the study of home range preferences
- animal location;

Such information is analyzed by the tool through a two-step process. The first one starts from primary data to come up with the definition of the analysis matrix. This data array can be

exported in txt or dbf file format to be used in other statistical software. In the second step is possible to run the ANOVA test directly on the newly created matrix, without leaving the ArcGIS© environment.

The flow chart of the program's work is highlighted in figure 1 and 2.

Figure 1. The flow chart of the first part of the tool



A dialog window containing 3 dropdown boxes (4 if the Fixes analysis is chosen) is displayed upon tool start-up. The purpose of each is explained to follow:

- 1<sup>st</sup> BOX: select the layer that represents the land use and the field that identify itself
- 2<sup>nd</sup> BOX: select the layer that represents the environment where the study takes place, that can be formed either by study area or by sampling on it
- 3<sup>rd</sup> BOX: select the layer that represents home ranges and the field that identify the animal
- (In the Fixes analysis) 4<sup>th</sup> BOX: select the layer that contains the fixes and the field that identify the animal

The tool creates a temporary directories where the original layers are copied, to preserve them from modifications, and the software works with the copies, named “layer for the analysis purpose”. Then it begins geoprocessing operations (Chang 2007; Gorr & Kurland, 2007):

- INTERSECT between the study area and the use of the land + DISSOLVE for the land use typology, with the creation of Land use layer only for the single study area (the related table shows a row for each typology of land use)
- INTERSECT between the home ranges and the land use + DISSOLVE for the identification of the animal + SPLIT for the identification of the animal, with the creation of a layer for each animal (the related table shows a row for each typology of land use present in the home range of that animal)

Hence on the tools works with the table layers and performs a JOIN between the single table of each animal and the table of land use UPDATING the tables of each animal so to have all the environmental typologies present in the study area, including the ones not present in the home range of the animal, that have no value.

Then the tool creates an empty matrix layer and APPENDs single tables to the MATRIX. After that, it executes a CALCULATION of the various matrix fields useful to obtain the field with logarithmic transformations of the connection between the land use present in the home range of each animal and the land use present in the study area, divided by animal. All this calculation outputs the Analysis matrix and the first stage ends.

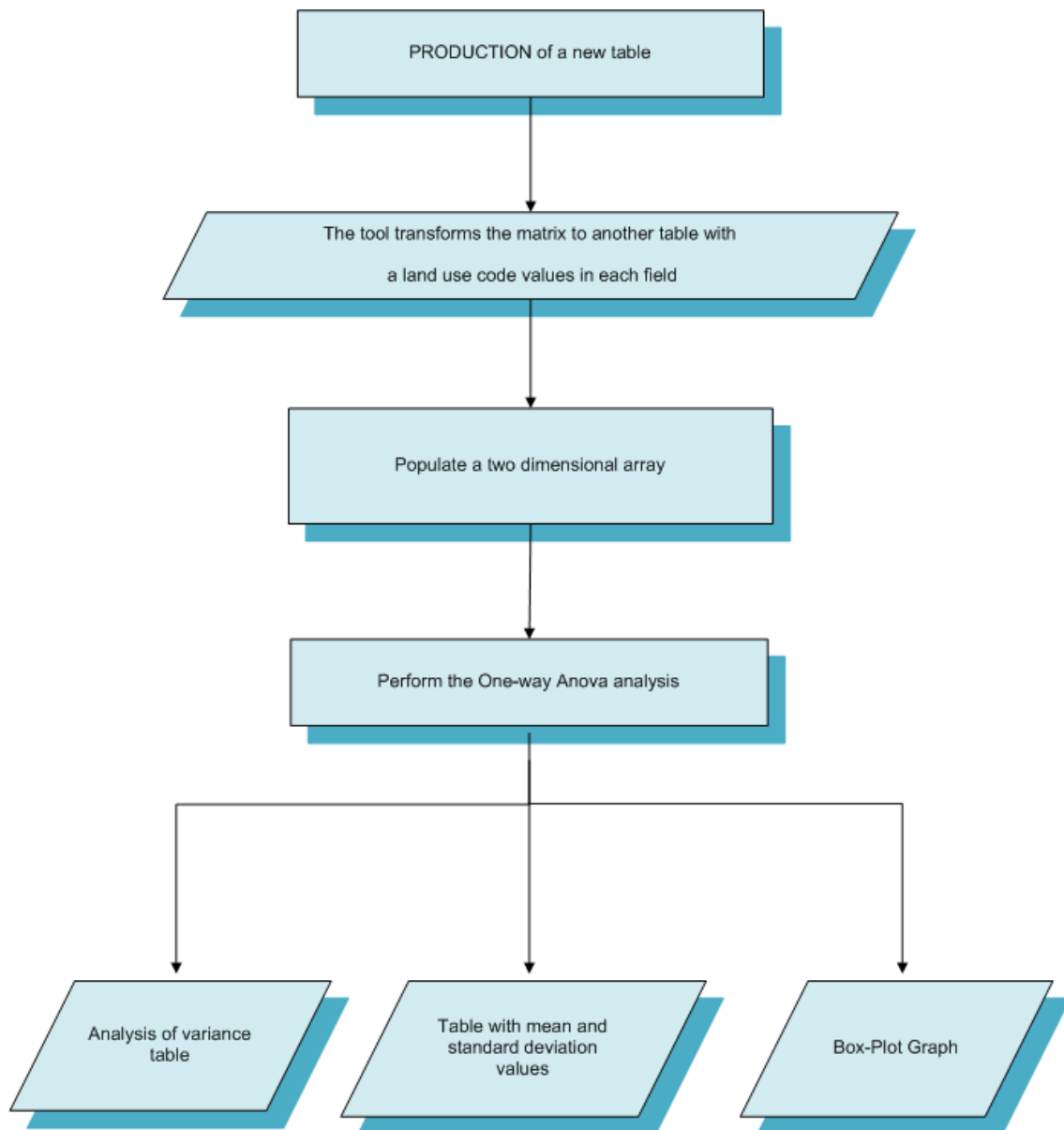
Then a new dialog window asks the user whether to continue the analysis in ArcGIS environment and to make a One-Way ANOVA or to export the matrix in txt or dbf format file and quit.

