DOI 10.13131/ARCHEOLOGICADATA-1Z7V-5405

Gabriele Gattiglia[A], Nevio Dubbini[A]

[A] Dep.of Civilisations and Forms of Knowledge, University of Pisa.

SPATIO-TEMPORAL NETWORK ANALYSIS APPLIED TO ROMAN TERRA SIGILLATA[1]

Abstract: This work concerns the analysis of data related to Terra Sigillata (TS), gathered integrating different sources, and carried out within the ArchAIDE project (www.archaide.eu). We analysed the data to disclose statistical relationships between the variables considered. Statistical techniques were used as explorative in order to summarise the main characteristics of data and identify outliers, trends or patterns. Specifically, we focused on Network Analysis and on the identification of significant temporal breaks in the data. The network structure is given by linking together locations where ceramics were produced to locations where the same ceramics were retrieved, getting 3853 locations forming its vertices throughout Europe, Middle East and North Africa. The 16820 different edges were built, joining 322764 different data. Network analysis allowed identifying communities in the network, i.e. groups of vertices being densely connected internally but poorly connected externally. Such communities can represent commercial routes adopted by producers or that established themselves by geographical or historical reasons. Temporal breaks were identified by an algorithm minimising the variance within intervals, while maximising the variance between intervals. Production and supply of ceramics have a specific relevance only in certain temporal intervals. We were able to distinguish four main periods, characterised by different production centres emerging and declining in the different phases (Italian, South-Gaulish, Rhine productions), and showing different production dynamics.This work also underlines how the availability of a high volume of data (unfortunately rare in Archaeology), joined with data analysis, allows new insight into archaeological research.

Keywords: Network analysis, Roman Ceramics, R, ArchAIDE.

Introduction

The aim of this work concerns the construction of networks and communities related to Roman Terra Sigillata (TS) and their changes across time. Roman Terra Sigillata – also called Samian or Arretine Ware – is a class of bright red glossy tableware. It was produced in a circumscribed number of centres distributed in the western part of the Roman Empire (Italy, Gaul, Spain, Germany and Britain) from the second half of the 1st century BCE to the 3rd century CE. TS was widely distributed all over the Roman Empire, and it was used for eating, drinking and serving food. Its standardised shapes (types) were dishes, cups and bowls mainly. From a technological point of view, TS is a slipped pottery. Vessels were formed on the fast-turning wheel, in a mould, and then dipped in a fine clay emulsion that gives the pots a glaze-like aspect due to the sintering process happening during the firing. Shapes were often decorated with figures in low relief, and more interestingly, name-stamps of workshops were impressed inside the vessel, on the bottom of the body (Van Oyen, 2016, p. 12). Stamps contain names of freeborn, freedmen or slaves, or more in general names of people responsible for the production of the vessels. They could be both the masters and dependants of a workshop, and work as potters themselves and/or function as small-scale workshop managers (Aubert, 1994, p. 220). Stamps were employed for the organisation of production phases such as firing (Peacock, 1982, p. 125; Pucci, 1990; 1992), when the products of various potters were fired all in one kiln, and for distribution reasons. They could also have acquired value for sales promotion, but in any case, this should be a secondary effect based on the success of first Arretine ware (Fülle, 2000, p. 34). To sum up, stamps are related to the producer or someone connected to him until pots are sold. How many vessels were stamped is a question without a solution; the percentage may change from different producers, types, periods and other variables. Nevertheless, they indicate the ownership, i.e. the workshop itself, and the relation between producers and consumers: we considered stamps as such in our study. Catalogues of workshop stamps were edited from the first half of the 20th century and continuously updated (Oswald, 1931; Oxé, 1933; Oxé & Comfort, 1968; Oxè, Comfort & Kenrick, 2000; Hartley & Dickinsion, 2008-12). TS is also supported by reliable typology and chronology based on datable contexts, a good number of excavated production sites (Bemont & Jacob, 1986; Bergamini, 2004; Bergamini & Manca, 2008; Bet, 1988; Bet, Delage & Vernhet, 1994; Desbat, 2001; Menchelli, 2013; Passelac, 1992; Passelac, Sabrié & Sabrié, 1986; Picon & Garmier, 1974; Pucci, 1990; Sforzini, 1987; Soricelli, 2004; Vaccaro, Capelli & Ghisleni, 2017), archaeometric analysis for determining the provenance (Menchelli et al., 2001; Olcese, 2004; Picon, 1973; Picon, 2002a; Picon, 2002b; Picon & Lasfargues, 1974), and the evidence of graffiti concerning the production organisation (Camodeca, 2006; Johnston, 1985; King, 1980; Marichal, 1988). These features have made it possible to build up a detailed knowledge of the production process, its distribution and consumption. For such reasons, TS is of fundamental importance for the comprehension and dating of archaeological contexts and even more crucial for understanding and interpreting dynamics of production, trade flows and social interactions. Following the stamps is a way for understanding networks as a process in constant flow. As Van Oyen pointed out (2016, p. 5), TS distribution maps were used to show the economic performance, spread of dining habits, values, and sign systems in a scenario in which economic effects in the standardisation of the production are taken for granted (increased efficiency, mass production, economic growth).

The present work, integrating different sources collected within the ArchAIDE project (2016), aims to propose a mathematical approach based on big datasets to disclose production and consumption dynamics taking into account TS as a whole quite homogeneous pottery production and producer as part of a *koinè*. Production and consumption have been analysed together from the broad point of view of networks composed of communities of workshop owners, potters (both freemen and slaves), merchants, buyers and consumers. This has been achieved by disclosing statistical relationships, extracting significant patterns, visualising data and discussing the results. The primary concern of this work relates to networks, described and handled as mathematical graphs, obtained by linking locations where ceramics are produced to locations where (the same) ceramics are retrieved. The interest of such networks is mainly due to the possibility of applying techniques of network theory, specifically concerning link analysis, classification and clustering (i.e. identify communities). Network Analysis, despite the importance of relationships in archaeology, is not a mainstream part of the discipline, but in the last decade, a growing number of researchers have begun to explore the possibilities offered by this analysis (Brughmans, Collar & Coward, 2016; Knappett, 2013). Roughly speaking, communities, or clusters, are defined as groups of vertices of the network having a higher probability of being connected than to members of other groups: this can be computed and checked in terms of network links. Identification of significant communities in the network draws attention to the principal 'import-export' systems and their dynamics. Distribution maps, for instance, allow representing areas where a particular ceramic type was found, i.e. produced or used, traded and discarded in the past. When a distribution map is analysed as a graph linking origin and destination sites, analysing and visualising a Big Dataset, we create a deep map (Gillings, Hacigüzeller & Lock, 2019) able to describe a process, in our case a community map delineating the main supply movement of pottery.

Chronology and assemblages changes across time allow disclosing the variation of the communities during centuries. Based on the disparity exhibited by the chronology of ceramics, we have also identified temporal intervals giving rise to different network behaviours. We then specialised the analyses in such temporal intervals separately. These analyses show the dynamics of the primary production sites and the main export areas, the increase and decrease of productions, and the spheres of influence of the significant production poles over time.

Finally, this work relates to how the availability and the re-use of a high volume of data (being unfortunately rare in Archaeology), joined with explanatory analysis, allows for new insight into archaeological research. In this case, data re-use means merging different datasets and analysing them in new and different ways. The merging of archaeological data does not only concern the availability of rich metadata, but it also involves issues regarding the quality, quantity and methodology of data collection and the real interoperability of data (data that appears similar, but, indeed, archaeologically different). In our case, data related to Gaulish-Rhenish productions are richer than Italian. In spite of this, we decided to work with a minimum set of comparable data to explore the overall production of TS, not only central European one.

1. Data and preprocessing

This section introduces data sources, data that are currently available and preprocessing operations done on such data. The software developed and the charts shown in this paper were obtained by running the code on data about TS stamps, coming from the sources listed below:

- ArchAIDE database (Anichini et al., 2020), including data about types of ceramics and stamps, geolocation and chronology. This data source has been populated through the digitisation of (i) Conspectus (1990) ceramics catalogue, a compendium of Terra Sigillata Italica and a standard in the archaeological practice and pottery classification; (ii) Oxé, Comfort & Kenrick (2000) catalogue, a corpus of potters and stamps related to Terra Sigillata Italica, and also contains 3D representations and dimensional information stored in .svg files.
- The website of The Roman-Germanic Central Museum (Samian Research, 2008), an international research institute for archaeology, that offers its database on Roman Terra Sigillata under modular Digital Peer Publishing Licence (m-DPPL), i.e. permitting a free use and re-use. This source contains data from different catalogues (mainly Hartley & Dickinsion, 2008-12, and Oxé, Comfort & Kenrick, 2000) and has been filtered to avoid duplication of data coming from ArchAIDE database.

In a preliminary phase, datasets have been cleaned with OpenRefine (OpenRefine, 2012) and harmonised in order to have consistent information. In this way, data can be effectively merged and prepared for the exploration and analysis phases. The harmonisation was particularly important for data about stamps, coming from different sources. For what concerns the present analysis, the main available information about stamps are:

- Workshop, i.e. the name of the freeborn, freedmen or slaves, or, more in general, the name of the person responsible for the production of the vessels.
- Origins, i.e. names for the locations where the pottery was produced, together with a unique ID of the location (name). Only exact locations were considered. Inaccurate locations (e.g. Central Italy, Gaul, etc.) were excluded.
- Origins longitude and latitude, i.e. geolocation of places of origin (production areas).

- Occurs, i.e. names for the locations where the pottery was retrieved, together with a unique ID of the location (name).
- Occurs longitude and latitude, i.e. geolocation of places of occurrences (where potsherds were found).
- External ID, i.e. a unique code used to connect stamps to morphological types. Note that one stamp can be connected to more than one type.
- Stamp ID, i.e. a unique code identifying the stamp.
- Chronology of the stamp, i.e. a range of dates identifying the temporal positioning of the stamp.
- Frequencies, i.e. the number of different stamps (not the quantity of each one) found.

When missing, coordinates of the places have been retrieved by using API services of Geonames (Unxos GmbH, 2018) for modern places and Pleiades (Pleiades, 2018) for ancient places. Calls to such API are available within the R software suite. Latitude and longitude use WGS84 (EPSG:4326) as their reference coordinate system.

2. Data Analysis

Statistical techniques were used as explorative to summarise the main characteristics of data, identify outliers, trends, or patterns. Specifically, we focused on Network Analysis and the identification of significant temporal breaks.

Network analysis has allowed identifying communities in the network, i.e. groups of vertices that are densely connected internally but poorly connected externally. From an archaeological point of view, such communities can represent commercial routes adopted by producers and merchants or commercial routes that established themselves by geographical or historical reasons. Temporal breaks are substantial from an archaeological and historical point of view because communities, productions, exporting and importing of ceramics have a natural context only in certain temporal intervals. For this reason, we applied an algorithm capable of selecting optimal temporal breaks.

