Please cite this article as: B. Zaragozí, P. Giménez-Font, Antoni Belda-Antolí & Alfredo Ramón-Morte. A graph-based analysis for generating geographical context from a historical cadastre in Spain (17th and 18th centuries), Historical methods, https://doi.org/10.1080/01615440.2019.1590269

A graph-based analysis for generating geographical context from a historical cadastre in Spain (17th and 18th centuries)

B. M. Zaragozí[a], P. Giménez-Font[b] A. Belda-Antolí[c] and A. Ramón-Morte[b]

[a]Dept. de Geografia, Universitat Rovira i Virgili, Vila-seca, Spain;
[b]Instituto de Geografía, Universidad de Alicante, Alicante, Spain;
[c]Dpto. Ciencias de la Tierra y Medio Ambiente, Universidad de Alicante, Alicante, Spain

ABSTRACT

The *cabreves* are notarial documents prepared between the 13th and 19th centuries in the Catalan and Valencian regions of Spain. These historical records were published before the first cadastral maps and contain geographical information that could help spatially reconstruct historical landscapes. However, these documents have not been used to their full potential mainly because of their semi-structured and complex nature. In this article, we propose a new graph-based interactive methodology for partially reconstructing historical landscapes. We have successfully applied this methodology for reconstructing the historical landscape of the Barony of Sella in the 18th century and the methodology has also helped us locate "*El Poblet*", a previously unknown archaeological site abandoned after the expulsion of the Moors in 1609.

KEYWORDS

cabreve; graph theory; historical geography; landscape reconstruction

1. Introduction

European states started producing detailed cadastral surveys of their territories around the 19th and 20th centuries. There is no adequate cartographic information on which to carry out studies related to the historical landscape prior to then, and this makes some relatively recent periods and territories difficult to delimit. For this reason, making maps based on the available textual information constitutes a crucial challenge for geo-historical researchers (García Juan, Álvarez Miguel, Camarero Bullón, and Escalona Monge, 2012).

1.1. Pre-cadastral sources of geographic information

During the transition from the Middle Ages to the Modern Age, more specifically before the 19th century, many Western European countries experimented a progressive and complex process of property consolidation (Congost, 2003; Hoofs, 2010). This process encouraged the creation of notarial documents and property registers now

CONTACT B. M. Zaragozí. Email: benito.zaragozi@urv.cat

kept in public and private archives. These documents stand out for their diversity and functionality and are fundamental sources of semi-structured geographical information (Domingo Pérez, 2018). Although there are remarkable precedents, such as the Domesday Book in England (11th century), the documentation generated between the 17th and 19th centuries in Europe is of special significance for the large number of records and the richness of detail. During this period, feudal rights were breaking down and being replaced by "fine-paying tenants" (*emphyteutic* tenants in Italy, Spain, and France – and copyholders in English), in line with new ideas of ownership that brought an end to feudal rights in Europe (Barbot, 2015). A specific example of this process in Spain is described in detail by Gómez Benedito (2018) in his paper "Administrative reforms in the Valencian domains of the House of Medinaceli during the second half of the eighteenth century".

We propose to evaluate in this paper the case of the Principality of Catalonia and the Kingdoms of Valencia, Aragon and Mallorca (former Crown of Aragon, that existed alongside the Kingdom of Castile in modern Spain), with predominance of the *emphyteusis or copyhold lease*. This legal formula was used to favour the colonization of the territory, settling population and expanding agriculture through the figure of the "*establiment*" or emphyteutic copyhold lease contract. This contract divided a property into two parts: (1) *dominium directum* (freehold) and (2) *dominium utile* (copyhold), held in perpetuity by the copyholders and their descendants in exchange for an annual payment (Congost, 2003). In the case of the Kingdom of Valencia (a Christian kingdom created in the 13th century by settlers from Aragon and Catalonia after the conquest of part of al-Andalus), the copyhold lease contract served to colonize the territory, of which 75% was under feudal jurisdiction (García-Borbolla, Gil Olcina, and Abad Navarro, 1988). However, its main implementation took place in 1609, after the expulsion of 130,000 Moriscos that caused the depopulation of many rural areas and the need to recolonize. The conditions offered to attract potential copyholders were beneficial for undertaking costly agricultural work, such as the construction of cultivation terraces or planting of trees. The prospect of perpetuity convinced settlers to initiate projects from which their descendants could benefit, something that did not happen in the time-limited land leases that prevailed in the south of the Iberian peninsula (Naranjo-Ramírez, 1992). After the expulsion of the Moriscos, agrarian growth towards the end of the 17th and the 18th centuries brought about the diffusion of the emphyteutic copyhold lease system so that many of the inhabitants of these territories held such contracts with feudal lords.

The significance in the Valencian and Catalan case is greater than that in other regions as historical land tenure regimes had an impact on the landscape and organization of the territory (Congost, 2007; Gil Olcina, 2012; Romero González, 1983). With the liberal revolution of the 19th century, the political and economic pressure of the *emphyteutic copyholders* and the weakening of the nobility's land claims caused a consolidation of the copyholds so that they eventually became freeholds, in correspondence with the massive redemption of lordly rights. In this way, hundreds of thousands of emphyteutic copyholders became freeholders, while their lands maintained the original parcel structure, only to be fragmented or multiplied by land sales, and subsequent hereditary partitions. As evidence of the development of this trend, it should be noted that the nobility had claim on significantly less land at the end of the 19th century than at the beginning, when they claimed three-quarters of the total land of the Kingdom.

Between the 13th and 19th centuries, emphyteutic copyhold leases in the Crown of Aragon were documented through historical documents of great value: the *cabreves* in

Spanish or *capbreus* in Catalan. These documents operated as procedures to control the economic production of the territory (Congost, 2003; Gil Olcina, 2012) and resulted in numerous notarial documents in Catalonia and Valencia (Congost, 2009; García Trobat, 2001). In these documents, the emphyteutic copyholders acknowledged the rights of their direct lords over the land that the former held in copyhold (Gil Olcina, 1998).

The *cabreves* contain very detailed and rigorous information because the economic interests demanded control of the land and manor houses (Domingo Pérez, 2018). The *cabreves* offer an exhaustive inventory of the emphyteutic copyholds, as well as the produce of each parcel. This is invaluable data from the spatial point of view and for landscape reconstruction (Benítez, 1984). The documents contain a description of the assets affected together with data about location, area, and limits, and also specify the parcel boundaries and tax obligations. Although *cabreves* can be considered a type of cadastre at the service of private individuals (owners) or the Crown, their main objective was to both control the assets and provide legal security to the owners of the freehold with respect to those copyholders who were using the land. For this reason, they lack cartographic references. The original research aim was to develop a series and study changes in ownership (structure, distribution and transmission), size of the plot, type and changes in crops, and toponymical inventory (Benítez, 1984; Domingo Pérez, 2018; Gómez Benedito, 2018). However, the quality of the data is such that it enables a reconstruction of the agrarian landscape based on the permanence of some structural elements and a comparison with the current territory.

1.2. Historical geographic information systems

In recent decades, historical geographic information systems (HGIS) have been receiving a great deal of attention from geographers and historians for their versatility. These technologies can be applied in different contexts, such as geo-referencing old maps, administrative boundary reconstruction, understanding natural processes, microdata geo-referencing, or historical-versus-current statistical data comparison (DeBats, 2008; Gregory and Healey, 2007; Hunter, 2010; Knowles, 2014; Turner, Bolòs, and Kinnaird, 2017). For example, Hin, Conde, and Lenart (2016) use HGIS to distinguish individuals who appear in more than one Egyptian census or register at the time of the Roman Empire, providing an appropriate interpretation of certain historical documents. In the study of the Roman world, the Orbis project, maintained by Stanford University, constitutes a high-quality digital resource for the study of connectivity through the transport network of the Roman Empire around 200 AD (Scheidel, 2014).

HGIS projects may vary in terms of scope, objectives, technological approach, and the type of document studied. For example, Schlichting (2008) describes case studies that range from a historical GIS of China (CHGIS) covering 2000 years of Chinese history, a study of the Dust Bowl in the United States, to even more specific projects such as an HGIS of the Battle of Gettysburg. Moreover, García Juan et al. (2012) present notable experiences and results in Spain: such as the study of the Marques de la Ensenada cadastre developed in the SIGECAH project (García Juan et al., 2012), and Femenia-Ribera, Benitez-Aguado, Mora-Navarro, and Martinez-Llario (2014) in which a methodology was proposed for recovering municipal boundaries in the Province of Valencia using historical cadastral maps (Femenia-Ribera et al., 2014).

