PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

A GIS open source application for cost distance modelling for site catchment analysis of prehistoric sites

L. Duarte, S. Godinho, A. Gomes, S. Rodrigues, A. C. Teodoro

L. Duarte, S. Godinho, A. Gomes, S. Rodrigues, A. C. Teodoro, "A GIS open source application for cost distance modelling for site catchment analysis of prehistoric sites," Proc. SPIE 12734, Earth Resources and Environmental Remote Sensing/GIS Applications XIV, 1273412 (19 October 2023); doi: 10.1117/12.2678993

SPIE.

Event: SPIE Remote Sensing, 2023, Amsterdam, Netherlands

A GIS open source application for Cost Distance Modelling for Site Catchment Analysis of Prehistoric Sites

Duarte, L.[a,b], Godinho, S.[c], Gomes, A.[c,d], Rodrigues, S.[c], Teodoro, A.C.[a,b]
[a]Department of Geosciences, Environment and Spatial Planning Faculty of Sciences of the University of Porto; [b]Institute of Earth Sciences, FCUP pole; [c]Department of Geography, Faculty of Arts and Humanities of the University of Porto; [d]Centre of Studies in Geography and Spatial Planning

ABSTRACT

Archaeological research employing least-cost analysis and Geographic Information Systems (GIS) has increased over the past decade. In the Portuguese archaeological research, several approaches were applied to estimating the cost-distance, describing the relationship between travel distances and their associated costs. Archaeologically, site catchment analysis (SCA) consists of a territorial analysis of the area surrounding a site that its inhabitants can exploit in terms of natural resources. There are a few generic least-cost path plugins available in QGIS software, but none of them incorporates an archaeology-specific methodology, as many GIS users combine QGIS algorithms with ArcGIS tools to achieve the desired results. This work's objective is to automate the cost-distance modelling for SCA at archaeological sites by creating a QGIS open-source application. Tobler's hiking function and a set of geographic information – slope, aspect, rivers - were implemented in the application, allowing the evaluation of archaeological hypotheses in a specific research study area. The application was tested for Northwest Portugal in a region comprised of the districts of Porto, Braga, and Viana do Castelo, where an inventory of archaeological sites from the second millennium BC until 500 BC was compiled. The creation of SCA maps for archaeological sites enabled the detection of a few differences in the composition of 1-hour walking territories. This suggests the existence of a set of similarities between these territories, despite their different topographic histories. For archaeologists, the application has benefits for automating the creation of specific maps and streamlining the process of making these models.

Keywords: archaeology sites, QGIS, open source, least cost path, least cost distance

1. INTRODUCTION

The application of Geographical Information Systems (GIS) in the archaeology area has allowed not only to store, manage, and analyse data but also to incorporate statistical analysis of archaeological sites. The creation of new types of information is also allowed, for instance, cost surface analyses, mapping of site catchment areas (SCA), and the definition of exploitation territories (ET).

As a fundamental theoretical-methodological reference in the creation of the ET, Vita-Finzi and Higss [1] promoted research on the territory, based on the principle that communities exploit the resources closer to the village, and the greater the distance of these resources from the habitat, the less likely they are to be exploited [2].

Ethnographic estimates suggest that ET in agrarian communities should be within a territory of 5 km, equivalent to about 1 hour of walking [1, 3-5]. The relationship between time and territories aims to create isochronous lines in mapping.

In this regard, this approach should be seen as one of the many procedures that allow the knowledge of the communities of the past, namely the physical characterization of the areas around their habitat spaces, regardless of the meanings that the territories may have assumed in the past. The demarcation of ET can be carried out both by manual drawing (using, for instance, charts, ruler and square) and through digital resources (for instance, GIS software). In addition to these two different forms, there are different methodologies for creating ETs, and, according to Osório and Salgado [6], the application of different methods to estimate the walking time, may lead to different results regarding the distance covered.

