Methodological perspectives for applying spatial point pattern analyses to Finnish Iron Age remote sensing data

Eljas Oksanen & Johanna Roiha

Abstract

Remote sensing technologies and spatial analytical approaches hold proven potential for generating completely new quantitative information on, as well as assisting the qualitative interpretation of, archaeological phenomena. They have been, however, under-utilized in Finland. This paper presents an illustrative case study of applying a sequence of well-established computational spatial analyses to data derived from a very precise LiDAR survey of the Nuuttilanmäki Iron Age site near the village of Kalho in Päijät-Häme. Previous fieldwork has identified a group of several stone cairns on a field and forested area near the settlement. Many more cairns, hidden by vegetation and almost invisible to fieldwork inspection, are revealed by the LiDAR data and add to our understanding of the area. Initial fieldwork assessment has suggested that the cairns in the field area and the cairns in the immediately adjacent forest have been produced by different historical processes, with the former presumed burial cairns and the latter the result of agricultural land clearance. An archaeological excavation to verify this interpretation would be prohibitively expensive and might yield uncertain results. But a computational examination of the cairn pattern shows that the spatial interactions between the two groups significantly differ from each, thereby
supporting the interpretation. Such methodological approaches will assist future archaeological surveys and fieldwork, and can be applied flexibly at multiply spatial scales of study.

Keywords: Iron Age, Päijät-Häme (Finland), spatial statistics, GIS, LiDAR, remote sensing, digital humanities

27.1 Introduction

In this paper we test the suitability of a suite of Geographic Information Systems (GIS) methods to produce data on and subsequently analyse Finnish archaeological site features. Though GIS has become a staple feature of fieldwork over the recent decades, the use of spatial approaches is still underexploited in Finnish archaeological research (however, see recently Roiha et al. 2021; Seitsonen & Ikäheimo 2021; Wessman & Oksanen 2022). As we will argue, remote sensing and spatial analytical methods can assist in generating survey data quickly, and in forming an initial interpretation or supporting a fieldwork-based interpretation of archaeological phenomena. In our case this is the study of a group of Iron Age (until AD 1200/1300 in Finland) cairns whose function and relationship to each other and the wider site within which they are located have remained unclear. Here our methods are deployed at a site level, but these point pattern analyses can be scaled to smaller or larger spatial dimensions, such as studying regional site distribution patterns, for instance from large-scale mapping or satellite produced remote sensing data.

The archaeological site of Nuuttilanmäki (or Nuuttilanmäki 1–4 in the Ancient Relics Register; Muinaisjäännösrekisteri 2023) is located in the municipality of Hartola, in the region of Päijät-Häme in southern Finland (Fig. 1a). Today, the area is sparsely populated countryside and the nearest settlement is the small town of Heinola. The historic Kalho manor and its courtyard are located within the area of the Nuuttilanmäki site. The small Kalhojoki River flows through the area and Lake Enojärvi is located less than one kilometre away from the site. The landscape is largely open, with
fields, meadows and pastures, but the site itself is partly covered by dense forest.

Nuuttilanmäki is a very large and complex archaeological site (Fig. 1b). It includes an Iron Age settlement site, possible Iron Age burials, a cup-marked stone and an exposed bedrock formation with a cup-mark probably dating to the Iron Age, the remains of a medieval village, and possible remains of previous slash-and-burn agricultural areas. Field surveys have recovered a number of quartz flakes, which have been interpreted as signs of a Stone Age settlement site. However, most likely these quartz flakes are related to

Figure 1. A – the Nuuttilanmäki site is located in southern Finland, in the area between Lake Päijänne and Lake Saimaa; B – the site is complex and includes a possible medieval village area, Iron Age features, cairns and a cup-marked features. The Kalho manor house is located in the northwest corner of the area. Illustration J. Roiha.
the Iron Age settlement, as they have also been found at other Iron Age sites in the region (Roiha & Holopainen in prep.). The archaeological complexity of the area makes it challenging to research and interpret. One previous find, a spearhead found in a field, connects the site to the Viking Age (AD 800–1050). The spearhead is held by a local museum, and it is not listed with other site finds in the Finnish Heritage Agency’s digital record. It is very likely that there were houses or even a village at Nuuttilanmäki already during the Viking Age. The earliest written reference to the village of Kalho is from 1405, and according to historical sources it was composed of ten houses in 1560 (Tikkala et al. 2012).

