



Col·lecció
INSTRUMENTA  60

CUANTIFICAR
LAS ECONOMÍAS ANTIGUAS.
PROBLEMAS Y MÉTODOS

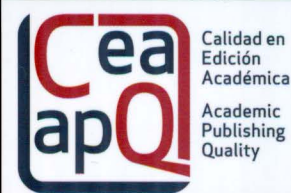
QUANTIFYING ANCIENT ECONOMIES.
PROBLEMS AND METHODOLOGIES

José Remesal Rodríguez, Víctor Revilla Calvo,
Juan Manuel Bermúdez Lorenzo (eds.)



UNIVERSITAT DE
BARCELONA

Edicions



© PUBLICACIONES I EDICIONES DE LA UNIVERSITAT DE BARCELONA, 2018
Adolf Florensa, 2/n; 08028 Barcelona; Tel. 934 035 442; Fax 934 035 446.
comercial.edicions@ub.edu

1ª edición: Barcelona, 2018

Director de la colección: JOSÉ REMESAL.

Secretario de la colección: ANTONIO AGUILERA.

Diseño de la cubierta: CESCA SIMÓN.

CEIPAC

<http://ceipac.ub.edu>

Sello de Calidad en Edición Académica. Promovido por la Unión de Editoriales Universitarias Españolas (UNE) y avalado por la Agencia Nacional de Evaluación de la Calidad y Acreditación (ANECA) y la Fundación Española para la Ciencia y la Tecnología (FECYT).

Unión Europea: *The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013/ ERC grant agreement n° 340828).*

Gobierno de España: DGICYT: PB89-244; PB96-218; APC 1998-119; APC 1999-0033; APC 1999-034; BHA 2000-0731; PGC 2000-2409-E; BHA 2001-5046E; BHA2002-11006E; HUM2004-01662/HIST; HUM200421129E; HUM2005-23853E; HUM2006-27988E; HP2005-0016; HUM2007-30842-E/HIST; HAR2008-00210; HAR2011-24593; HAR2015-66771-P (MINECO/FEDER, UE); HAR2017-85635-P (MINECO/FEDER, UE).

MAEX: AECI29/04/P/E; AECL.A/2589/05; AECL.A/4772/06; AECL.A/01437/07; AECL.A/017285/08.

Generalitat de Catalunya: *Grup de Recerca de Qualitat*: SGR 95/200; SGR 99/00426; 2001 SGR 00010; 2005 SGR 01010; 2009 SGR 480; 2014 SGR 218; 2017 SGR 512; ACES 98-22/3; ACES 99/00006; 2002ACES 00092; 2006-EXCAV0006; 2006ACD 00069.



The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013). ERC grant agreement n° ERC-2013-ADG340828.

European Research Council



Montaje: Ariane Capdevila Sanz.

Portada: Estela funeraria. *Viminacium* (hoy Kostolac, Serbia).

Impresión: Gráficas Rey, S.L.

Depósito legal:

ISBN:

Impreso en España / *Printed in Spain.*

Queda rigurosamente prohibida la reproducción total o parcial de esta obra. Ninguna parte de esta publicación, incluido el diseño de la cubierta, puede ser reproducida, almacenada, transmitida o utilizada mediante ningún tipo de medio o sistema, sin la autorización previa por escrito del editor.

Índice general

Prólogo. (José Remesal Rodríguez)	9
Quantificare la storia: questioni esegetiche intorno alla cultura materiale. (Giorgio Rizzo)	13
Contar para comparar: algunas reflexiones sobre el tema, con ejemplos desde la Lusitania. (Carlos Fabião)	33
Comptages, comparaisons, évaluations: l'exemple des amphores en Gaule. (Fanette Laubenheimer)	57
Algunos apuntes sobre la cuantificación: el protocolo de Sevilla (PRCS/14). (César Carreras Monfort)	71
A matter of formalities. A basic measure of social complexity in the past. (Dries Daems)	89
Using agent-based modelling to infer economic processes in the past. (Iza Romanowska)	107
Cuantificación arqueológica para la construcción de hipótesis históricas: métodos, protocolos (PCRS/14) y correctores estadísticos (MR y AC/CM). (Jaime Molina Vidal)	119
Testeo sobre sistemas de cuantificación en ceramología antigua. La aplicación de los protocolos de Sevilla sobre un contexto votivo ibérico de Iliberri (Albaicín, Granada). (Andrés María Adroher Auroux, Manuel Abelleira Durán)	137
Todo el pescado vendido. Una lectura cuantitativa de la producción púnica y romana de ánforas, sal y salazones en la Bahía de Cádiz. (Enrique García Vargas, Antonio M. Sáez Romero)	161
El valor estadístico de la epigrafía sobre ánforas Dressel 20. (José Remesal Rodríguez)	215
Provincias, sellos e hipótesis nulas: la identificación de rutas de comercio a través de medidas de distancia cultural. (Xavier Rubio-Campillo, Juan Manuel Bermúdez Lorenzo, Jean Marc Montanier, Juan Moros Díaz, Jordi Pérez González, Guillem Rull Fort, José Remesal Rodríguez)	237