In Portuguese archaeology (among other research areas), this type of analysis can be stated in other studies, for instance, Costa et al. [7], Fernandes [5], Tereso [8], Sanches [9], Martins [10], Jorge [11], among others. However, these authors used different methods, namely: i) the formula from Davidson and Bailey [3] was used by Sanches [9] and Martins [10];

Earth Resources and Environmental Remote Sensing/GIS Applications XIV, edited by Karsten Schulz, Ulrich Michel, Konstantinos G. Nikolakopoulos, Proc. of SPIE Vol. 12734, 1273412

ii) the Naismith formula was used by Tereso [8]; iii) the Tobler's formula [12] was used by Costa et al. [7]; and iv) the formula from Oubiña and Álvarez [4] was used by Fernandes [5].

GIS can provide potential imagery reconstruction and information regarding ET and agricultural/humanized landscapes, in order to contribute to a better understanding of landscapes as complex mosaics where human and natural factors are articulated [8]. Therefore, it is possible to conclude that for a better understanding of patterns at a spatial level, the articulation of various scientific components encourages the sharing of knowledge, which could define new directions and enhance discovery and innovation [13-14].

An ET is defined as the area delimited by time/walking distance around a community [1,15]. In this study, the creation of the ET consisted only of an isochronous line of 1 hour of walking around the study villages. This time margin is acceptable for possible interpretations of the management of plant resources and cultivated areas.

There are a few generic least-cost path plugins available in QGIS, but none of them incorporates an archaeology-specific methodology, as many GIS users combine QGIS algorithms with ArcGIS tools to achieve the desired results [16-18]. This work objective is to automate the cost-distance modelling for SCA at archaeological sites by creating a QGIS open-source application. Tobler's hiking function and a set of geographic information – slope, and aspect - were implemented in the application, allowing the evaluation of archaeological hypotheses in a research study area. In order to optimise the creation of SCA maps, a GIS open-source application was created under QGIS software. So, the application presented in this work can contribute to the open-source community to be used by several experts in archaeological areas. The application is free and open-source, so it can be modified according to the user's requirements and adopted for different crops and regions.

2. METHODS

2.1 Study area

To test the GIS application, a region on the Northwest of Portugal was selected as a study area (Figure 1), between the Minho river and extending to the Douro catchment [19]. This region is known as the “Entre-Douro-e-Minho” and it is characterized as a mountainous area, with deeply reliefs [20]. It is also characterised by being one of the most humid regions and having the highest precipitation rates in mainland Portugal. Ribeiro and Lautensach [19] refer to the entire coastline of the Northwest with rainfall indices greater than 1000 mm, which may exceed 3000 mm in mountainous areas. It is a temperate region, humid and super humid in the mountainous regions. However, it may vary according to the relief distribution [20].

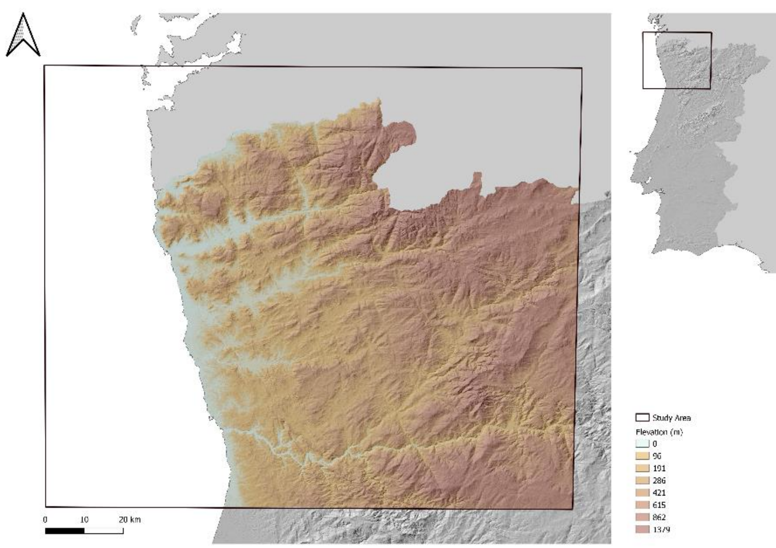

Figure 1. Location of the study area.

2.2 Data sources

In order to create the ET map, three main geospatial variables are required: i) a Digital Elevation Model (DEM); ii) the slope map; and iii) the aspect map. The DEM was obtained from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model Version 3 (GDEM 003), with 30 meters of spatial resolution. The slope and aspect maps was derived from the DEM using QGIS specific algorithms.