Data obtained from historical archives is usually related to places that can be found on a map. HGIS applications work on the premise that there is always data, such as postal addresses, reference cartography, or recognizable place names that can be tracked down in historical documents. Based on this premise, a geographic information system (GIS) becomes a powerful tool for the visualization, organization, management, and spatial analysis of historical geographical data (Carrion, Migliaccio, Minini, and Zambrano, 2016; Jiang and Hu, 2018). This process, commonly known as geo-parsing, has two steps. The first step consists in automatically recognizing the place-names within a historical document, and the second step pairs these place-names with spatial data from a gazetteer, such as the GeoNames database or The Getty Thesaurus of Geographic Names (Clifford, Alex, Coates, Klein, and Watson, 2016; Gregory, Donaldson, Murrieta-Flores, and Rayson, 2015; Manguinhas, Martins, and Borbinha, 2008; Mostern, 2008). In many cases, however, there may not be a proper gazetteer to which to refer. Over time, the urban fabric may have changed so much that even the toponyms may have been lost. This problem becomes more complex in historically less populated areas, in which the references can either be minimal or just mere conjectures. In such cases, reconstruction of the spatial pattern from the available information is vital, but the process of reconstruction is time and resource consuming. Therefore, automation of the process becomes highly desirable (Schlichting, 2008). Automation requires careful discrimination between what can run on its own and what aspects need expert mediation when dealing with the uncertainty associated with incomplete information and its ambiguities. Many authors prefer to use semi-automatic methods for the compilation of data in an HGIS, given that manual methods are very time-consuming, and completely automatic methods can generate false relationships (Hin et al., 2016).

The problem of data discontinuities in historical series is accentuated in pre-cadastral sources, such as the *cabreves*, where the percentage of identifiable place names and known references is much lower than in other historical documents. In such cases, common GIS methods, such as geo-parsing, are not expected to offer reliable results, and other techniques may be necessary to deal with a higher degree of spatial and temporal uncertainty. In this paper, we explore the use of graph theory to analyze the topological information hidden in the *cabreves*.

1.3. Modelling past landscape structure with graphs

Graph theory is the branch of mathematics concerned with the study of graphs. It dates to 1736 when Leonhard Euler published a well-known paper in which he studied the general case of the Königsberg Bridge problem and where his original graph consisted of a geospatial data problem (Robinson, Webber, and Eifrem, 2015). Exactly 200 years later the first book on graph theory was published (see Konig, 1936). Since then, graph theory has developed into an extensive and popular branch of mathematics that is applied to a wide range of problems in fields such as mathematics, social sciences, physics, biology, computer science, and geography (Abedin and Sohrabi, 2009; Donato, 2017; Heal, Bartlett, Wood, Thomson, and Woolfson, 2018; Katambi, Jiming, and Xiangyuan, 2002; Mackaness and Kate Beard, 1993).

The definitions of graphs vary, but put simply, a graph is a mathematical structure used to model pairwise relations between objects. Graphs are made up of nodes, which are connected by edges. Depending on the application context, the nodes are called vertices, points, or objects, while the edges can be referred to as arcs, lines, or relationships. Nodes and edges can be used to represent different object attributes (such as names and weights) and relation types (such as directed, undirected, and weighted). More specifically, the term network or spatial network is commonly used in GIS routing applications (Fischer, 2006) and there are very well-known GIS tools or tool extensions available based on graph theory (such as ESRI ArcGIS Network Analyst Extension, QGIS Road Graph plugin, pgRouting).

Graph theory techniques have been used in archaeological research for many decades, but their main use has been to visually compare results and explicitly address interactions between people, data, and places (Brughmans, 2013). Graph theory has been successfully applied to develop a wide range of diverse research topics in archaeology (Collar, Coward, Brughmans, and Mills, 2015); visualize archaeological data (Hart and Engelbrecht, 2012; Weidele, van Garderen, Golitko, Feinman, and Brandes, 2016); analyse complex databases and infer missing relationships (Amati, Shafie, and Brandes, 2018; Brughmans, 2010); test hypotheses (Brughmans and Brandes, 2017; Brughmans, Keay, and Earl, 2014; Knappett, Rivers, and Evans, 2011) describe changes in settlements and their morphology (Bevan and Wilson, 2013; Szmytkie, 2017); study street network changes over time (Barthelemy, Bordin, Berestycki, and Gribaudi, 2013; Lagesse, Bonnin, Bordin, and Douady, 2016); and analyse transportation networks (Ducruet, Cuyala, and El Hosni, 2018; Isaksen, 2008; Knappett, Evans, and Rivers, 2008; Knappett et al., 2011). A further relevant application has been to explore spatial relationships hidden in historical documents. For example, Kendall (1971) – almost five decades ago – calculated with graphs the hypothetical locations of villages. This approach is a variant of the geo-parsing method supported by a dataset containing the spatial location of known villages. The logic applied in this type of study is clearly explained by Collar et al. (2015). These examples, among many others, show that graph theory is a viable framework for modeling past settlement and landscape structure from incomplete evidence (Bevan and Wilson, 2013).

In this paper, we show that discontinuities in the historical record can be partially filled by extracting the abundant geographical and topological information that is hidden in pre-cadastral documents. We then propose a new methodology capable of extracting and analyzing the topological information in the *cabreves*. The following research objectives facilitate the achievement of this aim:

(1) Design an extensible graph data model for extracting implicit relationships from the *cabreves*.
(2) Propose a reproducible methodology based on free and open source software readily reusable in other projects.
(3) Test this methodology through a feasible case study by creating testable conjectures and hypothesis.
(4) Validate the proposed methodology analyzing either historical sources or performing fieldwork.

2. From the *cabreve* to a graph

In this section, we propose a new methodology to extract and analyze the geographical information contained in the *cabreves*. Our method consists in a semi-automated workflow where experts: a) interpret the *cabreves*; b) extract the contained geographical information; and c) code it in YAML format (from the English acronym "YAML ain't a markup language'). In this way, a YAML file is obtained that enables process automation: such as graph creation and visualization of the structure of property described in these historical documents. Finally, these graphs generate different con-

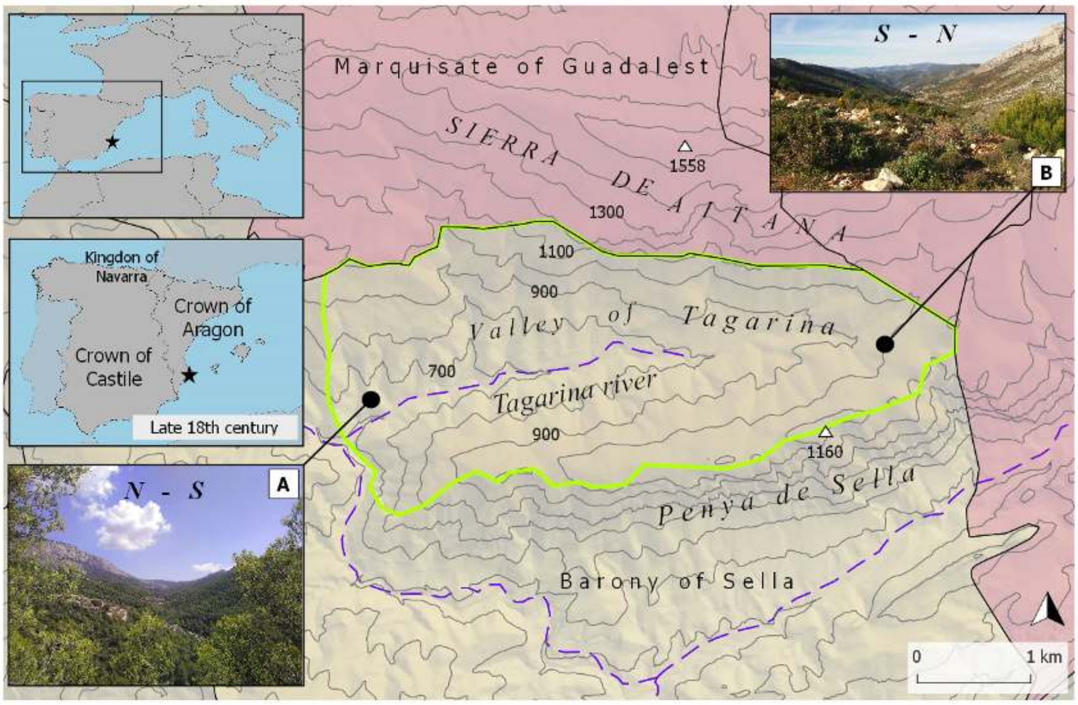

Figure 1. Location of the study area in the southeast of the Iberian Peninsula and some context for the Tagarina valley.

jectures about the structure of property and the intensity of the agrarian colonization of a territory when the *cabreves* were created.

2.1. The Tagarina valley cabreve (1726)

In our first case study, we have applied our methodology to a single rural district (*district*; see Table 1) in the Barony of Sella in 1726 (Kingdom of Valencia). We chose this study area because we obtained a copy of the original *cabreve* and digitized the transcription in the form of text files. The original document encompasses the entire Barony of Sella, but for practical reasons, we analyzed a smaller dataset, a more specific study area, where validation of our methodology is feasible.

The *Cabreve* of Sella dates back to 1726 and is guarded in the Provincial Historical Archive of Alicante (Spain). It contains 214 pages and the names of appear of 133 emphyteutic copyholds of the Barony of Sella (held by the Calatayud family). This family still holds the barony title, although following the previously mentioned abolition of feudal rights in the mid-19th century, they do not hold any land in our study area. We then limited our study to the rural district of the Tagarina valley, which comprises of 42 emphyteutic copyhold leases and 65 parcels.

