

PERCEPCIÓN Y CUANTIFICACIÓN DE LA VARIABILIDAD ARQUEOLÓGICA A ESCALA *INTRA-SITE*. UNA REVISIÓN DE LOS FORMATOS DE DATOS Y SUS IMPLICACIONES ANALÍTICAS

ALFREDO MAXIMIANO CASTILLEJO

Instituto Internacional de Investigaciones Prehistóricas de Cantabria. Universidad de Cantabria.
Edificio Interfacultativo. Avda. los Castros s/n, 39005 Santander, España

maximianoam@unican.es

RESUMEN

Los avances geo-computacionales recientes han permitido una revolución empírica y metodológica en Arqueología, lo cual ha supuesto una serie de notables avances, desde los procesos de recogida y gestión de datos -pasando por nuevas series analíticas- hasta las nuevas formas de representación de la información (por ejemplo, visualización analítica). Como punto de partida se considera que el formato de dato elegido (indistintamente al medio de captura empleado) para la caracterización de la variabilidad espacial arqueológica es el elemento que dirigirá el proceso de análisis, en términos de percepción de la variación y del tipo de test estadísticos que se puedan emplear. El objeto de este artículo es la descripción de diferentes formatos de datos (coordinados, frecuencia, geoestadísticos y espacio-temporales), junto a sus principales implicaciones analíticas en función de la problemática espacial a tratar y los medios de representación, la implementación de test geo-estadísticos y, por último, las posibilidades interpretativas acerca de la variabilidad espacial observada.

Palabras claves: datos espaciales y temporales; análisis espacial *intra-site*; semivarianza; STSS; STKDE

DETECTION AND QUANTIFICATION SPATIAL VARIANCE AT *INTRA-SITE* SCALE. REVISITED DATA FORMATS AND ITS ANALYTICAL IMPLICATIONS

ABSTRACT

Actually geo-computational advances have allowed an empirical and methodological revolution in Archeology, which has brought a number of significant workflows since processes of data collection and management, through new analytical series-to new forms of information representation (e.g. analytical view).

Starting point it is considered that chosen data format (either by capturing sensors employee) for the characterization of archaeological spatial variability is the element that will lead the process of analysis in terms of perception of the change and what kind of statistical test can be employed. The purpose of this article is the description of different data formats (coordinate, frequency, geostatistical and spatial), together with its main analytical implications: spatial issues to be addressed and the means of representation, test implementation of geo-statistical; and finally, interpretive possibilities about the spatial variability observed.

Keywords: spatial and temporal data; intra-site spatial analysis; semivariogram; STSS; STKDE

1. Formatos de datos y el análisis espacial arqueológico a escala micro

El trabajo de campo arqueológico, y por ende las posibilidades del análisis espacial, han cambiado ostensiblemente desde la década de los noventa de la pasada centuria gracias al aporte de nuevos medios en geo-procesamiento y computación (McCoy y Ladefoged, 2009, 264). Hoy en día, los medios de captura y gestión de datos arqueológicos abarcan un amplio conjunto de técnicas y metodologías que pueden "ajustarse a la carta" de acuerdo a las circunstancias particulares de cada yacimiento. Pero a veces no queda muy clara la conexión entre problemática, medio de registro de datos, representación de la varianza espacial (y espacio-temporal), serie analítica y los entornos de representación de resultados. Por consiguiente, se debe intentar definir un nivel adecuado de resolución del problema, en el que el formato de datos ocupe una posición clave en todo el proceso de investigación, desde los medios empleados en la detección de datos hasta la interpretación y difusión de los resultados.

Sin duda, no existen recetas mágicas y en bastantes ocasiones la no adecuación del formato de datos empleado, junto a la aplicación de tecnologías que, por lo general, han sido diseñadas e implementadas para aplicaciones distintas a la arqueológica, puede (y suele) derivar a un entorno de "caja negra". La Arqueología se ha integrado en el uso -en ocasiones abuso- de la tecnología y de ciertos recursos analíticos inherentes al análisis espacial en Ecología, Biología, Geología y, como no, Geografía (por ejemplo, K-Ripley, I de Moran, Semivarianza, etc.). Evidentemente, desde la Arqueología no se pueden entender ni enunciar los fenómenos espaciales y temporales como sucede en otras esferas del conocimiento, ya que el objeto de estudio de los arqueólogos/as es una fracción material de un conjunto extenso de acciones/procesos realizados en el pasado.

En esta tesitura, el objetivo que se pretende alcanzar mediante el análisis espacial *intra-site* consiste en establecer una relación coherente y estadísticamente significativa entre el probable uso/funcionalidad del espacio y la población espacialmente distribuida en él. Para este fin no sólo es preciso contar con un entorno analítico y algunas reglas que permitan relacionar categorías nominales con funcionalidad del espacio, además es necesario que el problema espacial se pueda caracterizar de manera adecuada. Para ello es necesario partir desde un formato de datos espaciales ajustado al tipo de variabilidad espacial objeto de estudio. En este trabajo se pretende acentuar los aspectos analíticos vinculados a cada formato de datos, los cuales deben estar en consonancia tanto con los medios de captura potenciales como con los test estadísticos empleados. Todo ello permite

elaborar una estrategia coherente en el proceso de investigación de la variación espacial arqueológica a escala *intra-site*.

2. Tipo de datos, caracterización e implicaciones analíticas

A continuación se exponen cuatro formatos de datos que han sido implementados, algunos con bastante intensidad, en análisis espacial *intra-site*. Además de su caracterización, se introducen algunos test estadísticos y se citan casos de estudio existentes en la literatura arqueológica a modo de ejemplos de aplicación.

2.1. Datos bidimensionales: Coordenados y frecuencia

Estos dos formatos de datos son los más utilizados en Arqueología. Su implementación es relativamente reciente, ya que la primera excavación arqueológica en la que se incluyó la georeferenciación de objetos arqueológicos fue realizada en 1954 (Laplace y Meroc, 1954).

En Arqueología, los formatos de datos espaciales más empleados son *coordenados* y *frecuencia*. Básicamente, con el primer formato se discretizan entidades en torno a sus valores geográficos (relativos o absolutos) en términos de x , y , z . Con la *frecuencia*, la cantidad de restos se vincula a una unidad de superficie normalizada (reticulado).

Al emplear datos coordenados, los estadísticos que más comúnmente han sido implementados son: *K-Medias* y el *vecino más próximo (NNA)*. Los principios teóricos de ambos algoritmos se basan en el análisis de la distancia entre pares de puntos. Mientras que *K-Medias* es un algoritmo partitivo que busca agrupaciones (*clusters*) dentro de la distribución sin determinar el patrón de la misma, *NNA* se ocupa de estimar cómo se encuentra distribuida la población (donde la hipótesis nula es la aleatoriedad espacial).

Estos dos test han sido aplicados en el análisis espacial arqueológico de patrón de puntos (*point patter analyses*) (Kinting y Amerman, 1982; Simek y Larick, 1983; Carr, 1984; Simek, 1984; Amerman *et al.*, 1987; Kintigh, 1990; Kvamme, 1990; Blankholm, 1991) y en la actualidad son herramientas consolidadas que se siguen empleando en el análisis *intra-site* (Orton, 2005, 304).