Once collected the data, we first built the networks. All analyses have been carried out with the R (R version 4.0.1, 2020-06-06) integrated suite of software facilities for data manipulation, calculation and graphical display. R is a language and environment for statistical computing and graphics, a standard in statistics and data science practice and research. It is available as Free Software under the terms of the Free Software Foundation (FSF, 2018) GNU General Public License (GNU, 1991) in source code form.

2.1 Network Construction

Analysis and visualisation of available data rely on the construction of a network, intended as a mathematical graph. The basic structure of the network is given by linking together locations where ceramics were produced with locations where the same ceramics were retrieved. More in detail, in the following we describe how the set of vertices and edges of the mathematical graph are defined.

- Vertices are locations, intended as unique locations, where ceramics were found. This is independent of the fact that ceramics were produced or simply retrieved in that specific location. Attached to vertices, there are attributes representing features available from data:
  - The name of the location;
  - Latitude and longitude of the location;
  - The number of edges starting from the vertex, i.e. the out-degree of the vertex;
  - The number of edges arriving in the vertex, i.e. the in-degree of the vertex.

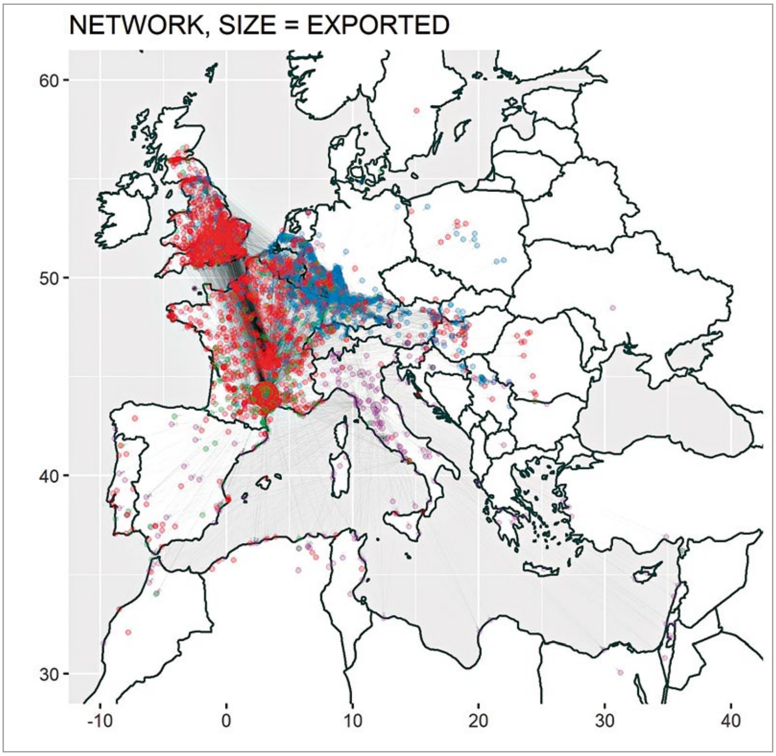

fig. 1. The whole network, obtained from 3372 different locations forming its vertices, and 322764 different data about stamps, giving rise to 15407 different edges. Vertices size is proportional to the quantity of exported stamps.

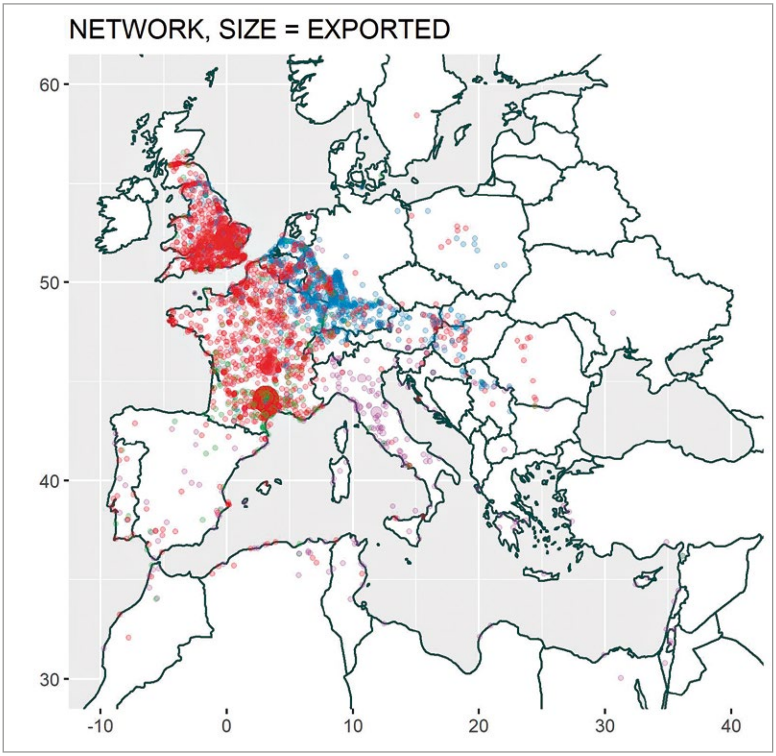

fig. 2. Vertices of the same network depicted in fig. 1, with size proportional to their out-degree.

- Edges between two vertices are directed. Each location where ceramic was produced is linked to all locations where the same ceramic was retrieved. Attached to edges, there are attributes representing features available from data:
  - Latitude, longitude and names of origins and occurs relative to the edges;
  - Workshop, indicating the workshop/potter having produced the ceramic;
  - An external ID, identifying the ceramic types when available;
  - Stamp ID, a unique code identifying the stamp;
  - The frequency, or the weight, of the edge, indicating how many different stamps belongs to the edge, i.e. produced in the place where the edge starts and found in the place where the edge arrives;
  - Chronology of the stamp, i.e. a range of dates identifying the temporal positioning of the stamp.

Note that the edges' weight is not the quantity of ceramics retrieved but the number of different stamps. The graph obtained considering all data (and all chronologies) together is depicted in fig. 1 and has 3372 different locations forming its vertices throughout Europe, the Middle East, and North Africa. There is also one location in Southern India, here not shown in order for the network to be more conveniently visualised. The network has 15407 different edges connecting the vertices, i.e. the locations. It is important to note that in the visualisation we have only one edge connecting each pair of vertices since the weight attribute of edges represents the number of stamps relative to the same edge. Therefore, the 15407 different edges come from a much higher quantity of data about stamps: the graph is based on 322764 different data. This network represents the basis for the development of the rest of the link analysis.

2.2 Network Analysis

Within the description and analysis of data, primary interest relates to its network structure. This is mainly due to the possibility of applying classification and clustering techniques. After the creation of the graph as described in Section 2.1, the focus is the identification of communities in the network, i.e. groups of vertices in the graph that share much more edges internally than externally. The archaeological hypothesis behind the detection of communities is the identification of commercial routes adopted by producers, or commercial routes that established themselves by geographical or historical reasons. When a distribution map is associated with a production (or origin) area, the distribution map represents the supply movement of pottery. Besides, some locations act as infrastructural nodes given their connectedness. These nodes represent well-connected centres for distribution. We refer to them as distribution nodes (see Section 3). Highlighting these nodes is of paramount importance for understanding points of turnover and transhipment in production's trajectory. In some cases, they may be located in critical administrative centres; in other cases, their spatial location may seem 'illogical' and should be explained (Van Oyen, 2015). Even if it is possible to enlighten the correlation between the origins and the occurrence site(s) for indicating the trade(s) route from production areas to the destination, data analysis on a larger scale than the one represented by the single site is necessary for a better understanding of the overall mechanism of the distribution process. In this way, as evidence grows, it is possible to create complex maps, allowing to understand the pottery supply and distribution. In case of quantitative information attached to points (e.g. the number of items on a site), we can create more complex distribution maps; on the other hand, this data needs to be handled with care because, in many cases, we have no information about proportion in an assemblage, or some site could be over-represented. Moreover, by working on the variation of the assemblages in time, it

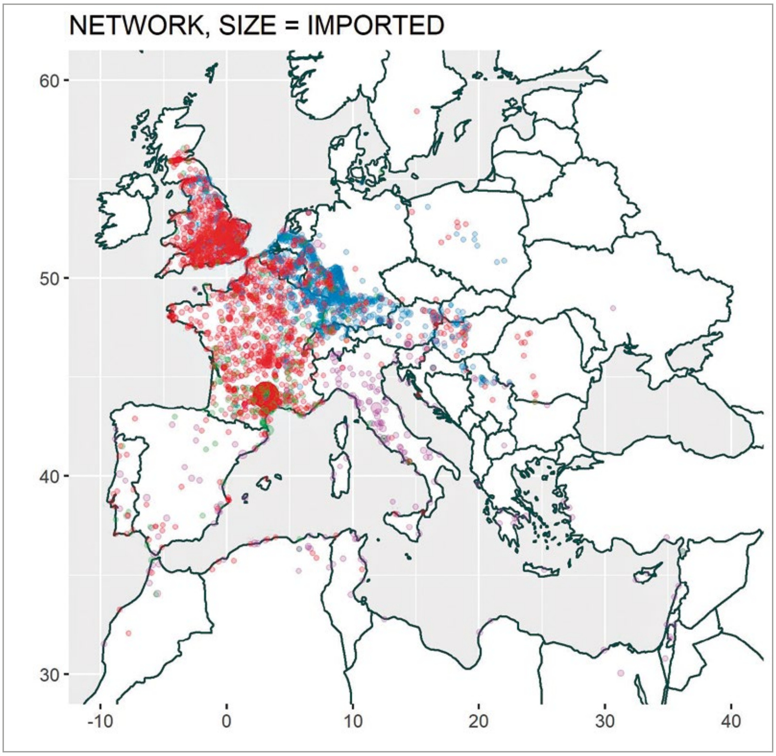

fig. 3. Vertices of the same network depicted in fig. 1, with size proportional to their in-degree.

