The ecology of Roman trade. Reconstructing provincial connectivity with similarity measures

Xavier Rubio-Campillo[a,*], Jean-Marc Montanier[b], Guillem Rull[c], Juan Manuel Bermúdez Lorenzo[d], Juan Moros Díaz[d], Jordi Pérez González[d], José Remesal Rodríguez[d]

[a] *School of History, Classics and Archaeology, University of Edinburgh, UK*
[b] *Barcelona Supercomputing Centre, Spain*
[c] *SIRIS Lab, Research Division of SIRIS Academic, Spain*
[d] *CEIPAC, Universitat de Barcelona, Spain*

Abstract

The creation of the Roman Empire promoted the connectivity of a vast area around the Mediterranean sea. Mobility and trade flourished over the Roman provinces as massive amounts of goods were shipped over thousands of kilometres through sea, rivers and road networks. Several works have explored these dynamics of interaction in specific case studies but there is still no consensus on the intensity of this connectivity beyond local trade.

We argue here that the debate on the degree of large-scale connectivity across the empire is caused by a a lack of appropriate methods and proxies of economic activity. The last years have seen an improvement on the availability of evidence as a growing amount of datasets is collected and published. However, data does not equal knowledge and the methods used to analyse this evidence have not advanced at the same pace.

A new framework of connectivity analysis has been applied here to reveal the existence of distinctive trade routes through the provinces of the Western region of Rome. The amphora stamps collected over more than a thousand sites have been analysed using quantitative measures of similarity. The patterns that emerge from the analysis highlight the intense connectivity derived from factors such as the spatial closeness, presence of military units and the relevance of the Atlantic sea as a main shipping route.

Keywords: Rome, trade, amphora stamps, MRPP, Jaccard

1. Introduction

The intensity of provincial connectivity is one of the most debated aspects of the Roman economy. Hypotheses oscillate between a unified market defined by a constant flow of goods through long-range trade to isolationist approaches based on autonomous regions with little contact, with some exceptions (Temin, 2001; Bang, 2007). Both archaeological and written sources indicate that there was a large diversity of scenarios as connectivity was not homoge-

*Corresponding author
Email address: xavier.rubio@ed.ac.uk (Xavier Rubio-Campillo)

neous and some regions were much more integrated than others. A key player of this integration was the Roman army as its supply required the import of vast quantities of products (Scheidel et al., 2007, 591). They were mostly produced in specialised provinces and required large-scale trade. A good example of this connectivity is the shipping of massive amounts of olive oil from the Baetican province to Britannia after its conquest (Remesal Rodríguez, 2011, 60). These basic goods were distributed amongst military garrisons but it seems probable that the trade network rapidly expanded to supply civilian settlements (Williams and Peacock, 1983). Other goods such as exotic foods were widely shipped to distant urban centres using non-military redistribution networks (Livarda and Orengo, 2015; Orengo and Livarda, 2016). However, the general question remains unanswered: how frequent and intense were these economical contacts beyond specific case studies?

The topic has a renewed interest as an increasing corpus of datasets including archaeological, epigraphical and written sources is becoming available. One example of this exciting explosion of evidence is the Orbis project which is focused on exploring the cost of mobility along the entire Roman Empire (Scheidel, 2015). Other initiatives such as the Pelagios project aims at aggregating tens of different databases to generate a multifaceted view of the classical world (Barker et al., 2016). This collection of evidence is a critical step towards understanding the Roman economy but its use also presents several challenges (Bowman and Wilson, 2009, 3-87). As other authors have pointed out this data is riddled with biases and uncertainty up to the point where it is difficult to find patterns beyond the noise (Bevan, 2014; Wilson, 2009). The datasets being merged often have diverse temporal and spatial dimensions and were collected by different formats and methods while the projects creating them use different theoretical approaches to the past (Bevan, 2015; Calvanese et al., 2016).

The aim of integrating datasets should be combined with the creation of methods able to tackle the complexities of the existing evidence (Brughmans and Poblome, 2016). Roman studies typically use descriptive statistics and linear regressions to analyse relations between variables (Wilson, 2009) but these generic approaches were not designed to face the uncertainty of archaeological data. First, our sample sizes are usually very low as they consist of tens or hundreds of data points for a vast region that did not remain stable over time. Second, the data points have a large degree of uncertainty which is badly captured by exploratory methods and require the use of probabilistic approaches to the past (Yubero-Gómez et al., 2016; Crema, 2015; Bevan et al., 2013a,b). Finally, the multiple biases generated by the archaeological process should be taken into account while analysing the existing evidence (Bevan, 2012; Rubio-Campillo et al., 2012).

This work presents a method to study provincial connectivity through the estimation of similarity indexes. The premise of this analysis is that regions that share trade routes should exhibit more similar cultural traits between them than with the rest of the empire. We reconstruct here the dynamics of provincial trade based on a well-tested proxy of long-range trade: the stamps found in amphorae containers found over the entire Roman Empire (Scheidel et al., 2007, 690). By applying a Null-Hypothesis Significance Testing Framework based on ecological methods we explore two specific research questions: a) was large-scale trade related to the provincial structure? and b) can we find patterns of connectivity between provinces beyond spatial closeness?

The next two sections define the dataset and the methods we used for this large-scale analysis. The fourth section presents the results of the analysis which are discussed and interpreted in section five. The text finishes with a summary of the method and its potential contribution within the current debates on the discipline.

2. Patterns of trade in the Roman empire

Clay amphorae are arguably the archaeological artefacts that best represent trade dynamics in the classic world (Bevan, 2014). These standardised containers were used to transport large quantities of liquids and other goods through a dense network of sea and river routes. Maritime shipping was the fastest and cheapest transport system so amphorae were massively distributed over the entire Roman empire. At the same time amphorae were functional and robust because they were designed to be transported aboard ships that may be crossing hazardous waters. This robustness and widespread use has allowed amphorae to survive in higher quantities and frequencies than containers serving a similar purpose such as wooden barrels (Tchernia, 1986).

The study of these containers plays a key role in our understanding of the Roman economy thanks to their visibility in the archaeological record (Greene, 1986, 162). The production of an amphora type is typically linked to a specific area and product so a trade link can be suggested between the production place of a type and the sites where the amphorae of this type are found. The aggregation of large volumes of findings reveals the degree of specialisation of certain provinces that shipped thousands of amphorae filled with a single product to distant consumption places; this dynamic can be seen in Baetica for olive oil (Remesal Rodríguez, 1998; Funari, 1996) and some areas of Italia for wine (Paterson, 1982; Loughton, 2003).

The use of this archaeological proxy also presents some challenges. Elsewhere has been argued that the information provided by amphorae findings can be potentially biased by reuse activities (Peña, 2007, pp. 61-208). These biases could affect distribution patterns at least in two different aspects: a) transportation to a new destination and b) refill with a different substance than the original.