Otras aplicaciones, con menor uso en Arqueología debido a la necesidad de una elevada potencia de cálculo, pero que están proliferando son *K de Ripley* (Ripley, 1976) y el *test de Mardia* (Mardia, 1980). Este test puede discernir entre los casos extremos de uniformidad y aleatoriedad espacial de la distribución. Su funcionamiento se basa en la detección de la aleatoriedad espacial en términos de la disminución de la *curtosis* (proporcionalmente) a medida que aumenta la *desviación estándar* de la distribución. El potencial de este test es importante, aunque actualmente no suele aplicarse en el análisis espacial arqueológico, aunque se considera su inminente implantación en Arqueología (Barceló y Maximiano, 2007).

El operador *K de Ripley* (Ripley, 1976; Ripley, 1981; Venables y Ripley, 1997) puede determinar el tipo, la intensidad y el rango del patrón espacial de la distribución (figura 1). Básicamente, la función *K* se compara con la línea $y = x$, que es una función esperada bajo la condición de aleatoriedad espacial: distribución *Poisson* (función que describe el número esperado de individuos al azar sobre una determinada área o región).

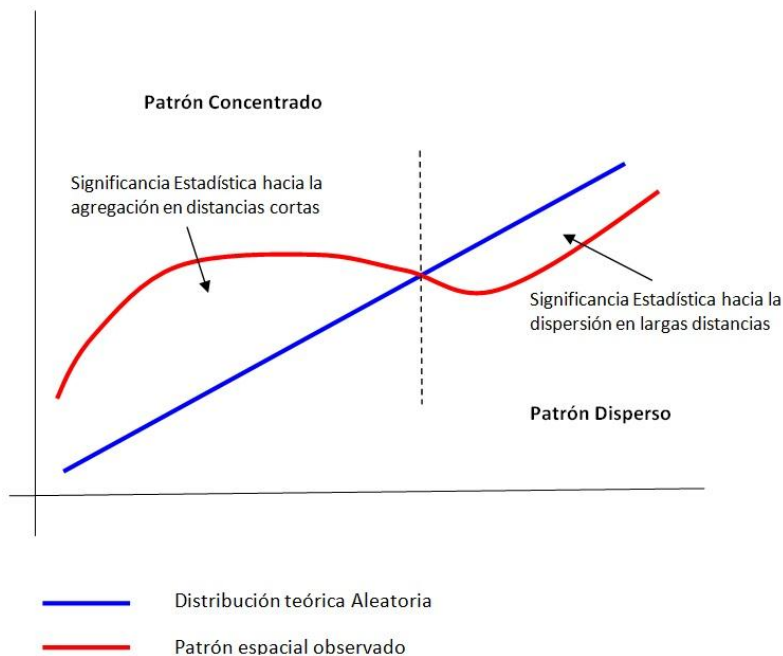


Figura 1: Representación esquemática de la función *K de Ripley* y sus consideraciones básicas.

Si la función empírica se sitúa encima de la función teórica *Poisson*, la distribución tiende a la agregación espacial. Si la función empírica se sitúa por debajo de la *Poisson*, entonces tenderá a la uniformidad espacial y si la distribución es igual a la función *Poisson*, la aleatoriedad será lo que defina a la distribución objeto de estudio.

Este test tiene muchas posibilidades en el análisis espacial arqueológico ya que la calidad del mismo ante las implicaciones de efecto barrera que inciden en determinadas distribuciones espaciales permite superar determinadas limitaciones inherentes al análisis de tipo clúster (*NNA*). Algunas aplicaciones interesantes se encuentran en las propuestas de Bevan y Conolly (2006); Conolly y Lake (2006); Schwarz y Mount (2006), entre otros, todas ellas realizadas a escala macro (*landscape*). En *intra-site* se pueden citar los trabajos de Orton (2004); Barceló y Maximiano (2007); Maximiano (2008), entre otros.

Otro aspecto significativo en la serie analítica de este formato de datos es la visualización de la distribución espacial. Probablemente la mejor representación de nubes de puntos se realice mediante la herramienta conocida como *Densidad Kernel* (KDE), la cual relaciona (a través de isolíneas o escala de color) los valores similares de la distribución en distintas localizaciones. Mediante este estimador se obtiene una visión rápida y completa acerca de cómo la población se encuentra distribuida (figura 2).

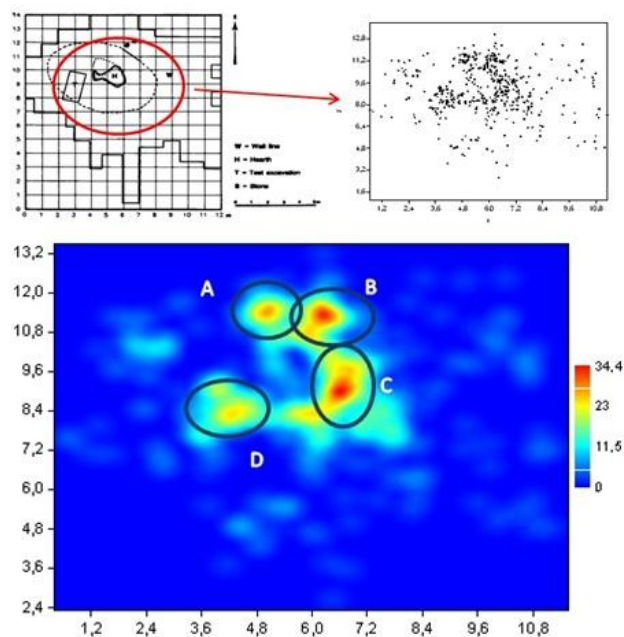


Figura 2: Parte superior composición de una distribución empírica (n = 470) parte inferior, imagen KDE. Datos del yacimiento *Barmose I* (Blankholm, 1991, en Maximiano *en prensa*).

La frecuencia espacial es otro de los formatos de datos más comúnmente empleados en el análisis espacial *intra-site*. Este formato se suele utilizar cuando la georeferenciación de la materialidad arqueológica, debido a diferentes factores, presenta un determinado grado de complejidad (por ejemplo, evidencias de pequeño tamaño que no se perciben a simple vista, escaso tiempo para la recolección de datos en el trabajo de campo, etc.). Básicamente, los datos de frecuencia responden al número de sucesos contenidos en unidades muestrales discretas, comúnmente conocidas como *grid* (figura 3).

Al igual que sucede con los datos coordenados, mediante el análisis de frecuencia se puede estimar la estructura espacial de la población en función a los valores empíricos que tienen cada una de las retículas en relación a los valores esperados bajo determinados supuestos teóricos. En

esencia, la estructura espacial se cuantifica mediante una tabla de contingencia en la que cada retícula -con su valor empírico- se compara con un valor numérico esperado en ese *grid*, de acuerdo con el número de evidencias contenidas en cada una de las retículas vecinas (Hawkins *et al.*, 2003). Con este sistema, y en función a las categorías en las que el arqueólogo/a haya clasificado la materialidad arqueológica, se realiza un muestreo pseudo-continuo de las n distribuciones existentes en toda la superficie del yacimiento, pudiendo así visualizar tanto la forma como la intensidad que adopta cada una de las distribuciones asociadas a cada una de las categorías.