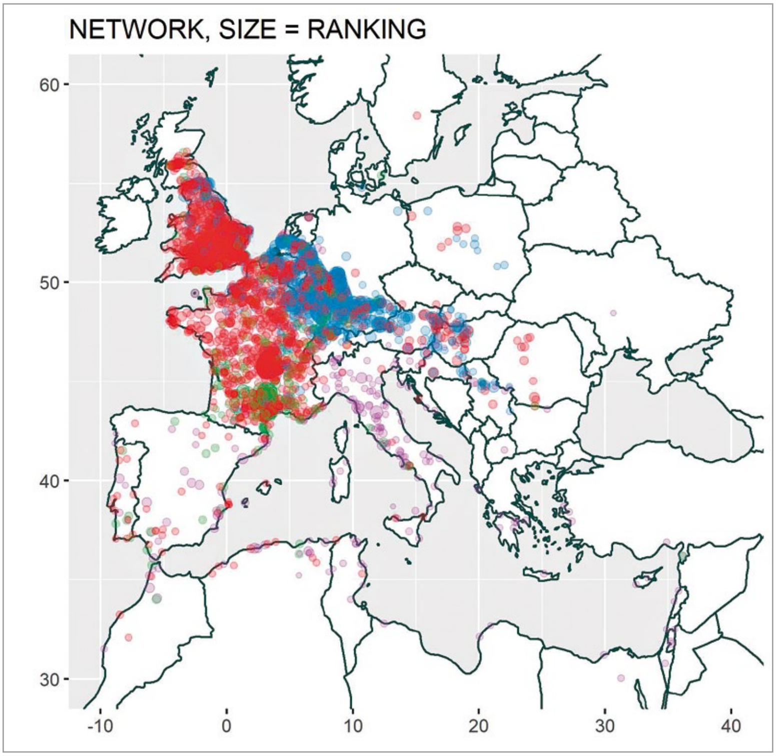

fig. 4. Vertices of the same network depicted in fig. 1, with size proportional to their eigenvector centrality scores.

is possible to enlighten the correlation between destination site and source site in order to visualise the variation of the main route of commercial exchanges during centuries.

These considerations led us to apply the leading eigenvalue clustering algorithm for community detection in the graph obtained. In the algorithm applied (Newman, 2006), the problem of detecting communities in networks is approached by the maximization of a benefit function known as 'modularity' over possible divisions of a network: modularity is the fraction of the edges that fall within the given groups minus the expected fraction if edges were distributed at random. Maximisation of modularity leads, therefore, to the identification of groups (communities) made of densely interconnected vertices that are only sparsely connected with the rest of the network. The maximisation of the modularity is obtained based on the eigenspectrum of the modularity matrix. Moreover, the algorithm scales well with the size of data.

After having applied the clustering, we added one more attribute to the vertices of the graph to indicate the community. Specifically, the attribute is given as a colour so that it can be visualised easily. For the sake of visualisation, the first four communities in terms of number of vertices are kept, being the most represented. Every other edge/vertex is associated with an additional (poorly structured) community, made by vertices and edges not belonging to the main four communities identified by the clustering. Considering again fig. 1, colours of vertices represent communities identified with clustering. Further considerations and interpretations on communities identified will be given in the Discussion (3), when considering communities obtained in narrower temporal intervals.

Another important feature concerning networks is the relative importance of the vertices. Which vertices are more important, central in the network, and why? For our setting, a measure of such importance can be the out-degree, i.e. the quantity of different ceramics stamps 'exported' from a specific location, and another is the in-degree, i.e. the quantity of different ceramics stamps 'imported' in a specific location. These measures give a view of places having produced or imported many ceramics, and important as such. However, networks often have complex structure; so many refined measures of importance have been derived. One of such measures is eigenvector centrality (Bonacich, 1987), corresponding to the values of the first eigenvector of the graph adjacency matrix. Scores can be interpreted as arising from a process in which the centrality of each location is proportional to the sum of centralities of connected locations. The underlying assumption is that the most important locations in the network are connected to many other locations, which are, in turn, connected to many locations. Eigenvector centrality scores were computed on our network, giving a further view of the most important vertices. By confirming the complex structure of the network, eigenvector centrality scores differ significantly from in-degree or out-degree. Eigenvector centrality scores were also added as attributes to the vertices. figs 3, 4 and 5 show the different measures of vertices importance in the network.

figs 2 and 3 show, respectively, in-degree and out-degree of vertices as a dimensional attribute (dimension of vertices is proportional to the degree). Comparing figs 2 and 3, we see that they tend to show similar importance patterns for two main reasons: production locations tend to be also places where many ceramics (also produced somewhere else) are found (e.g. Lyon); ceramics are in many cases exported close to their production location. fig. 4 shows the importance of vertices of the network as assigned by computing eigenvector centrality scores.

2.3 Temporal Breaks

From an archaeological and historical point of view, communities, productions, exporting and importing ceramics have relevance in relation to temporal intervals. For this reason, we applied an algorithm capable of selecting significant temporal intervals searching for optimal temporal breaks.

Given a vector of dates, which in our case were the lowest date in the chronology range, and the number k of desired breaks, such algorithm calculates the optimum breakpoints

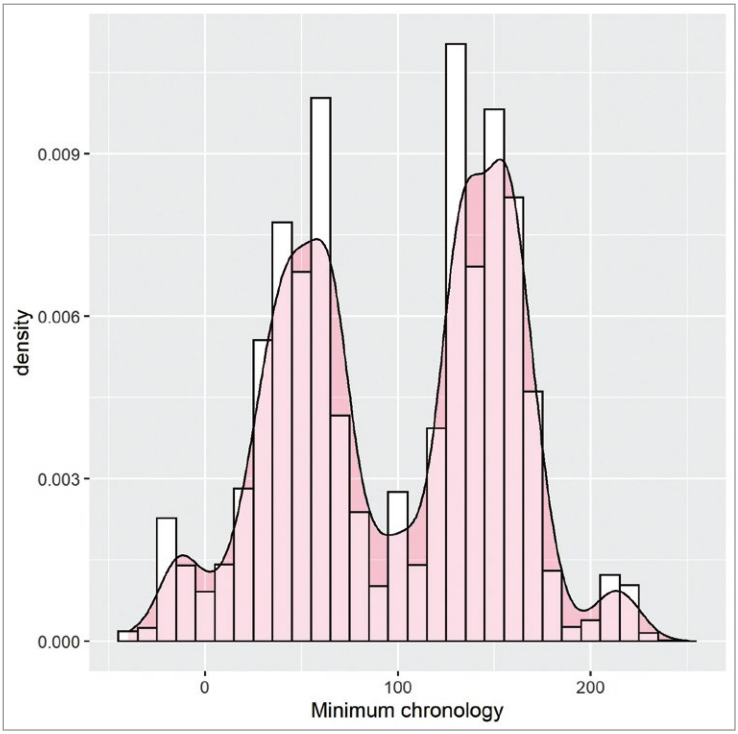

fig. 5. Histogram of the distribution of the minimum chronologies of stamps which shows 4 chronological peaks, suggesting 4 phases in the temporal interval considered.

using Jenks natural breaks optimisation (Conolly & Lake, 2006, p. 142). This algorithm works by binning the dates into k+1 categories, minimising the variance within categories while maximising the variance between categories.

By applying this algorithm, we have divided the data into four temporal intervals. The number of temporal intervals, which needs to be fixed in advance, has been chosen on the basis of the peaks of the distribution of the minimum chronologies of stamps, suggesting four phases in the temporal interval considered (fig. 5). The distribution of the maximum chronology of stamps looks the same.

Separate networks have been generated for each period. Network clustering and visualisation has been separately applied to the networks obtained in the four intervals. The four temporal intervals identified by the Jenks natural breaks optimisation are:

1. From -40 to 30;
2. From 31 to 90;
3. From 91 to 145;
4. From 146 to 300.

The first and last dates are the minimum and maximum dates of available data. The temporal division computed by the algorithm also has clear historical interpretation. Starting from the following charts representing only quantitative information, we see that the different temporal breaks identify clear changes in production patterns, specifically given by the variation of chronology range in the periods identified. Chronology extremities of each stamp can be interpreted as the time the production of that particular dye took place. Looking at this data, we can observe that:

- the first period (fig. 7) is characterised by an initial phase of many different production starting, with average lifetime increasing until first productions terminate (fig. 6);

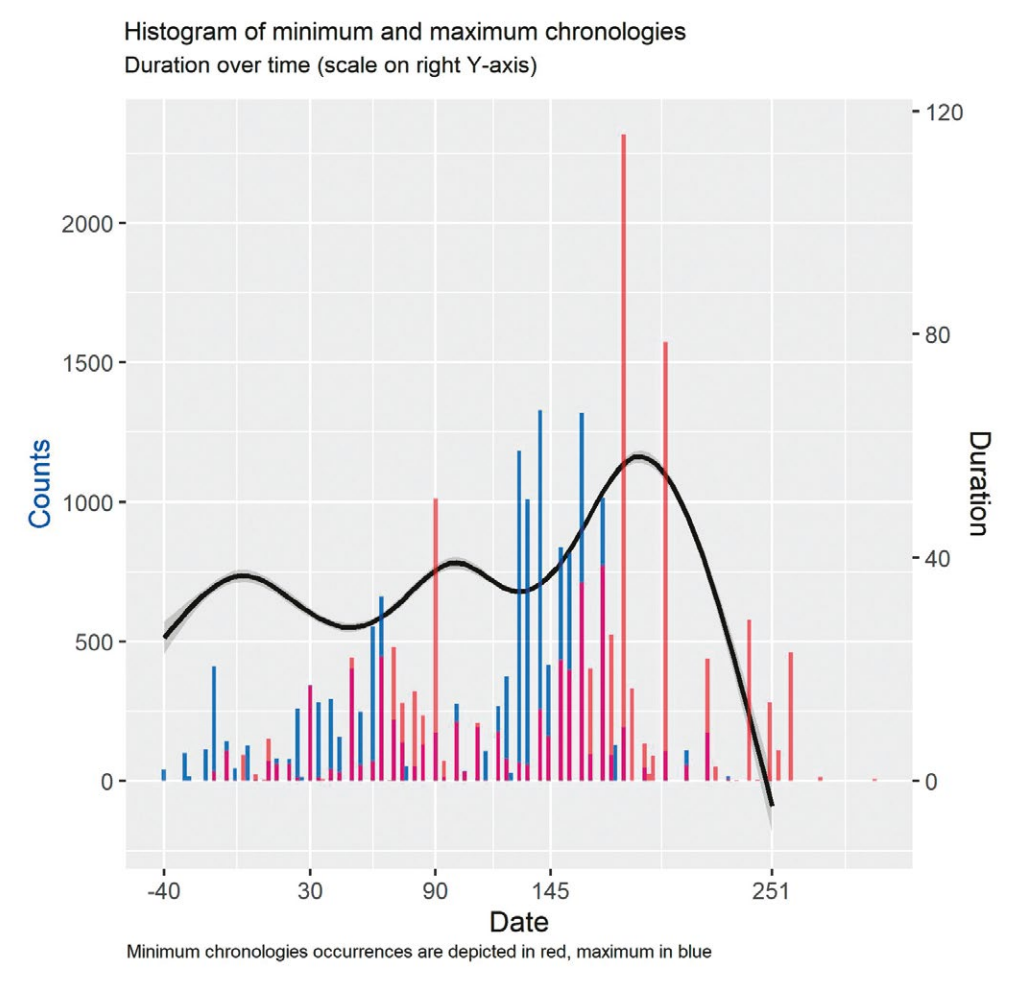

fig. 6**.** Histogram showing the frequencies of production starting (in blue) and closing (in red), plotted against dates of production starting, computed as the minimum in chronology range. Moreover, the curve approximating (smoothing) average ceramics production lifetime is shown in black, with units of measure on the right. It is computed as the number of years in the assigned chronology range (y axis), plotted against dates of production starting (x axis), computed as the minimum in the chronology range. Labels on the x-axis correspond to temporal breaks identified by the Jenks natural breaks optimisation algorithm.

