



UvA-DARE (Digital Academic Repository)

Sampling past landscapes

Methodological inquiries into the bias problems of recording archaeological surface assemblages

Waagen, J.

Publication date

2022

[Link to publication](#)

Citation for published version (APA):

Waagen, J. (2022). *Sampling past landscapes: Methodological inquiries into the bias problems of recording archaeological surface assemblages*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

CHAPTER 4.

SCRATCHING THE SURFACE: INTEGRATING LOW-VISIBILITY ZONES AND LARGE RURAL SITES IN LANDSCAPE ARCHAEOLOGY USING POINT SAMPLING³

ABSTRACT

Whereas archaeological field survey is relatively fast and effective for the mapping of surface finds in Mediterranean landscapes, two challenges limit its potential for reconstructing ancient settlement patterns. First, field survey usually excludes portions of the landscape that are inaccessible or present low ground visibility due to vegetation or the terrain, and second, including large settlements sites poses logistical problems, as these tend to produce unmanageably high frequencies of finds. In this paper, we explore the potential value of point sampling for integrating these areas in standard transect field survey projects. For this experiment, we sampled a large rural Archaic–Roman period site and its offsite environs in Molise, central-southern Italy. We present a systematic statistical and spatial comparison between data derived from both methods on the same areas. As such, the experiment contributes to the current debate on how to compare and integrate data from survey projects that apply different techniques.

4.1 INTRODUCTION

Mediterranean survey projects usually involve some kind of line-walking by teams of archaeologists across well-visible terrain, that is with no or limited vegetation, such as ploughed fields and fallow land. Surveys in the Mediterranean have a long history, and different practices of documenting and/or collecting finds exist and have evolved over time. Some projects document only ‘diagnostic’ finds, whereas others consider

³ This chapter is based on the article “Scratching the Surface: integrating low-visibility zones and large rural sites in landscape archaeology using point sampling” by Tesse D. Stek and Jitte Waagen, originally published (11 February 2022) in *Journal of Field Archaeology*, Early View (publisher Taylor & Francis) as an Open Access article distributed under the terms of the CC-BY license. This article is accessible through this link: <https://doi.org/10.1080/00934690.2022.2031466>.

This study has been executed in the context of the Tappino Area Archaeological Project (TAAP), coordinated by Tesse D. Stek (Royal Netherlands Institute in Rome). The TAAP is organized by Tesse D. Stek, Rogier A. A. Kalkers, Jesús García Sánchez, and Jitte Waagen. The work at Colle S. Martino was conducted with permission of the Soprintendenza Archeologia, Belle Arti e Paesaggio del Molise. The fieldwork would have been impossible without the support provided by the Comune di Jelsi, especially Michele Fratino, and our collaborative research center in Molise—the Centro Didattico di Studi Archeologici di Jelsi (CeDISA)—and the team itself, at Colle S. Martino, consisting of Tesse D. Stek, Rogier A. A. Kalkers, Lennart Kruijer, Koos Mol, and Max Caspers in the field and Lisa Götz, Lucia Lecce, Sheila Cherubini, Filippo Salamone, Jacqueline Splinter, and Rogier A. A. Kalkers for finds analyses. Particular thanks go to readers of the first drafts, that is, Vladimir V. Stissi, Jesús García Sánchez, and Rogier A. A. Kalkers.

all finds. The comparability of survey data collected using different sampling techniques is a key topic both within the projects themselves as well as between different survey projects, old and new (Attema et al. 2020). It is therefore imperative to increase our understanding of different sampling techniques and their specific analytical qualities. In this paper, we compare the ‘standard’ Mediterranean method of transect survey with point sampling, which are spots where, within a 1 m diameter, all finds are collected. We apply these methods to approach two important concerns in field survey, each regarding the extreme of the spectrum: on the one hand, the scarce data of highly vegetated areas in the landscape and, on the other, the very large rural sites that instead produce tons of finds. The experiment taps into a set of methodological issues that we shall briefly discuss.

4.1.1 SMEARING: ON SPATIAL DEFINITION

For standard transect survey, a well-established protocol exists in which typically 5–10 persons, distanced 5–10 m from one another, walk fields (as defined by modern topographical conditions) or regularized units (of standard measures, such as 50 × 50 m) along lines. Along the line they walk, they document all finds 1 m to their left and 1 m to their right, which results in a 20–40% coverage of the units. Collecting surface material per field (e.g. Voorrips et al. 1991) or per pre-established units of, e.g., 50 × 50 m or 100 × 100 m (e.g. Alcock et al. 1994) has the considerable advantage over smaller collection units in that it costs less time for collection and documentation, both in the field as well as in processing the finds post-collection. However, the risk here is that archaeological finds potentially relating to different archaeological or historical realities are lumped together, thereby losing spatial resolution but also involuntarily widening interpretative scenarios from offsite material (Millett 2000, Tol 2012). For instance, Hellenistic material from one end of a field (potential interpretation: offsite material of Hellenistic farm?) will, as a result of post-factum decisions (i.e. the putting together of all bags with collected finds by the separate walkers), be mixed with Roman period material from the other end of the field (potential interpretation: offsite material of the nearby Roman settlement?), thus losing interpretative precision (‘offsite Hellenistic–Roman material’).

With high find densities (‘sites’), these problems are somewhat limited, because finds are (besides per unit) documented within the boundaries of the high concentration area (the ‘site’). However, within sites, a smearing effect of the spatial attributes of the archaeological data also occurs. With regard to small or medium-sized rural sites, of dimensions from ca. 5–50 m in diameter, this is less problematic, especially for single period sites. For small and medium-sized rural sites, fuller or total collection strategies with a higher resolution of collection units is feasible in light of their limited size and overall find quantities. A whole range of possible methods is available for such sites, from the selection of only diagnostics (e.g. the methods of the *Forma Italiae* series; for an example, see Marchi 2016) down to the mapping of single finds with (D)GPS (e.g. García-Sánchez and Cisneros 2013). There is a clear trade-off between intensity of collection strategies and time investment, with an optimum between the extra information acquired on the one hand and the time investment on the other. This optimum is different in each situation, because it depends on the research questions and the character of the archaeological situation.

4.1.2 SAMPLING HIDDEN LANDSCAPES

Despite the advantages and proven merits of the transect method, a more important weakness is its bias favoring specific, presently arable, land-use conditions (cf. Attema et al. 2020: 5). As transect surveyors depend on recently ploughed or otherwise high-visibility portions of the landscape, terrains and soil types preferred in modern agriculture are likely to be overrepresented in most survey projects. In the Mediterranean, the low-visibility areas are often covered with forests and shrubs. As a result, these areas are substantially less, and often not at all, investigated. Consequently, we risk a systematic bias against ancient settlement for certain periods/societies in which the fertile terrains were avoided in order to cultivate them (cf. Van Leusen et al. 2007). Indeed, modern land use and vegetation patterns may diverge from ancient landscapes as a function of different technology, socio-economic and demographic regimes, or climatic conditions. Indeed, many areas settled or cultivated in antiquity are now hidden by vegetation (for the effect of reforestation on archaeological visibility in Molise, cf. Stek et al. 2015: 240). The character

of settlements may also differ according to soil conditions and geography, which means that we cannot simply extrapolate site numbers and site types from the fertile areas into non- or low-visibility areas in the research area. Sampling these low- or non-visibility zones with comparable methods is clearly essential, but precisely how to achieve this in a feasible way is the crux. For including low-visibility zones comes at significant cost: methods to sample low- or non-visibility areas are very costly in terms of time and energy. Therefore, these methods usually limit sampling to small areas (Van de Velde 2001, Caraher et al. 2014). This means that finding the right balance in sampling is of great importance—with the least redundancy, but enough statistical power to help interpret these hidden portions of the landscape.

4.1.3 SAMPLING LARGE RURAL SITES

Another key problem in standard field survey practice is the inclusion of large rural sites. Total collection is rarely an option in sites of over 5–20 ha in surface, as this would lead to the collection of hundreds to thousands of kilos of archaeological material, depending on the character of the site (cf. Attema et al. 2020: 30). An often-applied solution is to diverge from the standard collection strategy and be (very) selective in the collection of surface material, typically by only collecting finds that are deemed to be diagnostic in terms of periodization and function. But selective collection has risks, too. As we know, archaeological finds can turn out to be much more indicative if collected, washed, and studied by a specialist in the find lab, sometimes revealing phases or activities that were not recognized in the field by (often less-experienced) field walkers. Selection can thus lead to the systematic ignoring of entire chronological periods or functions of (part of) a site. In addition, these sites present differential visibility conditions within the sites themselves. Large sites can be located over many different plots of terrain, including non-visibility areas. Habitually, these areas are regarded as non-response areas with regards to the field survey and are thus excluded from analysis. In our view, however, this presents a considerable risk of losing important information. We think that trying to resolve the exclusion of non-visibility areas is in many cases as important as, if not more important than, attaining a high spatial resolution or total collection of finds in the high-visibility areas exclusively.