The 57 archaeological sites were selected from the Endovelic Portal (*Portal do Endovélico*; database of the Direção Geral do Património Cultural - https://arqueologia.patrimoniocultural.pt) corresponding to the chronology of Bronze Age. Then, only the sites categorized as "povoados" (settlements) were considered. Finally, the geographic coordinates of this selection were compiled to insert them into the software.

2.3 Cost-Distance and site catchment analysis (SCA)

Figure 2 presents the flowchart describing the steps implemented in the GIS open-source application. The cost-distance methodology implemented is composed of several steps: i) from DEM, a slope map was derived and then converted to radians; ii) Tobler's hiking function (see Equation (1)) was applied and slightly changed to calculate walking time per meter, because it is popular in raster-based cost-distance modelling or least cost analysis [21]; iii) the cost surface raster and the mandatory study sites were applied to the "Cost Distance" tool in QGIS, which aims to calculate the shortest accumulated cost-distance for each cell to the closest source over a cost surface; iv) then it is necessary to submit this cost surface raster to obtain the raster cells in minutes, multiplying the resulted raster by 60 (minutes); v) for the creation of the isochronous lines it is necessary to submit the raster (in minutes) to the tool "Contour" which allows choosing the interval of 60 for each line, representing the line of 1 hour.

$$h = 0.0001666 * (\exp(3.5 * (\mathrm{abs}(\tan(\mathrm{rad}(slope)) + 0.05)))) \quad (1)$$

where, h corresponds to the cost surface raster and the slope is the slope map derived from DEM.

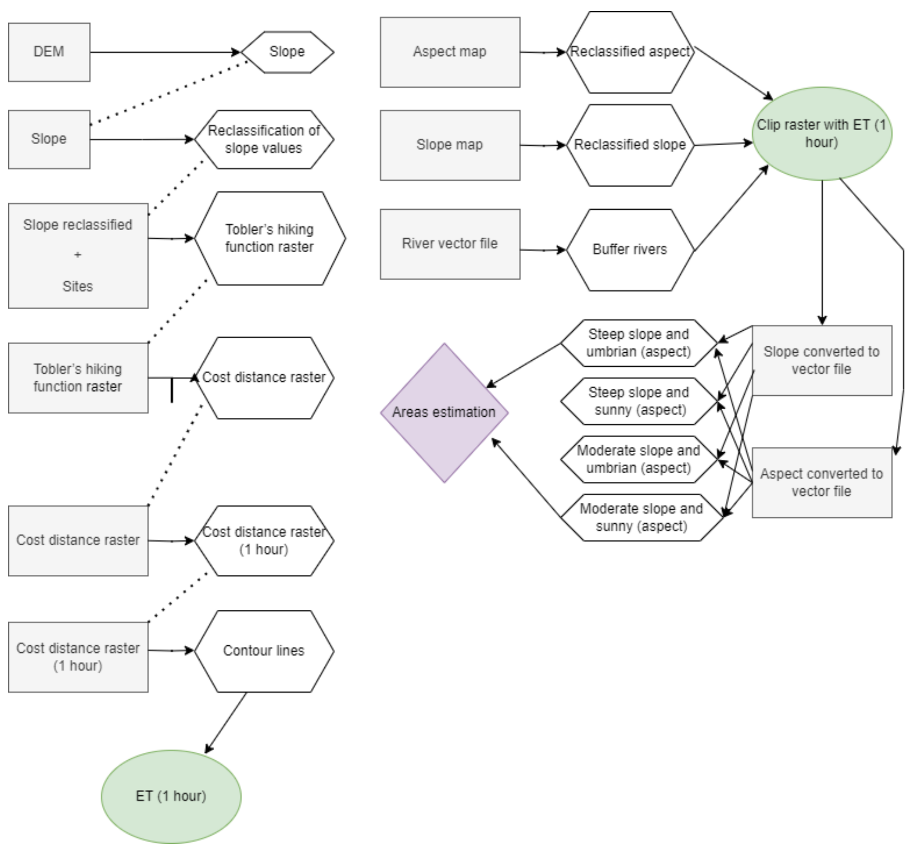

Figure 2. Flowchart of the methodology implemented in the GIS open source application.