- the second period (fig. 8) is characterised by a clear positive trend in lifetime of productions, together with an increase of productions starting ( );
- the third period (fig. 9) is characterised by the opposite trend, i.e. clear negative tendency in lifetime of productions with a decrease of productions starting (fig. 6);
- the fourth period (fig. 10) is characterised by an initial phase of many different production closing (fig. 6), followed by some events of 'isolated' production starting, with an average lifetime shorter and shorter.

The change in the dynamics of production and exportation of ceramics is strongly reflected in the network structure and clustering of networks relative to the single temporal intervals, which we show in the following, with some comments in the captions.

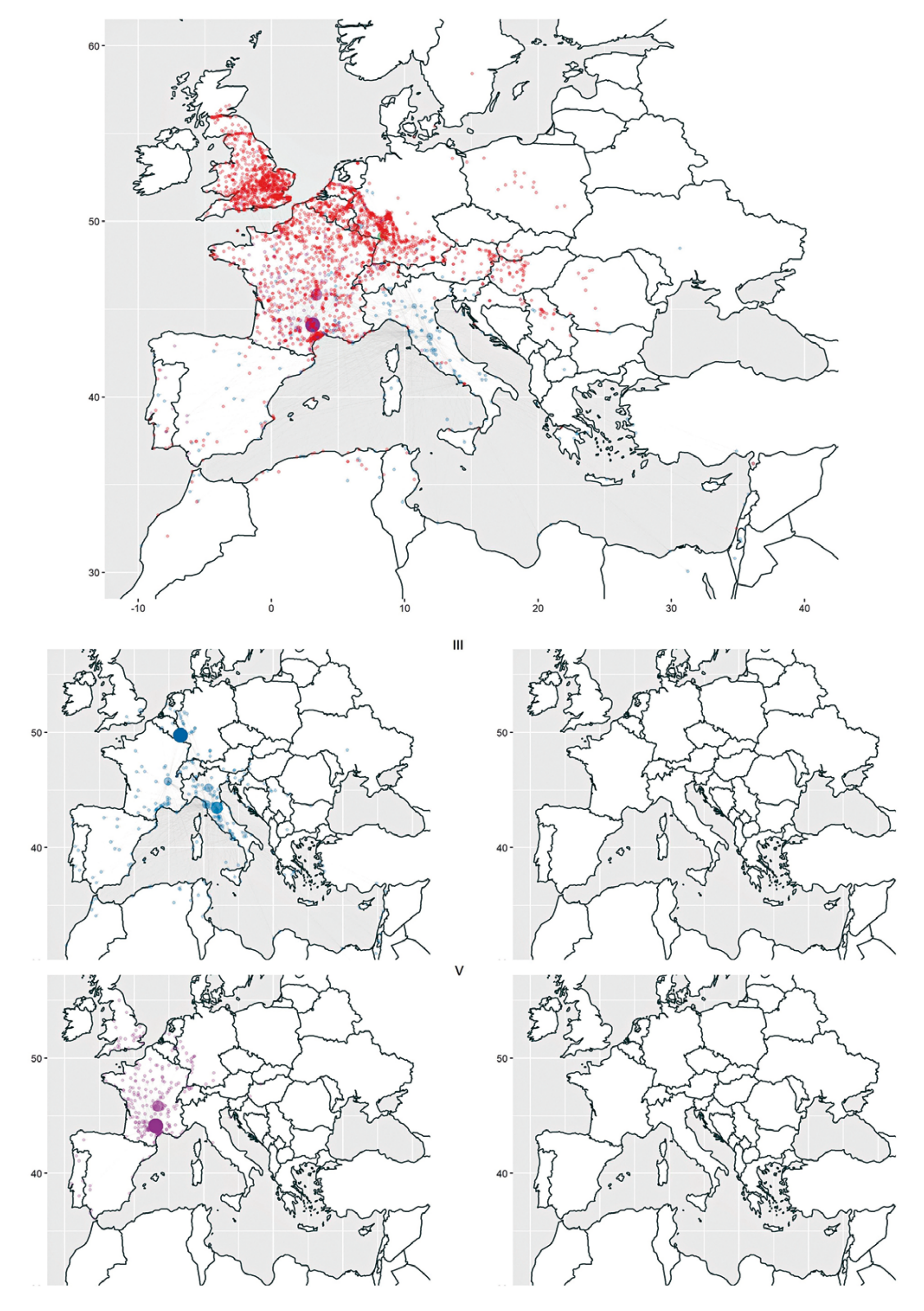

fig. 7. The network of stamps with chronology included in the first temporal interval (40 BCE, 30 CE), followed by the communities given by the clustering (modularity = 0.31). The first four communities are shown clockwise from top left.Taking into account the complete chronology of Terra Sigillata production from the second half of the 1st century BCE to the end of 3rd century CE, it clearly appears the importance of the Italian community in giving the start to the production of this type of pottery. This does not mean that the Italian production ends after 30 CE, rather that its community becomes weaker with respect to new emerging communities (it does not appear in the first 4 clusters).

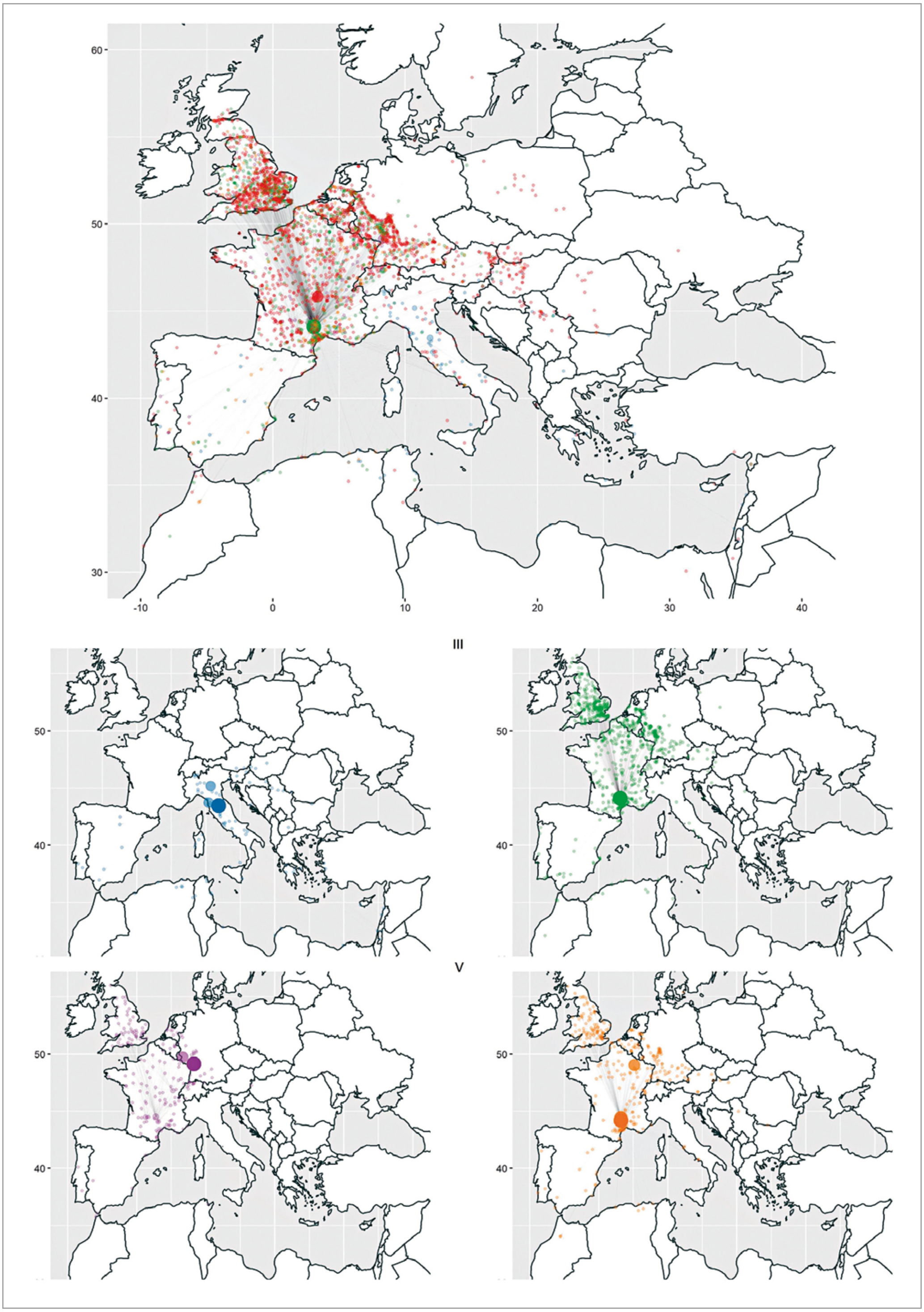

fig. 8**.** The whole network of stamps with chronology included in the second temporal interval (31-90 CE), followed by the clustering (modularity score = 0.38). The first four communities are shown clockwise from top left. This second period underlines the strong emergence of South Gaulish productions. Specifically, a wide network with robust connection in the western part of the Roman Empire has its fulcrum in La Graufesenque (orange community).