4.2 A SYSTEMATIC TEST OF POINT SAMPLING AND LINE-WALKING IN MOLISE

To try and develop a survey method to investigate lower and non-visibility areas in both offsite and onsite conditions, we created a research design to test the virtues of point sampling, as well as its comparability and integrative potential with other, more standard survey methods (Van de Velde 2001, Caraher et al. 2014, Attema et al. 2020). The area we selected for this test is located in the Tappino Valley in Molise, central-southern Italy (Stek 2018). The test formed part of the Tappino Area Archaeological project, based at Groningen University (2013–present), which builds on the Sacred Landscape Project (2004–2010). The aim of these projects is to shed light on settlement developments in the Classical, Hellenistic, and Roman periods in this area by applying a combination of intensive field survey strategies and remote sensing, as well as excavation. In antiquity, the area was inhabited by a people known in the ancient sources as the Samnites (and the area itself as Samnium). The Samnites are thought to have had a particular, non-urbanized settlement organization. Important research has been conducted in the area since the 1960s, especially, but there are still many unknowns (for Samnite history and archaeology, see e.g. Salmon 1967, La Regina 1989, Barker 1995, Dench 1995, Tagliamonte 1997, Jones 2004, Bispham 2007, Scopacasa 2015). Cult sites and hill-forts are the most visible remains of antiquity in the area. However, settlement sites are less documented and understood. Also, here, large rural sites such as villages are conspicuously absent in the overall picture. The Sacred Landscape Project was designed with a view to contribute to this debate. Its aim was to map the precise function and chronological development of rural cult sites (Stek and Pelgrom 2005, Stek 2009; Pelgrom and Stek 2010), along with methodological tests with the offsite data (Waagen 2014). One of the main results of the surveys around the sanctuaries of S. Giovanni in Galdo, Colle Rimontato and Gildone, Cupa was the documentation of very dense rural settlement around the sanctuaries, including village-type settlements. The Tappino Area Archaeological Project follows up on this project by investigating in more detail the pattern and development of Samnite and Roman settlement in the area and farther away from the sanctuaries and to investigate the character and development of both small and large

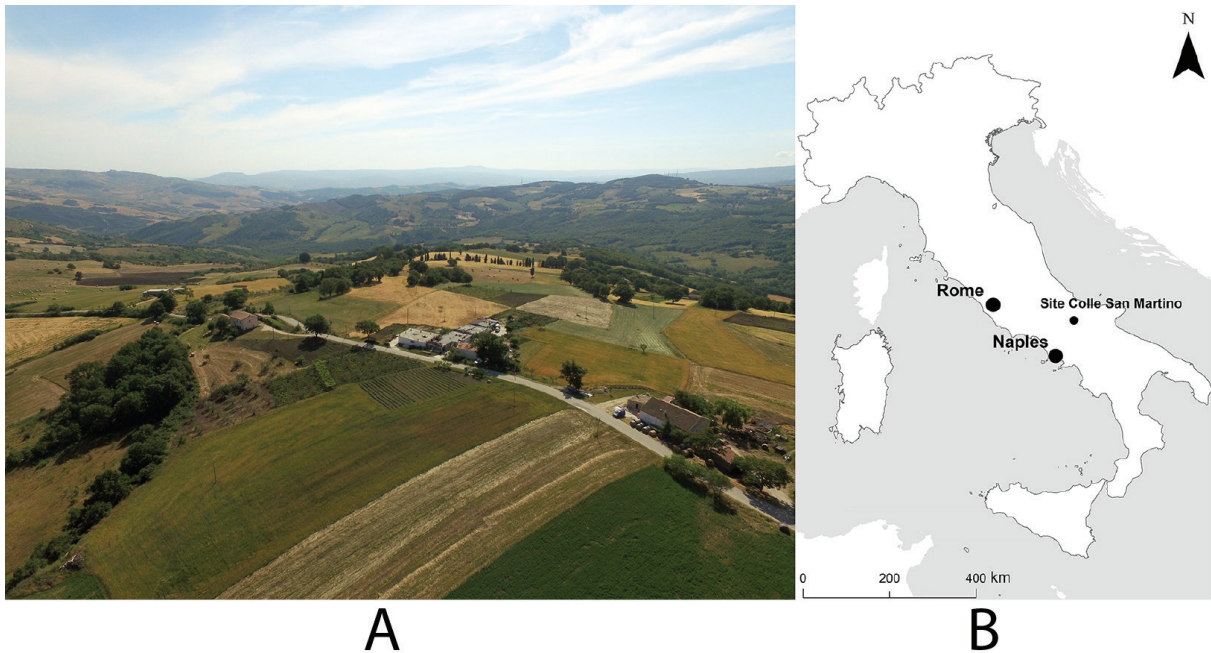


Figure 4.1. A) The site of Colle San Martino in Jelsi, (CB) Molise. Aerial photo (T. D. Stek) and B) location.

rural sites. As part of this on-going research project, we aim to develop methods to also include low- and non-visibility areas in our study. In the 2013 summer campaign, we experimented with the use of point sampling when we came across a large rural site in the territory of the modern Comune of Jelsi (fig. 4.1) on a hilltop called Colle (or Montagna) San Martino. In the 2014 campaign, we applied the same method on the large hill-fort site of Montagna di Gildone in the territory of the Comune of Gildone. Here, we describe the test design and first results of our research at Colle S. Martino.

4.3 THE SITE OF COLLE S. MARTINO AND THE SAMPLING DESIGN

The site of Colle S. Martino is positioned at the top of a long moderate ridge at an altitude of 660 m. While the site has long been known locally to produce archaeological finds (e.g. D'Amico 1953), it has never been subjected to detailed archaeological analysis. At a first visit to the site in 2012, we found evidence of a very large scatter of archaeological material, visible to various degrees in small plots of land, including within the orto (garden) of a farmhouse where we found, among many other finds, a 6th century BC black gloss (BG) rim. During a preliminary visit to establish our strategy in the field in August 2013, we encountered exceptional archaeological finds, in terms of both quantity and quality, in a recently ploughed field, where almost intact pots of BG pottery, as well as bronze items, were retrieved. In view of the size of the site (estimated over 8 ha), its exceptionally high find density, and its possible historical importance, we decided to apply an intensive survey strategy to the site, involving detailed surface survey, aerial photography using drones, and coring. The archaeological interpretation and historical embedding of the site will be presented elsewhere; for the current paper, it suffices to acknowledge that we are dealing with a large, rural settlement site with finds dating from the Archaic/Classical periods to the Roman imperial period. In view of the size and character of the site, we decided to select this site and its environs for our first test with point sampling. Later, we also applied point sampling at the large hill-fort site of Montagna di Gildone and at a large rural site of the Roman period near S. Giovanni in Galdo. We focus here on Colle S. Martino, but we will make some comparisons when useful.

Our aim was to establish the suitability of the method for two different (but related) aspects. First, to implement point sampling as a systematic sample method in comparison to standard transect survey in the same location and, second, to use point sampling as a way to overcome difficulties in the interpretation of the overall archaeological situation due to visibility constraints by sampling areas that were not sampled or with bad response in the transect survey.

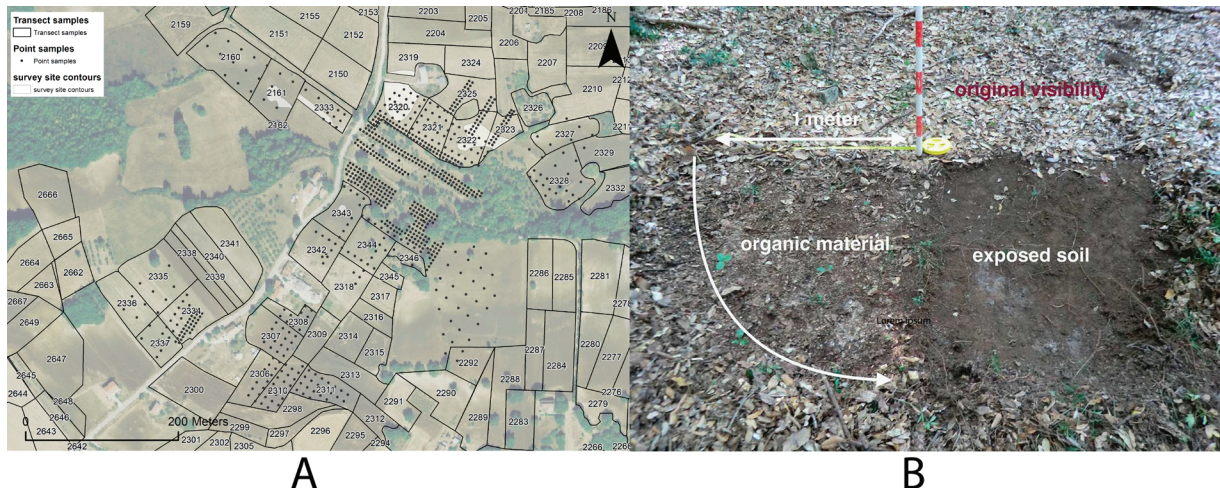


Figure 4.2. A) The site area of Colle San Martino, coverage of transect samples and point samples, and B) example of a point sample.

As to the first aim, the sheer size of the site and the find density encountered meant that it was clear that a selective sampling strategy was in order to avoid spending all our future field and find lab campaigns just on this site. On the other hand, we wanted to minimize the problems related to collecting in transects because of the smearing effect and the potential bias resulting from collecting only so-called diagnostics. A method applying total collection of small samples of the entire site therefore seemed appropriate and the use of point samples worth a try. As to the second aim, we noted considerable problems in the overall interpretation of the site as a result of limited visibility in portions of the area where the site seemed to be located. Both for establishing its internal organization and continuity, as well as for determining the extension of the site as such, point samples seemed a suitable additional tool.