The first scenario would see an empty amphora refilled and shipped to a different location. The archaeological record does not allow us to track the route of the amphora which will always be found in the last location it was shipped. This bias would not heavily affect large-scale analysis such as the one we present here because the evidence for long-range reuse is very scarce (Peña, 2007, p. 72). If short-range reuse was frequent then the amphorae found on nearby sites would be more homogeneous but it would not affect the role of the dataset as proxy of long-range trade.

The second scenario would break univocal ties between specific amphora types and their contents. While this bias does not affect the current work given our focus on stamps it is certainly a relevant barrier to improve our understanding of Roman trade and requires further exploration (probably through residue analysis techniques, see

Pecci et al., 2017).

A significant percentage of these amphorae were stamped on one of their handles with a code of letters and symbols. Most of these codes are *tria nomina* identifying an individual linked to trade activities, albeit it is difficult to know if this person was involved in the production of the container or its contents (Remesal Rodríguez, 1998; Funari, 1996). In any case these codes highlight the dynamics of trade because they were not unique: amphorae found in distant sites were stamped with the same code while containers found in the same place often exhibit a diversity of them. The study of the frequencies of codes has found interesting patterns on their spatiotemporal distribution, and for this reason they seem a good proxy for long-range trade in the classic world (Remesal Rodríguez, 1998; Berni Millet, 2008; Broekaert, 2015; Rubio-Campillo et al., 2017).

This long tradition of amphora stamps analysis has been mostly focused on single sites or provinces. Here we use this proxy to identify links within the Western part of the Roman empire by comparing the similarity of stamp codes found across thousands of Roman sites. The hypothesis to test can be defined as follows: sites receiving goods through different trade networks would be supplied by distinct producers, so we should find differences in the stamps found on these sites. In a majority of sites only a small number of stamps has been found, but if this hypothesis is correct then a large dataset should exhibit a pattern significantly distinctive from a random distribution of code stamps. In addition, if a group of provinces were more intensely connected because they shared trade routes then some code stamps should be more present in these provinces than in the rest of the areas.

The database used to test our working hypotheses is the Corpus of amphorae with Latin epigraphy compiled by the CEIPAC group over 30 years (Remesal Rodríguez et al., 2015). For each record in the dataset the following information was compiled: a) *id* of site where it was found, b) *province* where the archaeological site was located and c) stamp *code*. At present the Corpus contains 32.375 amphora stamps from which the amphorae collected in the city of Rome were removed for two reasons. First, the economic activities of the capital's supply were unique given its size and political role. Second, the amount of evidence collected in Rome is so large compared to the rest of the sites that the entire analysis would be biased towards this city. As a consequence the dynamics of the rest of the territory would be masked by the large weight of the capital. The remaining set of 24.092 stamps displayed 5.539 unique codes and is distributed over 1.278 sites covering a large percentage of Europe as depicted in Figure 1.

It is worth noting that around 25% of the stamps are not complete due to fragmentation or erosion. A previous study showed that the impact of this uncertainty in large-scale analysis was low (Rubio-Campillo et al., 2017). As a consequence we have integrated the fragmented stamps in the dataset without further issues.

The dataset contains a wide diversity of amphora types; nevertheless a majority of stamps has been found on Dressel 20 Baetican amphorae containing olive oil and Brindisian amphorae transporting olive oil or wine. The frequency distribution of the most popular amphorae types can be seen in Figure 2.

Figure 3 shows the heterogeneity of the

Figure 1: Spatial distribution of amphora stamps collected in the CEIPAC database. Most of the dataset comes from sites in the Western area of the Roman empire with the highest densities located at the Mediterranean coast and the provinces with strongest military presence (Britania and Germania)

sample both in terms of number of sites per province and number of stamp codes per site. Provinces such as Italia, Narbonensis, and the two Germania have a large quantity of stamps spread over several sites while in most provinces less than 100 stamps were collected. The sites with a higher number of findings are located in the provinces with larger sample size while the sites in the rest of the provinces typically show less than 10 code stamps. This pattern can be explained by a strong intensity bias as archaeologists working on some regions of Europe would have more interest in recording amphora stamps than areas where this type of studies is less common. The challenge then is to use appropriate methods able to detect spatial patterns despite this diversity of sample sizes.

3. Methods

The analysis of this dataset was performed in three steps: a) creation of a dissimilarity matrix between sites b) evaluation of province significance and c) identification of province clusters.