2.4 Exploitation Territories (ET) modelling methodology

For the creation of the maps allowing to reach the distribution of wood resources and agricultural/humanized landscapes, it was necessary to select a set of geospatial data: i) elevation (DEM map); ii) slope map; and iii) aspect map. The methodology also follows a flowchart presented in Figure 1. Table 1 presents the data acquired for ET modelling cartography.

Table 1. Data acquired for ET modelling cartography.

Description	Source	Spatial resolution (m)	Coordinate system
DEM	ASTER GDEM	30	World Geodetic System 1984 (WGS 84)
Slope	Derived from the DEM (in %)	30	World Geodetic System 1984 (WGS 84)
Aspect	Derived from the DEM	30	World Geodetic System 1984 (WGS 84)

The DEM used in the ET modelling was the same as that used to create the cost-distance and SCA. The slope map in the data modeling was used to detect which areas had the most appropriate slope for the farm spaces. In this way, the slope map makes possible the selection of preference areas for resource management, assuming that the less accentuated slopes should be chosen because they are better for the imposition of agricultural practices. The slope values were classified according to the current Portuguese legislation applied by the National Agricultural Reserve (*Reserva Nacional Agrícola* (RAN)) as "moderate" (reclassified with number 1), which is interpreted as areas favorable to the application of agriculture, while the steep slopes (reclassified with number 2) are no longer very stable, due to the slope of the land, which will require terraces to support crops. The latter were treated only for "steep slopes" (Table 2).

Table 2. Slope reclassification.

Slope (%)	Slope classification according RAN	Interpretation
0-15	Moderate (1)	Moderate slope or agricultural zones
> 16	Steep (2)	Steep slope or forest

The aspect can be considered an important variable for possible interpretations of vegetation growth in the landscape as well as for taken conclusions about the best agricultural spaces. It is responsible for the amount of light and heat received, with the slopes to the north being colder and darker (umbrian) and the slopes to the south warmer and sunny (sunny) [22]. In this study, the aspect classes considered were: Flat (-1-0°), North (315°-45°), South (135°-225°), East (45°-135°) and West (225°-315°) (Table 3). However, for a simplified interpretation, the aspect values were considered throughout the work as sunny (which face south and east, reclassified with value 3) and umbrian (north and west, reclassified with values 2 and 4). In the context of flat zones, it is possible to relate their meaning to valley bottoms, flats, and the top of elevations.

Table 3. Aspect map reclassification.

Aspect (º)	Classification	Interpretation
-1 - 0	Flat (1)	Flat
0 - 45	North (2)	Umbrian
45 - 225	East and South (3)	Sunny
225 - 360	West and North (4)	Umbrian

The final ET map intends to delimit the territory potentially explored by the communities of the studied sites, which is limited to 1 hour of walking on the ground, which allows the extraction of geographic data for comparison and extrapolation of information between the studied sites.

After obtaining the described information, the final map was clipped to each territory, converted to polygon shapefiles, and intersected in this way: the sunny zones with moderated slopes, the sunny zones with steep slopes, the Umbrian zones with moderate slopes and the umbrian zones with steep slopes. The flat zones were not considered. In the end, the area of the intersected zones was estimated. This process was performed for each territory (corresponding to each archaeological site).

2.5 GIS application

The GIS application (named *QCatchementAnalysis*) was developed in QGIS version 3.28 'Firenze', using *QtDesigner* for the creation of the Graphical User Interface (GUI) [23-24] (Figure 3).

The application was developed under QGIS software, using *Plugin Builder* to create the structure of the files [19]. The *QtDesigner* framework was used to create the GUI, composed of two tabs, one for the required inputs for cost-distance and exploratory territories and a second for the respective outputs (Figure 3).