The site was sampled using a standard transect method with 20% coverage of units with a maximum size of 50×100 m (which were used in the Sacred Landscape Project and in the Tappino Area Project; for explanation, see Waagen 2014). Point samples, circular sample units with 1 m diameter, were collected in those units at 10 m intervals and, in some cases, a 5 m interval. Point samples were also collected where transect survey was impossible due to heavy vegetation (fig. 4.2B). This resulted in a sample collection of 25 units with overlapping point samples, a total of 794 point samples in transects, as well as non-visibility areas, eventually amounting to 9255 individual sherds (fig. 4.2A). As a result, we have a considerable sample of transect sample units and point samples, partly overlapping, at various intervals. To enable direct comparison between them, we aggregated the data of the point samples on the basis of transect sample unit boundaries, so we have two sets of material which can be regarded as paired samples. By spacing our point samples systematically at equal intervals, we made sure that a larger unit would be covered by more point samples and thus avoid differential statistical power. This can be regarded as a spatial implementation of sampling with a Probability Proportional to Size (Orton 2000: 34). For our analysis, we selected only units with at least an overlap of four point samples to avoid limited sample size effects (Orton 2000, Waagen 2014). This resulted in a dataset of 18 transect sample units for which the data can be compared to overlapping point sample aggregates at 10 m intervals, with a minimum of four point samples (amounting to a total of 246). This is the dataset used for the comparison of the two sample methods, unless stated otherwise.

4.4 COMPARING POINT SAMPLES AND TRANSECT SAMPLES

To establish the usefulness and optimal implementation of point sampling, we first determine what type of information can be extracted from them and how that information compares to the information retrieved from standard transect survey. Techniques such as shovel-testing or test-pitting have been experimented with extensively (Krakker et al. 1983, Kintigh 1988, Lightfoot 1989, Chadwick and Evans 2000). The difference is that these collection methods are actually excavations, whereas point sampling is a less invasive surface examination. Evidently, the finds collected within small units, with archaeologists on their knees, result in a different and larger sample per examined area than that produced when archaeologists walk a

transect line and pick up finds (e.g. Burger and Todd 2006: 243). There are two sides to this difference: on the one hand, if point samples and transect samples produce roughly comparable information, this facilitates direct integration of both methods in one survey design. On the other hand, if a different type of information can be extracted from point sampling than from transect sampling, this may add new information on the archaeology of the area but means that straightforward comparison with transect sampling is more difficult.

We discuss the results along different parameters. We will start with a quantitative comparison, with the following key questions: is the density recorded in transect sampling and point sampling related and thus comparable? How does point sampling, where vegetation was removed, perform in relation to transect sampling? Is the proportion of finds different? Is there a different efficiency of point sampling in low or high find density areas? After this, we will perform a qualitative comparison, with the following key questions: does point sampling only add quantitatively (with more small undiagnostic finds) or also add new types of diagnostic finds? Does an increase in sample richness cause a shift in sample evenness? Then, we will turn to the issue of sample unit resolution, with the following key question: which resolution of point sampling (5, 10, or 20 m distance) is most effective to establish site density and record relevant information? We will then elucidate the potential of point sampling for archaeological interpretation, with attention to the key question: does point sampling lead to different archaeological interpretations, i.e. the spatial, chronological, and functional definition of the material assemblages? We will then finish with an assessment of cost estimation, with the key question: what is the resources investment difference between transect sampling and point sampling?

4.5 MATCHING POINT SAMPLES WITH TRANSECT SURVEY SAMPLES

To establish the reliability of point sampling, we need to assess to what degree point sampling results in a representative view of the material distributions on the site. Much work has been done on the comparison of different spatial sampling techniques, mainly concerning their relative efficiency from the perspective of statistical theory (e.g. Mueller 1974, Plog 1976, Plog 1978, Nance 1983). Ideally, an assessment of relative efficiency would involve comparing both sample types with a total collection sample with a 100% coverage or with a known population, such as in a seeding experiment (e.g. Ammerman 1985, Boismier 1997, Schon 2002). However, such practical field tests are often of limited scope. Therefore, we approach the analysis by performing an empirical comparison of both archaeological field survey methods as they are currently practiced in the Mediterranean.

Compared to transect survey, point sampling offers the advantage of close examination at an increased spatial resolution and, if combined with surface cleaning, neutralizing most negative surface visibility effects. If we substitute transect samples with point samples, however, it means that we sacrifice coverage, where transect sampling may be less sensitive to stochastic spatial variability. This stochastic variation must be understood as an error margin on the probability of getting a representative sample when placing a sample unit in a surface distribution of assumed density and spatial structure. Due to the decrease of total coverage on a site when point sampling, such risk increases. Whereas there is a large body of literature discussing the potential of estimating its statistical effects in prospective archaeology (Hodder and Orton 1976, Krakker et al. 1983, Nance and Ball 1986, Nance and Ball 1989, Kintigh 1988, Lightfoot 1989, Shott 1989, Orton 2000, Banning 2002, Verhagen and Tol 2004, Verhagen et al. 2012, Verhagen 2013), it is beyond the scope of this paper to deal with this explicitly. Nevertheless, the potential impact will be taken into account.

4.6 QUANTITATIVE COMPARISONS

Essential for quantitative comparisons in artifact studies is to establish the enumerator that is least affected by biases. Breakage (fragmentation of sherds) is a good example of such a bias. An extensive analysis (Waaagen 2021) shows that the variability in fragmentation within and between ceramic ware groups at Colle S. Martino has a strong effect on establishing abundance using simple frequencies. Although consistency in breakage patterns is suggested as a mitigation to this problem (cf. Molina Vidal 1997, Mateo Corredor

Code	Ceramic ware type
ARS	african red slip wares
BG	black gloss wares
CW	coarse wares
DO	dolium
GL	glazed wares
IBM	imbrex
IMP	impasto wares
IMPCW	impastoid coarse wares
ITS	italic terra sigillata
PMJ	protomaiolica
PW	plainwares
RPW	red painted ware (Late-Roman)
RTW	roman thin-walled wares
TIL	roof tiles
TS	terra sigillata (unidentified)

Table 4.1. Ceramic ware types used for the analysis in this paper.

and Molina Vidal 2016), samples from point sampling are likely to consist of generally smaller pieces of pottery, which creates a systematic bias in comparisons between the various sample types, even if general breakage would be uniform. Using only feature sherds is impossible because point samples and transect samples catch a very limited number of diagnostic sherds in comparison to excavation. Therefore, we use weight and weight densities alongside numeric densities. This means, however, that we should specify the ceramic ware types, because these have a different average weight. The ceramic ware types in table 4.1 are used in this analysis (for further specification, see Pelgrom and Stek 2010).

4.6.1 GENERAL CORRELATIONS VS. VISIBILITY AND MINIMUM COVERAGE

To establish whether point sampling yields a representative sample of surface distributions, it should show a degree of similarity in its results to transect sampling. We first look at the factor difference (*FD*) between point sampling and transect sampling, i.e. the ratio between total quantities of ceramics collected for both techniques expressed in weight densities (Caraher et al. 2014). The mean *FD* between transect samples and point samples is 1:22.8, meaning we find on average 22.8 times more finds per m² in point sampling as opposed to transect sampling. This is the result of closer inspection of the ground and the removal of vegetation in point sampling. In fact, if we compare with corrected transect sampling densities (weighted on relative detectability; see Waagen 2014: 422), we find a mean *FD* of 1:2.3. In other words, the difference in sherd density as collected between transect sampling and point sampling is between 2.3 and 22.8, depending on the validity of visibility-weighting (cf. Schon 2000, Schon 2002, and further below)

The *FD* for just coarse ware (*CW*) weight densities (uncorrected) is almost 1:75. Because these *FD*s are all higher for number than for weight density, we confirm that point sampling samples a smaller subset of the archaeological surface record than transect sampling does and that the difference in density between the two methods is considerable. To understand to what degree the two sampling techniques may still result in comparable trends, we use statistical regression. To avoid biases caused by ware-specific weight or by sample size (wares with too few items), we exclusively use coarse- and plain ware (*PW*) weight densities. The general comparison between *CW* and *PW* in both sample techniques shows rather weak correlations (respectively, $r^2 = .17$, $p = .000$ and $r^2 = .14$, $p = .000$) (fig. 4.3). This correlation is, however, stronger when

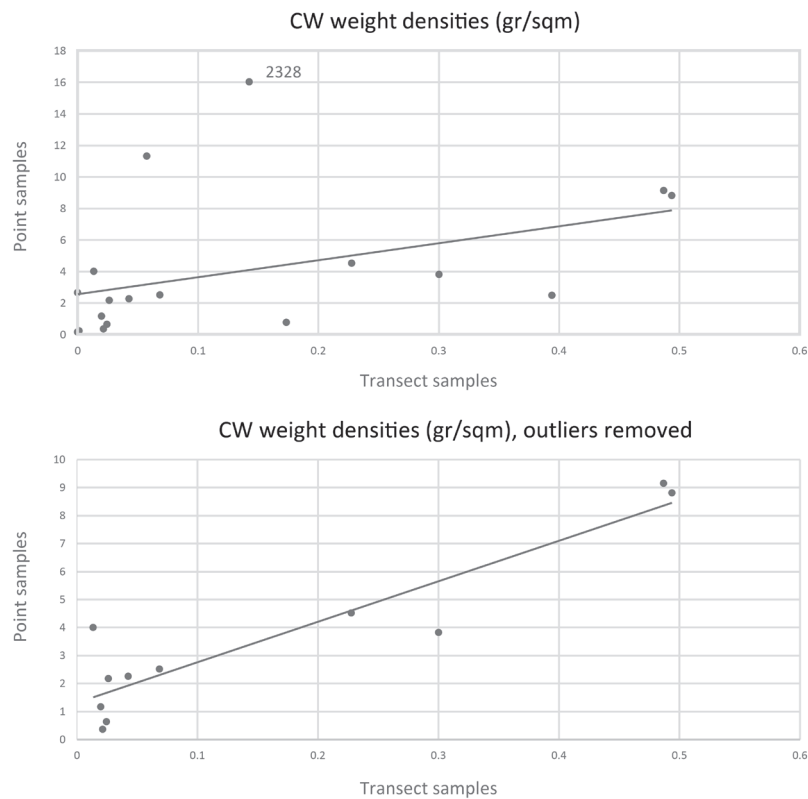


Figure 4.3. Linear regression between CW point sample and TS weight densities, before and after removal of outliers (fallow fields and transect samples covered by few point samples).

we factor in two of the point sampling method characteristics: 1) point sampling's power to detect archaeological finds in fallow, overgrown terrain, as opposed to transect survey and 2) the need for a minimum coverage by point samples in a certain terrain in order to detect the underlying archaeology.