The site was discovered by chance by the archaeologist Timo Sepänmaa in 1997 (Sepänmaa 1997). Further archaeological surveys were conducted in 2003 (Poutiainen 2003) and 2004 (Jussila 2004). Both of the surveys were related to the zoning of the area. In the spring of 2019, a new archaeological survey was carried out by one of the authors. The aim of the survey was to obtain more precise knowledge about the site and its different landscape features. As a result, new information about Nuuttilanmäki was produced and one new site (Peltola 1–3, consisting of a cup-marked stone, cairns and a historical building site) was discovered (Roiha 2019). Most likely these are related to settlement in the overall Nuuttilanmäki site and may have been part of a possible Iron Age village, significantly expanding its area.

27.2 Materials

The Nuttilanmäki site includes an interesting area of possible burial cairns. Most of the cairns are small, symmetrically round or oval-shaped with a diameter of 1–3 metres and a height of about 20–30 cm. The cairns are built of stones about 20 to 40 cm in diameter. More cairns have been found in the deeper forest area, but those cairns seem to lack symmetry and have been preliminarily interpreted as being related to slash-and-burn agriculture.
Overall, the cairns (or agricultural piles) are extremely difficult to visually spot on the ground because of the dense vegetation. In the pasture area, the cairns are covered with grasses and in the adjoining forest area by dense bushes and trees. The soil type in the area is rocky, and in some cases, cairns cannot be distinguished from natural rocks. Because of challenges in the field study and the interpretation of the area, drone-based ALS (airborne laser scanning) was tested as the site survey method.

In the autumn of 2019, the commercial company Geotrim Oy, together with one of the authors, conducted a field survey at the site with the GeoDrone X4L multicopter and YellowScan Surveyor LiDAR sensor. As part of
the field study the site was also photographed with the drone to produce a good quality orthophoto of the area. Geotrim Oy pre-processed the ALS data and produced the orthophoto. The results were published in 2021 (Roiha et al. 2021).

The following workflow was used to process and interpret LiDAR data: first, ground points were transformed into a 0.1 × 0.1 m raster-form Digital Elevation Model (DEM). The raster file can be conceptualized as akin to a digital image, where each of the individual pixels that form the image has been given a numerical value that indicates a height estimate of that location relative to the rest of the dataset. The hillshade effect (simulating sideways illumination of the ground) was then applied to the raster data to highlight the unevenness of the terrain and to make the cairns more visible. Unlike in the 2021 study, in this work we did not use orthophotographic material from the area for cairn interpretation because it could be used only in a very limited area (an open field not covered by trees and tall bushes).

From the LiDAR dataset it was possible to detect cairns in the field area and the forest cover area (Fig. 2). Cairns in the field area and at the edge of the forest are round or oval in form, which can also be seen in the LiDAR data (Fig. 3a–b). Those cairns were interpreted as graves. Cairns in the deeper forest area appear more as ‘spots’ in the LiDAR DEM data. In the field survey it was noted that many of these cairns seem to be more like piles of stones that have been thrown together (Fig. 3c–d). Furthermore, cairns in the forest area are taller and some are built around one larger stone. The area between these cairns appears to contain no, or at least far fewer, loose rocks, suggesting that the piles were produced as part of a process of clearing the area of the stones. In the area of burial cairns there remain more stones and rocky spots on the ground, which would have made any kind of cultivation difficult. The small area between possible burial cairns and agricultural cairns is covered with very densely growing spruce seedlings. Point density in that area is lower and it affects the DEM quality of that area (see Roiha et al. 2021). The same problem applies to the field survey. Very dense vegetation makes observations in the field difficult and can also hide archaeological features in the LiDAR data.
27.3 Methods