If the user chooses to continue, the tool transforms the matrix in another table with land use code values in each field. Then creates a two dimensional array e populates it. Finally the tool performs the One-Way ANOVA test, and creates the following information:

- Analysis of variance table
- Table with mean and standard deviation values
- Box-Plot Graph



Figure 2. The flow chart of the second part of the tool

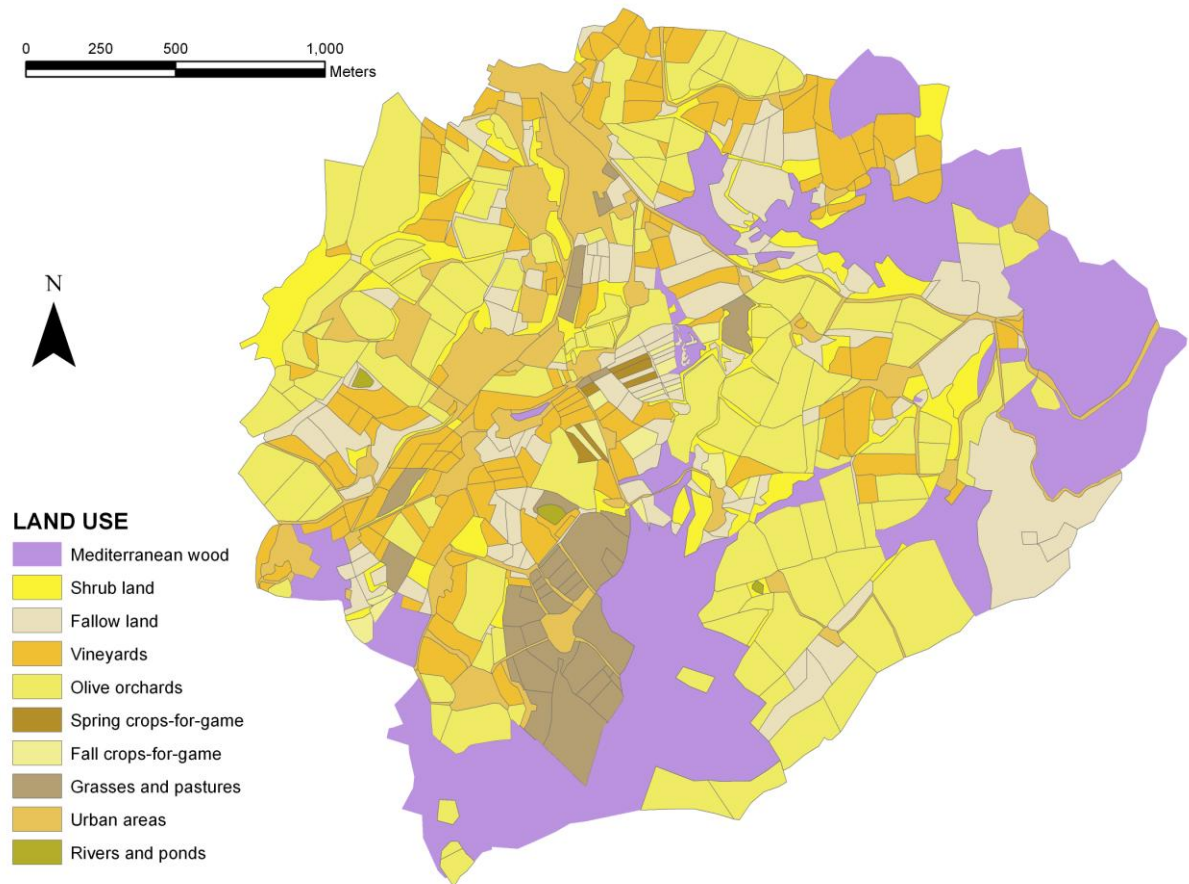


## SAMPLES OF TYPICAL PROGRAM RUNS

The current paragraph provides the mean to reach a better comprehension of the tool's workflow by describing its adoption in real data analysis. Sample information was taken from a study carried out between 2008 and 2009 on a group of pheasants (*Phasianus Colchicus*) released in a protected area (PA) near Florence (Tuscany – Italy). The interested area is used for agricultural activities and is characterized by a good habitat variety and fields of medium and small size. The crops mostly present are vineyards (*Vitis vinifera* L.), olive-yards (*Olea europaea* L.) and, secondarily, autumn cereals (mainly wheat: *Triticum aestivum* L. and oats: *Avena sativa* L.), alfalfa (*Medicago sativa* L.) and spring crops (sunflower: *Helianthus annuus* L., jowar: *Sorghum vulgare* L., and corn: *Zea mays* L.). In the PA some habitat enhancements, established essentially by the so-called “crops-for-game” (no-harvested crops) method, are also realized. These consist mainly in strips sown with mixtures of jowar, rape (*Brassica napus* L. var. *oleifera*) and sunflower in the spring and broad bean (*Vicia faba* L. var. *minor*) and wheat in the autumn. Both are left on the field without any weed control. The area is characterized by rich spontaneous vegetation.

‘Land use’ maps are a product of aerial photo interpretation and on-field survey. Obtained data are made available as a polygon shape file and loaded in ArcGIS as a layer.

Figure 3. Land use layer



The 'Land use' polygon layer is loaded in the ArcGIS Table of Content (TOC). Styling is performed for clarity as is not required for successful analysis (Figure 3).

A second layer holding the localizations of the pheasants is needed by the tool. Such data is retrieved by the study of the animal's movements which was possible since at the releasing moment they were equipped with a necklace radio-tag (Biotrak, TW3+ ½AA, 32.30 g) and an immovable leg tag.