As to 1), the strong potential of point sampling in countering low-visibility is illustrated with transect sample unit 2328 (for location of the transect samples see fig. 4.2A). Among the data points in both graphs, transect sample unit 2328 strongly overperforms, i.e. the point sampling resulted in more finds than predicted or explained by the regression line. This aligns with our field observations, because this field was directly next to the main site area, yet it was uncultivated and had not been ploughed for at least a couple of years. Leaving all transect samples with such conditions ('fallow' in our database) out of the equation, the correlations grow from weak to moderately strong (respectively, $r^2 = .62$, $p = .000$ and $r^2 = .65$, $p = .000$). A similar situation is documented by transect sample units 2318 and 2342 (contingent fallow fields), yielding CW weight densities of .15 and 2.65 g/m² for point sampling and 0 g/m² (for both) for transect sampling. Translated to sherd number densities, this amounts to .16 sherds/m² and .85 CW sherds/m² in fields where the standard transect sampling collection did not produce any finds. We conclude that point sampling is especially effective in overgrown, fallow terrain. Moreover, this did not happen the other way around: there were no transect samples that yielded archaeological finds during standard transect sampling, yet none by point sampling.

As to 2), there is a minimum coverage of a certain terrain with point samples for producing results that can be compared to transect sample collection. In the graphs, the cases in which point sampling underperforms are all explained by a very partial coverage with a relatively low number of point samples over the transect samples (e.g. in 2306, 2323, 2325, and 2337). Leaving such fields with low point sample coverage out of the equation, the correlation becomes stronger (respectively, $r^2 = .85$, $p = .000$ and $r^2 = .71$, $p = .000$; see fig. 4.3). Again, this aligns with our hypothesis that there is a general correlation between the two techniques, in that where point sampling coverage is low, the lumping of material in transect samples and/or the vulnerability of point sampling can cause deviation from the overall density.

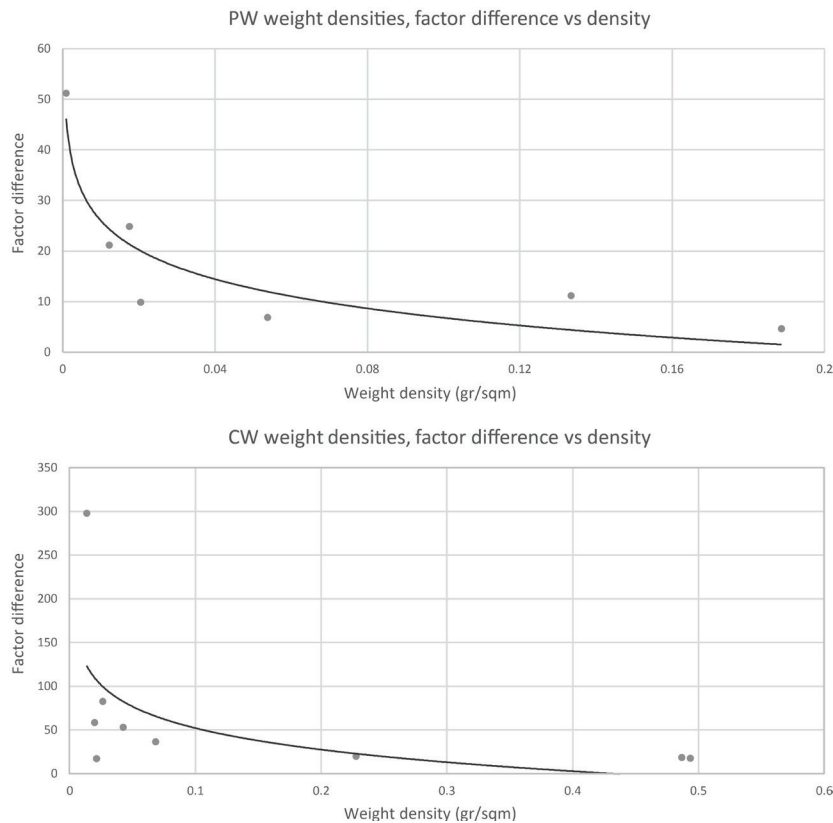


Figure 4.4. Logarithmic regressions between CW and PW FD and transect sampling weight densities.

We infer from this that at a certain coverage, point sampling accurately assesses the distributions of find densities and effectively combats the low-visibility problem. A too-low coverage of point samples, however, makes the samples vulnerable to variable spatial distributions and stochasticity.

4.6.2 RESIDUAL DIVERGENCES BETWEEN POINT SAMPLING AND TRANSECT SAMPLING

We can now turn to the remaining samples (i.e. with good point sample coverage and non-fallow) and discuss the residual divergences between point sampling and transect sampling. Differences in overall density is often raised as an important factor (cf. Wandsnider and Camilli 1992, Schon 2002). It is argued that, in general, *FDs* between point sampling and transect sampling are, *ceteris paribus*, larger in low densities than in high densities as a result of a greater effect of spatial stochasticity in low densities and a much more selective collection process. The latter is based on a rather impressionistic observation that low densities will have people kneeling down less, possibly paying even less attention, and, therefore, result in a less consistent scrutiny of the surface distribution. Indeed, we can establish this relationship for the *FD* by plotting point sampling weight densities against the transect sampling weight densities. The relation is best exposed as a logarithmic regression (fig. 4) for CW and PW, respectively, $r^2 = .535$, $p = .09$ and $r^2 = .359$, $p = .06$, although the effect is considerably stronger for PW than for CW. *FDs* grow increasingly bigger as the densities get lower, so the chance of collecting a sample through point sampling that shows much higher densities of material increases exponentially under low-visibility conditions (cf. Schon 2002).

Other factors that could explain the residual differences between point sampling and transect sampling are recorded visibility factors, soil humidity, light/shadow conditions, vegetation coverage, and presence of stones and recent material, as well as the final visibility indexation of the fields. It is elsewhere argued that the linear scale on which these are usually recorded is actually inaccurate as a method for bias correction (e.g. Schon 2000: 108–109; 2002, Caraher et al. 2014: 53) and thus unfit for simple explanation of density

variability. Indeed, the residuals tested against those factors did not result in any significant correlations: in other words, none of the visibility factors explain residual variation. Given that the influence of tillage conditions is instead very clear, we conclude that we should be careful to apply corrections using simple, linear scales based on these visibility factors.

We must note that by progressively excluding transect sample units from our batch, we reached a sample size of, respectively, $n=7$ and $n=9$ for both the last regression and the residual analysis, which for principles of robustness would not meet formal statistical procedure criteria. Notwithstanding this caveat, we argue that the data tentatively fall in line with the general theory of spatial sampling, as well as with our set of hypotheses and our impressions on artifact collection in low densities.

4.7 QUALITATIVE DIFFERENCES

It is important to understand whether point sampling actually leads to a qualitatively superior insight into the properties of the surface distribution. Although it has been argued that a more intensive investigation will increase representation for all classes equally (e.g. Wandsnider and Camilli 1992: 177, Schon 2002: 136), sample theory suggests that increased sample size leads to an increase of detection probability of more rare items, i.e. increasing sample richness. Due to the detection of smaller items, proportions can also be expected to change, which is known as sample evenness. It is important to evaluate such potential improvements in our data in the light of resource expenditure. The more intensive collection of materials carries the risk, and burden, of collecting large amounts of redundant data (cf. Caraher et al. 2014). Below, the qualitative properties of the point sampling pottery batches will be assessed in terms of qualitative properties for both proportional differences in wares and feature sherds, as well as periodization possibly affected by detection of more pieces of rarer pottery.

4.7.1 EVENNESS

As for the sample evenness, the main question is whether collection in transect samples is selective with regard to the larger fragments and if this therefore means that point sampling leads to the proportional increase of more fragmented fine ware categories. For comparing the proportions between wares, we leave out the very heavy material; it makes up 90% in the total sample weight density and thus skews any comparison. Although clearly there is a general increase in densities in both weight and number in the more fragmented wares, they are dwarfed by changes in CW and PW; CW increases its share from 51% to 64%, and the proportion of PW decreases from 24% to 11%. However small the effect for more fragmented fine wares, the actual increase in absolute numbers and weight density shows a notable effect. For the compared samples, in a total of 51,453 m² of transect sampling, we collected 19 BG sherds, while for a total of 773 m² of point sampling, we collected 12 BG sherds, where the point sampled surface area was 1/67th of the total area the transect survey investigated. In fact, the *FD* of the CW weight densities is 1:34, where the *FD* of the BG weight densities is 1:39.