The Tagarina valley is a popular area for researchers because it is rich in toponymy, communications, agricultural and historical livestock infrastructures, archival documentation, and oral memory (Doménech García, Giménez-Font, and Llorca Ibi, 2007; Giménez-Font and Marco Molina, 2017). Tagarina is an intra-montane depression located between the peak of the Sierra de Aitana (1,558 m a.s.l.) and the Penya de Sella (1,160 m a.s.l.). It covers approximately 1000 hectares, and channels a small river course fed mainly by karstic surges (see Figure 1). Land use has historically been restricted to the bottom of the valley and cultivated hillside terraces. Until the 17th century, the Moors occupied the Tagarina valley. They cultivated the valley floor with small irrigated terraces and used the slopes for livestock and forestry. After the expulsion of the Moriscos in 1609, it took several decades to reoccupy this territory with Christian settlers (emphyteutic copyholders) who were then governed by a new agreement with the baron. During the 18th and 19th centuries, population growth and feudal requirements forced the hillsides to be transformed into farmland wherever possible. The uncultivated lands were used for forestry and pasture. Peak agricultural expansion was reached in Tagarina at the end of the 19th century. In the middle of the 20th century, a massive abandonment of lands took place giving rise to pine and holm oak forests and natural regeneration (Giménez-Font and Marco Molina, 2017). Forests now dominate the formerly agrarian landscape, and the valley is environmentally important. The EU Habitats Directive lists it as a Site of Community Importance.

This small area shares the same history as most European Mediterranean mountainous regions (Agnoletti, 2014; McNeill, 1992), with the particularity of having experienced recent colonization, led by the emphyteutic copyholders, who generated a new agrarian landscape and new settlement patterns. Even though most of the Moorish farms were reallocated to new copyholders, a few were abandoned because of the new territorial organization. Some old Moorish settlements have remained relatively unchanged and become important archaeological sites (Eiroa, 2012; Torró Abad, 2003).

A study of the written documentation reveals that there were several settlements in the Tagarina valley. There is a reference from 1791 to an abandoned Moorish village (Galiana, 2015), but it lacks data about its precise location. The *Cabreve* of 1726 references the place-name as "*Poblet*" (i.e. a hamlet in Catalan) as an element adjacent to one of the 65 parcels located in the valley, so it would be possible to infer some sort of spatial reference. Based on this information, we considered that an attempt should be made to locate this settlement by means of a graph that would allow a spatial or visual analysis of the relations between the copyholders that gives an approximate location of the possible archaeological site.

2.2. Design of the labeled property graph data model

The *Cabreve* of 1726 is a semi-structured data source, which means that it requires pre-processing before providing useful answers to queries. Finding the correct answer to our queries depends on a well-designed data model. In this section, we present a data model for the *cabreve* structure which has been used for storing and querying selected information and for analyzing and allowing the definition of new nodes and relationships that must be inferred.

The labeled property graph is the most popular form of graph model. According to Robinson et al. (2015), a labeled property graph must have the following characteristics:

- It must contain nodes and relationships.
- Nodes must contain properties (key-value pairs).
- Nodes must be labeled with one or more labels.
- Relationships can be named and directed, and must then have a start and end node.
- Relationships can also contain properties.

There are other considerations that could be made about this type of model, in-

Table 1. List of node types considered and short description.

Stereotype	Node type	Description
Aggregation	Level2	District, name given to a rural district consisting of a cluster of former baronial properties (Level1).
	Level1	A cluster of plots defined as a former baronial property cluster, showing an inheritance relationship with a missing node (a hypothetical ancestor).
Plot	AgriculturalPlot	The minimal spatial unit described in the *cabreve*, delimited by its neighbours plots. In this model, the plot comprises an area unit of approximately 4358 square metres known as a "*jornal*"[a].
Person	Copyholder	Emphyteutic copyholders who hold the lease of one plot at least.
	Neighbor	A person not listed as an emphyteutic copyholder in this *cabreve* whose parcel borders an emphyteutic copyholder's parcel. A neighbor may hold the freehold of the parcel.
Limit	Administrative Hydrography Mountain Anthropic	Elements that delimit a plot. These can be of different types, so the list is open in case new alternatives appear.

[a] "*Jornal*" is a traditional land measurement unit with an average value of 4358 square meters. In the Barony of Sella, a "*jornal*" equals to 4804 square meters.

cluding some premises and common pitfalls to take into account, but we consider that graph modeling is a very expressive way of communicating complex problems. It is intuitive and easy to understand. In this paper, we present a brief example of a graph data model that contains the essential node and relation types for analyzing the *cabreves*.

We start by describing four different node types (see Table 1): 1) aggregation nodes represent different administrative levels of rural districts – increasing in number as required. These constitute Level1, Level2, among other levels (useful for creating a graph of more than one *cabreve*); 2) plot nodes can be identified as a Level0 aggregation, but for clarity purposes they are named "plots". Aggregation nodes need at least one plot to exist. 3) person nodes – there are two types in our study but more would be needed if representing legal issues, say freehold or shared ownership among heirs. 4) limit nodes represent plot edges rather than persons. In our case study, we distinguish four types of "limits": administrative (e.g. Marquisate of Guadalest); hydrographic (e.g. Ravine or Tagarina River); mountain (e.g. Manor Mountain); and anthropic (e.g. *Poblet* and *Assagador*).

Once these node types have been defined, they can be explicitly interconnected with a few relationship types (topological relationships in our study) described in the *Cabreve* of 1726. In this way, we consider that there can be binding relationships between copyholders and their plots, and also neighborhood relationships between plots and their neighbors. Since the rest of the relationships described in the *cabreve* are of a topological nature, we have used common GIS spatial predicates to name them (Strobl, 2017). By using these terms, limit nodes may be *"touching"* or neighboring plots, and aggregation nodes will *contain* one or more plots. This terminology may be extended to other spatial predicates if necessary, such as *equals*, *disjoints*, and *intersects*.

As mentioned above, in this subsection we show a readily extended graph data model with more relationships and node types. Figure 2 shows the most basic data model for the *cabreve* in our case study: it is a model for the explicitly written information

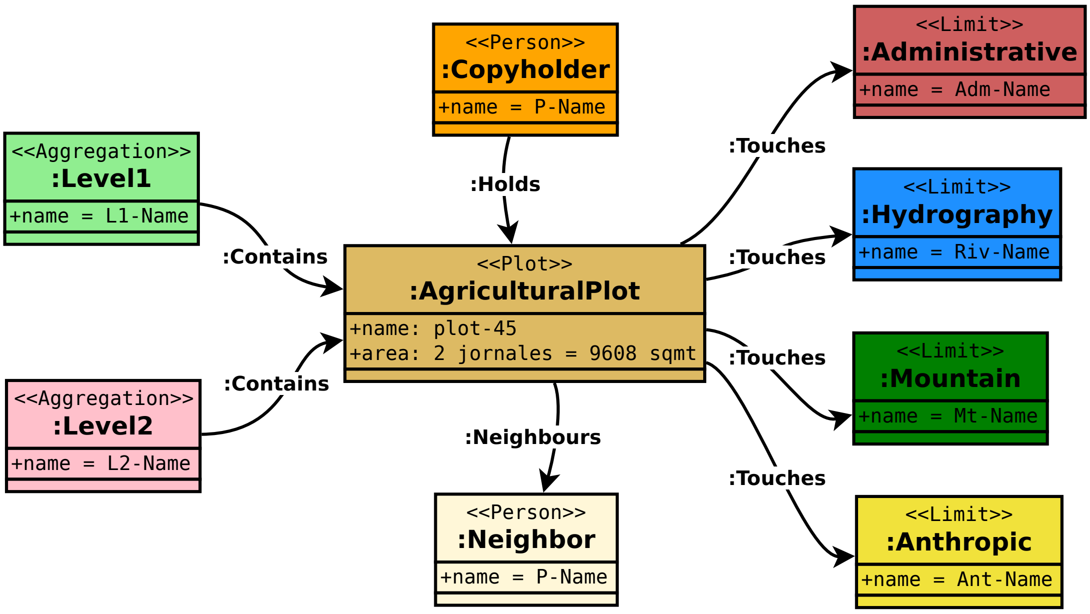

Figure 2. Graph data model for the explicit information contained in the *cabreves*.

contained in the source document. In the following subsections, we present methods and software tools for enriching this model with new implicit relationships, which have to be mined iteratively from these explicit relationships.

2.3. Software development and reproducibility

The original data model implementation and deeper analyses have been performed on computers with different hardware features, but always using GNU/Linux Ubuntu 16.04. This operating system has pre-installed versions of Make (GNU Make 4.1, x86_64-pc-linux-gnu) and Git (version 2.7.4).

Make is a dependency management tool widely used in software development. It deals with interpreting how and in what order a set of files should be compiled. However, Make can be used in many other scenarios in which it is necessary to update one or more files from another set. Make can run the programs that are necessary to achieve its objectives and compile the results, but it is necessary that those programs are available in the system. Git is version control software that helps keep track of changes in project files and coordinates the work of people who collaborate on a code project.