Reconstruir lo roto. Un método para vincular entre sí las inscripciones del Testaccio. (Jordi Pérez González, Mario Morvan, Luce Prignano, Ignacio Morer, Albert Díaz-Guilera, Juan Manuel Bermúdez Lorenzo, José Remesal Rodríguez)	251
Que compte-t-on et comment, sur les ateliers d'amphores Dr. 20 de la vallée du Guadalquivir. (Severine Corbeel, Quentin Desbonnets, Ivan González Tobar, Stéphane Mauné)	281
Sobre la capacidad de carga de los hornos romanos de ánforas: balance metodológico y reflexiones a partir de un horno de Dressel 20 del alfar de Las Delicias (Écija, Sevilla). (Charlotte Carrato, Ophélie Tiago, Stéphane Mauné, Enrique García Vargas, Oriane Bourgeon, Philippe Lanos)	295
Contexto arqueológico e historiográfico de las ánforas Crétoise 2 en <i>Pompei</i>: una crítica cuantitativa. (Daniel J. Martín-Arroyo Sánchez)	319
Ancient cliometrics and archaeological proxy-data. Between the devil and the deep blue sea. (Koenraad Verboven)	345
Riflessioni conclusive (Clementina Panella)	373

PROVINCIAS, SELLOS E HIPÓTESIS NULAS: LA IDENTIFICACIÓN DE RUTAS DE COMERCIO A TRAVÉS DE MEDIDAS DE DISTANCIA CULTURAL

X. RUBIO-CAMPILLO (School of History, Classics and Archaeology - University of Edinburg)

J. M. BERMÚDEZ LORENZO (CEIPAC – Universitat de Barcelona)

J.M. MONTANIER (Barcelona Supercomputing Centre)

J. MOROS DIAZ (CEIPAC – Universitat de Barcelona)

J. PÉREZ GONZÁLEZ (CEIPAC – Universitat de Barcelona)

G. RULL FORT (Siris Academic)

J. REMESAL RODRÍGUEZ (CEIPAC – Universitat de Barcelona)

1. INTRODUCCIÓN...O CÓMO SABEMOS SI ESTAMOS EQUIVOCADOS?

El estudio de la economía del Imperio Romano ha seguido una larga senda desde los primeros debates acaecidos a inicios del siglo XX. El interés en estos orígenes no era distinto del que podemos identificar en obras actuales: qué estructuras económicas generó el imperio? Eran parecidas a las que se han identificado en otras épocas? Qué diferencias y similitudes se pueden encontrar con sistemas actuales? La exploración de estas preguntas requería de dos componentes principales: a) indicadores o proxies que den información sobre las actividades económica y b) métodos capaces de revelar las dinámicas a través de las *proxies* mencionadas.

1.1 Evidencia y economía romana

En el primer caso el uso de fuentes escritas ha dado una información valiosa si bien escasa, dado el bajo número de registros y la magnitud del caso de estudio, tanto en el tiempo como en el espacio. Por ese motivo el campo ha ido incrementando el uso de fuentes arqueológicas, que si bien son más complejas de analizar proporcionan una visión más rica y plural de las dinámicas históricas, especialmente si se pueden combinar con otros tipos de fuentes. Sin embargo, la mencionada complejidad del registro arqueológico ha hecho de esta tarea un desafío que tan solo ahora estamos empezando a afrontar con las garantías adecuadas; gracias a la nueva ciencia de datos o *data science* es posible generar los *datasets* necesarios para capturar las dinámicas económicas del imperio romano desde una perspectiva multiescalar, más allá de casos de estudio específicos (Wilson, 2009; Bevan, 2014). Esta nueva corriente está mejorando la accesibilidad de la evidencia arqueológica a 2 niveles: a) tecnología y b) cultura de trabajo. Por una parte, la creación de nuevas tecnologías como las bases de datos semánticas o las infraestructuras geoespaciales han permitido por primera vez enlazar múltiples *datasets* con distintos orígenes, formatos y aproximaciones (Bevan, 2015; Scheidel, 2015; Calvanese et al., 2016). Por otra parte, la disciplina arqueológica ha ido aceptando el enorme potencial brindado por las políticas de acceso abierto tanto en resultados como en datos y análisis; especialmente destacables en este contexto son los beneficios en términos de replicación de estudios y agregación de datos a gran escala (Lake, 2012; Orengo, 2015; Marwick, 2016). Estas dos dinámicas han confluído en la agregación de decenas de *datasets* procedentes de todos los rincones del imperio y de todas las épocas susceptibles de ser utilizados para el tipo de estudio cuantitativo capaz de responder a las preguntas de investigación mencionadas anteriormente.

Sin embargo, el acceso a datos no es sinónimo de generación de conocimiento: la evidencia necesita ser analizada teniendo en cuenta el tipo de información que contiene, así como sus limitaciones. Es este uno de los grandes retos presentes en el estudio de la economía romana: ¿qué métodos necesitamos para responder a nuestras preguntas de investigación y avanzar en el debate actual?

1.2 Los desafíos del análisis cuantitativo

Uno de los puntos cruciales que hace de la arqueología una disciplina única es el tipo de evidencia que trata. La mayoría de ciencias estudian directamente su objeto de estudio, pero en nuestro caso no nos es posible observar la dinámica que queremos analizar. En lugar de observar el comportamiento social en el que estamos interesados no nos queda más remedio que aproximarnos a él a partir de *proxies* arqueológicas que *a priori* son la consecuencia de dicho comportamiento. Además de ello estas *proxies* fueron generadas por individuos concretos, mientras que en el registro arqueológico nosotros tan sólo podemos observar una agregación de materiales generados por una miríada de dichas actividades. Esta agregación temporal hace difícil discernir dinámicas en escalas menores a décadas o siglos, y casi imposible identificar a individuos excepto en casos muy específicos (Bayley, 2007).