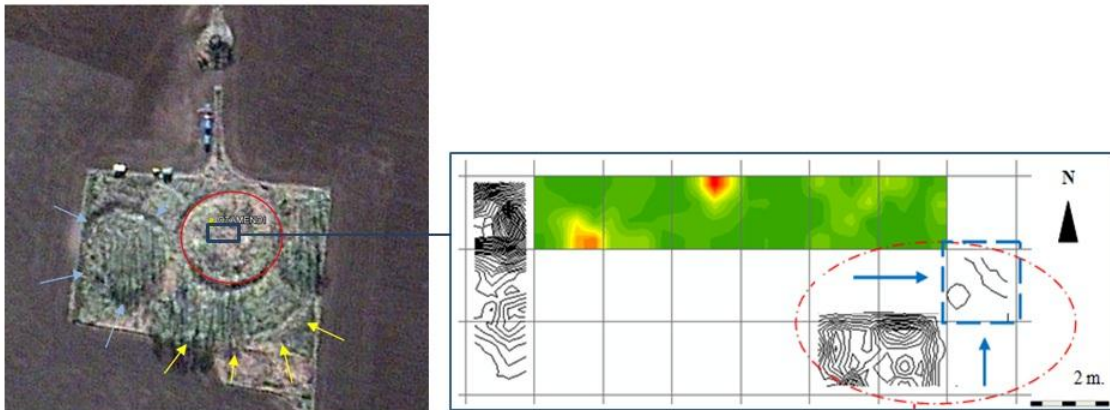


Figura 3. Ejemplo arqueológico de datos de frecuencia. A la izquierda foto satelital del yacimiento *Otamendi* (Gómez Romero y Maximiano, 2010) donde se ven los anillos perimetrales (3) y, enmarcada en azul, la zona excavada en sucesivas campañas. A la derecha imagen ampliada del área excavada, con la distribución de restos arqueológicos mediante frecuencia espacial (las retículas parten de un tamaño 2x2m segmentándose en subunidades excavadas de 0,5x0,5 m, resolución espacial de la distribución)

Básicamente, una distribución de frecuencias puede ser estructurada bajo tres supuestos:

1. Los valores de distribución son similares (bajos o altos) en las retículas próximas, y cambian paulatinamente a medida que existe un incremento/descenso de distancia entre pares de retículas, lo cual muestra una tendencia a la relación espacial *positiva* (ajustada al supuesto teórico de una *normal bivalente*).
2. La relación entre retículas será *negativa* cuando los valores de la distribución son diferentes en cuanto están más próximos entre sí (ajustada al supuesto teórico de una distribución *uniforme*)
3. *Ausencia de relación espacial*, donde los valores entre retículas próximas pueden ser, indistintamente, bajos y/o altos (ajustada al supuesto teórico de una distribución *Poisson*).

La principal operación en este tipo de análisis es la comparación de los pesos espaciales (cuántos sucesos por retícula) con la relación de covarianza entre retículas vecinas. Cuando la relación espacial es más *positiva* que la esperada por un proceso al azar estará indicando la agrupación de valores similares sobre una determinada porción del espacio, mientras que si se detecta una relación *espacial* negativa indicará que los valores existentes en las retículas vecinas son diferentes de lo esperado al azar. Para cuantificar el tipo de relación espacial se describen dos estadísticos adimensionales - *I de Moran*, *C de Geary*- y la función *Correlograma*.

La *I de Moran* (Moran, 1950) se comporta como un coeficiente de correlación de Pearson, oscilando en el rango de +1 a -1. Cuando $I = 0$ no hay relación espacial, es decir, la distribución es *aleatoria*. Cuando $I > 0$ la distribución muestra una correlación positiva, lo que significa una tendencia a la *agregación* espacial. Por último, si $I < 0$ la distribución espacial tiende a una correlación negativa (*uniformidad*) (Premo, 2004).

El test adimensional *C de Geary* (Geary, 1954) varía de 0 a 3 (Legendre y Fortin, 1989), de modo que si c tiende a 0 ($c < 1$) hay una preferencia a la *agregación* espacial. Cuando c tiende a 2 ($c > 1$) la distribución tiende a la *uniformidad*, mientras que el caso en donde $c = 1$ indica que la distribución tiende a la *aleatoriedad* espacial. En Arqueología existe una tendencia en la escasa implementación del test estadístico c frente al uso masivo de I .

Otro test estadístico para la frecuencia espacial es una función que estima la estructura de la distribución (similar a la K de Ripley): función *Correlograma*. Se representa mediante una gráfica que muestra la función empírica de la distribución comparada con las funciones teóricas de aleatoriedad y uniformidad para esa misma distribución (Sokal y Oden, 1978b; Upton y Fingleton, 1985). Esta función no se ha usado apenas en casos arqueológicos (Hodder y Orton, 1976; Maximiano, 2008), aunque su potencial es relevante.

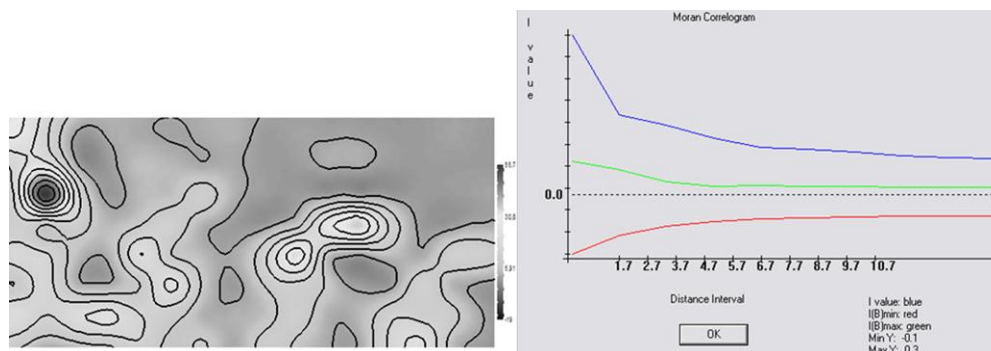


Figura 4: Izquierda, frecuencia de distribución espacial del yacimiento *Tunel VII* (desechos de *malacofauna* y *avifauna*). Derecha, la función *correlograma* de dicha distribución. La función empírica es la línea azul. El modelo teórico de agregación espacial se encuentra por encima de la línea verde (que es el supuesto contrastado). La condición de *aleatoriedad espacial* se sitúa entre la línea verde y roja. Si la distribución se situase por debajo de la línea roja, tendería a la uniformidad (Maximiano, 2008, 269).

Probablemente, el entorno más adecuado para la representación de los datos de frecuencia espacial sea el histograma 3D, desde el cual se pueden apreciar los valores de frecuencia para cada una de las retículas (figura 5). Por supuesto, esta herramienta es más robusta en su salida gráfica que un Kernel, pero, tanto su funcionalidad como el grado de representación real de la variación espacial, la convierte en un elemento analítico adecuado.

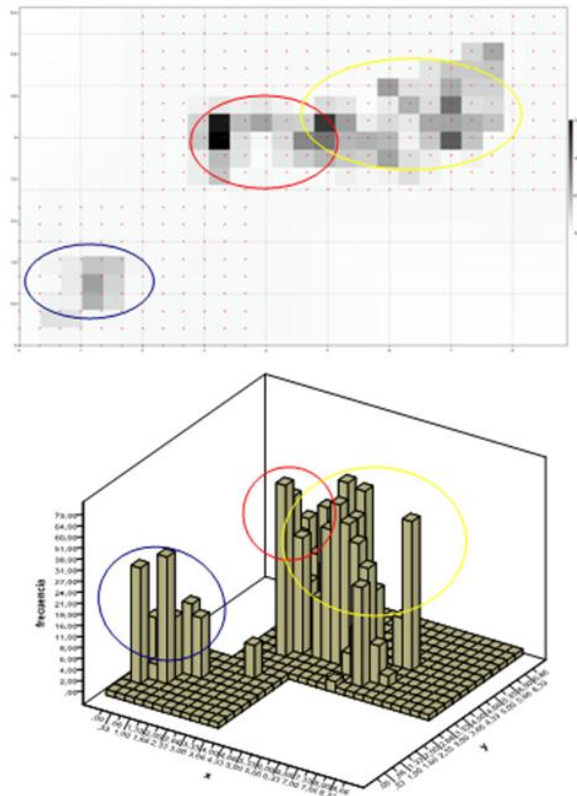


Figura 5: Un ejemplo teórico acerca de la visibilidad desde el histograma 3D. Arriba, la densidad de distribución espacial. Abajo, el histograma 3D. Los datos de frecuencia teórica de distribución (n = 250) elaborado *ex profeso* para este artículo.