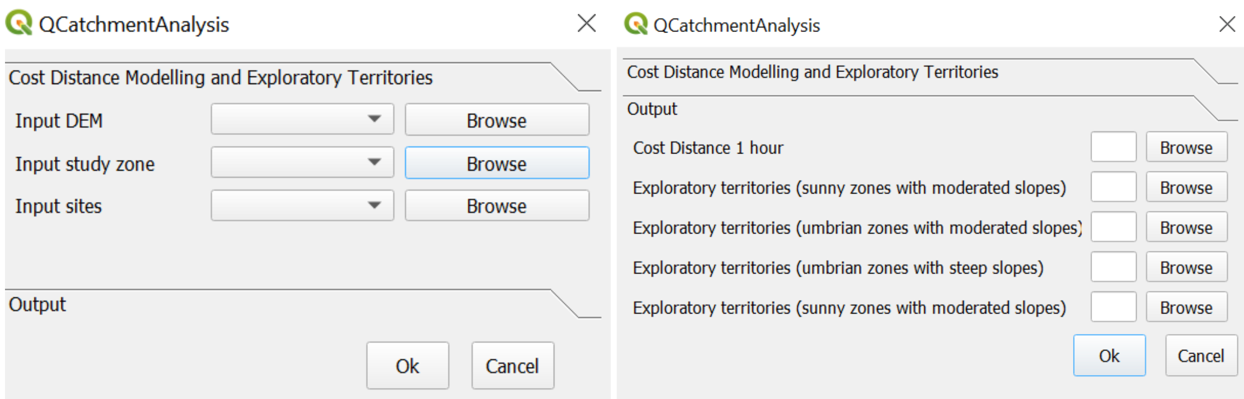

Figure 3. GUI of the developed GIS application: a) represents the Inputs tab and b) represents the Outputs tab.

The *Cost Distance Modelling and Exploratory Territories* tab allows inserting the DEM (in raster format), the study zone (in vector format), and the input sites (the archaeological sites in vector format). The application automatically clips the images to the study area extension. The *Output* tab is composed of four fields to create the cost-distance of 1 hour (vector format) and the exploratory territories (in vector format).

The developed code was developed based on the PyQt API and QGIS API libraries. Also, several algorithms from Processing Toolbox, specifically from GRASS, SAGA, and native QGIS algorithms, such as slope, aspect, raster calculator, contour, extract by attribute, and clip vector by polygon, among others were employed.

3. RESULTS AND DISCUSSION

Figure 4 presents the results obtained, namely the cost-distance and the exploratory territories (two combinations, the sunny zones with moderate slopes, defined as Exploratory Territories (1) and the umbrian zones with steep slopes, defined as Exploratory Territories (2) in the map of Figure 4) of the Coto da Pena (Viana do Castelo district) settlement. Through the results of the application, we verify that this settlement is located on a high spur, with the southern area characterised by mountainous terrain. In turn, these slightly steeper slopes could be forest or pasture areas. The exploratory territory has a total area of 31 km². The most suitable land for agriculture is found in the valley to the north, represented in the figure below by the green areas. Although these lands are mostly shady, there are sunny slopes, which would facilitate cultivation practices.