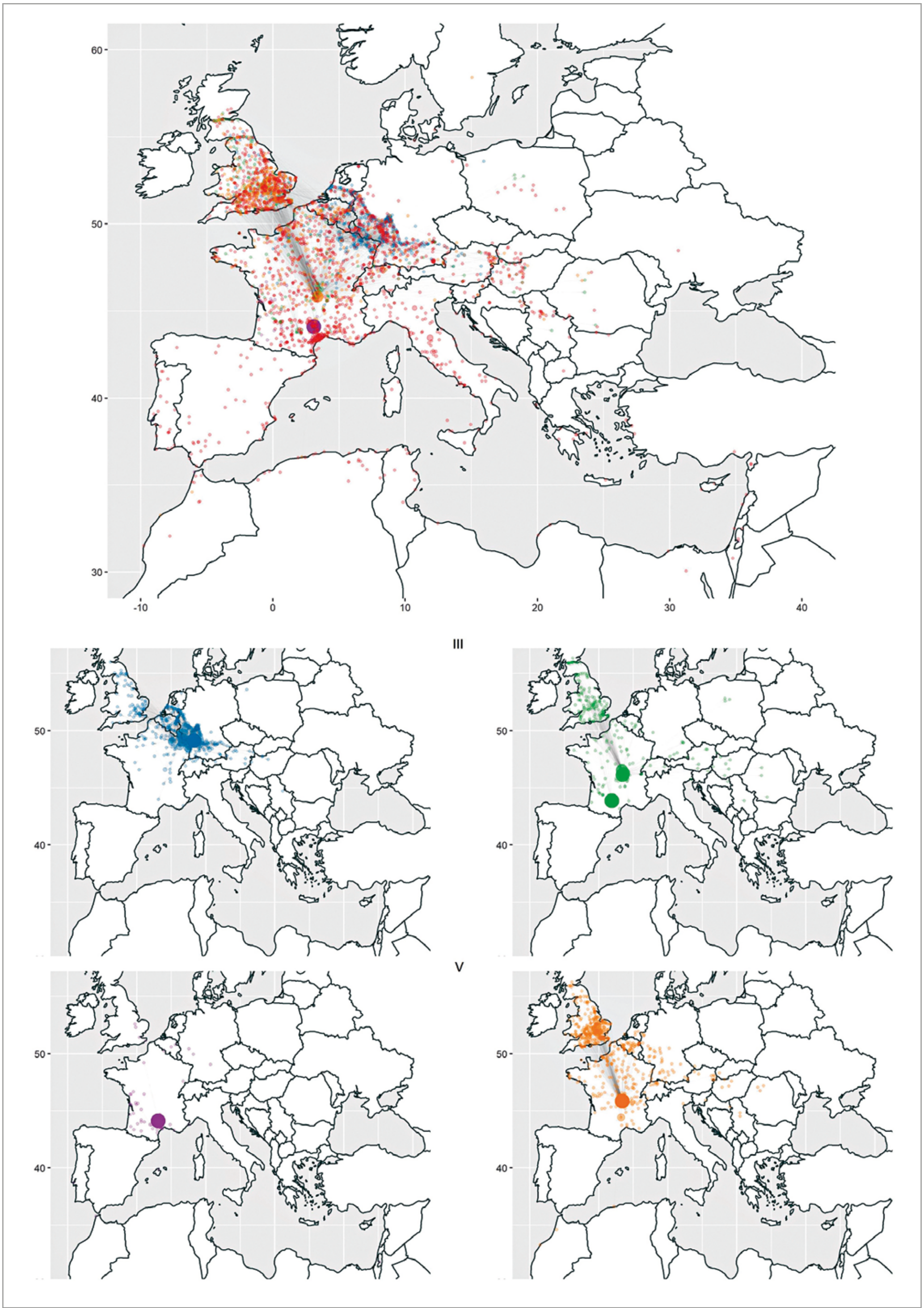

fig. 9. Network of stamps with chronologies in the third temporal interval (91-145 CE), followed by the clustering (modularity score = 0.39). The first four communities are shown clockwise from top left. This period underlines the emergence of Rhine productions (blue community), in an overall framework still characterised by the pre-eminence of South Gaulish productions, even if they start to grow weak.

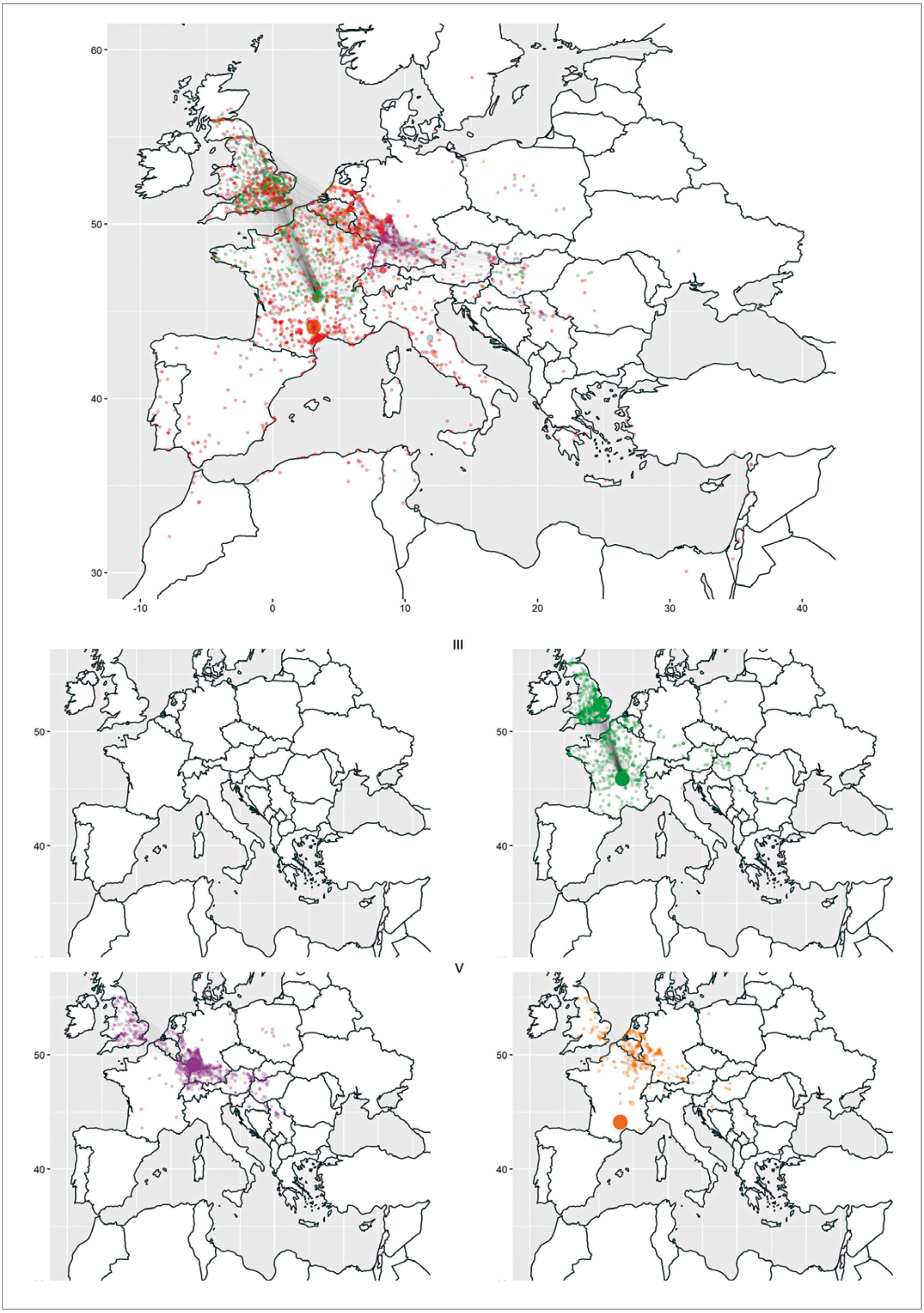

fig. 10. Network of stamps with chronologies in the fourth temporal interval (146-300 CE), followed by the clustering (modularity score = 0.36). The first four communities are shown clockwise from top left. This period underlines the differentiation and specialisation occurring among Rhine productions (green, purple and orange communities) previously (third period) aggregated in one network. The main community is referred to Gaulish production.

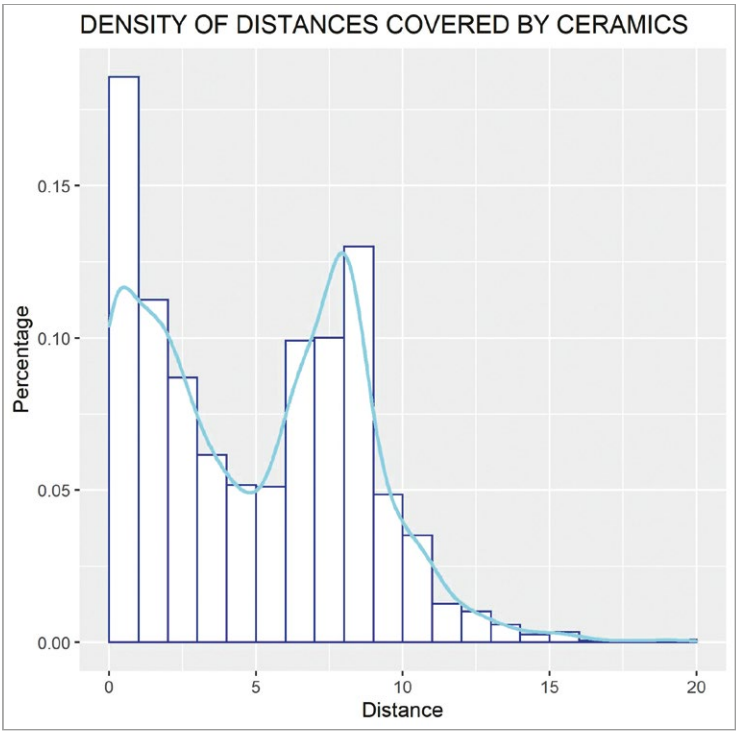

fig. 11. Histogram and estimated density of distance covered by ceramics. On the x-axis the distance in terms of latitude-longitude degrees. Mean is 4.9, median is 5, maximum is 82.3.

2.4 Covered distances. Analysis and discussion.

Data concerning distances covered by ceramics, from the production centres to the locations where they were found, has been analysed in order to highlight if these distances are related to political, geographical, or socio-economic barriers.

The analysis has been carried out on the overall data, so considering the complete chronological range. The histogram in fig. 11 shows the distribution of distances within the network. The reference system of spatial data is expressed in degrees (WGS 84). Transformation to kilometres has been made considering distance 1 as about 140 kilometres: approximately, 1-degree latitude equals 110.574 km, and 1-degree longitude equals 85 km (at this latitude). Distribution of distances accumulates around 0-300 km and 1100-1300 km. The average distance covered is about 680 km. The overall distribution of distances is strongly affected by data from southern France. Data related to Lezoux, La Graufesenque and Les Martres-de-Veyre represents 23% of the data.