Los problemas no acaban aquí, ya que la evidencia recogida es en todo caso una muestra reducida y no aleatoria de la población inicial; una serie de procesos tafonómicos eliminan la mayor parte de observaciones posibles a través de múltiples sesgos o *biases* que afectan la muestra desde su generación hasta su redescubrimiento a través de trabajos arqueológicos (Drennan, 2010 : 79-96).

Finalmente, a esos desafíos se añaden los inherentes al área de estudio: la exploración de la economía romana requiere de una aproximación amplia capaz de integrar múltiples escalas de análisis,

desde el yacimiento al imperio, y desde decisiones políticas concretas a dinámicas económicas que pueden abarcar siglos de duración (Remesal Rodríguez, 1986; Scheidel y Friesen, 2009).

Todas estas complejidades requieren métodos de análisis avanzados, capaces de integrar la incertidumbre de la evidencia arqueológica en el mismo análisis y generar resultados robustos. Este avance crucial no se ha realizado aún, ya que la mayoría de análisis cuantitativos sobre economía romana han adoptado una aproximación basada en el Análisis Exploratorio de Datos o *Exploratory Data Analysis* (EDA). La ventaja del EDA es su fácil comprensión, ya que sus métodos se basan en estadísticas descriptivas básicas (p.e. mediana, variancia, etc.) y visualizaciones de datos. Sin embargo, las complejidades mencionadas anteriormente tienen un impacto directo en la validez de aproximaciones basadas en EDA; como consecuencia podemos identificar patrones que en realidad no existen y solo están presentes a causa de los sesgos, mientras que otras veces las dinámicas de estas *proxies* no serán obvias al ojo humano a causa del ruido generado por estos mismos sesgos (Tukey, 1980).

Si queremos responder a nuestras preguntas de investigación necesitamos hacer un paso más allá y movernos del análisis exploratorio a la evaluación de hipótesis. Esta rama de la estadística se basa en evaluar hasta qué punto una posible explicación puede sostenerse en la evidencia recogida. Dicho de otro modo; si a pesar de todas las complejidades de los datos arqueológicos somos capaces de generar hipótesis plausibles entonces nuestras respuestas serán más robustas que cualquier resultado aportado por el ojo humano en base a un gráfico exploratorio.

1.3 La evaluación de hipótesis nulas

La aproximación más común en investigación científica para poner a prueba hipótesis es conocida como Prueba de Significación de Hipótesis Nula o *Null-Hypothesis Significance Testing* (NHST). El objetivo del NHST es evaluar si un patrón identificado en los datos a través de EDA pudo haber sido generado por pura casualidad. Esa es precisamente nuestra hipótesis nula: la evidencia recogida ha sido generada de manera aleatoria. La mayoría de tests estadísticos se basan en intentar rechazar esta hipótesis nula. Para ello se calcula la probabilidad que un valor igual o más extremo que el observado se hubiera podido generar de manera aleatoria; si esta probabilidad (llamada p-valor) es más baja que un umbral de confianza (típicamente fijado en 0.05 o 5%) entonces podemos rechazar dicha hipótesis nula.

El método es enormemente popular y efectivo siempre y cuando tengamos en cuenta sus limitaciones y asunciones. En primer lugar, NHST funciona correctamente cuando las hipótesis que evaluamos son mutuamente exclusivas. Esto es correcto en una prueba clínica donde se quiere evaluar si un nuevo fármaco es efectivo contra una enfermedad, pero es más discutible en el estudio de sociedades humanas donde las posibles respuestas pueden ser más variadas. Además de ello, el umbral del 5% debe entenderse como lo que es: una convención que no necesariamente es correcta; la hipótesis nula podría ser la correcta aunque el p-valor sea muy bajo, mientras que un p-valor alto no quiere decir necesariamente que la hipótesis nula sea correcta. Finalmente, la presión actual por tener resultados positivos en la academia aumenta el riesgo de *p-hacking*: estudios donde el investigador realiza múltiples análisis en busca del p-valor correcto. La comunidad científica es consciente de todos estos límites y por eso mismo en la actualidad se ha abierto un interesante y crucial debate sobre la validez de la aproximación y sus posibles alternativas (Anderson et al., 2000; Gliner et al., 2002).

1.4 Identificando rutas de comercio

Este trabajo presenta un caso de estudio donde se ha aplicado una aproximación basada en hipótesis nulas para evaluar la existencia de rutas de comercio de aceite de oliva durante el Imperio Romano. Hemos elegido como *proxy* de ruta comercial los sellos encontrados en las ánforas que servían como contenedor de dicho líquido. Como se ha ilustrado en otros trabajos la frecuencia de los códigos hallados en estos sellos es un excelente identificador de la estructura comercial subyacente que organizó la distribución de estos contenedores de líquidos (ver Rubio-Campillo et al., 2017).