3.1. Jaccard distance matrix

Dissimilarity between two sites was based on the number of stamp codes that were

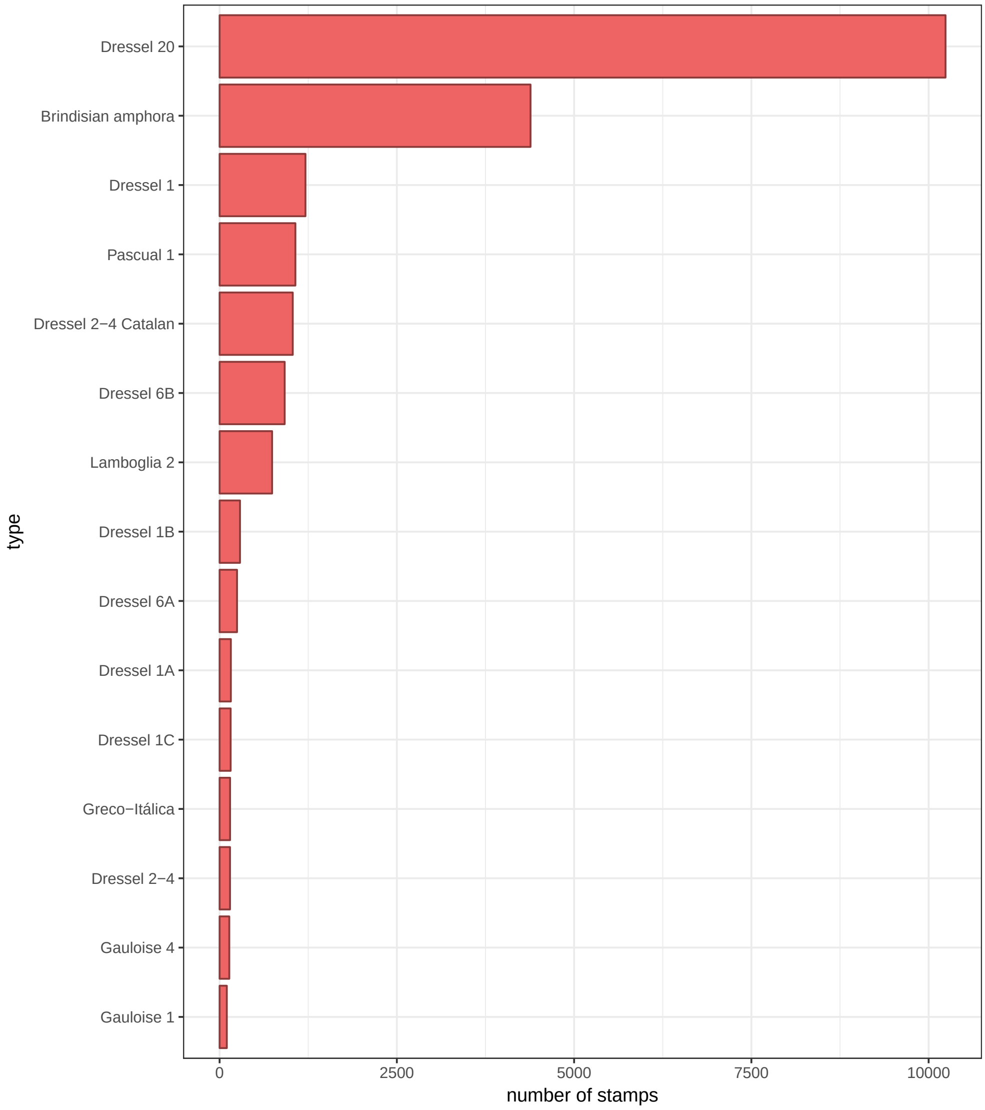

Figure 2: The CEIPAC database comprises a large diversity of containers with a total of 115 amphora types. This figure displays the frequency of the types having at least 100 stamps

present on one location and absent on the other one. This was quantified with a popular similarity measure known as Jaccard distance. The distance between the sets of codes c_i and c_j collected in a pair of sites i and j is defined as the ratio between the number of codes found in both sites and the number of codes found at least in one site as defined in Equation 1:

$$D_{Jaccard}(i, j) = 1 - \frac{|c_i \cap c_j|}{|c_i \cup c_j|} \quad (1)$$

The Jaccard distance is bounded between 0 (i.e. the sites have exactly the same stamp codes) and 1 (i.e. the sites do not share any code). The pairwise computation of this index for the entire dataset generated a squared dissimilarity matrix of 1.278 rows

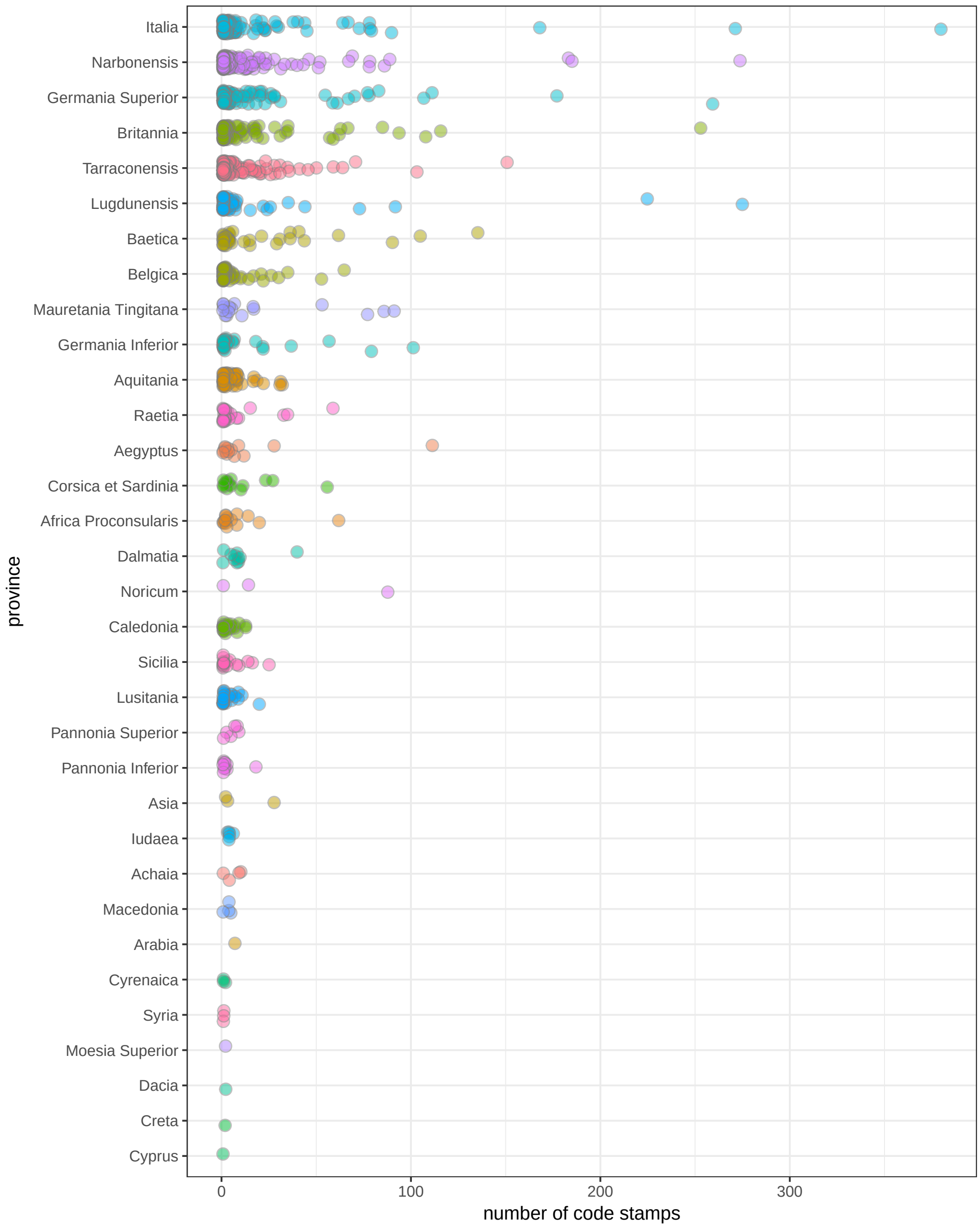

Figure 3: Distribution of sites based on its number of amphora code stamps (X axis) and province (Y axis). Each dot is a site and the provinces are sorted in decreasing order based on the total amount of stamps found on each province. Repetitions of the same stamp code on a site are counted only once

by 1.278 columns (i.e. number of sites). The average distance was close to 1 as most of the sites had a small number of unique stamps.

3.2. Multi-Response Permutation Procedure

The second step required a comparison of the Jaccard distance between sites against their province. We estimated the significance of the first hypothesis by evaluating the opposite *null* hypothesis: *The Jaccard distance between 2 sites is independent of their provincial attribution*. This is equivalent to compute the probability that two random sites from the entire dataset have a lower Jaccard distance than two sites randomly sampled from the same province; if this probability is low enough then we can reject the *null hypothesis*, thus suggesting that provincial structure played a role on trade routes.