It follows that under moderate effects, the shift with the biggest quantitative impact is in CW and PW densities. Although it is clear that the PW weight and number densities are also higher in the point samples, the increase in CW is so strong that the relative share of PW decreases. Although the increased presence of CW may be a more realistic reflection of the amount of kitchen and storage wares expected on a site such as this, it also increases the collection of redundant data. In the case of individual point samples, the shift in proportions can be assessed in more detail. In case of, e.g., transect sample unit 2320 (fig. 4.5), the spatially more precise mapping of proportions per point sample shows that the higher density area recognized as 'site' is now dominated by CW, where the lower density areas in the transect sampling (offsite) shows a more prominent presence of PW. This gives us a much crisper image of actual distribution and potential causes for this spatial configuration; is the high CW presence related to an area of storage and/or cooking and the area with more PWs to an area of consumption? This actually showcases the advantage for interpretation of the locational precision of the point sampling, countering the smearing effect in this field.

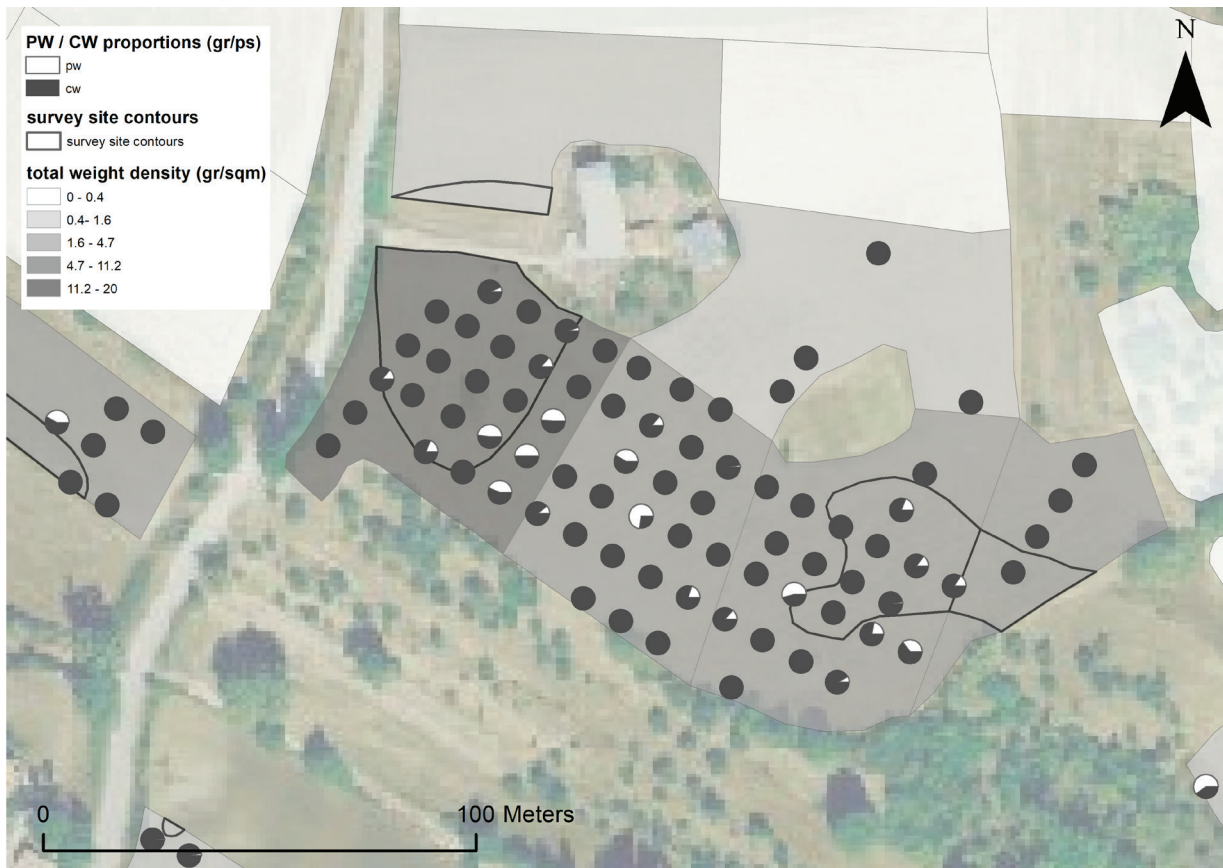


Figure 4.5. PW/CW proportions per point sample in weight density.

4.7.2 RICHNESS MEASURED IN FEATURE SHERDS

As for sample richness, for point sampling, we can expect that we find a higher variability of materials, i.e. more rare items, in the point sample, specifically for the fragmented fine wares. We can assess this by looking at feature sherds and less common ware classes. As for feature sherds, the question is whether it is reasonable to suggest that transect sampling will be more effective in collecting feature sherds, as they are generally heavier (larger) and have a more distinctive shape, making them easier to recognize during walking. Indeed, the difference between the mean weight of feature and non-feature sherds of 5.25 g is statistically significant ($t = -10.42, p = .000$).

The transect method overall retrieves more feature sherds: an average of 7.68 CW feature sherd per transect sample unit against 4.25 CW feature sherd per point sample aggregation. However, this is due to a larger coverage of transect sampling. Feature sherd density shows that point samples have markedly higher densities, with an average of 0.07 CW feature sherd per m^2 for point sampling against 0.003 CW feature sherd per m^2 for transect sampling: effectively 23 times denser. This is very likely an effect of full collection, i.e. a larger sample size. The proportional difference between feature and non-feature shows a decided increase in both weight and number of feature sherds for the point samples (fig. 4.6), with a *FD* of, respectively, 1:1.78 and 1:1.57 for the CW, which points towards a proportional increase in smaller feature sherds. For PW, the *FDs* are, respectively, 1:2.78 and 1:2.80, indicating an increase in similar sized feature sherds but a much larger increase than for the CW sherds. The most dramatic increase is in PW feature weight densities, as those increase from less than 25% to more than 50% (!), likely reflecting small rim and base fragments. The difference between PW and CW is explained by the generally smaller sizes of PW: \bar{x} sherd weight PW in transect sampling is 6 g and in point sampling is 4.4 g, and \bar{x} sherd weight for CW in transect sampling is 8.5 g and in point sampling is 5.1 g. The smaller sized material collected through point sampling evidently results in a dramatic increase in small PW feature sherds, whereas the effect on CW is smaller, probably because they are relatively more easily traced through transect sampling.

CW/PW feature/non-feature sherd proportions

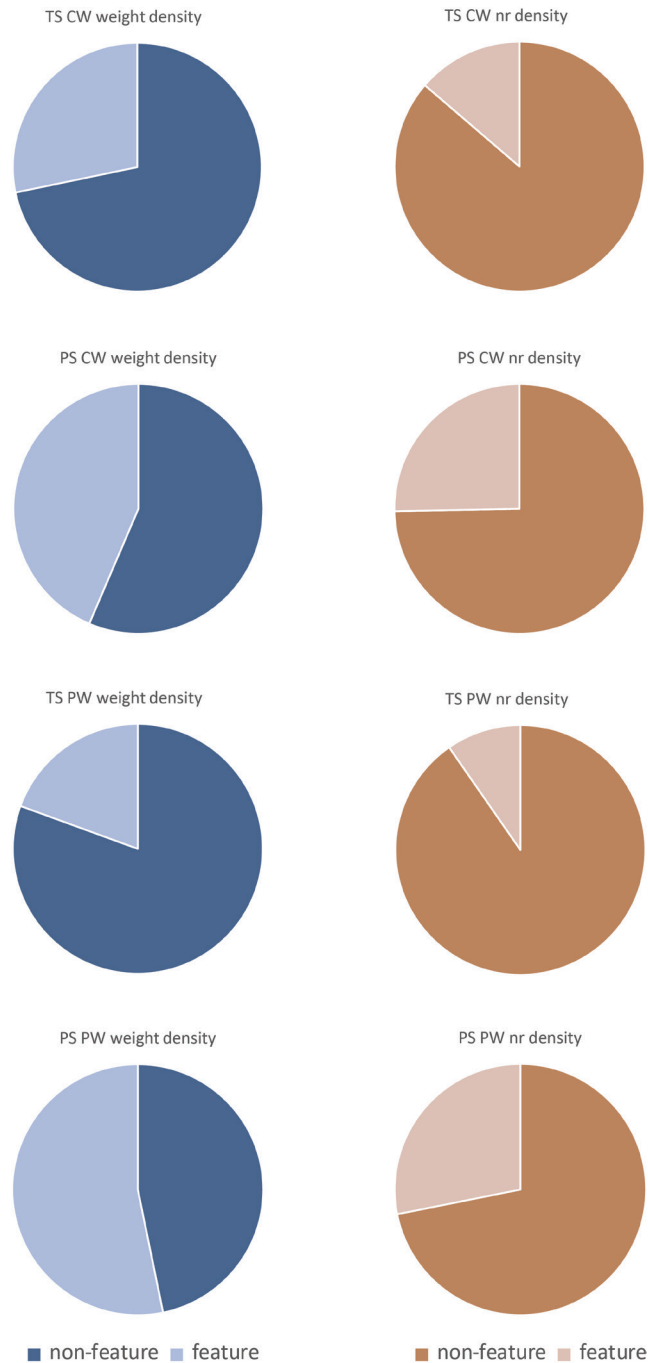


Figure 4.6. Proportions of feature and non-feature sherds: blue are weight densities and brown are numeric densities.