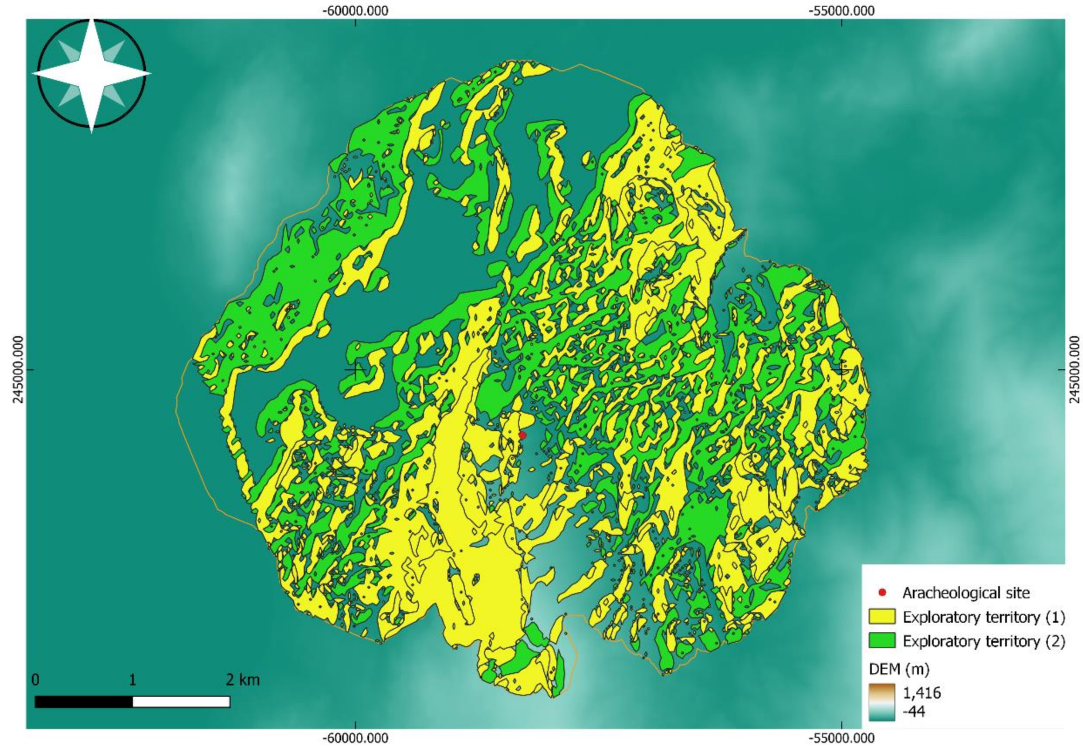

Figure 4. Coto da Pena (Viana do Castelo district) site and its exploratory territories.

The methodology implemented in the open-source application presents some advantages, such as: i) the application is free and open-source and can be adapted to any other region; ii) it is intuitive and allows to automatic detect the desired inputs or variables in the case study, reducing the steps required in the original process; iii) it performs the calculation as a mechanical process in the case of an abundance of study sites; iv) the results provide a spatial framework for understanding the relationships between archaeological sites and their surrounding landscapes, facilitating the comparison analysis.

It also presents some limitations that, in the future, will be improved, such as: i) using more high spatial resolution DEM data, we would achieve a more accurate representation of the terrain and movement; ii) limited integration of data from archaeological excavations, such as archaeobotanical information, to understand pardons between sites, or even if they existed in the surroundings; iii) lack of additional information, such as lithology maps, that can induce some bias into past interpretations; iv) and the long-term influence of human actions reflected in DEM´s.

4. CONCLUSIONS

A GIS application was created under QGIS version 3.28 'Firenze' which allows to creation of SCA maps for archaeological sites, enabling the detection of a few differences in the composition of 1-hour walking territories. The creation of SCA maps for archaeological sites enabled the detection of a few differences in the composition of 1-hour walking territories. This suggests the existence of a set of similarities between these territories, despite their different topographic histories. In the future, the hydrographic network can also be integrated as one of the study components, which is related to the possible growth of riparian vegetation and the availability of water for agriculture. This resource is also important because it is essential for human consumption. In this way, the hydrographic network can also be used for interpretations of the proximity of water lines in the study villages.

The application offers advantages for automating the production of specific maps and simplifying the process of creating these models for archaeologists. The application is free and open-source, with a wide variety of end-users, especially archaeologists. It can be modified, adapted, and improved with new algorithms. In the future, we intend to: i) continue developing ET maps to improve their representativeness by incorporating major rivers as barriers and creating zones with higher friction to reflect real-world conditions; ii) analyse topographical features like "viewsheds" to understand territorial control patterns; iii) create least-cost paths between areas to better understand communication networks. The application can be used by any user and in different regions. Since the application was developed under the open-source concept, it can be optimised and improved by any user and/or developer.

5. ACKNOWLEDGMENT

This work was funded through the Foundation for Science and Technology by FEDER funding through the COMPETE 2020 program and framed within the activities of the UIDB/04683/2020 and UIDP/04683/2020 - ICT and UIDB/04084/2020 - CEGOT financed through the European Regional Development Fund (COMPETE 2020), with ref. POCI-01-0145-ERDF-007690 and POCI-01-0145-FEDER-006891.