2.2. Datos de dimensión 2.5: datos geoestadísticos

Los datos de dimensión 2.5 definen otro tipo de eventos espaciales, los cuales se componen de datos coordinados que miden una variable cuantitativa, es decir, son datos continuos acerca de la variación de una determinada propiedad según sus distintas localizaciones (*variable regional*). Los datos 2.5 son la base de las técnicas geoestadísticas, aunque este tipo de datos y técnicas no suelen aplicarse comúnmente en Arqueología (Markofsky y Bevan, 2012, 428).

Mediante el análisis geoestadístico se cuantifica la dependencia espacial existente en una distribución en términos del grado de *autocorrelación*. Si los datos que caracterizan a una *variable regional* no son independientes entre sí, entonces la variable se encuentra *autocorrelacionada*. Concretamente, si los valores en lugares separados por una cierta distancia son más (*autocorrelación* positiva) o menos (*autocorrelación* negativa) similares de lo que cabría esperar según una distribución aleatoria de la misma (Maestre *et al.*, 2008 ,9).

Este *concepto* es clave en el análisis espacial, como se recoge en el trabajo de Pierre Legendre (Legendre, 1993, 1661) donde se cuestiona si *la autocorrelación es un problema o una nueva ventana de oportunidad para el conocimiento de la importancia de la estructura espacial en los procesos ecológicos*. Trasladar este argumento a los procesos espacio-temporales a los que los arqueólogos/as se enfrentan en el análisis *intra-site* abriría un interesante debate, en términos epistemológicos y ontológicos, acerca del uso/gestión del espacio por parte de nuestros antepasados, puesto que incluir el grado de *autocorrelación* de una/varias distribución/es se convierte en una oportunidad de enfocar mediante escenarios opuestos a los planteamientos de la estadística paramétrica, donde la independencia de los datos es un axioma.

Según esto, y en un hipotético caso arqueológico *intra-site* donde los diferentes componentes que lo constituyen se cumplieran los preceptos de ausencia de estructura espacial, ocurriría lo siguiente: la gestión del residuo sería al azar en todo el *suelo de ocupación*, no existiría relación entre el interior de una unidad doméstica y la presencia de una estructura de combustión, las concentraciones de componentes químicos (alcalinidad, fósforo, ácidos grasos, etc.) estarían en localizaciones disímiles a los lugares donde las acciones causales fueron realizadas, etc. Frente a ese panorama, que indirectamente es potenciado por un mal uso de los recursos estadísticos clásicos, la determinación de la *autocorrelación* proporciona tanto el enfoque como las herramientas cuantitativas para lograr un orden lógico-formal a la hora de determinar la interacción entre diferentes componentes de un sistema espacial y temporal.

El semivariograma $h(v)$ es la principal herramienta geoestadística, mediante la cual se calcula la medida de *autocorrelación* de la variable (figura 6). Básicamente, $h(v)$ responde a la siguiente pregunta: ¿Cuál es la distancia máxima que relaciona dos localizaciones pertenecientes a una misma distribución? La solución no significa que más allá de ese módulo de distancia no exista relación, pero lo que sí es cierto es que más allá de ese valor no existen garantías acerca de la *correlación espacial* entre pares de localizaciones. Si la distribución es tendente a la aleatoriedad, la *autocorrelación* será cero, en otros términos, ninguna localización de la distribución se puede relacionar con otra.

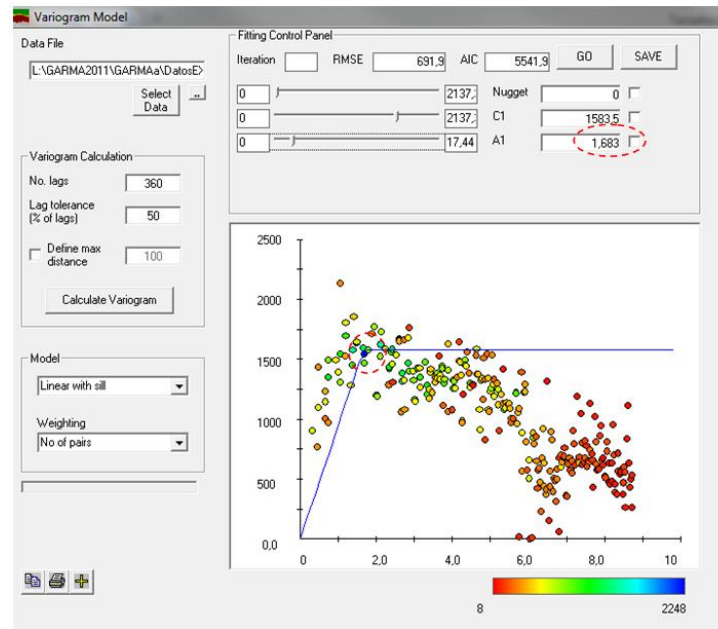


Figura 6: Semivariograma con una distribución teórica (n =1000) Bi-variente, donde la autocorrelación es de aprox 1.7 unidades de longitud (en este caso son metros), elaborado *ex profeso* para este artículo.

El uso de este tipo de datos y la potencialidad del análisis geoestadístico aún no se encuentra muy desarrollado en problemas espaciales arqueológicos a escala *intra-site*. Probablemente, el principal escollo se encuentra en que la mayor parte de las distribuciones de restos arqueológicos ni se definen ni son tratadas, como la varianza, en el espacio de una determinada propiedad (en el sentido estricto del término), puesto que se suele georeferenciar elementos discretos (por ejemplo, la distribución de fragmentos de cerámica romana tipo A, la distribución de restos de *megafauna* en el interior de una cueva, etc.) y no se intenta definir el tipo de continuidad espacial que pueda indicar la distribución de *artefactos arqueológicos*.

Probablemente ese sea el principal escollo entre análisis arqueológico *intra-site* y la aplicación de técnicas geoestadística, ya que la mayor parte de las distribuciones de restos no se ajustan a la formulación de cómo una determinada propiedad varía en el espacio porque, por lo general, el problema arqueológico se plantea en clave de distribuciones de *artefactos* carentes de una atribución espacial, es decir: ¿qué implica y qué puede significar el espacio de los fragmentos de cerámica romana tipo A?. Bajo esa "etiqueta" no se puede hacer un verdadero análisis espacial en clave de *variable regional*¹.

Los datos geoestadísticos son multiscales, aunque en Arqueología se ha empleado más a nivel de paisaje que a escala *intra-site* (algunos ejemplos en macroescala se encuentran en los trabajos de Zubrow y Harbaugh, 1978; Warren, 1990; Kvamme, 1992; Bevan y Conolly, 2009,

entre otros). A escala *intra-site*, el uso de la geoestadística ha sido escaso aunque la problemática tratada ha sido variada. Por ejemplo, en las propuestas de Loveluck (2004) y en Lloyd y Atkinson (2004) se realiza una aplicación sobre distribución de residuos químicos (fosfatos); en el trabajo de Dawson *et al.* (2007) se definen diferentes gradientes de luminosidad en múltiples posiciones para el interior de una cabaña paleolítica; en Maximiano (2008) y en Gómez Romero y Maximiano (en prensa) se cuantifica la *autocorrelación* en diferentes tipos de distribuciones que afectan a la organización espacial de actividades sociales en diferentes casos arqueológicos.