Therefore, we can affirm that point sampling does indeed lead to a more effective collection of smaller feature sherds that are otherwise missed in transect sampling. PW feature sherds from point sampling can significantly contribute to interpreting the site. The identification of the vessel shapes, i.e. table ware, light cooking gear, etc., in combination with the spatial precision of their collection, offers good potential to give us insight into the various types of human activity in the different zones of the site. Additionally, any datable material will help specify the chronological resolution of those activities. A preliminary finds analysis revealed that some of the plain ware feature sherds are probably very worn BG sherds. The likeliness of this association is strengthened because BG material is only found when there is PW in the sample, so they

<i>Transect unit</i>	<i>Note</i>	<i>PS contain</i>	<i>TS contain</i>
2318	Fallow field	CW	-
2328	Fallow field	BG, TS, TIL	-
2342	Fallow field	BG, CW, GL, IMB	-
2343	Fallow field	GL	-
2344	Fallow field	-	-
2306	Few PS (<6)	-	GL, ITS, PW
2323	Few PS (<6)	-	BG, IMB, PW, TIL
2325	Few PS (<6)	-	GL, IMB, PW
2333	Few PS (<6)	-	GL, IMB, IMPCW
2307	-	-	BG
2308	-	IMB	-
2310	-	BG, IMB, TIL	GL
2311	-	-	PW, TIL
2320	-	RPW	PMJ, GL
2321	-	ITS	BG
2322	-	IMP	ARS, IMPCW, RTW
2334	-	BG, GL	DO, TIL
2337	-	-	DO, IMB, PW, TIL

Table 4.2. Comparison of ware categories retrieved through either transect samples (TS) or point sample assemblages (PS) per transect. Given are differences; rest of the ware profile is similar (for abbreviations, see tab. 1).

are strongly spatially correlated (i.e. PW in 11 of 18 point sample aggregations, of which 9 also contained BG; there are no BG fragments in point samples without PW). If it is any indication, for all BG sherds collected through point sampling on the site, 72 sherds and 29 diagnostic shapes were identified, which resulted in 9 shape dates, which is ca. 1 in 3 (R.A.A. Kalkers, personal communication 2019). If this is the upper threshold for the identification of datable PW sherds, it could still add substantial information to our understanding of functional and chronological characteristics of the site.

4.7.3 RICHNESS MEASURED IN RARE WARE CATEGORIES

As for the less common ware classes, we can directly compare the assemblages as collected through the two sampling techniques (tab. 4.2). We can clearly differentiate the samples on the same principles as the quantitative comparisons above.

If we look at the fallow fields, clearly the point sampling uncovered wares that were not found in the transect sampling, whereas as if we look at the fields with few point samples (<6), transect sampling outperforms the point sampling. Assessing the remaining transect sample units, the comparison shows considerable variability.

The pattern confirms the effect of visibility and spatial stochasticity/coverage on both transect sampling or point sampling. For areas with reasonable visibility (again: not fallow) and a relatively large covering of point samples, the return in wares is variable, predominantly concerning the rarer ware classes (CW is not affected). Rare ware categories such as BG in point sampling consist of overall smaller pieces than in transect sampling (e.g. the BG sherds collected in the point samples of 2310 and 2334 measured 1, 1, 3, and 4 g, whereas those collected in transect sample units 2307 and 2321 measured 4, 4, 7, and 8 g), so, indeed, the point sampling targets the smaller materials much more effectively. Although point sampling is not necessarily better when it comes to retrieving rare ware categories, it surely collects a particularly useful, complementary dataset of items that are very likely missed in transect survey. For example, in 2342, two BG sherds were recovered, including one 4th to mid-3rd century BC skyphos fragment, and in 2328, 2310, and 2334, the point samples produced BG sherds where no BG was found in the transect sampling.

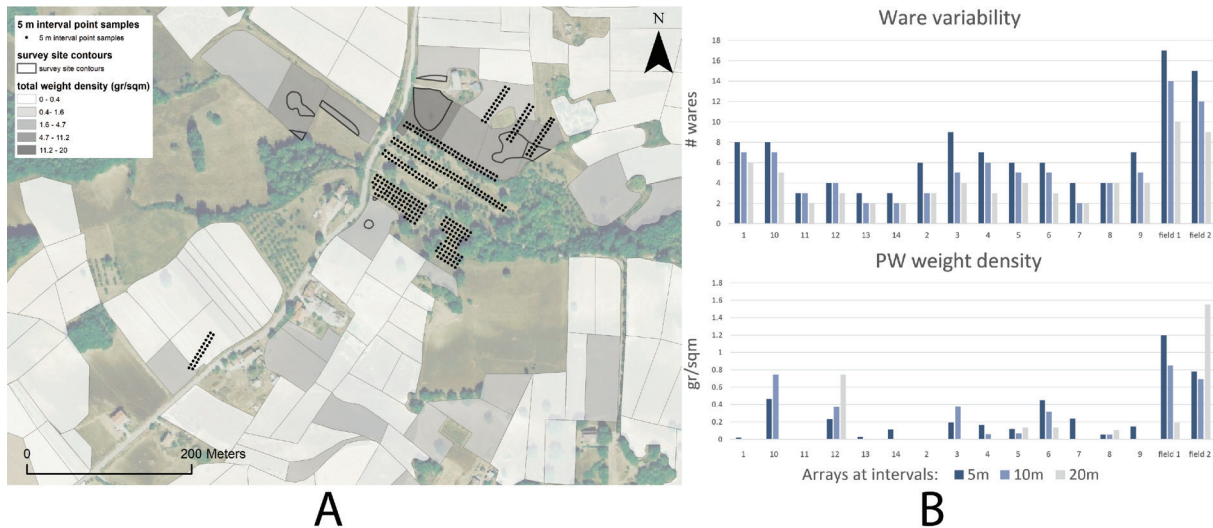


Figure 4.7. A) Point sampling at 5 m interval, arrays and fields; B) PW weight density per 5, 10, and 20 m selection and ware variability per 5, 10, and 20 m selection.

These were only generically datable but still add evidence of the spatial extent of human activity on the site in these periods. In fields where BG material was retrieved from both transect sample units and point samples, the point sampling assemblage added a narrower chronological frame, i.e. point samples in 2320 produced a skyphos fragment dating to 290–225 BC, while two BG sherds from the transect collection could only be generally dated to the 3rd century BC. In another example, in transect sample unit 2322, containing another scatter boundary, the point samples produced one BG sherd, outside of the scatter boundary, of a rare late 6th to early 4th centuries BC date.

4.8 POINT SAMPLING AT VARIOUS INTERVALS

For the methodological analysis for using point sampling in survey designs, optimizing the intensity of the sample grid is of paramount importance. The analysis we perform is not aimed at detectability of specific archaeological phenomena (cf. Banning 2002: 28), but at understanding the potential for retrieving a statistically representative sample of a continuous surface material distribution. We assess point sampling at different spatial intervals, namely 5, 10, and 20 m, to allow comparison between those intervals in terms of efficiency. Based on sample theory and our findings so far, the 20 m point samples are likely to result in too small a sample to adequately represent the surface distributions, whereas for the 5 m samples, we have theoretically higher probabilities to retrieve rare ware categories. In order to test these hypotheses, we assess richness and PW densities of the samples. Since the 5 m point samples were limited to two fields that have not been transect surveyed and several single rows (arrays), it is not very useful to compare the various spatial intervals to the transect sampling assemblages. Therefore, we opted for a comparison between the various point sampling intervals for a selection of arrays and the two aforementioned fields (fig. 4.7A).

Assessing weight density by creating subselections in the fields and arrays of 5, 10, and 20 m point sampling, we see that, overall, the 5 and 10 m point samples stay considerably closer together than either does to the 20 m point samples (fig. 4.7B). The linear correlation between the 5 and 10 m point samples ($r^2 = .74, p = .000$) is considerably stronger than the correlation between 5 and 20 m ($r^2 = .21, p = .07$) and between 10 and 20 m ($r^2 = .25, p = .05$). The deviations between the weight densities of the various intervals are positive as well as negative: again, the result of the sample size effect and spatial stochasticity. The increase in sample richness with sample size is shown by plotting the number of ware categories for every interval (fig. 4.7B). The increase appears to be rather linear, suggesting a proportional growth.

In conclusion, it is clear that the point samples at the 20 m interval show considerable deviation in relation to the point samples at the 5 and 10 m interval and are, at least in the case of tracking material densities on a large rural site, less adequate. The 5 m point samples are effective at improving probabilities of re-

trieving rare materials, with the obvious drawbacks in increased time investment and increasing numbers of abundant material, likely CW.