3. Discussion

The network analysis and the identification of temporal breaks performed permits new insights. Consequently, we present a short and general discussion of the leading communities (the first four) to show the potentiality of our method. It is essential to remind that (i) chronological extremities go from 40 BCE to 300 CE and represent the minimum and maximum dates of available data; (ii) each location is related to only one community in each period; this does not mean that pottery could not arrive in a specific location from production centres related to different communities, but that a specific location is preferably associated with the community in which is included; moreover, productions centres are always included in only one community for each period; (iii) the different temporal breaks identify changes in production patterns: communities arise, transform and fall; (iv) the communities identified by the community detection algorithm do not come with a natural ordering of their role. The place they occupy in the overall network is discussed based on (a) the number of links received by the locations reached by workshops (links to); (b) the amount of links spread by the production centres (links from); (c) the sum of all the links (total links); (d) the total number of locations (locations) belonging to the community (tab. 1); and (e) on further qualitative archaeological considerations. Nevertheless, it is important to state that following the number of stamps can lead to misinterpretation: (i) production centres with many small workshops could be overrepresented in comparison to centres with a few, big workshops (such as Arezzo with Ateius workshop); (ii) well-investigated production centres could be overrepresented than less archaeologically investigated ones. For this reason, a more articulated analysis of archaeological data and their meaning is envisaged in a future work.

Period/community	Links to	Links from	Total links	Locations
P1/community1	80.559	27.710	108.269	1004
P1/community2	97.966	173.854	271.820	501
P2/community1	8.002	11.961	19.963	151
P2/community2	22.252	20.487	42.739	840
P2/community3	144.263	41.633	185.896	881
P2/community4	44.924	106.974	151.898	1065
P3/community1	31.315	63.797	95.111	919
P3/community2	25.612	15.276	40.888	950
P3/community3	170.395	84.236	254.631	3030
P3/community4	33.961	94.741	128.702	92
P4/community2	137.517	128.693	266.210	1053
P4/community3	78.388	53.152	131.540	2176
P4/community4	44.401	65.661	110.062	2141

tab. 1. P1 goes from 40 BCE to 30 CE; P2 from 31 to 90 CE; P3 from 91 to 145 CE; P4 from 146 to 300 CE. "Links to" represents the number of links received by the locations reached by the workshops; "Links from" the amount of links spread by the production centres; "Total links" the sum of all the links; "Locations" the total number of locations belonging to the community.

The first period (40 BCE, 30 CE; fig. 7) seems characterised by the initial phase of Italian and South Gaulish productions. The algorithm identifies two primary communities. The Italian production centres in central (Arezzo, Pisa, Scoppieto, Vasanello, Ostia), in southern (Aquinum, Pozzuoli, Cales, and Venosa) and in northern Italy (Po Valley) mark the first (community 1), together with centres of Lyon and Vienne in Narbonensis, and German workshops in Haltern. This community underlines the importance of Arezzo and Pisa, which started the production of this type of pottery but also intercepted the Haltern production that had a short life. Analysis of distribution nodes put in evidence the presence of many medium-range nodes located along the coasts (e.g. Ampurias, Carthage, Corinth, Berenice, Alexandria, Aquileia), the crossroads of the main roads (e.g. Poitiers, Gergovia, Milan) and the Rhein limes (e.g. Xanten, Asciburgium, Neuss, Koln, Mainz); a few high-range nodes also correspond to some of the main production centres (Arezzo, Pisa, Lyon, and secondarily Pozzuoli and the Po Valley area). Overall, this community seems resting on the widespread distribution of equal nodes. The second community (community 2) is less addressed to the Mediterranean, and more to central Europe, Britain and partly the Iberian Peninsula. Production centres are gathered in South Gaul, in Narbonensis and Aquitania (Aspiran, Carrade, Crambade, Espalion, La Graufesenque, Le Rozier, Lezoux, the Lot Valley area, Montans, and Valery). Analysis of distribution shows that nodes are not widely distributed. Medium-range nodes are concentrated in southwestern Gaul (Narbonensis and Aquitania: Roanne, Limoges, Perigues, Jarnac, Bordeaux, Agen, Toulouse) around the main crossroads. High-range nodes are located in Paris, Amiens, addressing Belgium and northern Rheinland, in Basel dealing with southern Rhein and eastern limes, and in Spain both along the coast (Tarragona, Elche, Belo) and the interior (Merida). Interestingly, this production penetrates Britain before the Roman conquest, underlining the strength of Roman trade also outside of the border of the Empire and the Britain gravitation towards the Roman Empire before 43 CE. The structure of the two communities is different. Community 1 distinguishes itself by the double of connected location and the half of overall interactions in comparison to community 2. Community 1 is wider but weaker, considering the number of connections, than community 2.

The second period (31 to 90 CE; fig. 8) seems characterised by a clear positive trend in the lifetime of productions (together with an increase of production starting). The network analysis underlines the strong emergence of South Gaulish productions. Community 1 rotates around the workshops based in Etruria (Arezzo, Pisa and Torrita di Siena) and the Po Valley. Interestingly, productions from Torrita di Siena reach as many locations as the other production centres together. Despite the broader distribution of Period 1, Italian productions are specifically oriented to the Mediterranean and the eastern border (the Balkans). In this period, the area beyond the Alps is dominated by South Gaulish productions. Analysis of distribution nodes puts in evidence the presence of a widespread distribution network, which is characterised by many medium-range nodes (Luni, Naples, Siscia, Osijek, Corinth, Tharros, Ibiza, Cordoba e Constantine), and high-range nodes (Hieres, Valeria, Algiers, and Aquileia). Low-range nodes seem concentrated in Italy, Eastern provinces, and the southern Mediterranean. The other three main communities are located in South and central Gaul. The first (community 2) is related to the workshops in Banassac and Le Rozier (respectively in southern Aquitania and Narbonensis) with a similar distribution. The distribution of TS produced by this community is concentrated in the western part of the Empire, primarily, in Gaul, along the Rhine and western Danube border, and in Britain. The pottery produced by this community is also found in the Iberian Peninsula, Algeria and Morocco (Mauretania Caesariensis and Tingitana). Analysis of distribution nodes highlights the presence of a widespread distribution of medium-range nodes. The second (community 3) is centred around workshops in northern Narbonensis (Miliau - Rajol) and southern Aquitania (Lot Valley, Carrade, and Espalion) except for Saint-Sauveur in Lugdunensis (central Gaul). Espalion is the most connected centre. Distribution of this community is addressed to Central Europe beyond the Alps and Britain, with sporadic occurrences in the south of the Iberian Peninsula. Analysis of distribution nodes shows a concentration of medium and high-range nodes in Eastern France, along the German limes, and in Britain. Interestingly, Saint Sauveur, in addition to its role as a production centre, represents the most critical node in the community. The third one (community 4) is related to the workshops in La Graufesenque. The overall pattern of this community is similar to community 2, with a more widespread distribution, except for the node in Antakya, Syria. The role of Antakya should be further investigated because data show relationships only with community 1. Is this a mistake? Or is the algorithm able to predict the role of Antakya without the collection of other data in this region? To conclude, analysis of the structure of the communities shows a pre-eminence of communities 3 and 4 (see tab. 1), the high number of locations connected by community 2, but with a low number of interactions, and the weakness of the Italian community with a low number of locations and interactions.

The third period (91 to 145 CE; fig. 9) seems characterised by the opposite trend, i.e. negative tendency in the lifetime of productions with a decrease of productions starting. Network analysis underlines the emergence of Rhine productions, in an overall framework still characterised by the pre-eminence of South Gaulish productions, even if they begin to grow weak. The absence of a Mediterranean market, and the concentration of the market north of the Alps and Pyrenees, along the eastern border (including the newly conquered region of Dacia), and in Britain that begins to produce TS as well, needs further investigation. Italian production disappears from the four main communities. This can be seen as a weakness of Italian production centres, or this can be attributed to the structure of the data itself. In fact, the so-called Late-italic productions are less studied, and data is less exhaustive. In this case, the missed penetration of Gaulish productions in the Mediterranean scenario could be interpreted as due to the presence of a well-controlled market. Community 1 is related to production centres in the south and central Gaul (Banassac, Lezoux, Les Martres-de-Veyre, Nouatre, in Aquitania, Jaulges-Villiers-Vineux, in Lugdunensis), and in Germania Superior (Luxeuil-Les-Bains). Analysis of the distribution network shows the presence of a few medium-range distribution nodes in Lugdunensis (Dijon), Belgica (Liberches), Germania Inferior (Fallais, and Fouron-le-Comte) and Superior (Andernach and Vindonissa) to supply the Rhein limes, and Britain. Community 2 is related to workshops in Montans (Narbonensis), Terre-Franche, Lubie, Toulon-sur-Allier (Aquitania), and Pulborough (Britannia). The pattern of distribution is similar to community 1. Still, distribution nodes are more differentiated in medium and high-range nodes, and their allocation is widespread with presences in eastern and northern Gaul, along the eastern border in Noricum and Pannonia, and high occurrences in Britain. Community 3 registers the arising of Rheinland production and the parcelling out of workshops related to 25 different sites located in Aquitana (Queinon), Belgica, Germania Inferior and Superior (in order of importance: La Madeleine, Lavoye, Blickweiler, Chemery-Les-Faulquemont, Trier, Heiligenberg, Sinzig, Boucheporn, Haute-yutz, Mittelbronn, Argonne, Eschweilerhof, Ittenwiller, Le Pont-des-Remes, Rheinzabern, Avocourt, Hombourg-Budang, Gueugnon, Swabia, La Foret de Hesse, Dinsheim, Aachen-Schoenforst, Kraherwald, Waiblingen-Beinstein), and finally in Britannia (Colchester). Distribution is concentrated in the northeastern provinces of the Roman Empire and especially along the Rhein and western Danube limes. Medium-range distribution nodes are mainly focused along the German border. Of interest, the presence of nodes in London and York in Britannia. High-range nodes are mostly located in production centres, so highlighting their proximity to the potential market and the short-range market of this community. Community 4 is related to workshops in Narbonensis (La Graufesenque, Valery and Marseilles). La Graufesenque productions seem to lose importance in this period. Distribution appears scattered in the south-west of Gaul, south of Britain and northern Rhein, as well as its few medium and high-range nodes. To sum up, analysis of the structure of the network highlights the pre-eminence of community 3. North eastern productions reach more than three thousands occurrences, with strong connections. On the other hand, community 4 is the smallest one with only 92 locations connected and rests on the leading position of La Graufesenque workshops, which alone contributes three-quarters of the total links. Network 1 and 2 address quite the same number of sites, but community 1 outnumbers the number of connections.