Con el nivel actual de desarrollo epistemológico acerca del uso de datos geoestadísticos en Arqueología, se pueden considerar dos líneas de actuación interesantes:

- i. El análisis *intra-site* de *variable regional* acerca de residuos químicos (distribuciones químicas que están directamente asociadas a un determinado conjunto de actividades antrópicas como pueden ser: procesamiento de alimentos, del curtido de pieles, sitios de enterramiento, etc.).
- ii. Modelización de dinámicas de difusión en determinados procesos sociales a escala macro (por ejemplo, Hamilton y Buchana, 2007; Bocquet-Appel *et al.*, 2009; Francois *et al.*, 2010; Markofsky y Bevan, 2012, entre otros).

3. Nuevas oportunidades en Arqueología: Datos espacio-temporales.

En la década de los setenta del siglo pasado la modelización geográfica del espacio-tiempo, entendida como el estudio de geometrías que cambian en el tiempo, fue formulada en términos matemáticos y computacionales (Hagerstrand, 1970; Rucker, 1977). Pero esta formulación teórica no podía traducirse en una aplicación empírica, ya que la tecnología computacional al uso era el entorno CAD, el cual no estaba diseñado con los requerimientos necesarios de conectividad topológica para representar procesos espacio-temporales. En este sentido, la incorporación de la componente temporal en geo-computación ha sido un avance relativamente reciente, desde mediados de los años noventa del siglo pasado tras el desarrollo de los SIG (Yuan, 1996).

Los primeros trabajos sobre esta cuestión fueron editados a finales de la década de los ochenta (Armstrong, 1988; Langran y Chrisman, 1988), siendo a mediados de los años noventa el momento en el que se consolidó el análisis, la modelización y visualización espacio-temporal de fenómenos (algunas obras de referencia son Langran, 1993; Peuquet 1994; Yuan 1995, entre otros).

En este marco, la praxis arqueológica supone un buen entorno para poner en práctica las aplicaciones espacio-temporales, ya que numerosos eventos arqueológicos se pueden representar en términos de sucesión, extensión, localización y relaciones espaciales (Gutiérrez, 2007). Bajo estos parámetros y sus posibles relaciones, se podrá:

1. Conocer la trayectoria histórica de determinados fenómenos.
2. Recuperar la información espacio-temporal histórica sobre determinados objetos/eventos.

3. Realizar una predicción de la localización mediante el pronóstico de la posición futura de un objeto en base a su posición actual y a su patrón de movimiento.

Para realizar un análisis exhaustivo sobre algún proceso espacio-temporal accesible únicamente mediante la Arqueología (por ejemplo, proceso de difusión de la sedentarización humana: la *Neolitización*), se requiere resolver cuestiones relacionadas con los sistemas de posicionamiento y referenciación (por ejemplo, datos coordinados en 4D = x, y, z, t), establecer medidas funcionales acerca de las distancias, sucesiones y relaciones entre eventos y construir modelizaciones que establezcan vínculos en términos de *extensión*, *intensidad*, *durabilidad* y *sucesión* entre los eventos espacio-temporalmente distribuidos.

Para generar un entorno congruente de trabajo con datos espacio-temporales en contexto arqueológico, será necesario:

- i) Un *medio de representación* adecuado de las componentes *espacio*, *tiempo* y *materia*. Para ello, la visualización más ajustada es la que ofrece el *hipercubo*. Fue a finales de los años sesenta del siglo pasado cuando T. Hägerstrand (Hägerstrand, 1970) introdujo un modelo considerado como el inicio de los estudios de la *Geografía del tiempo*. En su aspecto básico, consiste en un cubo cuya base representa el espacio (ejes x, y), mientras que la altura del cubo (eje z) representa el tiempo.
- ii) *Arquitecturas de datos y lenguajes* de consulta ajustados a los eventos sociales espacio-temporales. Consiste en bases de datos que capten aspectos espaciales y temporales del fenómeno investigado, representando el mundo real en términos de *objetos* (características) y *propiedades* (atributos). Esta arquitectura debe permitir que los eventos espacio-temporales se encuentren limitados por un período de validez en el tiempo, indistintamente a su extensión.
- iii) *Significancia estadística* de las estructuras en el espacio-tiempo. El algoritmo que se presenta en este trabajo se denomina: *Scan Statistics Space-time* (STSS), desde el cual se puede detectar la presencia de clústeres dentro de un *hipercubo* (Kulldorff *et al.*, 1998). Esto es posible a través de una expresión de dominios espacio-temporales en forma cilíndrica, donde el radio del cilindro representa al espacio y la longitud del cilindro es la componente temporal. Su notación matemática es la siguiente:

$$Z = \{(x, y, t) \mid (x - x_c)^2 + (y - y_c)^2 \leq r^2, t_s \leq t \leq t_e\},$$

Donde (x_c, y_c) es el centro geográfico del dominio, r es el radio del dominio y $[t_s, t_e]$ es el dominio del periodo.

Maximiano Castillejo A. (2013) "Datos óptimos para la caracterización espacial y temporal de la variabilidad arqueológica a escala intra-site", *GeoFocus (Artículos)*, n°13-1, p. 131-153. ISSN: 1578-5157

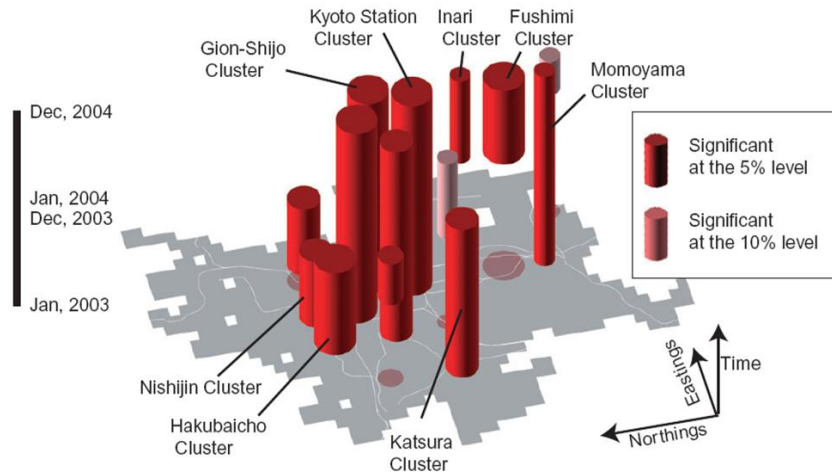


Figura 7: Imagen de un evento espacio-temporal (evolución del crimen) en donde se detectan agrupaciones significantes en el tejido espacio-tiempo. Los resultados han sido calculados mediante *Space-time Scan statistics*. En Nakaya y Yano (2010, 233). Aún no existe caso de estudio ni una aplicación arqueológica desarrollada acerca de este método.