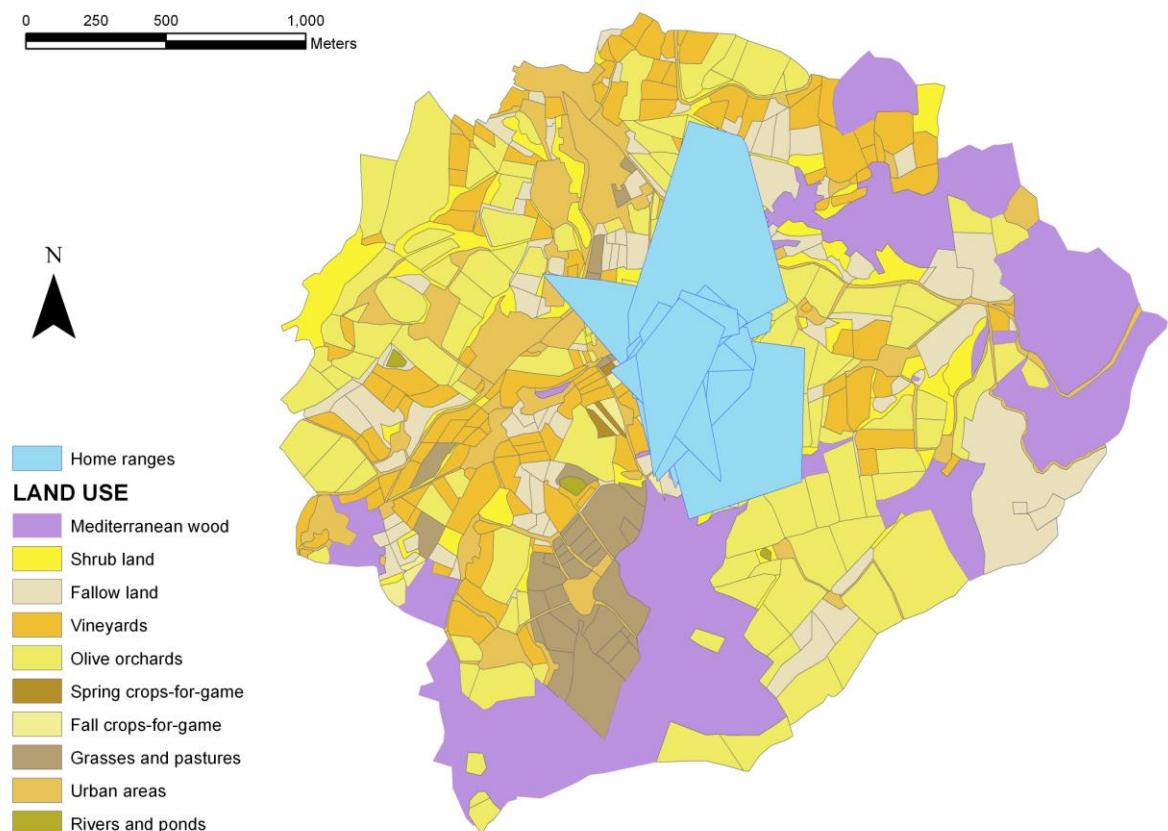
The pheasants were then localized and sighted during the day, at least 2-3 times a week in a period from October to April by two receivers equipped with four-element Yagi antennae, binoculars and a walkie-talkie (Kenward, 1993). The fixes (546) were obtained by two means. Either taking synchronous compass bearings from at least two different positions localized on the top of the hills at a range of 300-500 m or by direct sighting of the pheasants during the control of their presence in the triangulated patch. For each sight fix, the time and the land use

were registered, either on paper or on a GPS portable device (Garmin eTrex Legend navigator; 1-5 m location error). Subsequently, these data were transferred to ArcGIS with the software GPS-Utility Ltd. 1998-2006 or, for the points on paper, with the creation of a point shape file on the map. Both fixes from GPS and paper were merged in a unique points shape file. (Modern GPS collars directly output the stored position fixes as point layers, ready to be used in the GIS software. Every fix has a name and a reference to a single animal).

The home range polygon shape file of every single animal is generated with the Hawth Tool (<http://www.spatial ecology.com/htools/>).

This layer is hence added to the Table of Content (TOC) (Figure 4).