Durante el análisis usamos la epigrafía relacionada con el tipo de ánforas Dressel 20 para reconstruir las rutas por las que se redistribuía el aceite de oliva producido en la Bética durante los siglos I-III dC (Remesal Rodríguez, 1986; Funari, 1996; Remesal Rodríguez, 1998; Berni, 2008). La intención es contribuir al debate sobre qué rutas comerciales de largo recorrido eran más activas durante este período y, en concreto, caracterizar las rutas por las que el aceite de oliva producido en el valle del río Guadalquivir llegaba al resto del territorio occidental del imperio. Para ello pretendemos identificar si la división administrativa basada en provincias tenía algún impacto en la ausencia o presencia de códigos específicos o, dicho de otro modo, si los productos que llegaban a distintas provincias eran bienes fabricados por distintos agentes económicos. Evaluamos dos hipótesis de trabajo concretas:

1. los asentamientos de una misma provincia comparten sellos anfóricos similares, y
2. existen agrupaciones de provincias cuyos sellos son más similares que los encontrados en el resto del imperio.

La siguiente sección expondrá el dataset elegido para evaluar las hipótesis de trabajo así como los métodos estadísticos necesarios para el análisis mediante NHST. Los resultados son evaluados e interpretados en la tercera sección y el texto acaba con unas reflexiones finales sobre el método de trabajo y el futuro de los estudios de la economía romana.

2. MATERIALES Y MÉTODOS

Para evaluar nuestras dos hipótesis de trabajo necesitamos a) cuantificar la similaridad de los sellos encontrados en yacimientos y b) evaluar si, para estos yacimientos, existe una relación entre la provincia a la que pertenecen y los sellos que se encontraron.

2.1 Epigrafía anfórica y Dressel 20

Para este análisis hemos utilizado la totalidad de sellos identificados en ánforas Dressel 20 recogidos por la base de datos CEIPAC (Remesal Rodríguez et al., 2015). De estos 16.492 sellos se han eliminado 2 grupos para evitar sesgos:

- Sellos encontrados en el valle del Guadalquivir. Estas ánforas fueron fabricadas y rellenadas en este valle, por lo que se ha considerado adecuado eliminar la zona de producción y quedarse tan sólo con los yacimientos dónde se consumió el aceite.
- Sellos encontrados en Roma. La evidencia recogida en la capital ha sido eliminada porque su gran volumen causaría un sesgo muy elevado en el resultado final; además de ello en este caso

de estudio estamos interesados en las relaciones interprovinciales por lo que el abastecimiento a Roma ya se ha analizado de manera independiente en otro estudio (Rubio-Campillo et al., en prensa).

El dataset final de 8748 sellos de Dressel 20 comprende una totalidad de 1904 códigos distintos. Están distribuidos entre 689 yacimientos que cubren un gran porcentaje de Europa Occidental y Central como puede observarse en la Figura 1. Para cada sello se ha registrado su referencia geográfica así como la provincia en la que se ha encontrado y el código identificado.



Fig. 1. Distribución espacial de sellos Dressel 20 recogidos por la base de datos CEIPAC. Se puede observar cómo la mayoría de ánforas de esta tipología están localizadas en las provincias occidentales y en especial en las fronteras septentrionales (Germanias y Britania).

2.2 Distancia Jaccard

Como se ha mencionado el primer paso es cuantificar la similitud de los grupos de sellos encontrados en cada yacimiento. Para ello se ha elegido una popular medida de similitud estadística conocida como coeficiente Jaccard. Este índice de similitud entre 2 entidades puede definirse como la división entre a) el número de elementos comunes a las dos entidades y b) el número de elementos totales recogidos por las dos entidades.

Para cada yacimiento se ha registrado la ausencia o presencia de la totalidad de códigos del dataset; estos grupos se han usado para calcular una distancia entre cada pareja de yacimientos

definida como $1 - Jaccard$ y cuyo rango va desde 0 (los yacimientos tienen exactamente los mismos sellos) a 1 (los yacimientos tienen sellos completamente distintos). Las Tablas 1 y 2 muestran ejemplos de este cálculo. En el primer caso los dos yacimientos tienen 3 sellos en común de un total de 4 posibles y por ese motivo su coeficiente es $3 / 4 = 0.75$ y la distancia entre ellos es de $1 - 0.75 = 0.25$. En el segundo caso se observa el caso opuesto: los yacimientos tienen tan sólo 1 sello en común de un total de 3 presentes y por lo tanto su coeficiente es $1 / 3 = 0.33$. La distancia es $1 - 0.33 = 0.66$.

	PNN	QCM	VIRG	SRN
Emporiae	1	1	1	0
Massalia	1	1	1	1

Tabla 1 – Primer ejemplo de distancia Jaccard

	PNN	QCM	VIRG	SRN
Emporiae	1	1	1	0
Vindolanda	1	0	0	0

Tabla 2 – Segundo ejemplo de distancia Jaccard

Este proceso se ha realizado entre cada pareja posible de yacimientos existentes en el dataset. El resultado es una matriz de distancias *Jaccard* de 689 filas por 689 columnas en las que cada valor representa la disimilitud entre una determinada pareja de yacimientos dados por el índice de la fila y la columna. Esta matriz puede usarse para ver si la distancia media entre yacimientos pertenecientes a la misma provincia es menor que la distancia media entre yacimientos localizados en otras partes del imperio.