- iv) *Modelos* de datos para representar procesos espacio-temporales. La modelización espacio-temporal se lleva a cabo con *iso-superficie* (el equivalente en el 3D de las isolíneas en el 2D), mediante la cual se calcula y dibuja datos volumétricos sobre una estructura 3D que corresponde a series de puntos con un solo valor escalar de x , y , t ; y el de la variable cuantificada (figura 8). Destaca la propuesta de Brunson *et al.* (2007), donde proponen el algoritmo de densidad *Kernel* espacio-tiempo (STKDE) como una solución de modelización. Su notación matemática:

$$\hat{f}(x, y, t) = \frac{1}{nh_s^2h_t} \sum_i K_s\left(\frac{x-x_i}{h_s}, \frac{y-y_i}{h_s}\right) K_t\left(\frac{t-t_i}{h_t}\right),$$

Donde $f(x, y, t)$ es la estimación de la densidad en la posición (x, y, t) , n es el número de eventos, y h_s y h_t son, respectivamente, los anchos de banda espacial y temporal (las funciones K_s y K_t se definen usando el *Kernel de Epanechnikov*. Epanechnikov, 1969). Se hace hincapié en que el algoritmo STKDE no es sólo una representación gráfica de la distribución espacio-temporal, también proporciona una medida aproximada en términos de la concentración/dispersión e intensidad del fenómeno en el espacio-tiempo (por ejemplo, ¿dónde, cuándo y cómo se distribuiría el crimen en una ciudad con una nueva estrategia social y policial sobre la criminalidad?).

Maximiano Castillejo A. (2013) "Datos óptimos para la caracterización espacial y temporal de la variabilidad arqueológica a escala intra-site", *GeoFocus (Artículos)*, n°13-1, p. 131-153. ISSN: 1578-5157

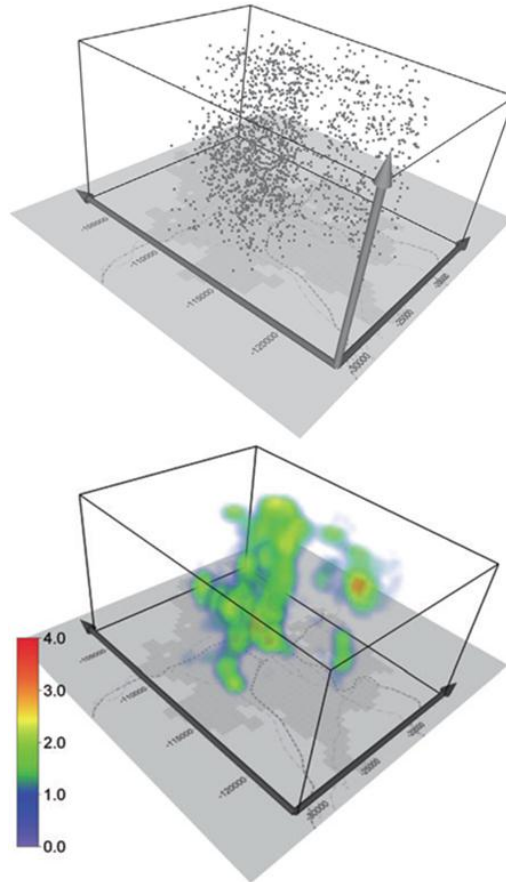


Figura 8: STKDE representación espacio-temporal de eventos relacionados con crímenes. En la parte superior se ve la nube de puntos (cada evento en sus coordenadas espacio-temporales), debajo el densidad Kernel 3D. Imagen tomada de Nakaya y Yano (2010, 232). Aún no existe caso de estudio ni una aplicación arqueológica desarrollada acerca de esta metodología.

Trasladar este tipo de perspectiva analítico-interpretativa acerca de determinados problemas arqueológicos es una vía útil, ya que trabajar desde este tipo de datos aportaría un cambio sustantivo en el modo de entender y resolver determinados problemas arqueológicos.

Otra posibilidad analítica, que solo se enuncia aquí, es el *3D vector mapa*. Esta herramienta permite cuantificar tanto la direccionalidad como la magnitud en la intensidad del proceso, facilitando la visualización de difusión/contracción.

Los datos espacio-temporales ya han sido implementados en casos arqueológicos, aunque no con los estadísticos descritos anteriormente. A nivel macro (*landscape*), cabría mencionar los trabajos de O. Huisman (Huisman *et al.*, 2009), en los que se desarrollan un prototipo de análisis

geo-visual basado en las propiedades del *Hipercubo*; y el trabajo de I. Feliciano (Feliciano, 2008), que trata sobre la visualización multivariante de eventos arqueológicos.

Otras propuestas que están en esta línea son los trabajos de Bocquet-Appel y Demars (2000); Hamilton y Buchanan (2007); Hamilton y Buchanan (2009); Bocquet-Appel *et al.* (2009); Crema *et al.* (2010), en los que se aplican herramientas geo-estadísticas para la caracterización de procesos de difusión espacio-tiempo de determinados procesos sociales.

A escala *intra-site*, la aplicación de estos datos se ha enfocado más hacia la caracterización y visualización de los -en ocasiones extensos- componentes que existen en una excavación arqueológica, que hacia cuestiones analíticas e interpretativas (por ejemplo en Katsianis *et al.*, 2008).

Esto se debe a la complejidad inherente a la hora de establecer las conexiones adecuadas entre una *semántica* coherente en gestión de datos y el modo de percibir y compilar las evidencias arqueológicas. Por tanto, es necesario reconsiderar determinadas técnicas arqueológicas aplicadas con éxito en el trabajo de campo bajo otros principios metodológicos, pero que no funcionan adecuadamente con estos nuevos entornos de análisis e interpretabilidad del registro arqueológico.

Sin duda, estas divergencias se empiezan a superar, siendo prueba de ello la propuesta de I. Feliciano (Feliciano, 2009) titulada: *Data visualization in Archaeology (intra-site applications)* donde se establece una interoperabilidad entre el registro arqueológico y las nuevas formas de percibir, cuantificar y visualizar resultados.

4. Conclusiones

Una de las problemáticas a la que se enfrentan los arqueólogos/as en la actualidad se centra en lograr una gestión coherente y eficiente dentro de la abundante oferta de recursos computacionales y analíticos que hoy en día brinda la tecnología espacial. Alcanzar un balance entre el problema y los entornos de solución pasa, necesariamente, por una elección proporcional y coherente en el formato o la combinación de distintos formatos de datos, que sea/n capaz/ces de caracterizar la fenomenología objeto de estudio. Por ello, incluso antes de abordar qué tipo de entornos de solución se puedan alcanzar mediante la aplicación de unos métodos u otros, deberá quedar claro el formato de datos que mejor se ajuste a una serie de requisitos en torno a la detección, captura, análisis, y lo más importante, a la interpretación de la variación espacial y temporal sobre las distribuciones de evidencias arqueológicas.

En la siguiente tabla 1, se expone una descripción básica entre diferentes componentes que intervienen en el proceso de elección de los datos, mostrando las posibilidades que aporta cada opción. De este modo, se puede contar con una información orientativa sobre los formatos de datos, las técnicas y los medios para afrontar el problema/hipótesis espacial con ciertas garantías de éxito.

Tabla 1: Descripción de formato de datos y su relación más coherente con otros componentes del proceso de investigación espacial.

Dato	Visualización	Análisis	Epistemología	Entorno SIG
COORDENADO	Kernel 2D, 3D	K-means; NNA; K de Ripley; Test de Mardia	Patrón espacial	Vector
FRECUENCIA	3D Histograma	I de Moran; c de Geary Correlograma	Patrón espacial	Raster
GEOSTADISTICO	Kriging	Semivariograma	Autocorrelación Predicción espacial	Raster/Vector
ESPACIO-TEMPORAL	STKDE	STSS	Percepción multidimensional de fenómenos	Vector/Raster Voxel

Mediante esta tabla se establece un orden acerca de las posibles relaciones considerando que la materialidad arqueológica objeto de estudio será la componente que determine, *in extremis*, el formato de datos más apropiado para su caracterización incidiendo con ello en aspectos de la captura, el análisis y representación de la variación espacial.