The fourth period (146 to 300 CE; fig. 10) seems characterised by an initial phase of many different productions closing, followed by some events of 'isolated' production starting, with an average lifetime shorter and shorter. The network analysis emphasises the pre-eminence of Rhine productions previously aggregated in one undifferentiated and robust community (third period). It also envisages specialisation that occurs among central Gaulish production centres. The fourth period is characterised by three communities. Community 2 is characterised by production centres that spread through Germania Inferior (Aachen-Schoenforst, Sinzig) and Superior (Baden), Belgica (Argonne, Avocourt, Blickweiler, Lavoye, Les Allieux, Les Pont-des-Remes, and Trier), and southwestern Britannia (Colchester). Analysis of distribution nodes shows the presence of medium-range distribution nodes in east Britannia and along the Adrian Wall, and then along the Rhine/Danube border. In contrast, high-range nodes are related to production centres in Belgica. This distribution network represents well the overall diffusion of this production along the northeastern provinces of the Roman Empire. This community involves a little bit more than one thousand locations, and it is highly connected, as demonstrated by the high number of links. Community 3 follows the same spatial distribution as community

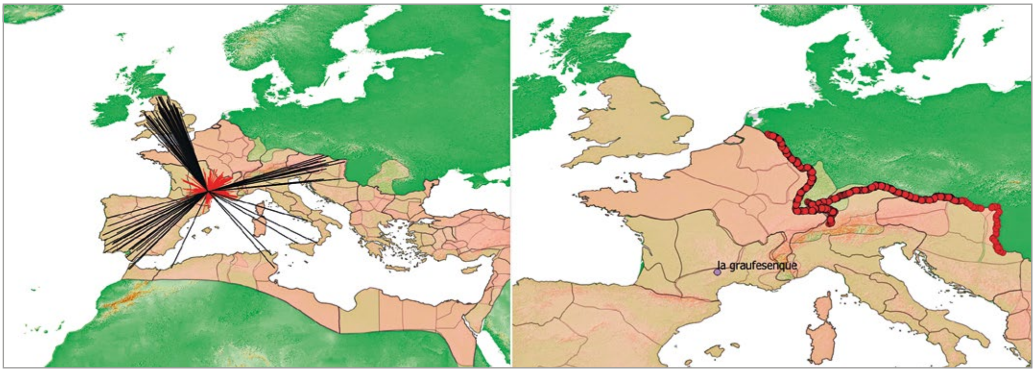

fig. 12. On the left, pottery distribution from La Grafesanque in the ranges 0-300 km, in red, and 1000-1300 km, in black. The vast majority of the black lines end at the Roman Empire's borders. Locations further than 1300 km (not represented in the map) are related to maritime distribution routes. On the right, it is visible the Rhein-Danube limes that stays within 1300 km. The overlapping of the two images well clarifies how the distance of 1300 km represents a physical barrier.

2 but is characterised by a southern distribution of production sites. They are located in Belgica (Chemery-Les-Falquemont, Eschweilerhof, Haute-Yuts, and La Madeleine), Germania Superior (Ittenwiller, Luxeuil-Les-Bains, Heiligenberg, Kraherwald, Mittlebronn, Rheinzabern, and Swiss workshops) and in Raetia (Schwabegg, Swabia, Waiblingen-Beinstein, and Westerdorf-Sankt Peter). Analysis of Distribution nodes shows a high number of them located along the southern Rhine and the Danube, until Singidunum in Moesia, in the inland of Germania Superior and Raetia, and Britannia. The number of locations involved in the community is double of community 2, but the number of overall links is the half. The concentration of sites shows, on the one hand, widespread dissemination, and on the other hand, a weaker community, given the lower number of interaction than community 2. Community 4 has its core in Gaul and Britannia, with sporadic dissemination along the Rhine/Danube limes, except for Pannonia Superior and Inferior. It also reaches Dacia. Production centres are concentrated in Central (and East) France: in Aquitania (Les-Martres-de-Veyre, Lezoux, Lubie, Terre-Franche, and Toulon-sur-Allier), and Lugdunensis (Gueugnon, and Cournon). Medium-range nodes are mainly located in Britannia, where one-fourth of the overall places take part in the communities; the other nodes are sporadically spread in Aquitania, Belgica and Noricum. The structure of the network (number of locations and links) is similar to community 3.

In addition, the analysis of covered distances deserves a short discussion. Ranges between 1000 and 1300 kilometres characterise the distance from southern France (Lezoux, La Graufesenque, and Les Martres-de-Veyre represent one-quarter of the overall data) to the northern, western, and eastern border of the Roman Empire. Such borders stand for, in the North and the East, the Rhein-Danube limes and the Antonine and Adrian Walls, i.e. the military and the political limits of the Empire, and in the West, the geographical border of the Atlantic Ocean in the Iberian peninsula. fig. 12, describing La Graufesenque distribution, well clarify these aspects. In some cases, these physical and political barriers were exceeded, and further distances covered. Nevertheless, this range represents a peak, i.e. a significant volume of pottery reached these distances. TS entered the areas where the Roman garrisons resided, and it is ubiquitous in Roman camps along the limes (Van Oyen, 2015). As pointed out by Willis (2005, 6.2, 6.5.1), south Gaulish TS in Britain is more represented on military and urban sites than on smaller settlements, rural sites or villae. The reasons for this high presence could be put in relation to orders made directly by the army or to the free market that developed around the garrisons. For Britain, Willis (2005, 6.3, 13.1.1) suggests that military sites were supplied by a separate network, even if the rich collection of tablets coming from the forts of the Adrian Wall rarely mention orders of vessels for the army (Vindolanda Tablets Online). The picture is more complicated, and different distribution mechanisms have probably coexisted. Nevertheless, a market related to military supplies seems more convincing in the immediate post-conquest period, whereas a free market could be a winner in the long run. Interestingly, the dramatic fall of the distribution for distances higher than 1300 kilometres seems related to the fact that further distances were less profitable. This can be seen by looking at the distribution along the Danube. Legionary camps along the river more than 1300 kilometres were not reached or rarely reached.

Ranges between 0 and 300 kilometres also represent a peak. In this case, the distribution seems explainable with a short-range trade.

4. Conclusions

The method we presented seems promising and allows us to emphasise some methodological, theoretical, and archaeological points arising from the present work.

The first point involves the quality, quantity and the real interoperability of data (data that appears similar, but, indeed, archaeologically different). The applied method shows how the availability of high volumes of data (being unfortunately rare in Archaeology), joined with explanatory analysis, allows new insight into archaeological research. Availability of data mainly depends on three factors: digitisation of legacy data, datafication and open access to research data. Digitisation is turning analogue information into computer-readable formats; datafication is the act of transforming objects, processes, etc. in a quantified format, suitable to be tabulated and analysed (Gattiglia, 2015); open access to research data allows to use, re-use and redistribute data. Without the vast amount of open data created through digitisation by Samian Research (2008) and ArchAIDE consortium (2019), the latter will be implemented through datafication (Anichini & Gattiglia, 2017), this work would not have been conceived. In this research, we worked with a dataset composed of more than 300000 data. Only a massive amount of data fits a Big Data approach and allows new forms of quantification and associated data mining techniques that permit more sophisticated mathematical analyses to identify non-linear relationships in the data. Nonetheless, the quality of data is of paramount importance. As enlightened in the paper, some aspects of TS production are less understandable due to the different levels of the studies. This means that data could be fully interoperable but not completely comparable. Moreover, an archaeological interpretation of the results is essential. «The use of Big Data does not imply the end of archaeological theory, or even the end of archaeologists: no matter how comprehensive or well analysed the data are, they need to be complemented by big judgment» (Gattiglia, 2015).

The second point concerns the use of machine learning models, such as network analysis, for exploring and understanding archaeological data. This kind of analysis, despite the importance of relationships in archaeology, is not a mainstream part of the discipline. This poor attention could be related to a view of mathematical analysis as a reprise of processual methods with all the negative aspects of deterministic relationships that were stigmatised from the 1980s. On the contrary, this analysis is useful for illuminating the correlations between data, making clear the patterns and offering archaeologists novel and invaluable insights. Identification of possible correlations does not mean deterministic reasoning; correlations do not imply causation. In other words, a correlation between two variables does not necessarily mean that one causes the other, or to use a logic argumentation is not a sufficient circumstance. As suggested by Tufte (2004, p. 4) «observed covariation is a necessary but not sufficient condition for causality, (...) but it sure is a hint». Indeed, correlation is used to infer causation; the critical point is that such inferences are made after correlations are confirmed as real, and all causational relationships are systematically explored (Aldrich, 1995; Bollier, 2010, p. 4; Pearl, 2009). Correlations offer pretty clear insights that help archaeologists in capturing a phenomenon not by recognising its inner workings but by «identifying a useful proxy for it» (Mayer-Schönberger & Cukier, 2013); in this way, correlations are useful for archaeological interpretation because archaeology, unlike the natural sciences, is further from the deterministic dualism of cause and effect. For this reason, mathematical approaches are effective since they inform rather than explain, and expose patterns for archaeological interpretation, providing the opportunity to test new hypotheses at many levels of granularity.

The third point deals with publishing the results of big data analysis. Present data have been published on the MOD repository, together with data visualisation tools. Visualisation is entirely part of the present work since web-based visualisation improves access to archaeological heritage and generates new understanding. In our case, visualisation (https://mappaproject.shinyapps.io/ArchAIDE/) is not displayed in the form of a 'classical' distribution map represented by dots and densities. Instead, it is conceived as a new approach in digital archaeological cartography that could provide exciting opportunities for the querying and analysing Big Dataset (Gillings, Hacigüzeller & Lock, 2019).

We finally concern the archaeological meaning of this work. In the last years, there have been attempts of using big data approaches and mathematical models to Roman Pottery and Roman Economy; among the others, the EU funded project MERCURY-SIMREC (2018), the research network, 'Big Data on the Roman Table' (Allison, 2018), and the work of Astrid Van Oyen (2016). Our research aims to suggest a highly specialised algorithm for analysing wide-scale archaeological records and demonstrate that this kind of analysis is of interest for understanding archaeological phenomena. Moreover, TS production can be seen as a proxy for the movement of other resources (Van Oyen, 2015). From this point of view, the algorithmic detection of communities together with a robust archaeological interpretation may help to build up a broad general picture of the movement of the staple goods around the Roman Empire.

Aknowledgements

This research was supported by EU Horizon 2020 grant agreement No. 693548. We thank all the members of the ArchAIDE (www.archaide.eu) team.

References

Allison, P. (2018). An Introduction to a Research Network: the rationale and the approaches. *Internet Archaeology 50*, https://doi.org/10.11141/ia.50.1

Anichini, F. & Gattiglia, G. (2017). Big Archaeological Data. The ArchAIDE project approach. In *Conferenza GARR_17. Selected Papers. The Data Way to Science*, Venezia 15-17 November 2017, 22-25.

Anichini, F. et al. (2020). Developing the ArchAIDE Application: A digital workflow for identifying, organising and sharing archaeological pottery using automated image recognition. *Internet Archaeology 52*, https://doi.org/10.11141/ia.52.7

ArchAIDE consortium (2019). ARCHAIDE Portal for Publications and Outputs [dataset]. York: Archaeology Data Service [distributor], https://doi.org/10.5284/1050896

ArchAIDE project (2016) - http://www.archaide.eu/. Retrieved 23 September 2020.