Except Make and Git, all the necessary software to reproduce our work was containerized and executed on the Docker platform, which automates the deployment of applications within software containers. It is a technology similar to that of virtual machines, but designed for a better use of resources. Docker (17.05.0-ce, build 89658) was installed and used to deploy all the other applications when needed. This means that it was not necessary to install additional software or worry about versions, configurations, or features. Make defines a workflow that runs on several Docker containers to produce the output files, while Git manages the development of the source code to be uploaded to a platform in the cloud.

Make, Git, and Docker execute a series of R scripts. R is a programming environment

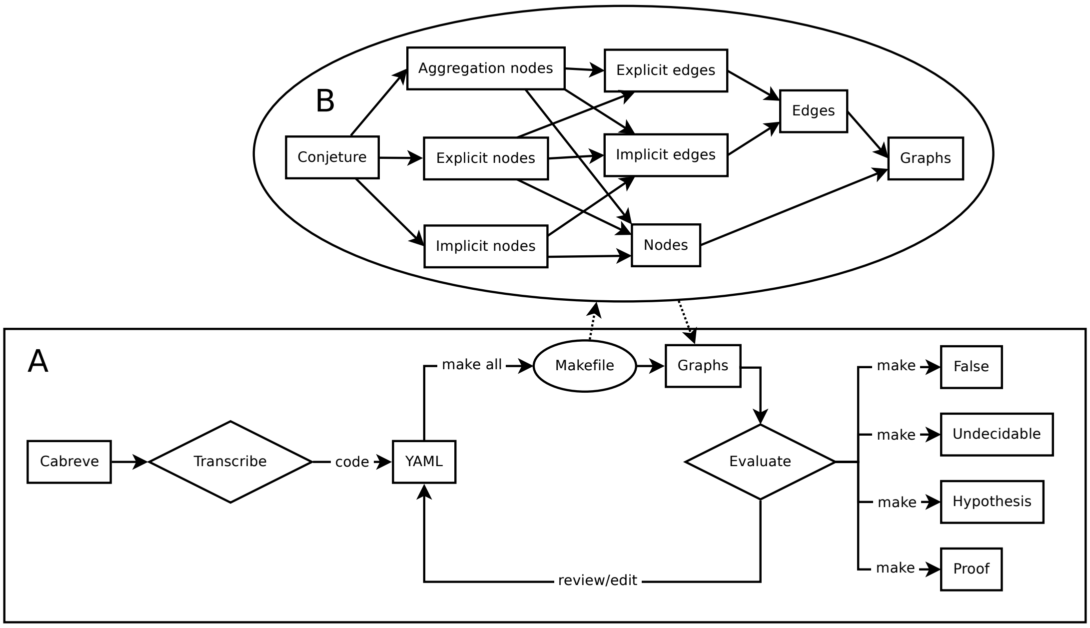

Figure 3. A) Flow chart diagram for the analysis of the *cabreve*. The "make" command executes those tasks that can be performed automatically. The execution of the Makefile (circle) involves several tasks programmed in Rscripts and the generation of multiple intermediate files, necessary for the creation of one or more interactive graphs. The other tasks (transcribe and evaluate) are mainly manual and involve decisions made by an expert. B) Automated workflow diagram with Make, Dockers and Rscripts. This diagram describes in a general way the automatic tasks orchestrated with Make. A conjecture is a structure of folders containing a copy of the current YAML, the intermediate files, and the resulting graphs of the interpretation of the *cabreve*.

for statistical and graphics analysis that is popular in statistics and data mining. R is distributed under the GNU GPL license and is available for Windows, Apple, Unix and Linux/GNU.

The developed scripts use a series of R libraries that facilitate the tasks of data importing (yaml, readr), pre-processing, and graph analysis (dplyr, reshape2, tidyr, igraph), as well as the generation of graphic outputs from pre-processed data (ggplot2, visNetwork, DiagrammeR). For example, the "YAML" library manages the semi-structured data import from text files in YAML format.

2.4. Semi-automatic workflow

The extraction of the geographic information contained in the *cabreves* can be done manually, but the process can be time-consuming and prone to error. To overcome these difficulties, we propose a semi-automatic workflow that leads the researcher through different manual and automatic phases, always supported by the previously described technological platforms (Make, Docker and R). Figure 3 provides a detailed workflow description.

Understanding *cabreves* entails transcribing, digitizing, and interpreting the geographic information contained. The extracted data identifies the copyholders, different levels of spatial aggregation (*disticts*, *clusters of former baronial properties*, and *agricultural plots*), the boundaries of each plot (neighbors, mountains, hydrography, administrative boundaries, etc.), as well as other data related to the productivity of each plot ("*jornal*" units). Figure 4 shows an example of the manual coding process. It can be said that the transcription process is done manually, since our aim is to eval-

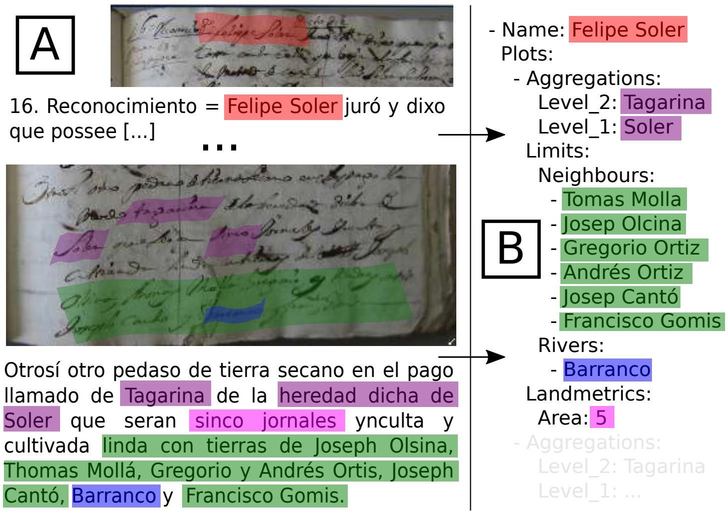

Figure 4. Example of digitizing the geographic information of the *cabreve* in YAML format. A) Excerpts from the *cabreve* highlighting geographical information. B) Structure and coding of the information in YAML format.

Table 2. Classification of node types depending on the degree of evidence found and their short description.

Node type	Description
Explicit	Elements mentioned in the *cabreve* that are not described with numeric or countable attributes (e.g. plots, rivers, mountains, buildings, etc.).
Aggregated	The aggregation nodes are also explicit nodes, but contain more geographic information to help contextualize the nodes in the territory. Only area summations are currently calculated (e.g. emphyteutic copyholders, and rural districts).
Implicit	Neighbors who do not appear as emphyteutic copyholders, i.e. these may be landowners, or appear in other parts of the *cabreve* not yet consulted).

uate the quality of the topological information contained in the *cabreves*. However, this part of the workflow could also be automated by applying the latest advances in handwriting recognition (Bouillon, Ingold, and Liwicki, 2018; Granell, Chammas, Likforman-Sulem, Martínez-Hinarejos, Mokbel, and Cîrstea, 2018).

We then encode the identified data in YAML format following a semi-structured model in which we rigorously establish the hierarchical relationships (see data model in Figure 2). For example, the statement, "*an emphyteutic copyholder has five plots*", is transcribed as five tenure relations. In a similar way, when the six plot neighbors are listed, we understand that there are six neighborhood relations. Figure 4.B shows an example of plot coding and related information.

The YAML file is then automatically analyzed with a series of R scripts executed by a Makefile. In Figure 3, we can distinguish a connector (oval shape) that refers to this third phase. The subdiagram in Figure 3.B is a subset of the Diagram 3.A. The analysis proceeds by extracting the elements of the *cabreve* represented as graph nodes with the characteristics precisely delimiting the written and the logically inferred information. The process requires three types of nodes: explicit, aggregated and implicit (see Table 2).

We then compile the nodes, and Make extracts of all the interesting relationships

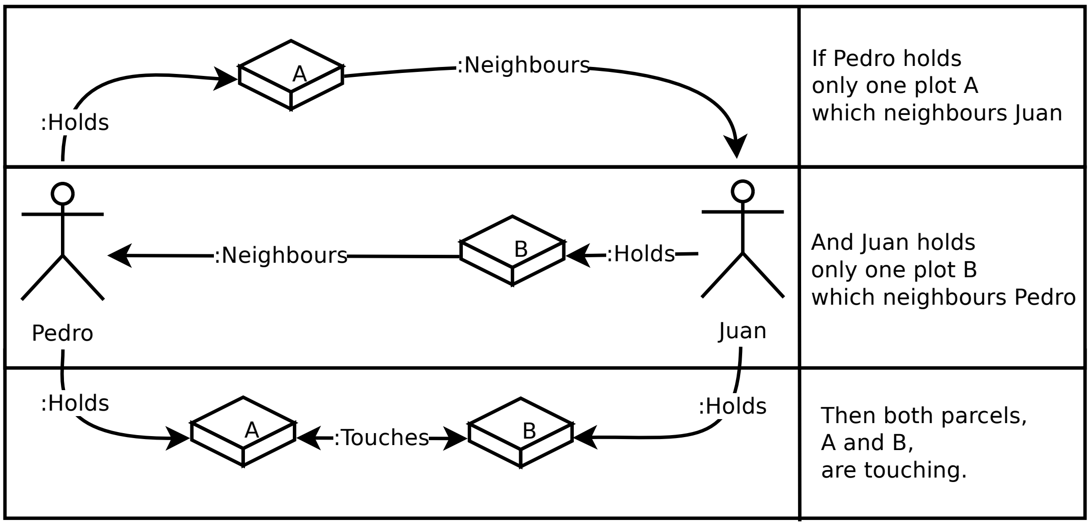

Figure 5. Diagram explaining how simple topological relationships can be obtained from *cabreve* data.

in these nodes. At first, only two types of relationships are distinguished (implicit and explicit), since the aggregation relationships are not useful for the reconstruction of the landscape structure (however, they are useful when a study of the settlement and other issues related to the emphyteutic copyholders is required). Aggregation relationships are those in which a neighbor relation is repeated, or those in which the two neighbors share a family name, for example.