As with our case study, much archaeological data is inherently spatial. Encoding observations, such as site or findspot distributions, as a point cloud with coordinate values and subjecting it to spatial statistical analysis forms a well-established suite of archaeological methods (e.g. Conolly & Lake 2006; Gillings et al. 2020; Nakoinz & Ritter 2016; Wheatley & Gillings 2002). Here we test two complementary computation methods on the presumed cairn data at the Nuuttilanmäki site. The first, kernel density estimation, is
probably familiar to most archaeologists who have practised GIS analysis; indeed, as the results can be expressed as a ‘heat map’ indicating spatial intensity of point locations it is easily interpretable by anyone. The second, a family of related techniques that examines spatial interactions within the data, is not as widely utilized. Together with a recent University of Helsinki master’s thesis (Ahvonen 2023), this appears to be the first time it has been deployed in a Finnish archaeological context. As we will demonstrate, these are useful exploratory techniques that can be flexibly applied to many different archaeological spatial datasets (see also Bevan 2012; Bilotti & Campeggi 2021).

The analysis was carried out with the widely used statistical software package R (version 4.2.2) using the development environment RStudio (build 576). Kernel density estimation was conducted with the package raster (Hihmans et al. 2023) and the spatial interactions analysis was conducted with the package spatstat (Baddeley et al. 2022). As described above, the point data was generated by visually examining features from a LiDAR scan of the Nuuttilanmäki area. In order to carry out the point pattern analyses, an observation window was described as a spatial polygon, generated using the QGIS (version 3.24.1-Tisler) plugin Concave Hull to drape boundaries around the full point cloud and around the two presumed subsets in the agricultural and burial areas.

Our first applied method is kernel density estimation (KDEN), a data smoothing technique that can be used in both spatial and non-spatial analysis (O’Sullivan & Unwin 2010). In its basic application for spatial analysis, the software moves a kernel shape of a predetermined bandwidth (or diameter) across the study area at set intervals. At each interval location the number of observations (or points) that fall within the kernel shape is counted and a density value calculated using a quadratic equation. As with the LiDAR scan, the final output is a spatial raster dataset. The results can be visualized using a colour range and are well suited for archaeological assessment and interpretation of distribution patterns (Baxter & Beardah 1997; Bevan 2020).

The granularity of the results is determined by the cell (‘pixel’) size and the bandwidth (‘kernel diameter’) value. In our example we contrast results
obtained using two bandwidths (Fig. 4). The lower bandwidth is set arbitrarily at \( \text{sigma}=10 \), although 10 metres is very close to the average distance between the estimated centre of each identified cairn and that of its nearest neighbour (10.62 metres) and therefore somewhat reflects the spatial structure of the distribution. At this range of analysis local high-density groups are usefully highlighted in both of the presumed subareas. It is, however, difficult to determine whether there are significant differences between the two subareas. If we increase the bandwidth to \( \text{sigma}=30 \), local variation is obscured but we can see more clearly that, overall, spatial density of the central cemetery group is greater than that of the cairns in the forested area.

The second applied method builds upon the classic Nearest Neighbour analysis formulated by Philip Clark and Francis Evans for examining animal and plant species distributions in ecology (Clark & Evans 1954). This approach envisages point patterns as divided into three idealized categories: dispersed (points repel each other), clustered (points attract each other into one or more clusters) or spatially random. The Nearest Neighbour test (also known as the Clark and Evans test) produces a numerical value \( (r) \) that falls

![Figure 4. Kernel density estimate based on the overall pattern of cairn locations using two different bandwidths: A – sigma=10 m; B – sigma=30 m. The thicker boundaries delineate the two subareas of presumed burial (northern) and agricultural (southern) cairns. Illustration E. Oksanen.](image-url)
on either side of 1. An $r$ value lower than 1 indicates a clustered point process and higher than 1 indicates dispersal. The closer to 1 the $r$ value tends, the closer it is to an idealized random distribution. But much as with conducting a KDEN only using a single bandwidth, a single result value for a full pattern can be misleading, as it obscures variation at different ranges of analysis; the observations (sites, monuments, findspots) may be governed by different processes at very close ranges than at larger scales.