The complex requirements of this test were met by the use of the Multi-Response Permutation Procedure (MRPP) (McCune and Grace, 2002; Mielke et al., 1976). MRPP was designed to analyse ecological datasets presenting similar challenges than the ones posed by archaeological data (e.g. fragmentation, noise, sampling biases). First, MRPP does not assume any specific distribution of responses unlike similar methods such as MANOVA. Second, this approach allows to give different values to the weight of each group in the final result. This was relevant in our analysis because the information of each province was not homogeneous as the number of sites per province was diverse (e.g. there were 3 sites in Syria for 182 sites in Narbonensis). Finally, MRPP accepted Jaccard as a distance metric so no further data transformations were required. Despite these advantages to our knowledge the method has only been applied once in archaeological research (Rodgers, 1987).

MRPP evaluates the *null* hypothesis by comparing the average distance for the entire dataset against the average distance for elements of the same group (i.e. province) weighted by its sample size. It does so by performing random permutations between elements and assessing changes in this distance.

For a set of provinces P and a set of stamp codes C_P the algorithm can be defined as follows:

1. compute the average Jaccard distance $\overline{D_p}$ between the sites of each province p in P.
2. compute the weight W_p of a province p based on the ratio between its number of sites s and the total number of sites in the sample:

$$W_p = \frac{s \in p}{\sum_{i=1}^{P} s \in i}$$

3. define an observed Delta value δ as the overall weighted mean of within-group means of distances:

$$\delta = \sum_{p=1}^{P} D_p W_p$$

4. permute the provinces associated with the different sites and compute δ again (this step is performed thousands of times).

The p-value is given by the percentage of permutations with δ lower or equal than the observed value computed in step 3. The algorithm also quantifies an effect size A suggesting the predictive power of the group (see McCune and Grace, 2002, for details).

3.3. Clustering

MRPP tests the statistical significance of the groups but it does not provide insights into the similarity between provinces. Our second hypothesis requires additional methods to group the provinces based on the stamp codes that can be found in their set of sites. This was achieved by creating a second matrix of mean within- and between-province distances from the results of the MRPP. The clustering algorithm *neighbour joining* was then used to group the provinces (Saitou and Nei, 1987). This algorithm was chosen because it generates an unrooted binary tree from a matrix of dissimilarities without making any assumptions on the existence of a root node (which did not exist in this case). The results could then be visualised using a cladogram as a means to evaluate what groups were created by the method.

4. Results

The results of these methods were organised by the two original research questions.

4.1. Significance of the provincial structure

The application of the MRPP algorithm generated the results that can be observed in Table 1:

	value
p-value	< 0.001
observed δ	0.9974
within-province distance	0.9939
A	0.0035

Table 1: Results of MRPP using the entire sample

The observed average within-province distance δ is consistently lower than the permuted δ values. As a consequence the *null hypothesis* that there is no relation between the province of a site and its stamps has a very low probability. However, the effect size A is low despite the high significance of the provincial structure.

The extreme diversity and sparsity of the dataset causes all Jaccard distances to be very high due to the low number of codes found in a majority of sites. Hundreds of sites have 5 or fewer codes so the probability that two of these locations share one code is very low, thus generating an A close to 0. This issue is summarised in Figure 4 where it is observed that the number of codes per site is not normally distributed. A majority of sites has one or two codes while the shape has a very long tail due to a small group of sites where hundreds of stamps were recorded.

This uneven distribution of codes is probably caused by excavation biases as most sites have not been fully excavated or have not published all their findings. The large number of sites with a small amount of codes is adding noise to the general picture by increasing the average Jaccard distances between sites.

This impact can be explored by repeating MRPP for filtered datasets of sites having at least a Minimum Number of Codes (MNC). An iterative process was performed with MNC values ranging from 1 to 100 (being MRPP with $MNC = 1$ equivalent to the previous analysis). Results can be seen in Figure 5

The signal given by A gradually intensifies as we discard sites with a low number of codes. It reaches a critical value an order of magnitude higher than the previous result when MRPP is computed on sites with at least 70 codes. The number of sites used in the analysis is rather low at this point and as a consequence A decreases again for $MNC \geq 75$.

4.2. Provincial similarity

A given MNC value was required to create the distance matrix of within- and between- province dissimilarities. The choice of MNC needed to balance the effect size A against the number of sites involved in the analysis; if the value of MNC

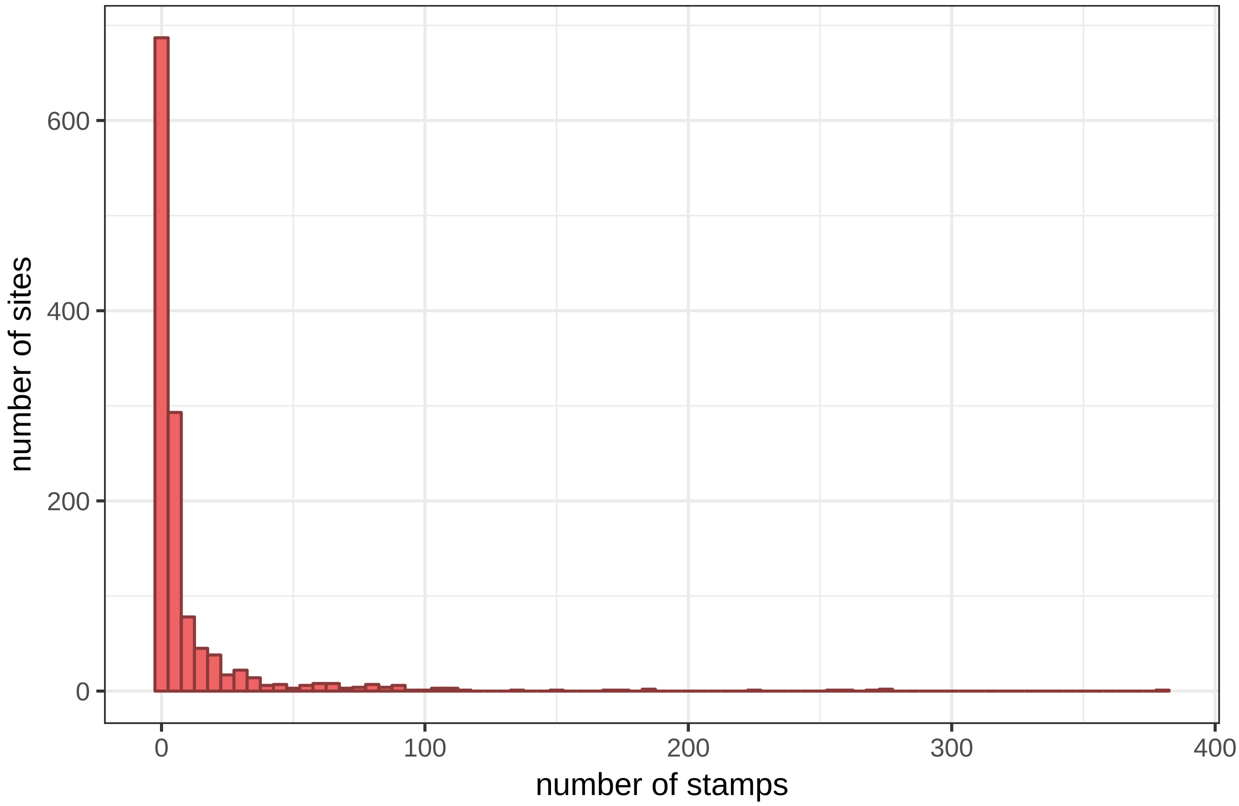

Figure 4: Histogram illustrating the number of code stamps per site with a bin width of 5. The distribution is extremely skewed with 800 sites containing less than 5 codes and a single site with 380 stamps (*Ostia Antica*)

was too low then it would contain a low signal-to-noise ratio while high MNC values would limit the number of provinces used in the clustering because at least 2 sites per province are required. $MNC = 20$ was chosen as a compromise because it had a good effect size $A = 0.016$ and a reasonable number of 154 sites. The resultant distance matrix can be seen in Figure 6.