The explicit relationships are read directly within YAML, and others are limited only by the detail of the information extracted from the *cabreve*. Following the data model (see Figure 2), all the relationships are presented in a more concise and unambiguous way in the R scripts. For example, a tenure relationship such as $\{copyholder \rightarrow: holds \rightarrow plot\}$ has been presented as a $copyholder - plot$ relationship in the code (other examples are $copyholder - level2$, $copyholder - level1$, $copyholder - neighbor$, etc.).

Implicit relationships are more interesting because they are harder to compile by hand. Figure 5 shows the rationale behind the specification of a simple topological relationship between agricultural plots. From the YAML structure, we learn that two emphyteutic copyholders ($Pedro$ and $Juan$) have one plot each (A and B), both surrounded by a certain number of neighbors $\{N1, N2, N3, ..., Nn\}$. If both emphyteutic copyholder plots are next to each other, then both are emphyteutic copyholders and neighbors. Furthermore, if there is a reciprocal relationship such that $\{Pedro \rightarrow A \rightarrow Juan\}$ and $\{Juan \rightarrow B \rightarrow Pedro\}$ and it is only registered once in the *cabreve*, we can say that plots A and B are juxtaposed agricultural plots. On the other hand, if this relationship is registered more than once, then the attributes of the parcels can be checked by taking into account the aggregation levels to which they belong. If this rationale is not enough, we would need to explore the relations of the neighbors with other emphyteutic copyholders. Following this line of logic, we distinguish three types of explicit relationships between plots depending on whether there are relations between:

- Plots of different parts of the *cabreve* ($plots - l3 - l3$).
- Plots of the same L2 (district) but different L1 (cluster of former baronial prop-

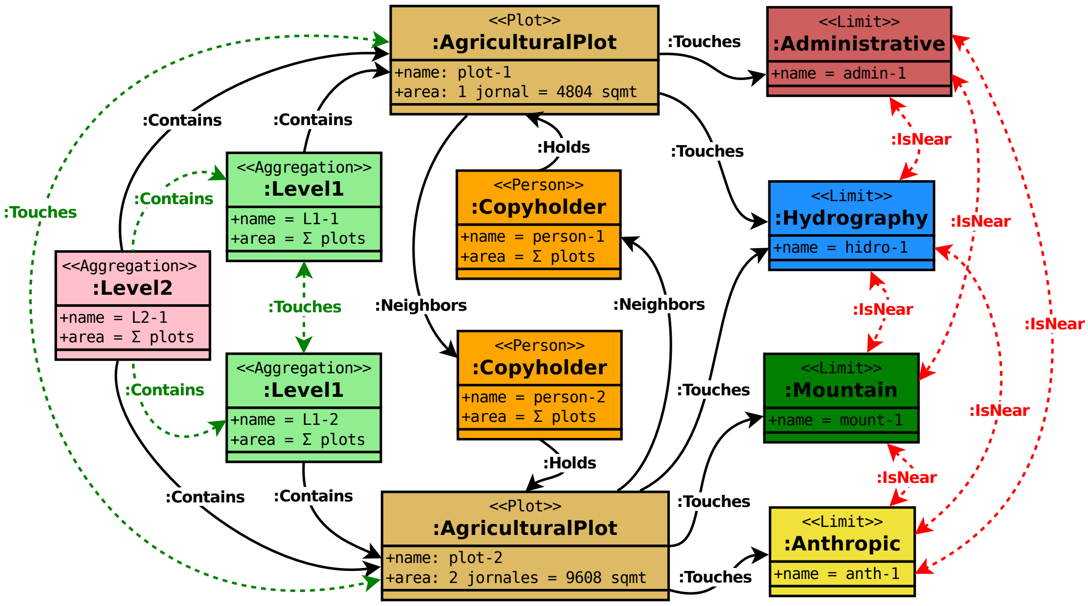

Figure 6. Graph data model for the explicit and implicit information contained in the *cabreve*. Dashed lines are for implicit relationships, green lines for the relationships studied in this paper, and red lines for other relationships that are not considered here.

erties) ($plots - l2 - l2$).
- Plots within the same L1 ($plots - l1 - l1$).

From the relationships between plots, we can establish further relationships between clusters of former baronial properties or between *districts*, which helps to reconstruct the landscape structure at different aggregation levels.

At the end of the extraction process of the different types of nodes and relationships, several R libraries can be used for both generating and visualizing the graphs. In this case, we have generated interactive graphs with the visNetwork library (`http://datastorm-open.github.io/visNetwork/`) that explores each conjecture in a very intuitive way. These widgets work with web components (HTML, JS and CSS) that allow the graph to be superimposed on a reference image or map if necessary (see Figure 7). The elements of the graph (nodes, edges, and extent) are interactive, allowing the nodes to be dragged without distorting the topological relationships. For example, in Figure 7, inheritance relationships under different territorial nodes (other inheritances, ravines, mountains within the Marquisate of Guadalest) are highlighted. In this case, since we know that Guadalest is located north of Tagarina, it becomes evident that this cluster of former baronial properties (L1) should be located north of the ravine. Once this inheritance is re-positioned, there will likely be other topological relationships demanding revision. Note that some relationships render the graph meaningless for spatial purposes, so the researcher must edit the information obtained, or review the YAML interpretation to adjust his or her conjecture. According to the diagram in Figure 3.A, the researcher can use Make to save a version of the current conjecture as: false; undecidable; hypothesis; or proof.

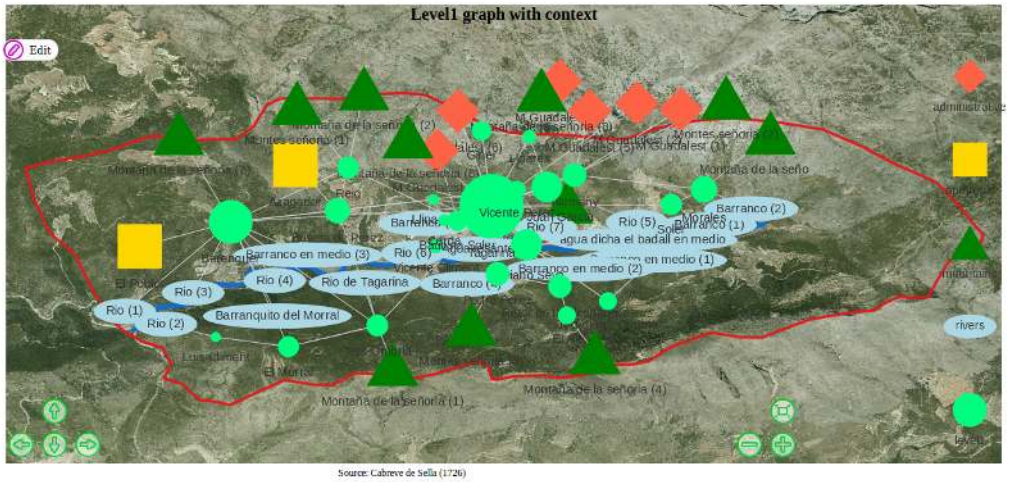

Figure 7. Example of a graph obtained from a conjecture. This image shows an interactive graph which represents clusters of former baronial properties with context nodes. Clusters of former baronial properties are represented by green circles, graduated according to the area in jornals (one jornal equals 0.48 ha). This graph has been manually rearranged according to various factors (e.g. hydrography and mountain nodes need to be placed in the center of the valley and in the peripheral areas respectively).

3. Results and discussion

The proposed methodology and tools have been applied to the case study of the Tagarina valley. In this section, we present the methodological and applied results. We make some considerations about the suitability of this methodology in terms of usability and reproducible research, and then we present a simple, but meaningful, test. Evidence is added to disentangle the precise location of the lost site of "*El Poblet*" (a potential archaeological site) (see Figure 8).