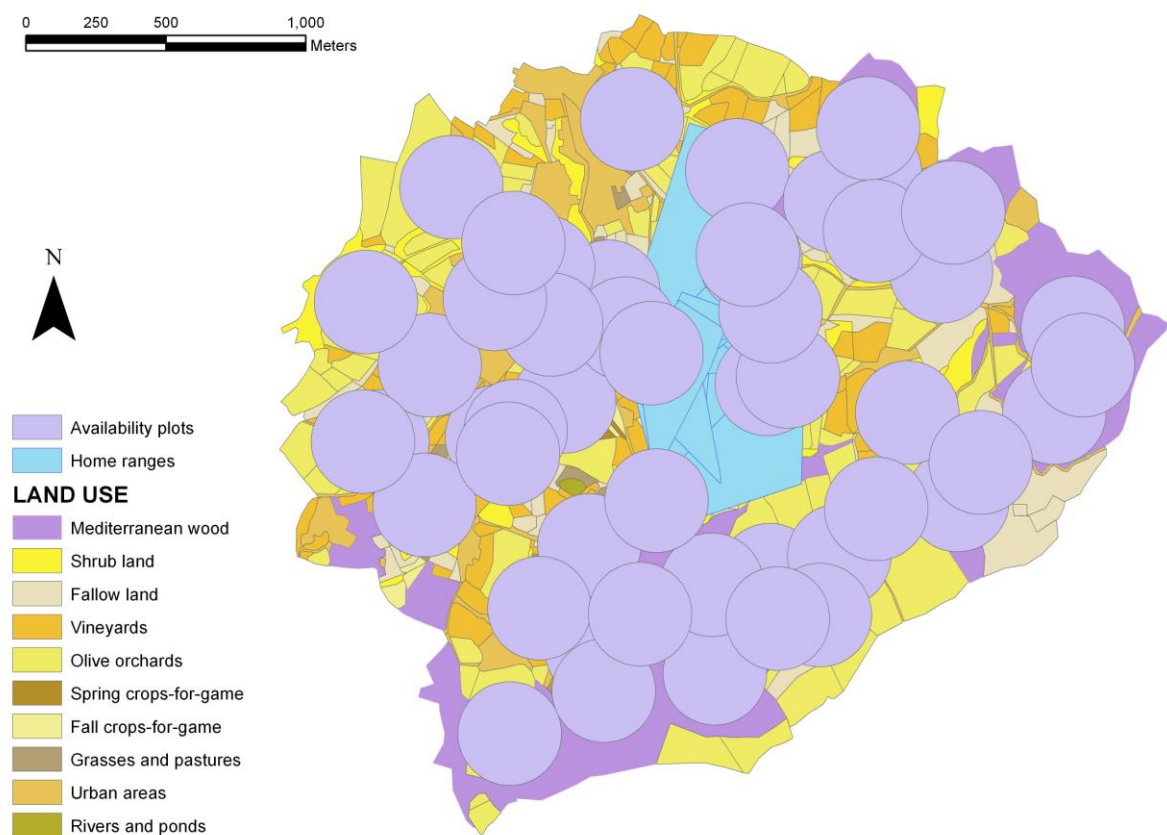
Figure 4. Land use and home ranges layers



To estimate the availability environment a number of 50 random circular plots must be chosen across the study area. Since the average home range size for the birds analysed in the study is approximately of 110 m, the same value is taken as the radius of the circular shapes. The random plots are generated using the Hawth Tool.

This layer is the added to the Table of Content (TOC) (Figure 5).