Aun así, es interesante tener presente una serie de consideraciones finales sintetizadas en los siguientes apartados:

1. Plantear los problemas a resolver y las hipótesis de trabajo de manera realista en función a la capacidad del registro arqueológico y de la calidad de los datos e información disponible y la potencialmente accesible.
2. Ser congruentes con las posibilidades analíticas de los datos, por ejemplo, no se puede estudiar la *correlación* en una variable que no sea regionalizada.
3. Entender los principios del análisis realizado y cómo opera el software seleccionado. Es necesario evitar dinámicas del tipo "caja negra" por el hecho de trabajar con aplicaciones más sofisticadas, que sin duda, si no sabemos *cómo* funciona, no nos aportará mejores resultados.

4. No hay recetas mágicas ni sistemas inmutables, en ocasiones trabajar con los datos más ajustados no garantiza mejores resultados. Con esto, se pretende reconocer la necesidad de dinamismo y espíritu crítico por parte de los/as que proyectan el proceso de investigación espacial.

Agradecimientos

Aprovecho para mostrar mi agradecimiento público a los revisores de este trabajo, gracias a sus comentarios y recomendaciones se ha podido mejorar y hacer más accesibles los contenidos aquí expuestos. Por otra parte, los errores y omisiones que puedan existir en este texto son únicamente responsabilidad del autor.

Referencias

- Armstrong, M. (1988): "Temporally in spatial dataset", en *Proceedings of GIS/LIS Conference*, San Antonio, Texas, pp. 880-889.
- Ammerman, H.; Kintigh, K. y Simek, J. 1987. "Recent development in the application of the K-means approach to spatial analysis", en Sieveking, G., and Newcomer, N. (Eds): *The Human Uses of Flint and Chert*. Cambridge: Cambridge Univ. Press.
- Barceló, J. A. y Maximiano, A. (2007): "Some notes regarding distributional analysis of spatial frequencies", en Posluschny, A.; Lambers, K. y Herzog, I. (Eds): *CAA 2007 Layers of Perception*. Proceedings of the 35th Annual Conference on Computer Applications and Quantitative Methods in Archaeology. Berlin, April 2007.
- Bevan, A. y Conolly, J. (2006): "Multi-scalar approaches to settlement pattern analysis", en Lock, G. and Molyneaux B. (Eds). *Confronting Scale in Archaeology: Issues of Theory and Practice*. New York. Springer, pp. 217-234.
- Bevan, A. y Conolly, J. (2009): "Modelling spatial heterogeneity and nonstationarity in artifact-rich landscapes", en *Journal of Archaeological Science* 36-4, pp. 956-964.
- Blankholm, H.P. (1991): *Intrasite spatial analysis in theory and practice*. Denmark: Aarhus University Press.
- Bocquet-Appel, J. P. y Demars, P. (2000): "Population Kinetics in the Upper Palaeolithic in Western Europe", en *Journal of Archaeological Science* 27: 551-570.
- Bocquet-Appel, J.P.; Naji, S.; Vander Linden, M. y Kozłowski, J. K. (2009): "Detection of diffusion and contact zones of early farming in Europe from the space-time distribution of 14C dates", en *Journal of Archaeological Science* 36, pp. 807-820.
- Brunsdon, C.; Corcoran J. y Higgs, G. (2007): "Visualising space and time in crime patterns: A comparison of methods", en *Computers, Environment and Urban Systems* 31, pp. 52-75.
- Carr, C. (1984): "The nature of organization of intrasite archaeological records and spatial analysis approaches to their investigation", en *AAMT* 7, pp. 103-222.

Conolly, J. y Lake, M. (2006): *Geographical Information Systems in Archaeology*. Cambridge: Cambridge University Press.

Crema, E. R.; Bevan, A. y Lake, M. (2010): "A probabilistic framework for assessing spatio-temporal point patterns in the archaeological record", en *Journal of Archaeological Science*, 37, pp. 1118-1130.

Dawson, P.; Levy, R.; Gardner, D. y Walls M. (2007): "Simulating the behaviour of light inside Arctic dwellings: implications for assessing the role of vision in task performance", en *World Archaeology* Vol. 39(1), pp. 17-35.

Epanechnikov, V. A. (1969): "Nonparametric estimation of a multivariate probability density", en *Theory of Probability and Its Applications* 14, pp. 153-58.

Feliciano Santiago, I. T. (2008): *Multivariate Geovisualization of archaeological events using Space-Time Cube*. PhD Dissertation. International Institute for Geo-Information Science and Earth Observation (ITC) Enschede, The Netherlands.

Feliciano Santiago, I. (2009): "Data visualization in Archaeology (intra-site applications)", disponible en: <http://datavisualizationarchaeology.blogspot.com/>.

Francois, O.; Currat, M.; Ray, N.; Han, E.; Excoffier, L. y Novembre, J. (2010): "Principal component analysis under population genetic models of range expansion and admixture", en *Mol. Biol. Evol.* 27(6), pp. 1257-1268.

Geary, R. C. (1954): "The Contiguity Ratio and Statistical Mapping", en *The Incorporated Statistician* 5 (3), pp. 115-145.

Gómez Romero, F. y Maximiano, A. (2010): "Avances en la aplicación del análisis geo-estadístico en arqueología histórica: el ejemplo del Fortín Otamendi (1858- 1869)", en *Revista de Arqueología Histórica Argentina e Latino-americana* Vol 4, pp. 85-99.

Gutiérrez Retamal, G. (2007): *Propuesta de un método de acceso Espacio-Temporales*. Tesis Doctoral. Universidad de Chile Facultad de Ciencias Físicas y Matemáticas. Santiago de Chile. (Disponible en: <http://users.dcc.uchile.cl/~mmarin/revista-sccc/sccc-web/Vol5/bd2.pdf>).

Hagerstrand, T. (1970): "What about people in regional science?", en *Papers of the Regional Science Association*. 24, pp. 7-21.

Hamilton, M.J. y Buchanan, B. (2007): "Spatial gradients in Clovis-age radiocarbon dates across North America suggest rapid colonization from the north", en *Proceedings of the National Academy of Sciences, USA* 104(40), pp. 15629-15634.

Hamilton, M.J. y Buchanan, B. (2009): "The accumulation of stochastic copying errors causes drift in culturally transmitted technologies: Quantifying Clovis evolutionary dynamics", en *Journal of Anthropological Archeology* 28, pp. 55-69.

Hawkins, A.L.; Stewart, S.T. y Banning, E.B. (2003): "Interobserver bias in enumerated data from archaeological survey", en *Journal of Archaeological Science* 30, pp. 1503-1512.

Hodder, I. y Orton, C. (1976): *Spatial Analysis in Archaeology*. Oxford: Univ. Press.

Huisman, O.; Feliciano Santiago, I.; Kraak, M.J. y Retsios, B. (2009): "Development of a geovisual analytics environment for investigating archaeological events based upon the Space-time Cube", en *Cartography and geographic information science* (36)3, pp. 225-236.

Katsianis, M.; Tspidis, S.; Kotsakis, K. y Kousoulakou, A. (2008): "A 3D digital workflow for archaeological intra-site research using GIS", en *Journal of Archaeological Science*. Vol, 35-3, pp. 655-667.

Kintigh, K. W. (1990): "Intrasite spatial analysis: A commentary on major methods", en A. Voorrips (Ed) *Mathematics and information science in archaeology: A flexible framework*. Studies in Modern Archaeology. Bonn Verlag. Vol. 3, pp. 165-200.