3.1. Methodology reproducibility

This methodology could be applicable to other *cabreves* and other similar historical documents. For this reason, we have developed our tools as free and open source software (FOSS) on platforms and environments that enhance reproducibility. These aspects of the research are not novel: reproducibility of scientific research is a major concern for many researchers (Munafò, Nosek, Bishop, Button, Chambers, Percie du Sert, Simonsohn, Wagenmakers, Ware, and Ioannidis, 2017; Singleton, Spielman, and Brunsdon, 2016; Zaragozí, Belda, Linares, Martínez-Pérez, Navarro, and Esparza, 2012; Zaragozí, Giménez, Navarro, Dong, and Ramón, 2012). Reproducibility constitutes an additional and rarely valued effort that allows experts to reproduce results and evaluate the methodology with minimal effort. The reproduction of all this work involves the installation of two FOSS license software (Make and Docker) together with the execution of a single command (make all) in the terminal, which can be achieved with basic knowledge of the GNU operating system/Linux.

The choice of YAML as a data entry format in the software has been a success. There are markup languages that yield similar results. However, YAML is easier to write by hand and less verbose for researchers. In addition, with the help of text editors such as Notepad ++ or gedit, YAML proves to be more readable than other alternatives, such as JSON or XML.

The software developed and the data used during this first experience has been published in its own repository within the Github platform (https://github.com/deja-visite/capbreu-builder), and downloading the entire project, or working collaboratively with Git (git clone), can be quickly achieved. All the code has been released under general public license 3 (GPL3), which allows any researcher to use, study, share, copy and modify the software at will. The use of the software components is discussed in the documentation available in this repository. The code includes more than 30 R scripts that reproduce the automated workflow described in Figure 3.B and several Make files that organize the proposed workflow. The code has been developed in this way to enhance scalability and reusability. The automated part of this workflow consumes only seconds for the case study explained in Section 2.1, which is insignificant, especially when compared to the time needed to perform the pairings manually and the risk of committing errors during the process.

The graph obtained in Figure 7 is the result of an example conjecture for which a couple of decisions were taken to disambiguate a family name and a place-name. In first place, two copyholder names, “widow of Francisco Soler” and “heirs of Francisco Soler Bautista”, were finally considered to be the same copyholder, after locating them indiscriminately as copyholders and adjoining properties. Secondly, the concepts of “river” and “ravine in the middle” are assumed to be the main channel of the valley (*Riu de Tagarina*) which acts as the only bisector of this area. After considering these coincidences the interactive graph becomes easier to rearrange. Table 3 summarizes the results obtained for this particular conjecture. As mentioned above, the implicit elements are those that would have been the most time-consuming to identify by hand. In an area as small as Tagarina, we have been able to create a main graph with hundreds of nodes and different types of relationships. Of course, such a large graph would be difficult to rearrange by hand, so it is better to work with partial graphs filtered by node types. For example, the graphs used in the case study proposed in this work only include clusters of former baronial properties and geographic context nodes (see Figure 8). This means that other graphs and applications could be developed from the same information. The counts presented in Table 3 may vary according to the conjecture made by the researchers to disambiguate names, or to answer a specific question, but it provides a valuable insight as to the amount of work these methods could save.

The implicit context of the *cabreve* is shown in the extended data model (see Figure 2). In the extended model, all nodes of *:Neighbor* class are promoted to *:copyholder* class only if they are recognized as copyholders in this *cabreve*. Thus, we have found only 13 neighbors who are not referred to as copyholders in the Tagarina valley. The reason may be that we have incomplete source documents, or that there are transcription mistakes. It may also be the case that these neighbors hold freeholds to the plots.

The methodology and the software developed facilitate conjectures on the past structure of the landscape. The case study developed in this paper provides an example of this, but it may well be applied to the study of other historical documents that contain semi-structured geographic information, similar to the *Cabreve* of Sella of 1726. For example, for the same study area, these could include emphyteutic copyhold leases derived from the town charters of the 13th, 14th, and 17th centuries, land ownership records of the 19th century, notarial records, or the modern Spanish property registry. Furthermore, there are also other similar documents available for other regions of Spain, such as the Ensenada Cadastre.

Table 3. Summarized adjacency matrix obtained from an example conjecture. Each node type is represented by a row and column; the presence of relationships between node types is denoted by integer values. At the moment, only those relationships necessary to obtain a graph of clusters of former baronial properties are calculated, so there are numerous relationships that are not yet exploited.

nodes	Districts (3)	Former baronial properties (23)	Copyholders (42)	Plots (64)	Admin. (1)	Anthropic (2)	Mountains (11)	Hydrography (7)	Neighbors (13)
Districts (3)	3	-	43	64	-	-	-	-	-
Former baronial properties (23)	-	45	62	64	4	2	11	16	-
Copyholders (42)	43	62	197	64	-	-	-	-	-
Plots (64)	64	64	64	102	6	2	11	19	-
Admin. (1)	-	4	-	6	-	-	-	-	-
Anthropic (2)	-	2	-	2	-	-	-	-	-
Mountains (11)	-	11	-	11	-	-	-	-	-
Hydrography (7)	-	16	-	19	-	-	-	-	-
Neighbors (13)	-	-	-	-	-	-	-	-	-

Edges/relationships

Aggregated
Explicit
Implicit
- Not calculated

We have carried out an analysis of the topological and geographic information contained in a section of a *cabreve*, which, after all, is the first step to creating an HGIS of periods for which there is no cartographic description of the territory. According to Gregory and Healey (2007), the cost in time and resources involved in the construction of an HGIS database is one of the main difficulties for applying GIS to historical studies. Considerable effort is needed before a researcher can start to analyze the information and obtain results. It is for this reason that a semiautomatic method must be developed if the structuring of the *cabreve* information is to be simplified, and the workflow defined to indicate the precise phases in which the researchers need to intervene. The viability of the analysis of the *cabreves* is then enhanced and the information they contain may be potentially combined with other GIS databases.

3.2. Location of "El Poblet"

One of the objectives set out in this research, the construction of the graph, has enabled us to locate lost material elements described in the source document. A summary of the process is shown in figure 8. Specifically, we have been able to find the remains of an archaeological site, "*El Poblet*", which appears cited in a document of the late 18th century and in the *Cabreve* of 1726 as depopulated (as described by Francisco Giner, a emphyteutic copyholder).

The "depopulated Moorish settlements" are considered an archaeological and ethnographic heritage of great value in Spain (Torró Abad, 2003), and most are documented and protected. In this way, the location of this settlement is a relevant contribution to our knowledge of the Moorish habitat in the Valencian Mountains, one of the main areas inhabited by Spanish Moors and Moriscos until their complete expulsion in 1609.

According to the different relationships provided by the graph, we have defined the location of the copyholder Francisco Giner and his agricultural plots as belonging to a cluster of former baronial properties at Berenguer, at the entrance of the valley (two kilometers from the "*Baseta de Pérez*"), near the river and oriented south. An approximate location of the cluster of former baronial properties at Berenguer is highlighted in Figure 8.A which shows an arranged graph of property clusters, and Figure 8.B (a close-up of the nodes located at the entrance of the valley). This level of detail has facilitated the delimitation of an acceptable search area to inspect using remote sensing or even fieldwork. Given the dense tree coverage in that area, the photointerpretation of common aerial photographs does not reveal the remains of any archaeological site. However, airborne LiDAR can be used for the detection of archaeological features under woodland canopies (Devereux, Amable, Crow, and Cliff, 2005). Figure 8.C shows building footprints from high-resolution (0.5 pts/m^2) LiDAR data (LiDAR-PNOA ceded by ©National Geographic Institute of Spain). We can see an anthropic structure located in the same place where, after querying the graph, we could interpret that "*El Poblet*" of the *Cabreve* of 1726 is located. This structure does not seem to be a cultivation terrace and complies with the previously studied parameters of a Moorish mountain habitat: proximity to a river, above small terraces of cultivation associated with an irrigation system, and with visual control over a large area (Glick, 1995). Figure 1.A shows a picture taken from this observation point.

Finally, fieldwork allowed the hypothesis of departure to be verified. We found pottery from the 15-16th centuries and remains of an oven along with housing structures (see Figure 8.D). An oral interview with two people born in the valley in the 1920s verified the correspondence of the site with the toponym "*Poblet*", along with others related to the Islamic past, such as "the Moor's terrace" (*bancal del moro*), together with the detected structures.