Figure 5. Land use, home ranges and availability plots layers

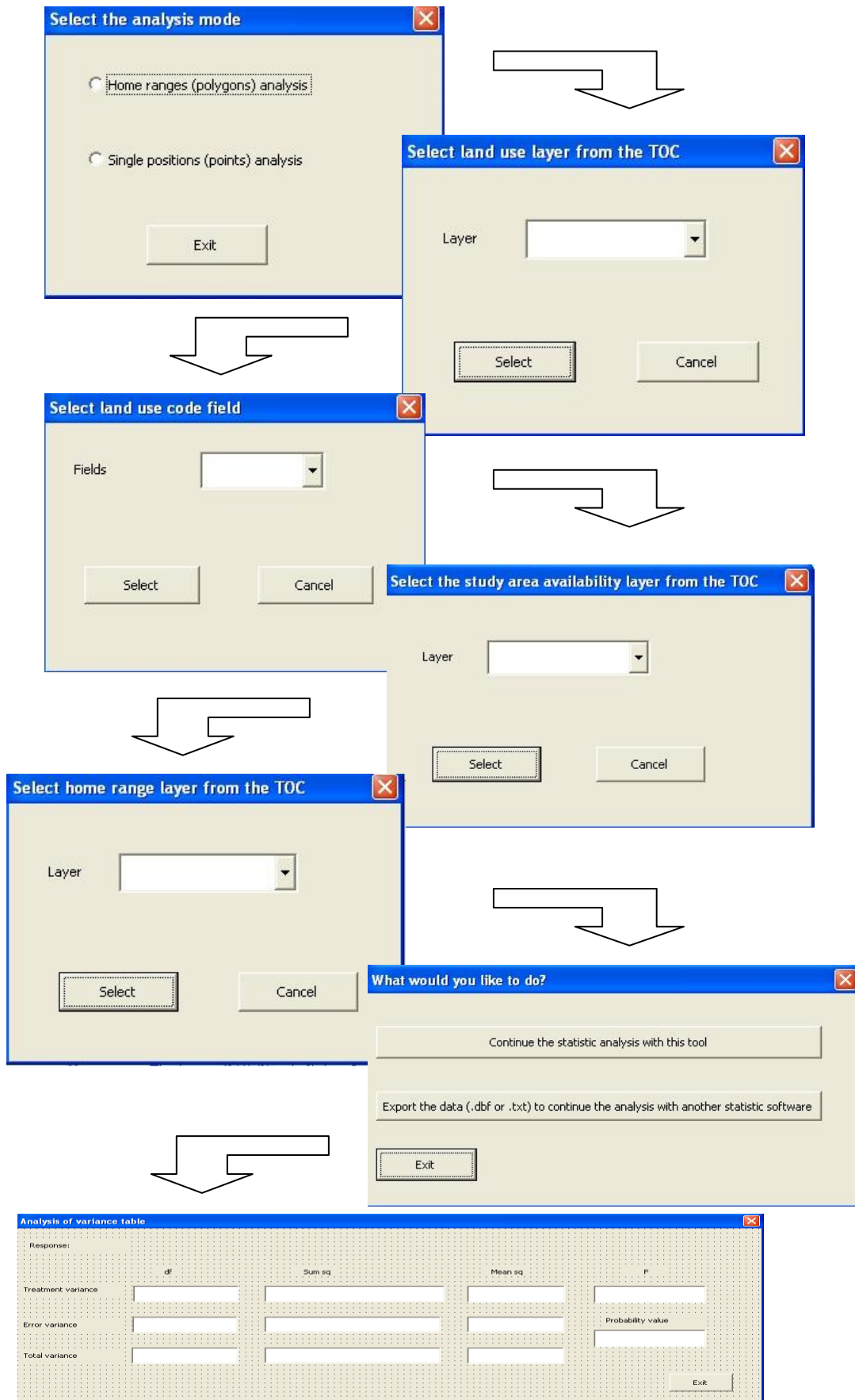


Once the 3 described layers are available in the ArcGIS environment, the Appennino Tool can be started from the toolbar. At first, the analysis mode must be chosen between ‘home ranges’ and ‘fixes’ level. For the second to be carried out, the point layer of the fixes must be added in the TOC before starting the tool.

Once the choice is done, a window containing 3 combo boxes (4 if the fix analysis is chosen) is displayed (Figure 6). Its form must be completed selecting from the drop-down boxes the appropriate layer present in the TOC. The field containing the 'land use' attribute of the land use layer and the field storing the 'animal code' of the home ranges layers must be specified.

Once all information is correctly provided, the calculation can be started by pressing the 'Start' button.

Figure 6. The FORMS of the tool



A complex matrix is given as a result, after some time of calculation activity depending on the workstation computational power. Such matrix can be exported in .dbf or .txt file and can be used for the Compositional Analysis inside a statistic software (Pendleton et al. 1998, Smith 2004) for comparison purposes with the tool's results.

The Compositional Analysis is performed during the second step of the tool workflow, which outputs an ANOVA table as result (Figure 7). This table contains averages and standard deviation values which can be automatically plotted on a chart for the user's purposes. In the described sample data, the 'land use' codes which give maximum values of occurrence are "6" (Spring crops-for-game) and "7" (Fall crops-for-game). The less utilized are the land with code "10" (Rivers and ponds) and "9" (Urban areas) (Figure 8).

In a future version of the tool, a post-hoc test will be made available along with the calculation algorithm (Brown 2005).