2.3 Procedimiento de Permutación Multi-Respuesta

Una vez se ha creado la matriz de distancias Jaccard con todas las combinaciones posibles es posible evaluar la primera hipótesis. Siguiendo la aproximación NHST deberemos rechazar la siguiente hipótesis nula: “la distancia Jaccard entre 2 yacimientos es independiente de su pertenencia a la misma provincia”. Para evaluar esto debemos calcular la probabilidad que, si cogemos 2 yacimientos aleatorios de la muestra, la distancia sea menor que si elegimos 2 yacimientos de la misma provincia. Esta probabilidad es nuestro p-valor: si es menor que el umbral de 5% entonces rechazaremos la hipótesis nula.

Cabe destacar aquí que no queremos comparar los sellos entre provincias, sino las similitudes entre los grupos de sellos encontrados en yacimientos ubicados en dichas provincias. Esta distinción es esencial para integrar en el análisis el componente multiescalar de los datos; si comparáramos directamente sellos y provincias no podríamos identificar casos donde, por ejemplo, los yacimientos de una misma provincia recibieron ánforas de dos rutas separadas distribuyendo dos grupos de productores completamente independientes.

Para ello aplicamos un método conocido como Procedimiento de Permutación Multi-Respuesta (Multi-Response Permutation Procedure, MRPP). Este algoritmo evalúa la distancia media entre elementos de una categoría (en nuestro caso yacimientos de una provincia) y la compara con la distancia media entre

toda la muestra para evaluar la hipótesis nula mencionada anteriormente (Mielke et al, 1976; Rodgers, 1987; McCune y Grace, 2002).

2.4 Agrupamiento jerárquico

La segunda hipótesis requiere evaluar si existen agrupaciones de provincias cuyos yacimientos contienen sellos similares. Para ello se calculan las distancias intraprovinciales (i.e. disimilitud media entre los yacimientos de una provincia) e interprovinciales (i.e. disimilitud media entre yacimientos de cada pareja de provincias). Esta nueva matriz de distancias entre provincias puede finalmente usarse para identificar grupos de provincias que tengan sellos más similares de lo que sería predecible observando el dataset completo. Mediante un proceso conocido como agrupamiento jerárquico o *hierarchical clustering* es posible no tan solo generar esos grupos, sino también ver qué grupos se parecen más entre ellos. Finalmente toda esta dinámica de similitud entre grupos puede visualizarse mediante un dendrograma: un árbol donde las ramas más cercanas pertenecen a entidades o provincias con materiales más parecidos; del mismo modo, el número de ramas entre 2 provincias servirá como indicador de la distancia existente entre los sellos de dichas provincias (van Sickle, 1997).

3. RESULTADOS Y DISCUSIÓN

El resultado de los métodos mencionados en el dataset de sellos ha generado los siguientes resultados.

3.1 MRPP para toda la muestra

La aplicación del algoritmo MRPP a la totalidad de la muestra genera el resultado observado en la Tabla 3.

variable	resultado
n. yacimientos	684
n. códigos	1904
p-valor	0.01
distancia media	0.992
distancia intraprovincial	0.988
efecto	0.004

Tabla 3: Resultados del MRPP para la totalidad de la muestra

En primer lugar, la hipótesis nula queda rechazada con una alta significación debido a que el p-valor es 0.01. Este resultado quiere decir que la probabilidad que la diferencia observada entre a) la disimilitud media entre yacimientos de una provincia y b) la disimilitud media entre todos los yacimientos sea debida a la casualidad es menor que un 1%.

Por otro lado, cabe destacar que el efecto de la provincia en los sellos encontrados es bajo: 0.004. Esto quiere decir que si bien los yacimientos de una misma provincia tienden a tener sellos más similares que el resto, siguen siendo muy distintos entre ellos.

Esto se debe a la distribución de sellos entre yacimientos como se puede observar en la Figura 2. El gráfico muestra que la inmensa mayoría de yacimientos tiene muy pocos sellos. Esto genera un nivel de incertidumbre muy alto ya que si en dos yacimientos tan sólo se ha recogido un sello entonces las probabilidades que ese sello tenga el mismo código son muy bajas.