A suite of scale-sensitive techniques, or functions, has been developed to overcome this problem. Here, we deploy the three most commonly used in archaeological investigations: the K-function, the L-function, and the Pair Correlation Function (Bevan et al. 2013: 32–42, Bevan 2020, Nakoinz & Ritter 2016: 135–44). The oldest is the K-function which was developed by B. D. Ripley (thus sometimes also referred to as Ripley’s K: Ripley 1977). The method calculates the spatial intensity of interactions around each point at increasing intervals. For example, we would start by selecting a point, counting the number of other points at 0–1 metre distance, then adding to that number the number of points at 1–2 metres, then at 2–3 metres, and so forth. The average intensity at each spatial interval (an $r$ value) across the whole population of points is then calculated using a formula and plotted as a line. This is contrasted against the idealized spatially random Poisson distribution demarcated as a second line. Much as with the Nearest Neighbour test, at the ranges where the line representing the real observations dips below the Poisson line the pattern can be said to be dispersed, and where it rises above the line the pattern is clustered (Fig. 5). The L-function is simply a variation on the K-function, with the Poisson line straightened so that the details of the observed pattern can be more easily visually interpreted.

As discussed, the way that the K- and L-functions calculate the intensity of interactions is cumulative: the number of other observations around each point is added together as the range increases. This is useful if we wish to establish how the overall structural tendency (dispersed or clustered) within each pattern develops, but we may also wish to examine the pattern in yet finer detail at specific spatial ranges. The Pair Correlation Function (PCF)
Figure 5. K-function, L-function and Pair Correlation Function analysis: A – the full point pattern; subdivided into B – the presumed burial and C – agricultural groups. It would appear that the burial pile pattern exhibits clustering from ranges of c 10 metres onwards where it reaches above the grey-coloured probability envelope. Illustration E. Oksanen.
achieves this by calculating the intensity of interactions not cumulatively, but separately for each spatial interval; that is to say, points at each increasing range are not added together but counted separately. With PCF we gain a better understanding of interactions at specific intervals. As a matter of good practice, it is always best to run all the different types of analyses in order to develop a fuller understanding of pattern structures and their underlying processes.

As a final point, this method can be further refined to assess the possibility that a pattern is statistically random. As noted, the Poisson line indicates idealized complete spatial randomness. But we know an observed distribution does not need to exactly dovetail with the Poisson line to be random, which raises the question of how close to the Poisson line must it be for this interpretation to hold true. We can construct a probability envelope around the Poisson line using a popular probability calculation technique called the Monte Carlo simulation (Robert & Casella 2013; Bevan 2020). Here it works as follows: we generate a number of points equal to the real observations and randomly place them within the observation window. A function graph is then calculated as usual. This process is repeated a number of times, with 999 simulations typically considered sufficient. The simulated results are laid on the same graph as that of the real observations, with the top 2.5% and lowest 2.5% deleted to remove statistical outliers. The resulting space between the top and bottom results can be visually depicted as a grey zone and comprise an area within which the pattern is considered random with a high degree (95%) of probability (Fig. 5).

27.4 Results

Both KDEN and the function analyses support the initial fieldwork survey interpretation that these two cairn areas are the product of different historical phenomena. The KDEN surface calculated with a lower bandwidth
(sigma=10 m) readily highlights cairn clusters and differences between the distributional densities of smaller groups that would be difficult to pick out by simple visual observation of the point pattern (Fig. 4a). The propensity of the human brain for visual pattern recognition may sometimes lead us astray, as we may detect false patterns in complex data. Here KDEN serves as a useful means of simplifying and clarifying the spatial structure of the cairns. Conducted with a larger bandwidth (sigma=30 m), the analysis shows that the clustering of cairns in the northern subarea is significantly more intense than in the southern, where, despite a number of tight small clusters that may produce the illusion of relative density, the overall character of the pattern is more dispersed (Fig. 4b). This would accord with the interpretation that the southern cairns are the product of local land clearance as people were piling nearby loose rocks together in order to move them out of the way of agricultural activity.