Figure 7. Analysis of variance table

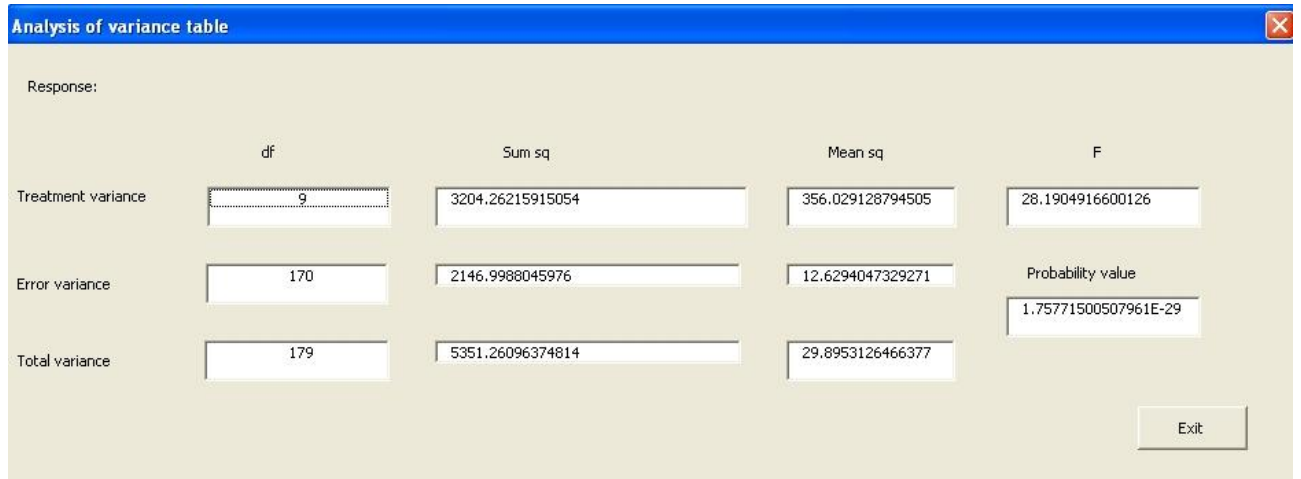
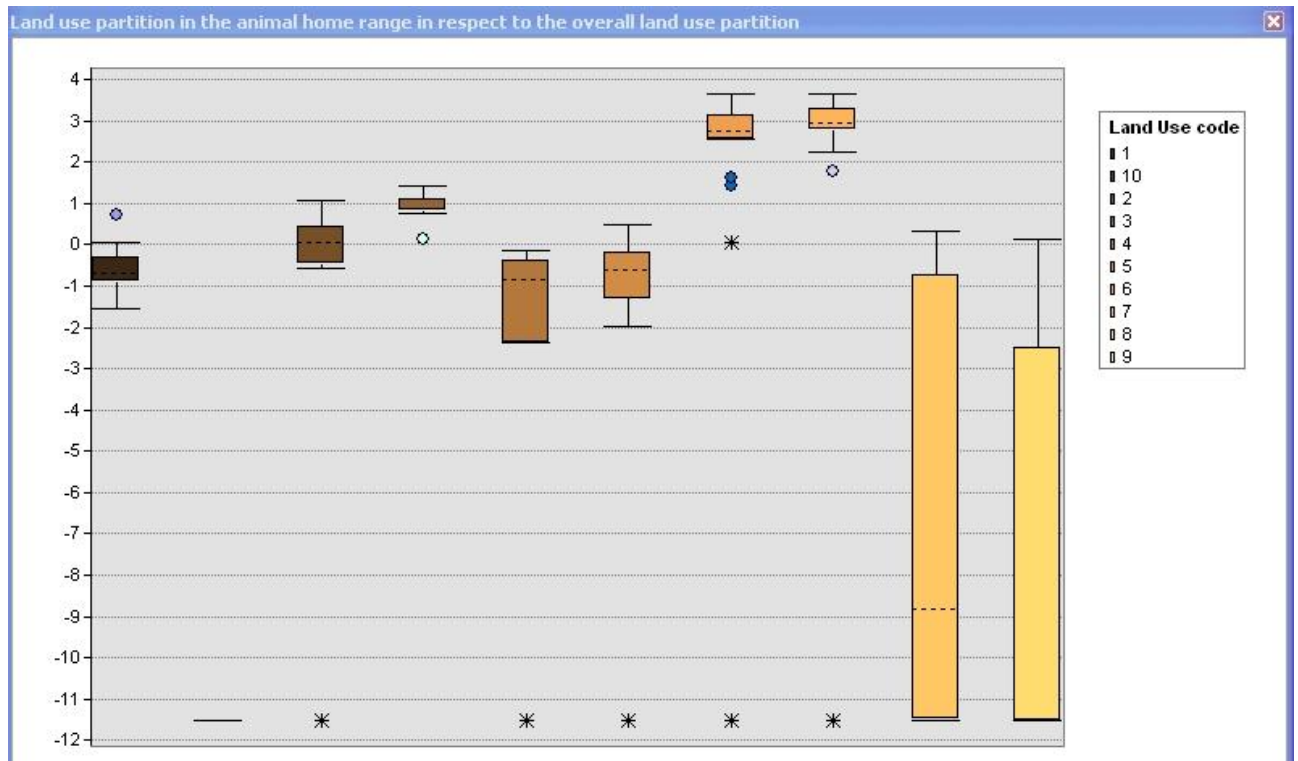