The matrix was used to generate the cladogram seen in Figure 7. Two clusters emerge from this visualisation: a group of tightly linked Mediterranean provinces and a second group comprising the northern *limes* of the Empire.

Additional analysis were conducted to explore the impact of parameter MNC and dataset variations in the final results. Supplementary Figure 1 shows the comparison of the cladograms reconstructed for different MNC values. This parameter exploration was performed from the complete dataset with $MNC = 1$ to the highest A value at $MNC = 75$. Supplementary Figure 2 displays cladograms on two different subsets of the original dataset: a) stamps found on Dressel 20 amphorae and b) stamps found on other types of amphorae. This exploration was required to assess if the predominance of Dressel 20 in the dataset was responsible for the similarity patterns.

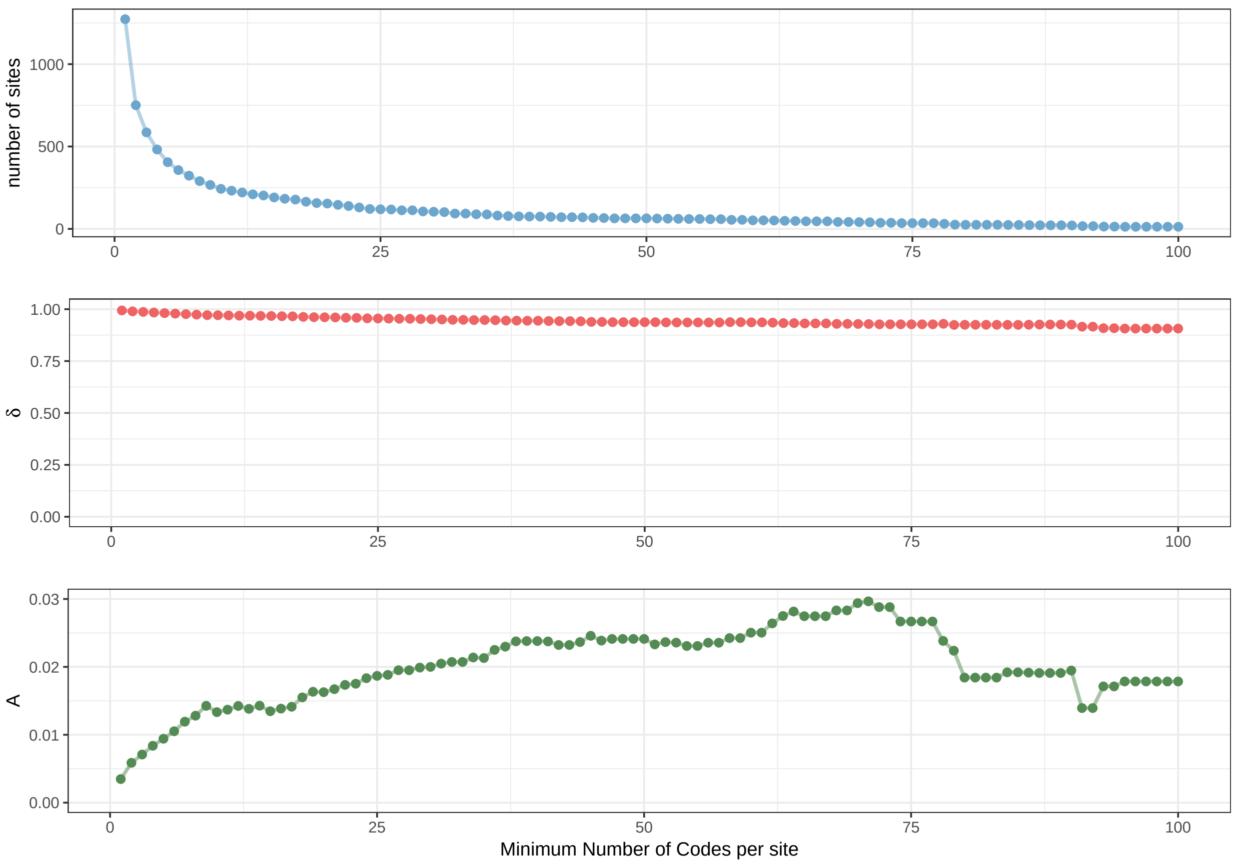

Figure 5: MRPP applied to filtered datasets. An MRPP has been performed to the collection of sites given by each MNC (X axis). The Y axis of the three graphs show: top) number of sites with given MNC, middle) average within-group distance or δ and bottom) effect size or A. The p-value for the entire dataset is consistently below 0.01

Insert Supplementary 1 here Insert Supplementary 2 here

5. Discussion

The analysis revealed non-trivial patterns of distribution and these results confirm that amphora stamps are good proxies of long-range trade. First, provincial structure played an important role on the distribution of liquid goods. Second, provinces that were supplied through the same network exhibit higher similarity of stamp codes. Finally, the approach provides insight into a diversity of factors including the impact of military units on logistics or the relative intensity of multiple trade routes.

The first test suggests that the stamps found in a site are related to the province where this site was located. It does not mean that trade was organised independently on every province, but it shows how distant regions of the Empire were supplied by different trade networks based on their code stamps. It is worth noting that a large percentage of the dataset is made by containers produced in specialised locations such as the Dressel 20 olive oil amphorae in the Guadalquivir river (Mattingly, D.J.,