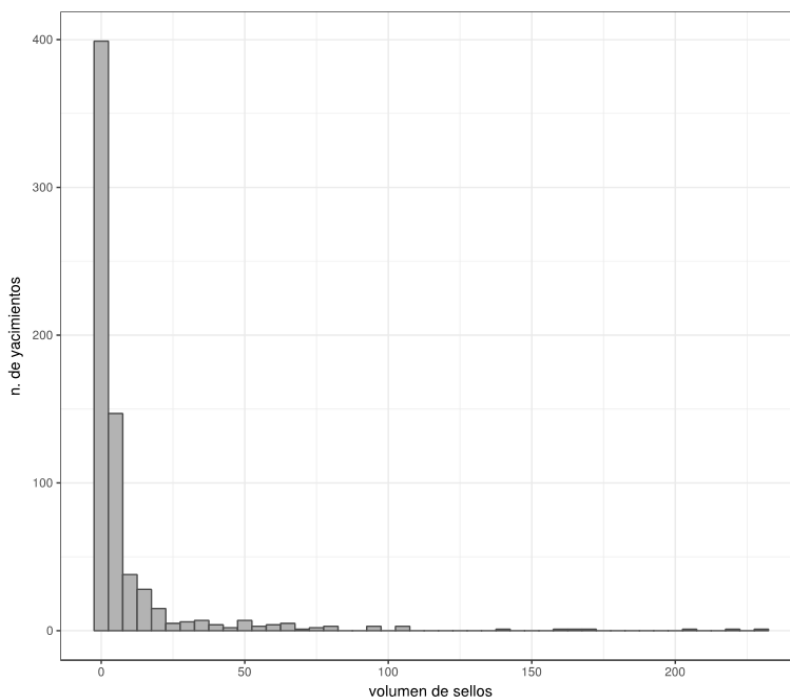


Fig. 2. Histograma de distribución de frecuencias de sellos por yacimiento. El eje Y muestra el número de yacimientos con el número de sellos definido por el eje X. Se revela en esta imagen la heterogeneidad de la distribución de epigrafía anfórica; unos pocos yacimientos contienen un número de sellos mayor que 50 mientras que en la inmensa mayoría (unos 400) tan solo se han catalogado uno o dos sellos.

Como vemos los sesgos de la evidencia arqueológica tienen un impacto directo en el análisis. Afortunadamente y pese a estas dificultades los resultados son suficientemente robustos para que podamos rechazar con confianza la hipótesis nula. Así, podemos afirmar que la demarcación provincial de un asentamiento tenía relación con los productores de aceite de oliva que lo abastecían.

3.2 MRPP selectivo

Para evaluar el impacto de la heterogeneidad de la muestra hemos realizado un nuevo análisis MRPP con subsets de los yacimientos con mayor muestra de sellos. En este caso hemos realizado un proceso iterativo en base a un umbral de sellos entre 1 y 100; para cada valor se ha realizado el MRPP con los yacimientos en los que se han encontrado como mínimo tantos códigos como el mencionado umbral. Los resultados se pueden ver en la Figura 3.

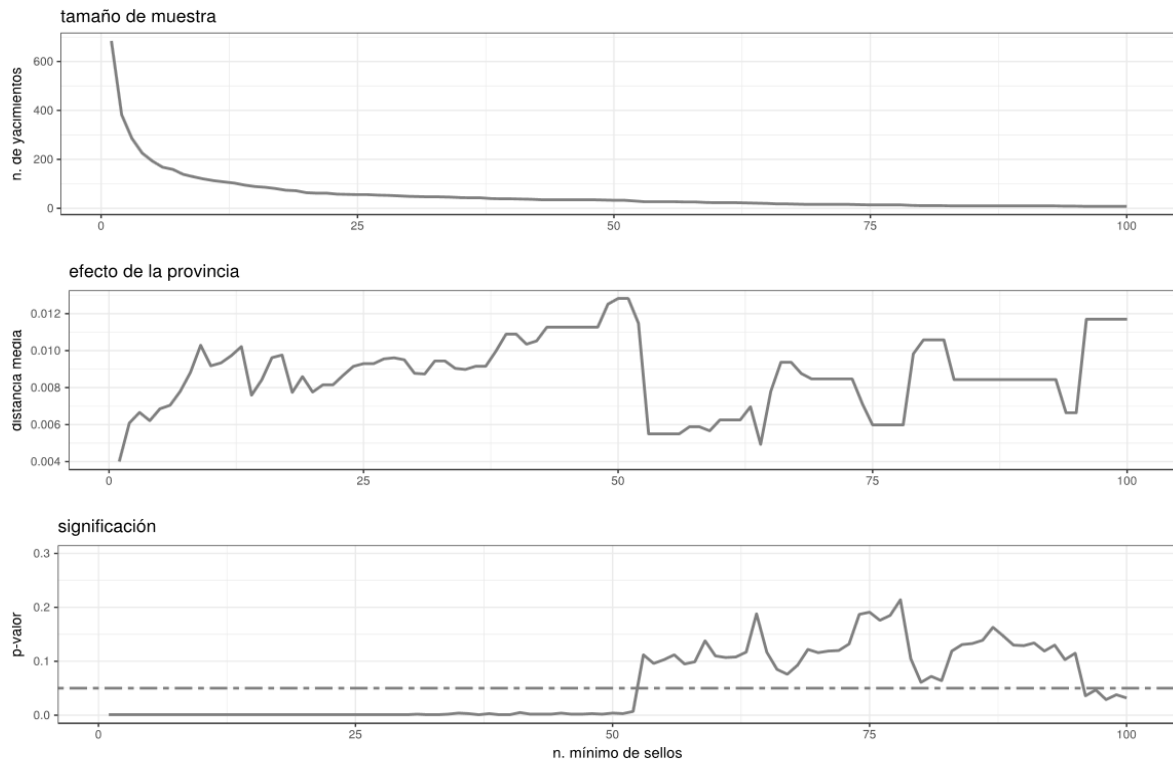


Fig. 3. Resultados del MRPP en base al número de sellos por yacimiento. Por cada experimento se ha aplicado MRPP al grupo de yacimientos con un número de sellos igual o superior al umbral (eje X). Las tres gráficas muestran en el eje Y: a) el tamaño de la muestra, b) el impacto de la provincia en el resultado y c) la significación del p-valor.

En primer lugar, el gráfico de arriba permite visualizar cómo decrece el número de yacimientos usado en el análisis a medida que aumentamos este límite de sellos y, por ejemplo, si seleccionamos los yacimientos con 25 o más códigos la muestra se reduce a unos 50 yacimientos.