Figure 8. Box-plot graph



*Acknowledgements*

- Daniel S. Soper, Ph.D. Assistant Professor Department of Information Systems and Decision Sciences Mihaylo College of Business and Economics California State University, Fullerton
- Province of Pistoia
- Ambito Territoriale di Caccia Firenze 5

## REFERENCES

- Aebischer, N.J., Robertson, P.A., Kenward, R.E. 1993: Compositional analysis of habitat use from animal radio-tracking data. - *Ecology* 74: 1313-1325.
- Allredge, J.R., Thomas, D.L., McDonald, L.L. 1998: Survey and Comparison of Methods for Study of Resource Selection.- *Journal of Agricultural, Biological, and Environmental Statistics* 3: 237-253.
- Bradley, J.C., Millspaugh, A.C., 2003: *Programming in Visual Basic.net*. McGraw-Hill Irwin, New York, USA, 627 p.
- Brown, A.M. 2005: A new software for carrying out one-way ANOVA post hoc tests. - *Computer Methods and Programs in Biomedicine* 79(1): 85-89.
- Burke, R. 2003: *Getting to know ArcObjects: programming ArcGIS with VBA*. ESRI. 422 p.
- Chang, K.T. 2007: *Programming ArcObjects with VBA: a task-oriented approach*. CRC. 342 p.
- Fearer, T. M., Stauffer. D.F. 2004: Relationship of ruffed grouse *Bonasa umbellus* to landscape characteristics in southwest Virginia, USA. - *Wildlife Biology*, 10: 81-89.
- Ferretti, M., Paci, G., Porrini, S., Galardi, L., Bagliacca M. 2010: Habitat use and home range traits of resident and relocated hares (*Lepus europaeus*, Pallas). – *Italian Journal of Animal Science* 9 (3): 278-284.
- Gorr, W.L., Kurland, K.S.: *GIS tutorial: workbook for ArcView 9*. ESRI. 374 p.
- Hooge, P. N., Eichenlaub, B. 1997: *Animal movement extension to arcview*. ver. 1.1. - Alaska Science Center Biological Science Office, U.S. Geological Survey, Anchorage, AK, USA.
- Hurlbert, S.H. 1984: Pseudoreplication and the design of ecological field experiments. - *Ecological Monographs* 54:187-211.
- Kenward, R. E. 1992: Quality versus quantity: programmed collection and analysis of radio-tracking data. - In: Priede, I. G., Swift, S. M.; *Wildlife telemetry: remote monitoring and tracking of animals*. Ellis Horwood. Chichester, United Kingdom 231–246.

Kenward, R. 1993: Wildlife radiotagging. Equipment, field techniques and data analysis. - Academic Press, London.

Manly, B. F. J. 1993: Comments on design and analysis of multiple-choice feeding-preference experiments. - *Oecologia* 93: 149–52.

Manly, B. F. J., McDonald, L. L., Thomas, D. L., McDonald, T. L., Erickson, W. P. 2002: Resource selection by animals: statistical design and analysis for field studies. Second edition. Kluwer Academic Publishers, Boston, Massachusetts, USA, 221 p.

Pendleton, G.W., Titus, K., Degayner, E., Flatten, C. J., Lowell, R.E. 1998: Compositional Analysis and GIS for Study of Habitat Selection by Goshawks in Southeast Alaska.- *Journal of Agricultural, Biological & Environmental Statistics* 3: 280-295.

Smith, P. G. 2004: Automated log-ratio analysis of compositional data: software suited to analysis of habitat preference from radio tracking data. - *Bat Research News* 45(1): 16.