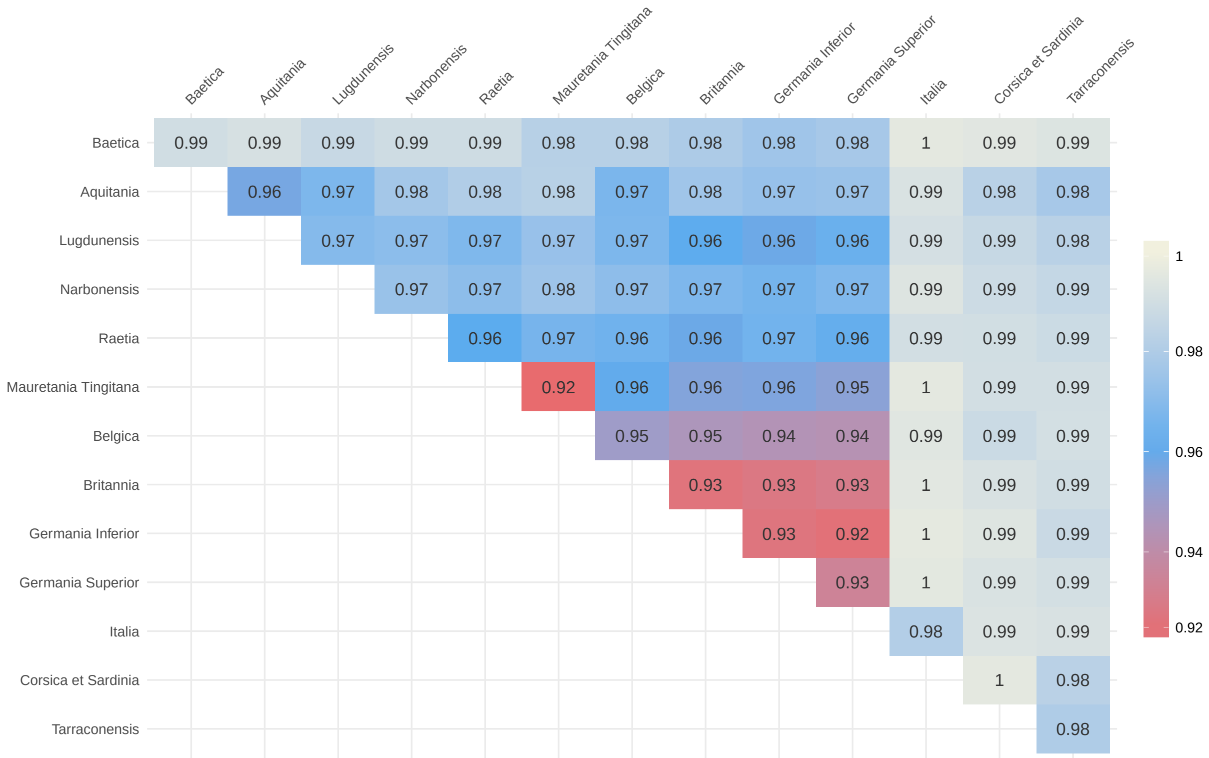

Figure 6: Distance matrix of mean within- and between- provincial distances for $MNC = 20$. Provinces with higher similarity are coloured in red tones while differences close to 1 are depicted in white

1988; Remesal Rodríguez, 1998). The workshops where these amphorae were made are located on a small segment of 20 kilometres along the riverside but they regularly shipped olive oil to different locations based on the code that was stamped on them. The patterns found on the distribution of stamps can only be caused by strong links between the workshops where the amphorae were made and the consumption places where they were found; if this relation did not exist then the result would be a random distribution over the Empire (so the *null* hypothesis would not be rejected).

The degree of provincial connectivity found in our dataset can mostly be explained by geographical distance. Good examples of this pattern are observed in the clustering of most Western Mediterranean regions (Corsica and Sardinia, Italia, Tarraconensis and Baetica) and Northern Gallia (Lugdunensis and Aquitania). It is worth noting that this spatial structure has been derived from stamp similarity and no spatial data was used as input of the analysis.

This relation between spatial closeness and stamp similarity can be explained by a combination of similar trade routes and intense local trade, including the reuse and reshipment of amphorae from local trade (Foley et al., 2012; Pecci et al., 2017). If future studies assess the relevance of amphorae reuse in Roman shipping then we will be able to improve even further our understanding of these similarities. In any case, the result appropriately captures the impor-

Figure 7: Similarity cladogram generated with neighbours joining algorithm for sites with $MNC = 20$

tance of Tobler's First Law of Geography in the case of economical dynamics: *everything is related to everything else, but near things are more related than distant things* (Tobler, 1970).

Spatial closeness is the main driver of connectivity but the cladogram also shows two interesting exceptions to this general rule. The first one is the similarity of provinces with strong military presence. The supply of the legions controlling the boundaries of the Roman Empire was one of the most important centres of demand for the empire (Scheidel et al., 2007, 575-576). The analysis shows the relevance of this factor as it breaks the pattern of spatial adjacency by grouping provinces with intense military activity. Germania Superior, Germania Inferior and Britannia form the more homogeneous group of the entire dataset which suggests that the military units deployed there received their olive oil and wine supply from the same producers (Carreras Monfort and Funari, 1998). The cluster is directly linked to Mauretania Tingitana; this province is thousands of kilometres apart from the northern group but it is more similar to the provinces of the German limes than to any other province, including the entire Mediterranean. Mauretania was considered a frontier province due to constant clashes between the Roman army and local seminomad groups so military units deployed here could have shared the same trade network than the legions stationed 3000 kilometres north (Knight, 1991; Cravioto, 2002; Pons Pujol, 2009).

The second pattern breaking Tobler's law is the similarity of provinces over the Atlantic-Rhine route. The four provinces located along the course of the Rhine river (Raetia, Germania Superior and Inferior) are linked to Britannia and Belgica. In contrast, Germania Superior and Germania Inferior are distant from Gallia Narbonensis which is the province where the Rhône river ends its course into the Mediterranean sea. This difference between the two rivers suggests that the Atlantic route had more intense long-range trade than the Rhine-Rhône river route. This result provides some insight into the current debate on the route network that supplied the legions garrisoning the German limes. On the one hand, several authors point out that the Atlantic route was too dangerous for the ships of this period while major harbour structures have been found in these two rivers (Fulford, 1992; Marlière, 2001). On the other hand, the Atlantic route would have been used if ships were able to safely move through the hazards of the ocean (Remesal Rodríguez, 1986, 2008, 2010). If this option was possible then the option would have been significantly cheaper in terms of cost and time (Greene, 1986, 39-41). Recent archaeological works are strengthening this hypothesis as they have discovered evidence for large-scale harbour facilities in the Atlantic facade (Carreras Monfort and Morais, 2012; Morillo et al., 2016). Our result supports this new perspective by highlighting the similarity of code stamps found in the Atlantic provinces in contrast with the low similarity between Germania and Narbonensis. This result implies that the Rhine-Rhône was not frequently used for long-range trade but other authors have highlighted the relevance of the Rhine-Rhône route for wine barrels (Marlière, 2001). This opposite results could be explained by the lack of statistical testing methods or the fact that different containers may have followed a diversity of routes.

The relevance of Baetican Dressel 20 amphorae in our dataset may also indicate that this was the main product being shipped through the Atlantic. Supplementary Figure 2 shows that this may be the case as Narbonensis is close to Germania Superior when the analysis excludes stamps found in Dressel 20. However, this result is heavily affected by sample size which forces us to discard some key provinces without the minimum number of sites (i.e. Germania Inferior). In any case, our analysis suggests that amphorae containing olive oil and wine arrived to Germania and Belgica more frequently through the Atlantic ocean than through the Rhône-Rhine route.