Kintigh, K. W. y Ammerman, H. J. (1982): "Heuristic approaches to spatial analysis in archaeology", en *American Antiquity*. Vol. 47, n° 1, pp. 31-63.

Kulldorff, M.; Rand, K.; Gherman, G.; Williams, G. y DeFrancesco, D. (1998): *SaTScan v2.1: Software for the Spatial and Space-time Scan Statistics*. Bethesda, MD, National Cancer Institute.

Kvamme, K.L. (1990): "Spatial autocorrelation and the Classic Maya collapse revisited: refined techniques and new conclusions", en *Journal of Archaeological Science* 17, pp. 197-207.

Kvamme, K. L. (1992): "A predictive site location model on the high plains: an example an independent test", en *Plains Anthropologist*, 37 (138), pp. 19-40.

Langran, G. y Chrisman, N R (1988): "A framework for temporal geographic information", en *Cartographica* 25, pp. 1-14.

Langran, G. (1993) *Time in Geographic Information Systems*. Taylor & Francis. Bristol

Laplace-Jauretche, G. y Méroc, L. (1954): "Application des coordonnées cartésiennes à la fouille d'un gisement", en *Bulletin de la Société Préhistorique Française* 51 (7), pp. 58-66.

Legendre, P. (1993): "Spatial autocorrelation: Trouble or new paradigm?", en *Ecology* 74(6), pp. 1659-1673.

Legendre, P. y Fortin, A.M.J. (1989): "Spatial pattern and ecological analysis", en *Vegetation* 80, pp. 107-138.

Lloyd, C. D. y Atkinson, P. M (2004): "Archaeology and geostatistics", en *Journal of Archaeological Science*, 31, pp. 151-165.

Loveluck, C.P. (2004): "Terres noires and early medieval rural settlement sequences: conceptual problems, descriptive limitations and deposit diversity", en: Verslype, L. and Brulet, R. (eds.): *Terres noires: dark earth: Actes de la Table-ronde Internationale Tenue à Louvain-la-Neuve*, les 09 et 10 novembre 2001, Louvain-la-Neuve, Université Catholique de Louvain, pp. 86-96.

Maestre, F. T.; Escudero, A. y Bonet, A. (eds) (2008). *Introducción al Análisis Espacial de Datos en Ecología y Ciencias Ambientales*. Editorial Dykinson

Mardia, K. V. (1980): "Tests of univariate and multivariate normality", en P.R. Krishnaiah

(Ed.), *Handbook of Statistics*, vol 1, North Holland, 1980, pp. 279-320.

Maximiano Castillejo A. (2013) "Datos óptimos para la caracterización espacial y temporal de la variabilidad arqueológica a escala intra-site", *GeoFocus (Artículos)*, n°13-1, p. 131-153. ISSN: 1578-5157

Markofsky, S. y Bevan, A. (2012): "Directional analysis surface artifact distributions: a case study from the Murghab Delta, Turkenistan", en *Journal of Archaeological Science*, 39, pp. 428-439.

Maximiano Castillejo, A. (2008): *Teoría geoestadística aplicada al análisis de la variabilidad espacial arqueológica intra-site*. Ph.D. Dissertation. Universidad Autónoma de Barcelona. Barcelona.

Maximiano Castillejo, A. (en prensa): "Experimentando con distribuciones espaciales en CPU: Simulación estocástica y Visualización como complementos analíticos en el Análisis espacial arqueológico", en *Actas del III Congreso Internacional de Arqueología Experimental*. (Banyoles, 2011).

Mc Coy, M. y Ladefoged, T. N. (2009): "New developments in the Ue of Spatial Technology in Archaeology", en *Journal of Archaeological Reseach*, 17, pp. 263-295.

Moran, P.A.P. (1950): "Notes on Continuous Stochastic Phenomena", en *Biometrika*, 37, pp. 17-33.

Nakaya, T. y Yano, K. (2010): "Visualising Crime Clusters in a Space-time Cube: An Exploratory Data-analysis Approach Using Space-time Kernel Density Estimation and Scan Statistics", en *Transactions in GIS 2010*, 14(3), pp. 223-239.

Orton, C. (2004): "Between Time and Space the geographical dimension of knowledge", en *Archaeologia e Calcolatori* 15, pp. 299-315.

Orton, C. (2005): "Point pattern analysis revisited", en *Archaeologia e Calcolatori*. 15, pp 299-315.

Peuquet, D. J. (1994): "It's about time: a conceptual framework for the representation of temporal dynamics in geographic information systems", en *Annals of the Association of American Geographers*, 84(3), pp. 441-462.

Premo L. S. (2004): "Local spatial autocorrelation statistics quantify multi-scale patterns in distributional data: an example from Maya lowlands", en *Journal of Archaeological Science* 31, pp. 855-866.

Ripley, B. D. (1976): "The second order analysis of stationary point process", en *Journal of Applied Probability*, 13, pp. 255-266.

Ripley, D. B. (1981): *Spatial Statistics*. New York: John Wiley and Sons.

Rucker, R. von B. (1977): *Geometry, relativity, and the Fourth Dimension*. Dover Books. New York.

Schwarz, K. R. y Mount, J. (2006): "Integrating Spatial Statistics into Archaeological Data Modeling", en Mehrer, M.K. and Wescott, K.L. (eds.): *GIS and Archaeological Site Location Modeling*. Boca Ratón Taylor and Francis, pp. 167-189.

Simek, J. (1984): "Integrating pattern and context in spatial archaeology", en *Journal of Archaeological Science* 11, pp. 405-420.

Simek, J. y Larick, R. (1983): "Recognition of Multiple spatial Pattering: A case of study from the French Upper Paleolithic", en *Journal of Archaeological Science* 10, pp. 165.

Maximiano Castillejo A. (2013) "Datos óptimos para la caracterización espacial y temporal de la variabilidad arqueológica a escala intra-site", *GeoFocus (Artículos)*, n°13-1, p. 131-153. ISSN: 1578-5157

Sokal, R. R. y Oden, D. L. (1978): "Spatial autocorrelation in biology. 2. Some biological implications and four applications of evolutionary and ecological interest", en *Biol J Linn Soc.* 10, pp. 229-249.

Upton, G. J. G. y Fingleton, B. (1985): *Spatial Data Analysis by Example. Volume 1: Point Pattern and Interval Data.* New York: John Wiley.

Venables, W.N. y Ripley, B.D. (1997): *Modern Applied Statistics with S-PLUS*, New York: Springer.

Warren, R.E. (1990): "Predictive modelling of archaeological site location", en Allen, S. Green, and Zubrow E.B.W. (Eds.): *Interpreting Space: GIS and Archaeology.* London: Taylor and Francis.

Yuan, M. (1995): "Modeling semantical, temporal, and spatial information in geographic information systems", en *Progress in Trans-Atlantic Geographic Information Research*, First ESF-GISDATA and NSF-NCGIA Summer Institute in Geographic Information. Taylor & Francis, Bristol, PA.

Yuan, M. (1996): "Temporal GIS and Spatio-Temporal Modeling", en CD-ROM by the National Center for Geographical Information and Analysis.

Zurbrow, E. B. y Harbaugh, J. W. (1978): "Kriging and simulation", en I. Hodder (Ed.) *Simulation Studies in Archaeology.* Cambridge: Cambridge University Press.

¹En algunas propuestas como Lloyd y Atkinson (2004); Maximiano (2008), entre otros pocos casos, se ilustran los principios de la teoría geoestadística y su relación con los problemas arqueológicos *intra-site*.