4.9 POINT SAMPLING IN SPATIAL SITE ANALYSIS

The extensive analysis of point samples, quantitatively as well as qualitatively, results in a characterization of their unique and/or complementary qualities in relation to transect sample units. With at least six point samples per transect sample unit at a maximum distance of 10 m, point samples provide very useful find information. The point sampling technique is effective in counteracting visibility problems due to vegetation and/or lack of tillage and results in much more accurate samples in those fields. The point samples effectively collect the more fragmented part of the surface material distributions, which results in the proportionally increased presence of smaller fine wares, specifically plain ware feature sherds, which appears to be correlated with the presence of BG and may be worn BG vessel parts. Point sampling also appears particularly effective in low densities; the increase in weight densities is stronger for low-density areas where transect sampling underrepresents offsite densities. Finally, the locational precision of point sampling has been demonstrated to potentially mitigate the smearing effect of transect sampling.

The advantage of higher locational precision can be further bolstered by a commonly held perception that larger objects travel further along the length of the ploughing direction than small objects (Lewarch and O'Brien 1981). Since we have shown point sampling to collect the more fragmented and smaller surface materials, such an effect could have quite an impact on the added value of localized point sampling. Although the evidence for this size effect based on field experiments is neither overwhelming nor uncontested (Odell and Cowan 1987, Yorston et al. 1990), it must be noted that these dealt with collections of stone artifacts and a different type of terrain and never surpassed (to the authors' knowledge) a duration of six years (Yorston et al. 1990). Computer simulations have shown considerable effects over longer periods (Yorston et al. 1990, Boismier 1997). Based on the latter, it seems reasonable to suggest that point sampling targets material that is less likely to be displaced once present on the surface. In so far as the location of that material can be considered meaningful, e.g. in the case of a site, point sampling would provide a more reliable representation of the spatial configuration of the site.

Based on the technical analysis of point sampling, the question is whether point sampling leads to different archaeological interpretations, i.e. the spatial, chronological, and functional definition of the material assemblages. The potential of point sampling is here demonstrated by three examples: site extents, offsite interpretation, and intra-/extra-site functional zones and chronology.

4.9.1 SITE EXTENTS

In order to assess the site extents, i.e., defining the area of sustained human activity related to the site, using point sampling in terms of conventional site density thresholds, it must be recognized that the weight densities are rather abstract representations of quantity. They do not relate to what field walkers usually perceive as site density, namely >5 sherds/m² (which is, of course, only a general indicator; cf. Waagen 2014). In order to express relative weight densities in a metric that is closer to dispersion of numerical quantities, we propose a sherd weight density equivalent (*SWDE*): the weight density divided by the mean sherd weight (after removal of building material, which is, of course, also a site indicator) to estimate the number of sherds the weight density would, on average, represent. Caution is warranted when comparing the *SWDE* to the 5 sherds/m² threshold which is used for site identification in the field. Since point sampling sheds quite a different light on overall densities of material, which turn out to be substantially higher than those perceived through transect sampling, any threshold will have to be re-evaluated in the context of point sampling research designs. In this paper, we will forego any attempts to align these, but we will use the *SWDE* as a heuristic for critical contextual assessment.

Overall, by mapping the point sampling *SWDE*, a tentative definition of the site extent can be projected (fig. 4.8). Through the detailed and visibility-mitigating looking glass of the point samples, there are some differences and/or refinements to the original conception (also fig. 4.9–4.12, for location of the transect

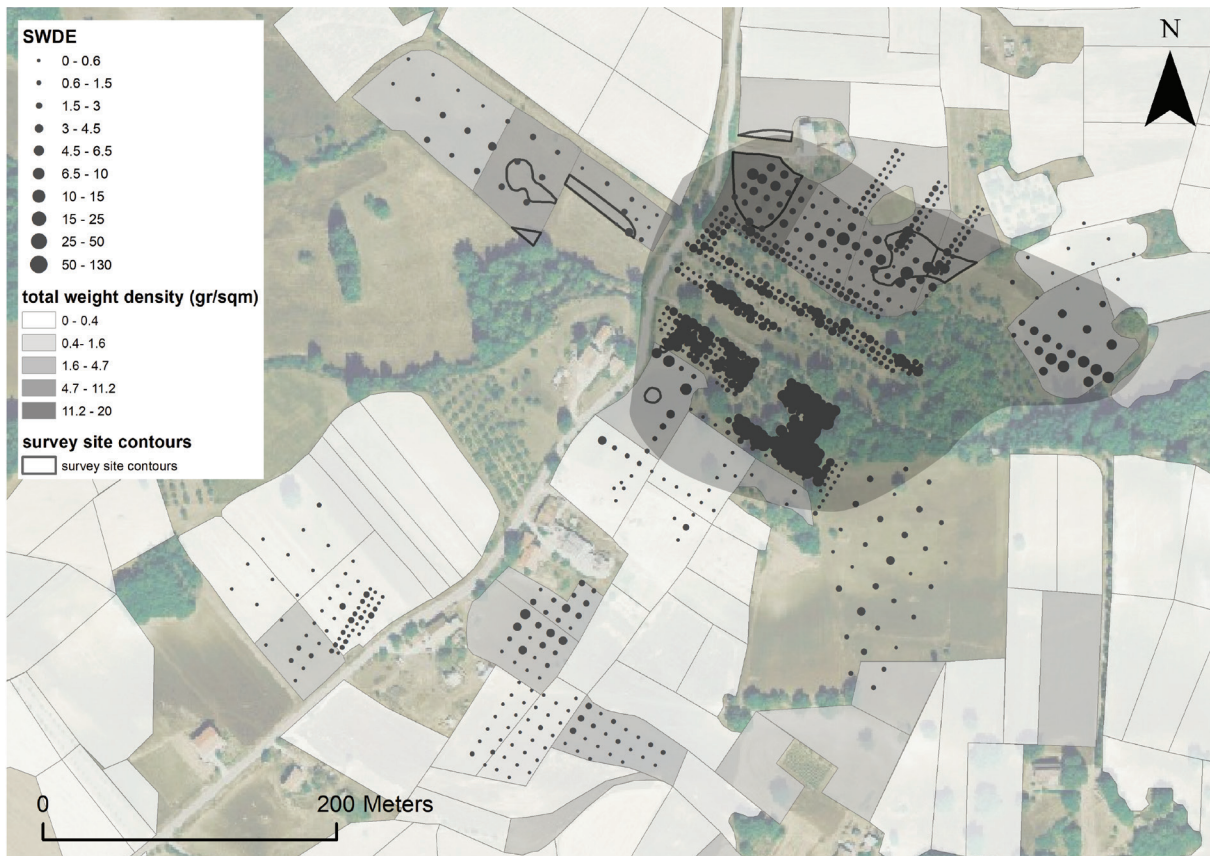


Figure 4.8. Approximate site boundaries based on the sherd weight density equivalent resulting in a total area of 5.85 ha.



Figure 4.9. Transect sampling weight densities, point sampling A) CW and B) PW diagnostics.

samples see fig. 4.2A). Transect sample unit 2328 is very likely part of the main site area; it reaches *SWDE* site density in seven of the point samples and approaches regular transect sampling site density in others, it shows a higher ware variability in the point samples, and it features relatively frequent diagnostic PW and CW. Transect sample unit 2321 shows evidence of continuity between 2320 and 2322, based on the *SWDE* and CW diagnostics. Although intersected with few point samples, transect sample unit 2343 appears to continue *SWDE* site density, rendering the identified nucleus in the center of that transect sample unit part of a continuous distribution. Again, CW diagnostics and weight densities corroborate these patterns.



Figure 4.10. Transect sampling weight densities and point sampling ware variability.



Figure 4.11. Transect sampling weight densities, point sampling A) CW and B) PW weight densities.

To the northwest (transect sample units 2160 and 2161) and southeast (no transect sampling), notwithstanding the sparse distribution of point samples, the samples do attest *SWDE* hovering just under site density. Their material compositions are primarily made up of building materials and CW, with some PW. The site nuclei drawn based on transect sampling are quite different based on the point sampling *SWDE*, with the exception of 2333, where BG comes from the point samples. Based on the spatial dispersion of the densities and ware variability (see fig. 4.10), the sharp contrast between high and low densities turns into a more gradual transition, as expected. Indeed, with mapping the point sampling *SWDE*, we recognize a gradual decrease in find densities and can put a finer point on the overall pattern, whereas in transect sampling, the site definition is much more binary.



Figure 4.12. Transect sampling weight densities and point sampling tile:pot weight density ratio.

4.9.2 OFFSITE INTERPRETATION

The point sampling allows a more detailed scrutiny of the offsite areas. This is an important issue in archaeological field survey, as, often, less dense and more dispersed find patterning eludes easy interpretation (Waagen 2014). Here, by example, it is shown how point sampling allows for a more detailed analysis of offsite find patterning in the direct surroundings of Colle S. Martino.

Two pairs of contingent samples, transect sample units 2334 and 2337, as well as 2307 and 2308, are conspicuous. These samples are situated in close proximity to the modern farms near the road, which is often a reason to be cautious about increased CW weight densities. These could be affected by recent farm refuse, which is difficult to differentiate from ancient material. On the other hand, such a general interpretation should not defer a critical assessment of individual cases. Here, in both areas, point sampling recovered BG and PW, as well as CW diagnostics, which may link them to the main site as activity zones. Transect sample unit 2334 is of particular interest: it recovered low weight density and no BG, but the point samples show an increase from 1.5 g/m^2 to 35 g/m^2 , corresponding to less than $.05 \text{ sherds/m}^2$ SWDE based on transect sampling to more than 1.5 sherds/m^2 based on point sampling of quite fragmented pottery. The SDWE of one of the point samples located centrally in the concentration goes up to 2.5 sherds/m^2 , the equivalent of ca. $7.5 \text{ sherds/point sample}$. Point sampling produced a substantial amount of CW, with relatively large pieces (feature sherds) with high variation in breakage among them, which contrasts with adjacent field 2337. The BG material from the point samples consists of extremely fragmented pieces, and two small PW sherds have been found, likely representing worn BG. The data from the point sampling give it the appearance of a site from which material has been dispersed over the field due to ploughing and other post-depositional processes, a clear example of the smearing effect masking a pattern only visible at an increased resolution.