6. Concluding remarks

The method presented here was able to answer the research questions while tackling the challenges posed by the complexity of the dataset. The combination of similarity indexes with a statistical permutation test has provided valuable insights into the dynamics of trade within the Roman Empire. This result also confirms the utility of amphora stamps as archaeological proxies of economic activity despite its high levels of uncertainty. For example, the ties between Mauretania Tingitana and the German limes was found despite the low volume of information currently available for Northern Africa (Teichner and Pons Pujol, 2008). This result showcases how this new approach is able to detect relevant signals of trade amongst the noise of fragmented archaeological data.

Nevertheless, the approach has limitations as any other method. First, the analysis is effective for large-scale resolutions and it would provide unreliable results in case of being applied at a lower scale, due to the need of large sample sizes. Additionally, we do not have any temporal information beyond amphora classifications and for this reason this analysis cannot be used to track change over time. This limit is defined by the dataset; the method could be extended by introducing temporal dynamics based on probabilistic approaches (Yubero-Gómez et al., 2016). However, the coarse scale of chronologies based on amphora classifications would possibly decrease the statistical robustness of the results.

Our approach could also be potentially complemented by spatial analysis for comparisons between stamp similarity and spatial closeness. This could strengthen the results but initial attempts based on Mantel tests suggested that spatial patterns were too sensitive to the uncertainty of this dataset. A quick glance to Figure 1 suggests a major impact of intensity and excavation biases over the analysed territory that would heavily affect any result obtained by pure spatial techniques.

The method presented here generates new opportunities for our understanding of the Roman economy despite these limitations. The same approach could be applied to identify differences in trade routes beyond provincial affiliation such as the existence of unique logistic networks for military or civilian sites, the use of different river routes over the territory or distinctive distributions of amphora types. The combination of similarity distances with statistical tests also opens the door to comparative analysis between the routes followed by different goods and types of containers. The use of other proxies is important because amphora is the most visible archaeological proxy for long-range trade and by its very own nature its importance can overemphasize provincial connectivity (Woolf, 1992). The analysis of goods more vulnerable to postdepositional processes such as textiles may provide a different perspective by highlighting the complexity and diversity of Roman trade (Greene, 1986, 13-15).

The last years have seen a dramatic increase in the quantity and quality of datasets on the Roman empire, but data does not automatically transform into knowledge. This work illustrates the current need for hypothesis testing in the study of past economies. Quantitative analysis is not as common in Roman archaeology as it is in other archaeological fields where statistics have been regularly applied for decades (Thomas, 1978). The study of Roman economy needs appropriate quantitative methods able to tackle the challenges posed by archaeological evidence in order to identify meaningful patterns in complex datasets (Bevan, 2015). The field needs to move forward from basic exploratory data analysis towards the use of new frameworks able to test specific working hypotheses. This combination of new datasets and methods is the only way to answer the big questions of the field and advance in our understanding of the complexities of the classic world.

Acknowledgements

Funding for this work was provided by the EPNet project (European Research Council Advanced Grant - 340828). GR was funded by the Torres Quevedo Grant PTQ-15-07419. We would like to thank María Coto, Simon Carrignon and two anonymous reviewers for their comments on previous versions of the manuscript. Province boundaries were based on the work of Ancient World Mapping Center (2016). All analysis have been performed in R using the libraries *vegan ggtree* and *ggplot2* (Oksanen et al., 2007; Yu et al., 2017; Wickham, 2009).

Source code and datasets are available under Open licenses and can be freely accessible from `https://github.com/xrubio/ecologyStamps`.

Research was conceived by JMM, XRC and JRR. Dataset was collected by JMBL, JMD and JPG and processed by GR. JMM and XRC performed the analysis. XRC wrote the paper.

References

Ancient World Mapping Center, 2016. Roman empire provinces ad 200. `http://awmc.unc.edu/awmc/map_data/shapefiles/political_shading/roman_empire_ad_200_provinces.shp`, accessed: 2016-11-21.

Bang, P. F., 2007. Trade and Empire In Search of Organizing Concepts for the Roman Economy. Past & Present 195 (1), 3–54.

Barker, E., Simon, R., Isaksen, L., de Soto Cañamares, P., 2016. The pleiades gazetteer and the pelagios project. In: Berman, M., Mostern, R., Southall, H. (Eds.), Placing Names: Enriching and Integrating Gazetteers. Indiana University Press, pp. 97–109.

Berni Millet, P., 2008. Epigrafía anfórica de la Bética: nuevas perspectivas de análisis. Vol. 29 of Instrumenta. Edicions Universitat Barcelona, Barcelona.

Bevan, A., 2012. Spatial methods for analysing large-scale artefact inventories. Antiquity: a quarterly review of archaeology 86 (332), 492–506.

Bevan, A., 2014. Mediterranean Containerization. Current Anthropology 55 (4), 387–418.

Bevan, A., 2015. The data deluge. Antiquity 89 (348), 1473–1484.

Bevan, A., Conolly, J., Hennig, C., Johnston, A., Quercia, A., Spencer, L., Vroom, J., 2013a. Measuring chronological uncertainty in intensive survey finds: a case study from Antikythera, Greece. Archaeometry 55 (2), 312–328.

Bevan, A., Crema, E., Li, X., Palmisano, A., 2013b. Intensities, interactions and uncertainties: some new approaches to archaeological distributions. In: Bevan, A., Lake, M. (Eds.), Computational approaches to archaeological space. Left Coast Press, pp. 27–52.

Bowman, A., Wilson, A., 2009. Quantifying the Roman economy: methods and problems. Vol. 1. Oxford University Press.

Broekaert, W., 2015. Sticky Fingers. The Investment Structure of the Spanish Oil Business. Cahiers Mondes anciens 7.

Brughmans, T., Poblome, J., 2016. Roman bazaar or market economy? Explaining tableware distributions through computational modelling. Antiquity 90 (350), 393–408.

Calvanese, D., Liuzzo, P., Mosca, A., Remesal Rodríguez, J., Rezk, M., Rull, G., 2016. Ontology-based data integration in EPNet: Production and distribution of food during the Roman Empire. Engineering Applications of Artificial Intelligence 51, 212–229.

Carreras Monfort, C., Funari, P. P. A., 1998.

Britannia y el Mediterráneo: estudios sobre el abastecimiento de aceite bético y africano en Britannia. Vol. 5. Edicions Universitat Barcelona.

Carreras Monfort, C., Morais, R., 2012. The atlantic Roman trade during the principate: new evidence from the western façade. Oxford Journal of Archaeology 31 (4), 419–441.

Cravioto, E. G., 2002. Tumultos y Resistencia indígena en Mauretania Tingitana (siglo II). Gerión. Revista de Historia Antigua 20 (1), 451–485.