4. Conclusions and future work

Our methodology has proven to be useful in our proposed case study. We have been able to collect new evidence to identify an archaeological site as "*El Poblet*". Combining this graph-based method with oral sources and remote sensing, it appears that we have identified the location of "El Poblet" with a high degree of certainty. In addition, through this case study, the *cabreves* have proven to be a reliable source of topological information. However, it is worth emphasizing that this work is a first attempt at solving the problem, and that there are several methodological and technical aspects that should be examined in further studies. The proposed methodology is flexible and the proposed graph data model provides a solid foundation from which to tackle new types of nodes and relationships. Regarding the technical aspects, three fundamental issues should be solved for the results of the analysis to be imported to a GIS environment: (1) Spatial context. The resulting graphs fail to show a spatial context, which means that the nodes do not have geographic coordinates assigned. Despite this drawback, the proposed methodology provides answers to vital research questions and creates new working hypotheses. (2) Automation. The automation of other aspects of research in historical geography needs to be considered. For example, given the dependence of recognizable toponyms in the area, a geoparsing strategy could be applicable to automatically extract toponyms and geolocate them from custom databases. Once the historical document has been transcribed, there are different text mining strategies that could be applied, and (3) Migration. We are working on the idea of migrating the functions developed to a web application deployed on an R Shiny server (`https://shiny.rstudio.com/`), an application server developed by RStudio. In this application, the workflow would be completely available through a

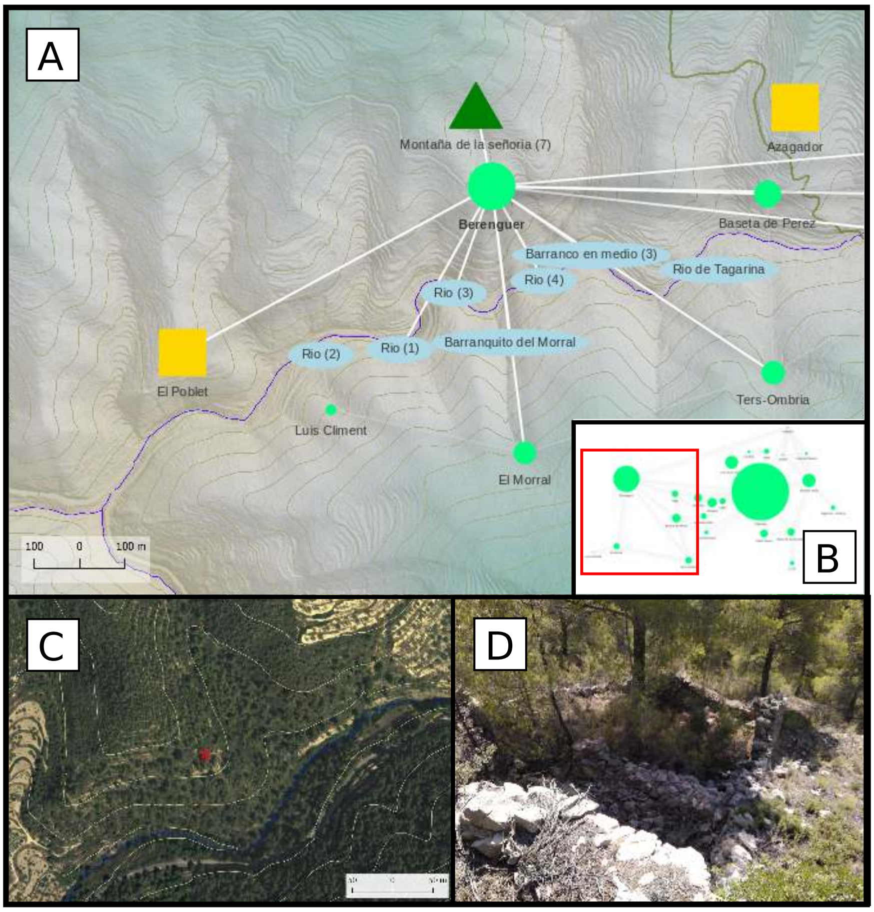

Figure 8. Evidence found about the location of "*El Poblet*". A) Superposition of an ordered graph on the map of the area; B) Spatially arranged graph of clusters of former baronial properties; C) Building footprints from LiDAR data; D) Photograph of the remains found.

web page and resulting in an even more transparent, more intuitive, and more visual experience.

This work brings new interesting research topics that could be approached from our starting point. In the first place, it would be necessary to extend the study of the *cabreve* to the rest of the municipality of Sella, beyond the Tagarina valley, and see if new relationships exist. Secondly, we can extend this methodology for studying this area diachronically using several available sources. We have explored the structure of the emphyteutic copyhold leases from 13th, 14th, and 17th centuries, land ownership documents from the 19th century, as well as cadastral cartography and property registers from the 19th century to present, and we consider that these documents have structures that could be easily adapted to the methodology proposed here. The older documents have a tree structure similar to the one described for the *Cabreve* of 1726, and the process of extracting a topology graph from recent cadastral polygon GIS layers seems straightforward. Finally, there is an apparent continuity in the family names found in the Tagarina valley. Some family names, such as Cerdá or Giner, have been present in the valley for almost five centuries and it would be interesting to check if a spatial pattern could be observed in time. If so, the changes (or lack of) in the patterns of land ownership in recent centruries would become evident.

5. Acknowledgements

We are grateful to the editor and reviewers for their help in improving this paper. This work was supported by grants from the Spanish Ministry of Economy and Competitiveness, project SIOSE-INNOVA (CSO2016-79420-R AEI/FEDER UE).

References

Abedin, B. and B. Sohrabi (2009). Graph theory application and web page ranking for website link structure improvement. *Behaviour and Information Technology 28*(1), 63–72.

Agnoletti, M. (2014). Rural landscape, nature conservation and culture: Some notes on research trends and management approaches from a (southern) European perspective. *Landscape and Urban Planning 126*, 66–73.

Amati, V., T. Shafie, and U. Brandes (2018). Reconstructing Archaeological Networks with Structural Holes. *Journal of Archaeological Method and Theory 25*(1), 226–253.

Barbot, M. (2015). When the history of property rights encounters the economics of convention. Some open questions starting from European history. *Historical Social Research 40*(1), 78–93.

Barthelemy, M., P. Bordin, H. Berestycki, and M. Gribaudi (2013). Self-organization versus top-down planning in the evolution of a city. *Scientific Reports 3*.

Benítez, R. (1984). Formas de transmisión de la propiedad en el País Valenciano (siglos XVII y XVIII). Aproximación metodológica a un nuevo tratamiento de una fuente notarial clásica: los cabreves. In *La documentación notarial y la historia*, pp. 353–370. Universidad de Santiago de Compostela.

Bevan, A. and A. Wilson (2013). Models of settlement hierarchy based on partial evidence. *Journal of Archaeological Science 40*(5), 2415–2427.

Bouillon, M., R. Ingold, and M. Liwicki (2018). Grayification: A meaningful grayscale conversion to improve handwritten historical documents analysis. *Pattern Recognition Letters 0*, 1–6.

Brughmans, T. (2010). Connecting the Dots : Towards Archaeological. *Oxford Journal of Archaeology 29*(3), 277–303.

Brughmans, T. (2013). Thinking Through Networks: A Review of Formal Network Methods in Archaeology. *Journal of Archaeological Method and Theory 20*(4), 623–662.

Brughmans, T. and U. Brandes (2017, aug). Visibility Network Patterns and Methods for Studying Visual Relational Phenomena in Archeology. *Frontiers in Digital Humanities 4*(August), 1–20.

Brughmans, T., S. Keay, and G. Earl (2014). Introducing exponential random graph models for visibility networks. *Journal of Archaeological Science 49*(1), 442–454.

Carrion, D., F. Migliaccio, G. Minini, and C. Zambrano (2016). From historical documents to GIS: A spatial database for medieval fiscal data in Southern Italy. *Historical Methods 49*(1), 1–10.

Clifford, J., B. Alex, C. M. Coates, E. Klein, and A. Watson (2016). Geoparsing history: Locating commodities in ten million pages of nineteenth-century sources. *Historical Methods 49*(3), 115–131.

Collar, A., F. Coward, T. Brughmans, and B. J. Mills (2015). *Networks in Archaeology: Phenomena, Abstraction, Representation*, Volume 22. New York: Springer Science+Business Media.

Congost, R. (2003, nov). Property Rights and Historical Analysis: What Rights? What History? *Past & Present 181*(1), 73–106.

Congost, R. (2007). *Tierras, leyes, historia: estudios sobre "la gran obra de la propiedad"*. Barcelona: Editorial Critica.

Congost, R. (2009). *Dels capbreus al registre de la propietat. Drets, títols i usos socials de la informació a Catalunya (segles XIV-XX).* Girona: Documenta Universitaria.

DeBats, D. A. (2008). A tale of two cities: Using tax records to develop GIS files for mapping and understanding nineteenth-century U.S. cities. *Historical Methods 41*(1), 17–38.

Devereux, B. J., G. S. Amable, P. Crow, and A. D. Cliff (2005). The potential of airborne lidar for detection of archaeological features under woodland canopies. *Antiquity 79*(305), 648–660.

Doménech García, A., P. Giménez-Font, and F. X. Llorca Ibi (2007). *Als peus de l'Aitana.* Ajuntament de Sella.

Domingo Pérez, C. (2018, jul). Fuentes agrarias pre catastrales valencianas. Su utilidad geográfica. *Estudios Geográficos 59*(231), 225.

Donato, V. (2017). Towards design process validation integrating graph theory into BIM. *Architectural Engineering and Design Management 13*(1), 22–38.

Ducruet, C., S. Cuyala, and A. El Hosni (2018). Maritime networks as systems of cities: The long-term interdependencies between global shipping flows and urban development (18902010). *Journal of Transport Geography 66*(August), 340–355.

Eiroa, J. A. (2012). Past and present of the archaeology of the alquerías. *Imago Temporis - Medium Aevum 6*, 49–72.

Femenia-Ribera, C., E. Benitez-Aguado, G. Mora-Navarro, and J. Martinez-Llario (2014). Method of recovering municipal boundary lines in Province of Valencia (Spain) by means of historical cadastral maps. *Survey Review 46*(337), 255–266.

Fischer, M. M. (2006). GIS and network analysis. *Spatial Analysis and GeoComputation: Selected Essays* (January 2004), 43–60.