The function graphs further support this interpretation (Fig. 5). If we consider the K- and L- function analyses of the full point pattern, we observe that its overall tendency is towards clustering: the function line rises above the Poisson distribution line and then clears the probability envelope at about the 10 m range. This is to say that a statistically significant clustering phenomenon occurs, on average, within the structure of the pattern from 10-metre distances onwards. As noted, this is approximately the average nearest neighbour distance between the centre points of cairns, and it would appear that clustering is an important feature in the overall composition of the pattern.

If we split the full point pattern into the two subsections, however, we see a clear difference emerging. The K- and L-functions show that the clustering is entirely a feature of the northern presumed burial area. If we consider the southern data in isolation, the function line never leaves the grey probability envelope, which indicates that the point pattern is spatially random. It would seem that the southern cairns were not built with reference to each other's locations. These results are further clarified by the PCF graphs. The results similarly indicate that the southern pattern is random and the northern clustered, but we can further see that the range of most significant clustering in
the presumed burial area occurs between circa ten and twenty metres, perhaps extending to thirty metres. This could mean that when new cairns were built their location was considered with reference to existing cairns within this range, as will be discussed below.

27.5 Discussion

The LiDAR data used in this study is remarkable and unique in Finnish archaeology for its high accuracy. Ground point density average is 69.22 points/m\(^2\) on open land and 12.35 points/m\(^2\) in the dense forested area. By contrast, the new nationwide LiDAR dataset that has been produced since 2020 by the National Land Survey of Finland (NLS) has a point density of 5 ground points/m\(^2\). The older generation (2008–2019)\(^1\) of NLS data with full coverage over Finland is only 0.5 ground points/m\(^2\), and during the field survey of 2019 it was the only other available LiDAR dataset for the area. Until the new national LiDAR survey is completed, this is still the most accurate data available for many regions.

Cairns and other small archaeological features simply cannot be reliably spotted in the old 0.5 points/m\(^2\) LiDAR dataset, and even the new data may not be accurate enough for such structures. A 2018 study at Ellinniitty in Rauma (Muinaisjäännöskisteri ID 684010137), in the region of Satakunta, used the 0.5 accuracy NLS dataset (so over 100 times fewer ground points per square metre than in the data used for this study) for an initial survey of the site, followed by fieldwork excavations. The study showed that only five percent of the cairns in the study area had been detected from the LiDAR data (Lehto 2018). Of course, even with the means to produce high accuracy data there are challenges that may not be easy to overcome. It is possible that Nuuttilanmäki also hides many more underground cairns and burials than have been identified in this study. We elected to create our study area and its sub-areas directly with reference to the existing point observations since we could

not securely delineate the boundaries of the archaeological activity site. As noted, the gap between the northern and the southern distributions (Fig. 1) is the product of a very dense thicket of young saplings that block LiDAR. We may presume that the area of cairns extends into it, but it is not possible to evaluate this. It is, of course, similarly challenging to examine the terrain in the thicket through fieldwork.

By combining qualitative (fieldwork assessment) with quantitative (spatial data and analysis) approaches, the results of this study indicate that the cairns in the field and the cairns in the deeper forest area are archaeologically different features. Most likely they were made for different purposes and perhaps at different times. Cairns in the field and at the edge of the forest are carefully constructed, symmetrical and clearly clustered structures. These structures and the field area resemble other well-known Iron Age burial sites in Finland, for example, Päivääniemi in Lempäälä (Muinaisjäännö斯rekisteri ID 418010001). Cairns in the forested area are most likely related to agricultural land use. Another possibility is that the area was cleared of stones during forest management, in which case the piles might not even be prehistoric.