El segundo gráfico muestra cómo el efecto de la provincia en los sellos encontrados se multiplica por 3 si descartamos los yacimientos con menos de 9 sellos. Este gráfico debe usarse en conjunción con el tercero que nos muestra el p-valor; si este p-valor está por encima de 0.05 entonces la muestra es demasiado variada para el reducido número de yacimientos y, por lo tanto, es preferible descartar el resultado.

Este análisis otorga mayor robustez a la hipótesis que la provincia tiene efectivamente un impacto en el comercio, y además sugiere que los patrones del dataset son más visibles al filtrar el ruido generado por yacimientos cuya muestra de sellos es excesivamente baja.

3.3 Relaciones entre provincias

Para evaluar la segunda hipótesis se ha realizado un análisis de distancias intra- e inter-provinciales según indicado en el apartado anterior. Los resultados siguen siendo estadísticamente significativos y generan las agrupaciones mostradas en la Figura 4.



Fig. 4. Clustering jerárquico de provincias en base a la similitud de su epigrafía anfórica. El análisis se ha realizado con todas las provincias en las que hay dos o más yacimientos con un mínimo de 9 sellos distintos.

En el gráfico las ramas muestran dos indicadores distintos: a) la distancia entre provincias a partir del número de bifurcaciones entre ellas y b) la distancia intraprovincial en el tramo final de la rama (e.g. irá para arriba si una provincia es más diversa que el grupo al que pertenece). El resultado genera ciertos patrones interesantes:

- La mayoría de provincias con ramas comunes son geográficamente adyacentes entre ellas (e.g. Aquitania y Narbonensis; las 2 Germanias). Este patrón subraya la robustez del método aplicado y sus resultados ya que se debe recordar que no se ha usado la distancia geográfica para generarlos. Es decir, se ha identificado una estructura espacial en la distribución de sellos pese a que no se ha usado ningún tipo de información espacial durante el análisis. Esta correlación entre distancia espacial y similitud de sellos tiene sentido en tanto que las provincias cercanas probablemente compartían rutas de comercio y, por ese motivo, los códigos encontrados en ellas deberían ser más similares.
- Hay una agrupación clara de provincias con fuerte presencia de guarniciones militares. Este patrón rompe la mencionada correlación entre disimilitud de sellos y distancia geográfica. En concreto las Germanias, Britania y curiosamente Mauretania Tingitana comparten sellos pese a la enorme distancia espacial entre la última provincia y el resto. Si el abastecimientos de las legiones romanas fuera organizado de manera centralizada entonces esperaríamos este resultado, ya que los sellos encontrados en estas provincias serían más similares que la media pese a no proceder de zonas adyacentes.

- Las provincias atlánticas tienen sellos similares. Las dos Germanias han sido agrupadas con provincias atlánticas como Britania, Bélgica y Lugdunensis. Este resultado es relevante porque indica que el aceite que llegaba a las Germanias procedía mayoritariamente de una ruta de comercio que reseguía la costa Atlántica. Este patrón contrasta con la poca similitud entre Germania Superior y Narbonensis, si tenemos en cuenta que ésta sería la otra ruta posible desde la Bética al limes germano a través de los ríos Ródano y Rin (Marlière, 2001).

4. CONCLUSIONES

Los resultados de este caso de estudio confirman que la estructura provincial tenía un peso relevante, si bien no decisivo, en la organización del comercio de aceite de oliva. Particularmente importante es el patrón de similitudes entre las provincias Atlánticas; creemos que esta contribución puede ayudar a interpretar los recientes hallazgos arqueológicos relacionados con las rutas de comercio de largo alcance en el imperio romano (Carreras y Morais, 2012; Morillo et al., 2016). Los patrones identificados aquí no implican que no se suministrara aceite de oliva a través del Ródano, ni que esta ruta no fuera más frecuente para otros productos. Sin embargo, la evidencia sugiere que, en el caso del aceite procedente de la Bética, el mayor volumen del producto llegaba a Germania a través del océano Atlántico.

Queremos también destacar que el análisis presentado es tan sólo un ejemplo del tipo de métodos que es posible aplicar al estudio de la economía romana. Como se ha mencionado anteriormente la aproximación NHST tiene sus problemas y límites, y en determinados casos podría no ser la mejor opción. Una alternativa a este método es la aplicación de inferencia bayesiana que permite evaluar la plausibilidad de múltiples hipótesis de trabajo de manera más robusta. Los resultados aportados por el creciente número de trabajos que usan estadística Bayesiana para estudiar el pasado están ya revelando el enorme potencial de esta innovadora metodología (Crema et al., 2014; Rubio-Campillo, 2016; Rubio-Campillo et al., 2017).

Este trabajo ha explorado el potencial que la evaluación de hipótesis tiene para el estudio de la economía romana. Los problemas de la aplicación de estadística en arqueología mencionados aquí no son nuevas (Thomas, 1978), pero es precisamente ahora que tenemos acceso a grandes bases de datos cuando necesitamos más que nunca aplicar métodos cuantitativos adecuados (Bevan, 2015). La disponibilidad y volumen de datos sobre las dinámicas de Roma ha aumentado exponencialmente en los últimos años, pero debemos ser conscientes de los desafíos que su análisis conlleva. Necesitamos aplicar nuevos métodos cuantitativos capaces de afrontar los sesgos de la evidencia arqueológica y aún así generar resultados positivos; tan sólo esta combinación de nuevos datos y nuevos métodos nos permitirá responder a preguntas planteadas ya hace décadas y avanzar en la comprensión de la complejidad económica y social del mundo clásico.