Galiana, A. (2015). *La Marina Baixa en 1791. Informe del Obispo Fabián.* Associació d'Estudis de la Marina Baixa.

García-Borbolla, A., A. Gil Olcina, and C. Abad Navarro (1988). Declive y ocaso de la enfiteusis señorial valenciana. *Agricultura y sociedad* (49), 293–318.

García Juan, L., A. J. Álvarez Miguel, C. Camarero Bullón, and J. Escalona Monge (2012). Generación de una metodología para la gestión y recreación cartográfica a partir de información del Catastro de Ensenada. *Revista Internacional de Ciencia y Tecnología de la Información Geográfica* (12), 268–302.

García Trobat, P. (2001). A Forgotten Result of the Spanish War of Succession: the Cadastre and its Fiscal Effects on the Crown of Aragón. In N. V. BADEN-BADEN (Ed.), *Cadastre and Modern State in Italy, Spain and France (18th c.)*, pp. 193–216.

Gil Olcina, A. (1998, jul). Control del dominio util y salvaguarda del directo en la enfiteusis señorial valenciana: Los Cabreves. *Estudios Geográficos 59*(231), 201.

Gil Olcina, A. (2012). *Singularidades del Régimen Señorial Valenciano. Expansión, declive y extinción de la Señoría Directa.* Alicante: Publicaciones Universidad de Alicante.

Giménez-Font, P. and J. A. Marco Molina (2017). La dinámica del paisaje en la Serra d'Aitana (Alacant, España): síntesis de transformaciones históricas en una montaña mediterránea (1600-2010). *Estudios Geográficos 78*(282), 105.

Glick, T. F. (1995, dec). *From Muslim Fortress to Christian Castle: Social and Cultural Change in Medieval Spain.* Manchester: Manchester University Press.

Gómez Benedito, V. (2018). Las reformas administrativas de la Casa de Medinaceli en sus dominios valencianos durante la segunda mitad del siglo XVIII. *Hispania 77*(257), 763.

Granell, E., E. Chammas, L. Likforman-Sulem, C.-D. Martínez-Hinarejos, C. Mokbel, and B.-I. Cîrstea (2018). Transcription of Spanish Historical Handwritten Documents with Deep Neural Networks. *Journal of Imaging 4*(1), 15.

Gregory, I., C. Donaldson, P. Murrieta-Flores, and P. Rayson (2015). Geoparsing, GIS, and Textual Analysis: Current Developments in Spatial Humanities Research. *International Journal of Humanities and Arts Computing 9*(1), 1–14.

Gregory, I. N. and R. G. Healey (2007). Historical GIS: Structuring, mapping and analysing geographies of the past. *Progress in Human Geography 31*(5), 638–653.

Hart, J. P. and W. Engelbrecht (2012). Northern Iroquoian Ethnic Evolution: A Social Network Analysis. *Journal of Archaeological Method and Theory 19*(2), 322–349.

Heal, J. W., G. J. Bartlett, C. W. Wood, A. R. Thomson, and D. N. Woolfson (2018). Applying graph theory to protein structures: an Atlas of coiled coils. *Bioinformatics* (August), 1–8.

Hin, S., D. A. Conde, and A. Lenart (2016). New light on Roman census papyri through semi-automated record linkage. *Historical Methods 49*(1), 50–65.

Hoofs, K. (2010). Property Rights in Legal History. *Property Law and Economics*, 5–30.

Hunter, R. (2010). Methodologies for reconstructing a pastoral landscape. *Historical Methods 43*(1), 1–13.

Isaksen, L. (2008). The application of network analysis to ancient transport geography: A case .... *Digitial Medievalist 4*, 1–22.

Jiang, N. and D. Hu (2018). GIS for History: An Overview. In *Comprehensive Geographic Information Systems*, pp. 101–109. Elsevier.

Katambi, S., G. Jiming, and K. Xiangyuan (2002). Applications of graph theory to gross error detection for gps geodetic control networks. *Geo-Spatial Information Science 5*(4), 26–31.

Kendall, D. G. (1971). Maps from marriages: an application of non-metric multi-dimensional scaling to parish register data. In F. R. Hodson, F. R. Kendall, and P. Tutu (Eds.), *Mathematics in the archaeological and historical sciences.*, Edinburgh:, pp. 303–318. Edinburgh University Press.

Knappett, C., T. Evans, and R. Rivers (2008). Modeling Maritime Interactions in the Aegean Bronze Age. *Antiquity 82*(February), 1009–1024.

Knappett, C., R. Rivers, and T. Evans (2011). The Theran eruption and Minoan palatial collapse: New interpretations gained from modelling the maritime network. *Antiquity 85*(329), 1008–1023.

Knowles, A. (2014). The contested nature of historical GIS. *International Journal of Geographical Information Science 28*(1), 206–211.

Konig, D. (1936). *Theorie der endlichen und unendlichen Graphen; kombinatorische Topologie der Streckenkomplexe.* New York: AMS Chelsea Publishing.

Lagesse, C., P. Bonnin, P. Bordin, and S. Douady (2016). Méthodologie de modélisation et de caractérisation des réseaux spatiaux. Application au réseau viaire de Paris. *Flux N 105*(3), 33.

Mackaness, W. A. and M. Kate Beard (1993). Use of graph theory to support map generalization. *Cartography and Geographic Information Systems 20*(4), 210–221.

Manguinhas, H., B. Martins, and J. Borbinha (2008, nov). A geo-temporal Web gazetteer integrating data from multiple sources. In *2008 Third International Conference on Digital*

Information Management, Volume 4, pp. 146–153. IEEE.

McNeill, J. R. (1992). *The Mountains of the Mediterranean World.* Cambridge: Cambridge University Press.

Mostern, R. (2008). Historical gazetteers: An experiential perspective, with examples from Chinese history. *Historical Methods 41*(1), 39–46.

Munafò, M. R., B. A. Nosek, D. V. M. Bishop, K. S. Button, C. D. Chambers, N. Percie du Sert, U. Simonsohn, E.-J. Wagenmakers, J. J. Ware, and J. P. A. Ioannidis (2017, jan). A manifesto for reproducible science. *Nature Human Behaviour 1*(1), 0021.

Naranjo-Ramírez, J. (1992). Enfiteusis en Andalucía: la villa de Fernán Núñez (Córdoba). *En: El medio rural español : cultura, paisaje y naturaleza : homenaje a don Ángel Cabo Alonso / Valentín Cabero Diéguez ... [et al.].*

Robinson, I., J. Webber, and E. Eifrem (2015). *Graph Databases* (2 ed.). O'Reilly Media, Inc.

Romero González, J. (1983). *Propiedad agraria y sociedad rural en la España mediterránea. Los casos valenciano y castellano en los siglos XIX y XX.* Madrid: Ministerio de Agricultura, Pesca y Alimentación (Servicio de Publicaciones Agrarias).

Scheidel, W. (2014). The shape of the Roman world: modelling imperial connectivity. *Journal of Roman Archaeology 27*(Paris 1966), 7–32.

Schlichting, K. (2008). Historical GIS: New Ways of Doing History. *Historical Methods: A Journal of Quantitative and Interdisciplinary History 41*(4), 191–196.

Singleton, A. D., S. Spielman, and C. Brunsdon (2016). Establishing a framework for Open Geographic Information science. *International Journal of Geographical Information Science 8816*(February), 1–15.

Strobl, C. (2017). Dimensionally Extended Nine-Intersection Model (DE-9IM). In S. Shekhar, H. Xiong, and X. Zhou (Eds.), *Encyclopedia of GIS*, pp. 470–476. Cham: Springer International Publishing.

Szmytkie, R. (2017). Application of Graph Theory To the Morphological Analysis of Settlements. *QUAESTIONES GEOGRAPHICAE 36*(4), 65–80.

Torró Abad, J. (2003). Arqueologia de la conquesta. Registre material, substitució de poblacions i transformació de l'espai rural valencià (segles XIII-XIV). In M. Barceló, G. Feliu, A. Furió, M. Miquel, and J. Sobrequés (Eds.), *El feudalisme comptat i debatut. Formació i expansió del feudalisme*, pp. 153–200. Valéncia: Universitat y Museu d'Història de Catalunya.

Turner, S., J. Bolòs, and T. Kinnaird (2017). Changes and continuities in a Mediterranean landscape: a new interdisciplinary approach to understanding historic character in western Catalonia. *Landscape Research 6397*, 1–17.

Weidele, D., M. van Garderen, M. Golitko, G. M. Feinman, and U. Brandes (2016). On graphical representations of similarity in geo-temporal frequency data. *Journal of Archaeological Science 72*, 105–116.

Zaragozí, B., A. Belda, J. Linares, J. Martínez-Pérez, J. Navarro, and J. Esparza (2012, may). A free and open source programming library for landscape metrics calculations. *Environmental Modelling & Software 31*, 131–140.

Zaragozí, B., P. Giménez, J. T. Navarro, P. Dong, and A. Ramón (2012). Development of free and opensource GIS software for cartographic generalisation and occupancy area calculations. *Ecological Informatics 8*, 48–54.