Aubert, J.-J. (1994). *Business Managers in Ancient Rome. A Social and Economic Study of institores 200 B. C.-A.D. 250.* Columbia studies in the classical tradition 21.

Bémont, C., & Jacob, J.-P. (Eds.) (1986). *La terre sigillée gallo-romaine. Lieux de production du Haut-Empire. Implantations, produits, relations.* Documents d'Archéologie française 6, Cambridge University Press.

Bergamini, M. (2004). Scoppieto (Terni). Scavo di un complesso produttivo di età romana (anni 1995-1998). *Notizie degli scavi di antichità 15/16*, 7-88.

Bergamini, M., & Manca, M.L. (2008). La produzione di terra sigillata nella media valle del Tevere. Indizi per la localizzazione dell'officina di L. Nonius. *Rei Cretariae Romanae Fautorum Acta 40*, 347-358.

Bet, P. (1988). Groupes d'ateliers et potiers de Lezoux (Puy-de-Dôme) durant la période gallo-romaine. In *SFECAG Actes du congrès d'Orange*, 221-241.

Bet, P., Delage, R., & Vernhet, A. (1994). Lezoux et Millau. Confrontation d'idées et de données. In S*FECAG Actes du congrès de Millau*, 43-61.

Bonacich, P. (1987). Power and Centrality: A Family of Measures. *American Journal of Sociology, 92*, 1170-1182.

Brughmans, T., Collar, A., & F. Coward (Eds.) (2016). *The Connected past. Challenges to Network Studies in Archaeology and History*. Oxford University Press.
Connolly, J., & Lake, M. (2006). *Geographical Information Systems in Archaeology.* Cambridge University Press.
Conspectus (1990). *Conspectus Formarum Terrae Sigillatae italico modo confectae*. Dr. Rudolf Habelt Gmbh.
Camodeca, G. (2006). Graffito con conto di infornata di sigillata tardo-italica da Isola di Migliarino (Pisa). In S. Menchelli & M. Pasquinucci (Eds), *Territorio e produzioni ceramiche: paesaggi, economia e società in età romana* (pp. 207-216). Pisa University Press.
DBpedia (2018). https://wiki.dbpedia.org/. Retrieved 30 September 2020.
Desbat, A. (2001). L'artisanat céramique à Lyon durant l'époque romaine. *Rei Cretariae Romanae Fautorum Acta 37*, 17-35.
FSF - Free Software Foundation (2018). https://www.fsf.org/. Retrieved 30 September 2020.
Fülle, G. (2000). *The organization of mass production of terra sigillata in the Roman Empire.* University of Oxford.
Gattiglia, G. (2015). Think big about data: Archaeology and the Big Data challenge. *Archäologische Informationen 38*, 113-124.
Gillings, M. Hacigüzeller, P. & Lock, G. (2019). On maps and mappings. In M. Gillings, P. Hacigüzeller & G. Lock (Eds.), *Re-Mapping Archaeology. Critical Perspectives, alternative mappings* (pp. 1-16). Routledge.
GNU General Public License, Version 2 (1991). https://www.r-project.org/COPYING. Retrieved 30 September 2020.
GNU project (2018). http://www.gnu.org. Retrieved 30 September 2020.
Hartley, B.R. & Dickinson B.M. (Eds.) (2008-12). *Names on Terra Sigillata. An Index of Makers' Stamps and Signatures on Gallo-Roman Terra Sigillata (Samian Ware), 9 vols*. Institute of Classical Studies, University of London.
Johnston, A. (1985). A Greek graffito from Arezzo. *Oxford Journal of Archaeology 14*, 119-124, https://doi.org/10.1111/j.1468-0092.1985.tb00236.x
King, A. (1980). A Graffito from La Graufesenque and 'samia vasa'. *Britannia 11*, 139-143.
Knappett, C. (2013). *Network Analysis in Archaeology: New Approaches to Regional Interaction.* Oxford Scholarship Online, https://doi.org/10.1093/acprof:oso/9780199697090.001.0001
Leaflet (2018). https://leafletjs.com/. Retrieved 30 September 2020.
Mercury - SIMREC (2018) https://projectmercury.eu/. Retrieved 30 September 2020.
Marichal, R. (1988). *Les Graffites de La Graufesenque*. Supplément á Gallia 47.
Menchelli, S., Capelli, C., Del Rio, A., Pasquinucci, M., Picon, M., & Thiron-merle V. (2001). Ateliers de céramiques sigillées de l'Etrurie septentrionale maritime: données archéologiques et archéométriques. *Rei Cretariae Romanae Fautorum Acta, 37*, 89-105.
Menchelli, S. (2013). Produzioni di Terra Sigillata nella Valle dell'Arno: evidenze archeologiche di un'economia in espansione. In G. Schörner (Ed.), *Leben auf dem Lande.: "Il Monte" bei San Gimignano: Ein römischer Fundplatz und sein Kontext* (pp. 299-255). Phoibos Verlag.
Newman M.E.J. (2006). Finding community structure in networks using the eigenvectors of matrices. *Physical Review E 74 (3),* 036104.
Olcese, G. 2004: Italian terra sigillata in Rome and the Rome area. Production, distribution and laboratory analysis. In J. Poblome, P. Talloen, R. Brulet & M. Waelkens (Eds.), *Early Italian Sigillata. The Chronological Framework and Trade Patterns. Proceedings of the First International ROCT-Congress Leuven, May 7 and 8, 1999* (pp. 279-298). Peeters Publishers.
Open Refine (2012). http://openrefine.org/. Retrieved 30 September 2020.
OpenStreetMap Foundation (2018). https://www.openstreetmap.org. Retrieved 30 September 2020.
Oswald, F. (1931). *Index of Potters' Stamps on Terra Sigillata ("Samian Ware")*. Margidunum.
Oxé, A. (1933). *Arretinische Reliefgefässe vom Rhein*. Joseph Baer.
Oxé, A., & Comforth, H. (1968). *Corpus Vasorum Arretinorum. A catalogue of the Signatures, Shapes and Chronology of Italian Sigillata.* Dr. Rudolf Habelt Gmbh.
Oxé, A., Comforth, H., & Kenrick, Ph. (2000). *Corpus Vasorum Arretinorum. A Catalogue of the Signatures, Shapes and Chronology of Italian Sigillata*. Dr. Rudolf Habelt Gmbh.
Passelac, M. (1992). Formes et techniques italiques dans les productions céramiques augustéennes du bassin de l'Aude. Mise en évidence d'un groupe d'ateliers. *Rei Cretariae Romanae Fautorum Acta 31(2)*, 207-229.
Passelac, M., Sabrié, R., & Sabrié, M. (1986). Centre de production de Narbonne. In C. Bémont & J.-P. Jacob (Eds.), *La terre sigillée gallo-romaine. Lieux de production du Haut-Empire. Implantations, produits, relations* (pp. 52-55). Documents d'Archéologie française 6.
Peacock, D.P.S. (1982). *Pottery in the Roman world: an ethnoarchaeological approach*. Longman.
Picon, M. (1973). *Introduction à l'étude technique des céramiques sigillées de Lezoux.* Université de Dijon, Faculté des sciences humaines, Publication du Centre des recherches sur techniques gréco-romaines 2.

Picon, M. (2002a). À propos des sigillées, présigillées et imitations des sigillées. Questions de “coûts” et de marchés. In *SFECAG Actes du congrès de Bayeux*, 345-356.
Picon, M. (2002b). Les modes de cuisson, les pâtes et les vernis de la Graufesenque. Une mise au point. In M. Genin & A. Vernhet (Eds.), *Céramiques de la Graufesenque et autres productions d'époque romaine. Nouvelles recherches. Hommages à Bettina Hoffmann* (pp. 139-163). Archéologie et Histoire romaine 7).
Picon, M., & Garmier, J. (1974). Un atelier d'Ateius à Lyon. *Revue archéologique de l'Est et du Centre-Est 25,* 71-76.
Picon, M., & Lasfargues, J. (1974). Transfert de moules entre les ateliers d'Arezzo et ceux de Lyon. *Revue archéologique de l'Est et du Centre-Est 25*, 60-69.
Pleiades (2018). https://pleiades.stoa.org/. Retrieved 30 September 2020.
Pons, P., & Latapy M. (2006). Computing Communities in Large Networks Using Random Walks. *Journal of Graph Algorithms and Applications, 10 (2)*, 191-218.
Pucci, G. (1990). A sigillata kiln in Valdichiana (central Etruria). *Rei Cretariae Romanae Fautorum Acta 27*, 15-23.
Pucci, G. (1992). La fornace di Umbricio Cordo. L'officina di un ceramista romano e il territorio di Torrita di Siena nell'antichità. All’Insegna del Giglio.
R version 3.4.3 (2017). https://www.r-project.org/. Retrieved 30 September 2020.
Samian Research (2008). https://www1.rgzm.de/samian/home/frames.htm. Retrieved 30 September 2020.
Sforzini, C. (1987). Vasai “aretini” in area falisca: l'officina di Vasanello. In G. Maetzke (Ed.), *La civiltà dei Falisci. Atti del XV convegno di studi etruschi et italici, Civita Castellano - Forte Sangallo, 28-31 maggio 1987* (pp. 251-281). Leo Olschki Editore.
Soricelli, G. (2004): La produzione di terra sigillata in Campania. In J. Poblome, P. Talloen, R. Brulet & M. Waelkens (Eds.), *Early Italian Sigillata. The Chronological Framework and Trade Patterns. Proceedings of the First International ROCT-Congress Leuven, May 7 and 8, 1999* (pp. 299-307). Peeters Publishers.
SPARQL Query Language for RDF (2008). https://www.w3.org/TR/rdf-sparql-query/. Retrieved 30 September 2020.
Unxos GmbH, GeoNames (2018). https://www.geonames.org/. Retrieved 30 September 2020.
Vaccaro, E., Capelli, C., &Ghisleni, M. (2017). Italic sigillata production and trade in the countryside of central Italy: new data from the Excavating the Roman Peasant Project. In T. De Haas & G. Tol (Eds.), *The Economic Integration of Roman Italy. Rural Communities in a Globalizing Economy* (pp. 231-262). Brill https://doi.org/10.1163/9789004345027_011
Van Oyen, A. (2015). The Roman city as articulated through Terra Sigillata. *Oxford Journal of Archaeology 34 (3)*, 279-299.
Van Oyen, A. (2016). *How Things Make History. The Roman Empire and its Terra Sigillata pottery*. Amsterdam University Press.
Vindolanda Tablets Online http://vindolanda.csad.ox.ac.uk/.Retrieved 30 September 2020.
Willis, S. (2005). Samian Pottery, a Resource for the Study of Roman Britain and Beyond. The results of the English Heritage Funded Samina Project. An E-Monograph. *Internet Archeology 17* https://intarch.ac.uk/journal/issue17/1/toc.html