The methods used provide perspectives especially into the spatial relationship of cairns with one another. That their positions seem to be in relation to other immediately neighbouring cairns is probably not surprising, but we would further suggest that the formation of the overall pattern was governed mainly by second-order geospatial processes (relationships internal to the pattern, which could be related to, for example, kinship relations, social status, or temporal vicinity in construction) as opposed to first-order processes (patterns influenced by external factors such as the landscape). It should be noted that there is only one large cairn that is visible – thanks to a large top-stone boulder – from a long distance away. A viewshed analysis has shown it is visible along the river and all the way to the lake, making it a stand-out feature in the local travel landscapes (Roiha et al. 2021). This cairn, however, appears to be a clear exception with a location right at the edge of the group. The rest of the cairns are low and difficult to spot at a distance. Visibility, at least, does not appear
to have been a concern in their construction. Such tentative hypotheses could guide the planning of future fieldwork.

It is naturally not necessary to subject all archaeological observations amenable to being represented as point patterns to these approaches. But it is generally recognized – and we have found this to be true in our research – that they critically assist in investigating complex spatial data and in studying patterns where variation and structure is difficult to detect visually. There is considerable unexploited potential for applying these (and other powerful spatial statistical techniques) in archaeological investigations in Finland, and we highlight this site in particular as a test case for combining different GIS-led approaches: LiDAR scanning to quickly generate archaeological data and observations, and statistical methods to perform exploratory analyses upon them. Such workflows can be used to perform rapid initial assessments and interpretations of any number of potential sites, especially ones that are large or difficult to access. In the future, as LiDAR survey accuracy continually increases and equipment costs come down, we may be able to generate a great deal of new archaeological GIS data in Finland: a project to identify larger archaeological features from the new 5 points/m² NLS LiDAR data in Lapland has already pointed the way (Seitsonen & Ikaheimo 2021). These approaches have been adopted to automatically detect Finnish archaeological sites using machine learning in the currently ongoing LIDARK project by the Finnish Heritage Agency and the University of Oulu².

As a suite of non-destructive approaches, spatial analyses can be deployed to produce information about site features, generate preliminary hypotheses, and support fieldwork interpretations. Site excavations are expensive and destructive, and should be carefully considered, justified and targeted. In the context of our case study, a future excavation of one cairn of both types, as indicated by the analysis, could give much-needed information about these poorly known structures. Similar types of ambiguous features can be found at many places around the Päijät-Häme region. In some cases, the sites have been classified as burial sites and in others as related to agriculture, but these interpretations remain uncertain. A similar type of site with mysterious cairns

and cup-marked stones was excavated in 2003 in Hartola, Kotisalo (Muinais-
jäännösrekisteri ID 81010011; Lahelma 2003). Unfortunately, but perhaps
not unrepresentatively, this small-scale excavation did not produce clear re-
sults and the Päijät-Häme cairns continue to keep a tight hold on their secrets.
A combination of traditional fieldwork approaches with new computational
studies may, we would argue, open new interpretative possibilities.

In conclusion, we would note that point pattern analyses scale flexibly
and can be used with a wide variety of spatial datasets. They can be deployed
equally in micro-level studies, such as studying distribution within excava-
tions pits, or for investigating large-scale relationships at regional and su-
pra-regional levels. Supplemented by KDE, we have used function analyses
on archaeological features in a Finnish Iron Age site-level study, but these
have also been tested to study object distribution patterns across a settlement
site in Late Chalcolithic northern Mesopotamia (Bilotti & Campeggi 2021),
and modern and ancient settlement patterns on the Greek island of Kythera
(Bevan & Connolly 2006). This case study has considered the underexploited
potential of spatial statistical approaches in Finnish archaeology, and there are
also many more furrows to plough internationally. Moreover, where consider-
ations relating to increasing costs, environmental perspectives or simply a lack
of safety in travel destinations may cause archaeologists to reassess the basis for
regular long-distance fieldwork travel (e.g. Scerri et al. 2020), greater use of
these methods can provide significant methodological and logistical support
for future international scientific endeavours.

Acknowledgements

Eljas Oksanen’s work on this project has been made possible by funding from the
European Union’s Horizon 2020 research and innovation programme under the
Marie Sklodowska-Curie grant agreement No 896044. Johanna Roiha’s work
on this project has been funded by Kone Foundation grant No 202006680.
References

Archival sources


Unpublished sources


Literature


Bevan, A. 2020. Spatial points patterns and processes. In M. Gillings, P. Hacigüzeller & G.