REFERENCES

[1] Vita-Finzi, C. Higgs, E.S. “Prehistoric Economy in the Mount Carmel Area of Palestine: Site Catchment Analysis,” Proceedings of the Prehistoric Society, 36, pp. 1–37 (1970). doi: 10.1017/S0079497X00013074.
[2] Renfrew, C., Bahn, P. G. “Archaeology: theories, methods, and practice”. 6nd ed. London, Thames and Hudson. (2012).
[3] Davidson, I., Bailey, G. N. “Los Yacimientos, sus Territorios de Explotation y la Topografia,” Boletín del Museo Arqueológico Nacional (Madrid), II (February 2016) (1984).
[4] Oubiña, C. P., Álvarez, P. F. “Diseño metodológico para el análisis localicional de asentamientos a través de un SIG de base Raster,” La aplicación de los SIG en la arqueología del paisaje, pp. 69–89 (2006).
[5] Fernandes, J. M. P. “Contributo para o estudo da metalurgia de base cobre dos povoados fortificados da bacia média do Cávado no I Milénio a. C,” (2017).
[6] Osório, M., Salgado, T. “Um Sistema de Informação Geográfica Aplicado na Arqueologia no Município do Sabugal,” pp. 9–22 (2007).
[7] Costa, A., Pacheco, E., Soares, L., Tavares, L. “O uso inteligente do território para a mobilidade na romanização.” (2014).
[8] Tereso, J. P. V. (2007) “Paleoetnobotânica do povoado romano da Terronha de Pinhovelo (NE transmontano),” p. 188 (2007).
[9] Sanches, M. J. “Pré-História Recente de Trás os Montes e Alto Douro’. Porto: Sociedade Portuguesa de Antropologia e Etnologia,” (1997).
[10] Martins, M. “O Povoamento Proto-Historico e a Romanização da Bacia do Curso Médio do Cávado’. Braga: Cadernos de Arqueologia – Monografias,” (1990).
[11] Jorge, S. O. “O povoado da Bouça do Frade (Baião) no quadro do Bronze Final do Norte de Portugal,” Monografias Arqueológicas, 2. Porto: GEAP (1988).
[12] Tobler, W. “Three Presentations on Geographical’, National Center for Geographic Information and Analysis,” (1993).
[13] van Raan, A. F. J. “The use of bibliometric analysis in research performance assessment and monitoring of interdisciplinary scientific developments,” Technikfolgenabschätzung-Theorie und Praxis/Technology Assessment-Theory and Practice, 1(12), pp. 20–29 (2003).
[14] Soares, L., Costa, A., Gomes, A. “Geografia, arqueologia e sistemas de informação geográfica: exemplos prospectivos de articulação,” Variações Paleoambientais e Evolução Antrópica no Quaternário do Ocidente Peninsular (2010).
[15] Vita-finzi, C. and Higgs, E. “Archaeology: the key concepts,” (2005).
[16] Duarte, L., Espinha Marques, J., Teodoro, A.C., “An Open Source GIS-Based Application for the Assessment of Groundwater Vulnerability to Pollution,” Environments, 6, 86 (2019).
[17] Duarte, L., Teodoro, A., et al., “Distributed Temperature Measurement in a Self-Burning Coal Waste Pile Through a GIS Open Source Desktop Application,” ISPRS Int. J. Geo-Inf., 6, 87 (2017).
[18] Duarte, L., Teodoro, A.C., Sousa, J.J., Pádua, L., “QVigourMap: A GIS Open Source Application for the Creation of Canopy Vigour Maps,” Agronomy, 11, 952 (2021).
[19] Ribeiro, O., Lautensach, H. “Geografia de Portugal,” 1a edição. Edited by E. J. S. da Costa. Lisboa (1987).
[20] Bettencourt, A. M. S. “A Paisagem e o Homem na Bacia do Cávado durante o II e o I milénios AC,” U.Minho (1999).

[21] Becker, D. de Andrés-Herrero, M., Willmes, C., Weniger, G.-C., Bareth, G., "Investigating the Influence of Different DEMs on GIS-Based Cost Distance Modeling for Site Catchment Analysis of Prehistoric Sites in Andalusia," ISPRS International Journal of Geo-Information, 6(2), p. 36 (2017). doi: 10.3390/ijgi6020036.
[22] Manuel, C. "Modelação de nichos eco-culturais para o Paleolítico Médio" U. Porto p.58 (2015).
[23] Qgis, A Free and Open Source Geographic Information System, 20 April 2022, < https://qgis.org/en/site/> (20 April 2022) https://qgis.org/en/site/.
[24] QtDesigner, 20 April 2022, < https://doc.qt.io/qt-6/qtdesigner-manual.html> (20 April 2022) https://doc.qt.io/qt-6/qtdesigner-manual.html.