Crema, E. R., 2015. Time and Probabilistic Reasoning in Settlement Analysis. In: Barceló, J. A., Bogdanović, I. (Eds.), Mathematics and Archaeology. CRC Press Boca Raton, pp. 314–334.

Foley, B. P., Hansson, M. C., Kourkoumelis, D. P., Theodoulou, T. A., 2012. Aspects of ancient Greek trade re-evaluated with amphora DNA evidence. Journal of Archaeological Science 39 (2), 389–398.

Fulford, M., 1992. Territorial expansion and the Roman Empire. World Archaeology 23 (3), 294–305.

Funari, P. P. A., 1996. Dressel 20 inscriptions from Britain and the consumption of Spanish olive oil. BAR.

Greene, K., 1986. The archaeology of the Roman economy. Univ of California Press.

Knight, D., 1991. The movements of the auxilia from Augustus to Hadrian. Zeitschrift für Papyrologie und Epigraphik, 189–208.

Livarda, A., Orengo, H. A., 2015. Reconstructing the Roman London flavourscape: new insights into the exotic food plant trade using network and spatial analyses. Journal of Archaeological Science 55, 244–252.

Loughton, M. E., 2003. The distribution of republican amphorae in france. Oxford Journal of Archaeology 22 (2), 177–207.

Marlière, E., 2001. Le tonneau en Gaule romaine. Gallia, 181–201.

Mattingly, D.J., 1988. Oil for export? A comparison of Libyan, Spanish and Tunisian olive oil production in the Roman Empire. Journal of Roman Archaeology 1, 33–56.

McCune, B., Grace, J., 2002. MRPP (Multi-response Permutation Procedures) and related techniques. Analysis of Ecological Communities: MjM Software Design, Gleneden Beach, Oregon, USA, 188–197.

Mielke, P. W., Berry, K. J., Johnson, E. S., 1976. Multi-response permutation procedures for a priori classifications. Communications in Statistics - Theory and Methods 5 (14), 1409–1424.

Morillo, A., Fernández Ochoa, C., Salido Domínguez, J., 2016. Hispania and the Atlantic Route in Roman Times: new Approaches to Ports and Trade. Oxford Journal of Archaeology 35 (3), 267–284.

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M. H. H., Oksanen, M. J., Suggests, M., 2007. The vegan package. Community ecology package 10, 631–637.

Orengo, H. A., Livarda, A., 2016. The seeds of commerce: A network analysis-based approach to the Romano-British transport system. Journal of Archaeological Science 66, 21–35.

Paterson, J., 1982. 'Salvation from the Sea': Amphorae and Trade in the Roman West. Journal of Roman Studies 72, 146–157.

Pecci, A., Clarke, J., Thomas, M., Muslin, J., van der Graaff, I., Toniolo, L., Miriello, D., Crisci, G., Buonincontri, M., Di Pasquale, G., 2017. Use and reuse of amphorae. Wine residues in Dressel 2–4 amphorae from Oplontis Villa B (Torre Annunziata, Italy). Journal of Archaeological Science: Reports 12, 515–521.

Peña, J. T., 2007. Roman pottery in the archaeological record. Cambridge University Press.

Pons Pujol, L., 2009. La economía de la Mauretania Tingitana (s. I-III d. C.): aceite, vino y salazones. Vol. 34. Edicions Universitat Barcelona.

Remesal Rodríguez, J., 1986. La annona militaris y la exportación de aceite bético a Germania. Editorial Complutense.

Remesal Rodríguez, J., 2008. Provincial interdependence in the roman empire: an explanatory model of roman economy. Bar International Series 1782, 155.

Remesal Rodríguez, J., 2010. De baetica a germania, consideraciones sobre la ruta y el comercio atlántico en el imperio romano. In: Viajeros, peregrinos y aventureros en el mundo antiguo. Universitat de Barcelona, pp. 147–160.

Remesal Rodríguez, J., 1998. Baetican olive oil and the Roman economy. Journal of Roman Archaeology - Suppl. series 29, 183–200.

Remesal Rodríguez, J., 2011. La Bética en el concierto del Imperio romano. Real Academia de la Historia.

Remesal Rodríguez, J., Aguilera, A., García Sánchez, M., Martín-Arroyo, D., Pérez González, J., Revilla Calvo, V., 2015. Centro para el Estudio de la Interdependencia Provincial en la Antigüedad Clásica (CEIPAC). Pyrenae.

Rodgers, P., 1987. Multi-Response permutation procedures. Computer and Quantitative Methods in Archaeology 7957, 45–54.

Rubio-Campillo, X., Cela, J. M., Hernàndez Cardona, F. X., 2012. Simulating archaeologists? Using agent-based modelling to improve battlefield excavations. Journal of Archaeological Science 39 (2), 347–356.

Rubio-Campillo, X., Coto-Sarmiento, M., Pérez, J., Remesal Rodríguez, J., 2017. Bayesian analysis and free market trade within the Roman Empire. Antiquity, in press.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular biology and evolution 4 (4), 406–425.

Scheidel, W., 2015. Orbis: The Stanford Geospatial Network Model of the Roman World. SSRN Electronic Journal.

Scheidel, W., Morris, I., Saller, R. P. (Eds.), 2007. The Cambridge economic history of the Greco-Roman world. Cambridge University Press, Cambridge, UK; New York.

Tchernia, A., 1986. Le vin de l'Italie romaine. Essai d'histoire économique d'après les amphores. Vol. 261. Persée-Portail des revues scientifiques en SHS.

Teichner, F., Pons Pujol, L., 2008. Roman amphora trade across the straits of Gibraltar: an ancient Antieconomic practice? Oxford Journal of Archaeology 27 (3), 303–314.

Temin, P., 2001. A market economy in the early Roman Empire. Journal of Roman studies 91, 169–181.

Thomas, D. H., 1978. The awful truth about statistics in archaeology. American Antiquity, 231–244.

Tobler, W. R., 1970. A computer movie simulating urban growth in the Detroit region. Economic geography 46 (sup1), 234–240.

Wickham, H., 2009. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Williams, D., Peacock, D., 1983. The importation of olive oil into Iron Age and Roman Britain. In: Blázquez Martínez, J. M., Remesal Rodríguez, J. (Eds.), Producción y comercio del aceite en la antigüedad. pp. 263–280.

Wilson, A., 2009. Approaches to quantifying Roman Trade. In: Quantifying the Roman Economy. Oxford Studies in the Roman Economy, pp. 213–249.

Woolf, G., 1992. Imperialism, empire and the integration of the Roman economy. World Archaeology 23 (3), 283–293.

Yu, G., Smith, D. K., Zhu, H., Guan, Y., Lam, T. T.-Y., 2017. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution 8 (1), 28–36.

Yubero-Gómez, M., Rubio-Campillo, X., López-Cachero, J., 2016. The study of spatiotemporal patterns integrating temporal uncertainty in late prehistoric settlements in northeastern Spain. Archaeological and Anthropological Sciences 8 (3), 477–490.