AGRADECIMIENTOS

Este trabajo ha recibido financiación del proyecto EPNET del European Research Council Advanced Grant (340828). El análisis ha sido realizado con las librerías *vegan* (Oksanen et al., 2007) y *ggplot2* (Wickham, 2009) sobre el paquete estadístico R. El código y dataset están disponibles bajo licencias abiertas y son accesibles desde <https://github.com/xrubio/mrppProvincias>

BIBLIOGRAFÍA

- ANDERSON, D.R., BURNHAM, K.P., THOMPSON, W.L., (2000), Null hypothesis testing: problems, prevalence, and an alternative, *The journal of wildlife management*, 912–923.
- BAILEY, G., (2007), Time perspectives, palimpsests and the archaeology of time. *Journal of Anthropological Archaeology* 26, 198–223.
- BERNI, P., (2008), *Epigrafía anfórica de la Bética: nuevas perspectivas de análisis*, Edicions Universitat Barcelona. ed, Instrumenta. Barcelona.
- BEVAN, A., (2014), Mediterranean Containerization. *Current Anthropology* 55, 387–418.
- BEVAN, A., (2015), The data deluge. *Antiquity* 89, 1473–1484.
- CALVANESE, D., LIUZZO, P., MOSCA, A., REMESAL, J., REZK, M., RULL, G., (2016), Ontology-based data integration in EPNNet: Production and distribution of food during the Roman Empire. *Engineering Applications of Artificial Intelligence* 51, 212–229.
- CARRERAS, C., MORAIS, R., (2012), The atlantic Roman trade during the principate: new evidence from the western façade: the Atlantic roman trade during the principate. *Oxford Journal of Archaeology* 31, 419–441.
- CREMA, E.R., EDINBOROUGH, K., KERIG, T., SHENNAN, S.J., (2014), An Approximate Bayesian Computation approach for inferring patterns of cultural evolutionary change. *Journal of Archaeological Science* 50, 160–170.
- DRENNAN, R.D., (2010), *Statistics for archaeologists: a commonsense approach*, 2nd ed, Interdisciplinary contributions to archaeology. Springer, New York.
- FUNARI, P.P.A., (1996). *Dressel 20 inscriptions from Britain and the consumption of Spanish olive oil*. BAR 250, Oxford.
- GLINER, J.A., LEECH, N.L., MORGAN, G.A., (2002), Problems with null hypothesis significance testing (NHST): what do the textbooks say? *The Journal of Experimental Education* 71, 83–92.
- LAKE, M., (2012). Open archaeology. *World Archaeology* 44, 471–478.
- MARLIÈRE, É., (2001). Le tonneau en Gaule romaine. *Gallia* 181–201.
- MARWICK, B., (2016), Computational Reproducibility in Archaeological Research: Basic Principles and a Case Study of Their Implementation. *Journal of Archaeological Method and Theory* 1–27.
- 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, 1409–1424.
- MORILLO, Á., 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, 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., (2015), “Open source GIS and Geospatial Software in Archaeology: Towards their Integration into Everyday Archaeological Practice”, en: WILSON, A.T., EDWARDS, B. (Eds.), *Open Source Archaeology Ethics and Practice*. Walter de Gruyter GmbH & Co KG, pp. 64–82.
- REMESAL RODRÍGUEZ, J., (1986), *La annona militaris y la exportación de aceite bético a Germania*. Editorial Complutense.
- REMESAL RODRÍGUEZ, J., (1998), Baetican olive oil and the Roman economy. *Journal of Roman Archaeology - Suppl. Series* 29, 183–200.
- 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., (2016), Model Selection in Historical Research Using Approximate Bayesian Computation. *PLOS ONE* 11, e0146491.
- RUBIO-CAMPILLO, X., COTO-SARMIENTO, M., PÉREZ-GONZALEZ, J., REMESAL RODRÍGUEZ, J., (2017), Bayesian analysis and free market trade within the Roman Empire. *Antiquity* 91(359), 1241-1252
- SCHEIDEL, W., FRIESEN, S.J., (2009), The Size of the Economy and the Distribution of Income in the Roman Empire. *Journal of Roman Studies* 99, 61.
- SCHEIDEL, W., (2015), Orbis: The Stanford Geospatial Network Model of the Roman World. SSRN *Electronic Journal*.
- VAN SICKLE, J., (1997), Using mean similarity dendrograms to evaluate classifications. *Journal of Agricultural, Biological, and Environmental Statistics* 370–388.
- THOMAS, D.H., (1978), The awful truth about statistics in archaeology. *American Antiquity* 231–244.
- TUKEY, J.W., (1980), We need both exploratory and confirmatory. *The American Statistician* 34, 23–25.
- WICKHAM, H., (2009), *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.
- WILSON, A., (2009), Approaches to quantifying Roman Trade, in: *Quantifying the Roman Economy*. *Oxford Studies in the Roman Economy*, pp. 213–249.