4.9.3 INTRA- AND EXTRA-SITE FUNCTIONAL ZONES AND CHRONOLOGY

For understanding sites, especially large rural sites, it is crucial to understand which function sites, or parts of sites, represent in which chronological period. The examples below showcase how the detailed information from point sampling supports a finer analysis of these aspects.

BG is very important for chronological anchoring of the site. Point sampling recovered BG finds where they were not found during regular transect survey, namely in four transect sample units (2328, 2342, 2334, and 2310). In two of those, other artifactual clues for human activity are absent, and the finds may be related to offerings from destroyed burials. In the other two cases, they add to our understanding of the extent of the site and related areas in the Hellenistic period. In the site's center, there is a huge density of BG, CW, and building material. *Italic terra sigillata*, on the other hand, is found in the adjacent field to the west (to the north of 2343), which could indicate a shift of focus of the site over time. PW weight densities are higher in the west of 2320 and in 2322 (see fig. 4.11), and the finds are heavier and more variably sized in 2322, as well as spatially correlated with concentrations of PW feature sherds. It is likely that these variations reflect a difference in the buried archaeological remains, with the western parts of 2320 and 2322 potentially relating to domestic structures. Dolium sherds appear almost exclusively in the center of the site, though two pieces have been found within the drawn nuclei boundaries in transect samples 2320 and 2322. Lastly, glazed pottery is found over the main site area, as well, and provides further chronological details.

The tile:pot ratio can be helpful in differentiating between built-up areas of a site and zones in which outside activities took place. The ratio projected based on point sampling shows a diffuse patterning (see fig. 4.12), but the scatter in 2334 has a clear relative abundance of pottery over tile: BM is present, but very fragmented. In the cluster on the other side of the road, on the other hand, 2307 and 2308, BM dominates. Whereas taphonomic processes can create such differences, the difference may be related to a different function of the area in the past. In both the nuclei in 2320 and 2322, BM dominates, whereas in 2321, pottery is more abundant. This clearly contributed to the establishment of the site boundaries in the field, but the area in between may be of a different functional profile more connected to 2322 that also features a more similar PW profile. The total dominance of BM in the southeastern nucleus of the central site must surely be indicative of a considerable building at that location.

4.10 COST ESTIMATION

Evidently, point sampling clearly has valuable qualities; however, for applying a point sampling research design instead of, or as a complement to, other types of sampling, the increase in information should be weighed against effort investment. Estimating relative efficiencies of different sample types in survey has been demonstrated to be extremely context dependent (Mueller 1974, Plog 1976, Plog 1978). Effort investment has regrettably not been recorded for the point sampling fieldwork at Colle S. Martino. A reasonable proxy is the cost effort estimation based on point sampling efforts at the large rural site called TAP03 and the large hilltop settlement on the Montagna di Gildone (MdG), both in the same research area. TAP03 represents a Roman rural settlement that is located in a small valley across a number of level agricultural fields. MdG is a very different kind of site in terms of chronology, location, and visibility conditions. It boasts a large prehistoric–Iron Age site with impressive fortification walls, as well as Classical–Hellenistic/early Roman period agglomerations visible as high-density scatters in often heavily vegetated areas. With some caution, these numbers averaged can be taken as a general indication of effort spent (tab. 4.3). The cost estimations for both include setting up a dGPS for localization, setting out the point samples, and cleaning and collecting the samples, as well as the full field documentation, but exclude breaks.

For TAP03, this amounted to 2.32 point samples per person hour, arriving at an estimated cost of .43 person hours per point sample. For Montagna di Gildone, we sampled .8 point samples per person hour, arriving at an estimated cost of 1.24 person hours per point sample (see tab. 4.3). Averaging these scores, we would arrive at a mean .83 person hours per point sample. Compared with the standard intensive transect sample collection method, which was measured over the course of three days, we surveyed .37 ha per person hour, arriving at an estimated cost of 2.71 person hours per ha. This includes navigation, adminis-

<i>TVS area</i>	<i>PS/hour</i>	<i>Hour/PS</i>	<i>HA/hour</i>	<i>Hour/HA</i>
<i>TAP03, little cleaning</i>	2.32	.43	.37	2.71
<i>MdG, heavy cleaning</i>	.8	1.24	.37	2.71

Table 4.3. Comparative cost investment. Hours are person hours; transect samples (TS) are measured in ha/hour; point samples (PS) are measured by their count/hour.

tration, and the occasional short displacements by car, but excludes breaks and was aggregated over several team leaders. Obviously, for the transect sampling, to compare the effort investment to point sampling, the person hours need to be multiplied by the number of members in a single team: in our case, five people. If we project these numbers onto the compared samples of 18 transect samples (roughly 4.5 ha) and 246 point samples (at 10 m intervals), we arrive at 61 person hours for the transect sampling and 204 person hours for the point sampling, so roughly a 3.3 times larger investment. Due to the heavy cleaning required at MdG, this is probably quite a conservative estimation, and it is thus advised to take the lower threshold of .43 person hours per point sample at TAP03 into account for future calculations.

4.11 CONCLUSIONS

In this paper, we have discussed the compatibility of point sampling and standard transect survey in Mediterranean landscape archaeology. Naturally, many of the issues discussed depend on the specific type of archaeology under study, and in different landscapes, with different types and quantities of finds, different decisions should be made. Still, our case study in the Tappino Valley in Molise presents landscape conditions that are quite common for late prehistoric to historical archaeology in inland Mediterranean areas. Our results from Molise show that the method of point sampling can be meaningfully compared with more standard field survey methods using transects and larger collection units. This means that intensive point sampling can be adequately integrated within the research design of landscape archaeological projects. We observed a clear correlation in density values and variability of artifacts between transect and point sampling methods. The latter has the additional benefit of offering a higher resolution image of fluctuating densities both off- and onsite.

The added value of point sampling methods is particularly clear through its ability to shed light on low-visibility portions of the landscape that often remain out of view in standard survey methods and thus can lead to the detection and better understanding of otherwise poorly detectable activity zones ('sites'). Detection of large activity zones may also be feasible even with relatively large distances between point samples, as we also see a rough correlation in the 20 m distanced samples.

Besides detecting archaeology where it is difficult to trace with standard transect survey, point sampling can add to our functional and chronological understanding of such activity zones, under reasonable time and energy investments. In particular, the systematic inclusion of patches with (very) low-visibility within the activity zones allows a better definition of the extension and character of the activity zones. This effect was especially clear in fallow portions of the landscape. Moreover, the precise spatial information combined with the intensive collection offered by point samples allow a detailed spatial analysis of find distributions within activity zones, which can also be useful when surface finds are to be compared with remote sensing data or excavation.

In areas with good visibility and large quantities of finds, we believe that point sampling could be usefully complemented with diagnostic sampling within collection units that are larger than the point samples and smaller than transect/field units, for instance using the center of four point samples to establish blocks of 5 × 5 or 10 × 10 m. An option to mitigate redundancy is to establish criteria for subselections, i.e., collect CW only every three point samples instead of during all of them.

We conclude that point sampling can help to refine the crude dichotomy between off- and onsite distinctions on the one hand and conceptions of the archaeological landscape as a continuum on the other. In this sense, we hope that our approach can help to develop a more holistic view of the archaeological landscape across different landscape conditions.

BIBLIOGRAPHY

- Alcock, S., Cherry, J., and Davis, J. 1994. "Intensive survey, agricultural practice and the classical landscape of Greece." In: Morris, I. (ed.) *Classical Greece: ancient histories and modern archaeologies*. Cambridge University Press, 137-170.
- Ammerman, A. 1985. "Plow-Zone Experiments in Calabria, Italy." *Journal of Field Archaeology* 12-1, 33-40.
- Attema, P., Bintliff, J., Van Leusen, P.M., Bes, P., de Haas, T., Donev, D., Jongman, W., Kaptijn, E., Mayoral, V., Menchelli, S., Pasquinucci, M., Rosen, S., García Sánchez, J., Luis Gutierrez Soler, L., Stone, D., Tol, G., Vermeulen, F., and Vionis, A. 2020. "A guide to good practice in Mediterranean surface survey projects." *Journal of Greek Archaeology* 5, 1-62.
- Banning, E.B. 2002. *Archaeological Survey*. Kluwer Academic Publishers.
- Barker, G. 1995. *A Mediterranean Valley: Landscape Archaeology and Annales History in the Biferno Valley*. Leicester University Press.
- Bispham, E. 2007. "The Samnites." In: Bradley, G. J., Isayev, E., and Riva, C. (eds.) *Ancient Italy. Regions Without Boundaries*, 179-223.
- Boismier, W.A., 1997. *Modelling the effects of tillage processes on artefact distribution in the ploughzone: a simulation study of tillage-induced pattern formation*. British Archaeological Reports, British Series 259, Oxford: Archaeopress.
- Burger, O., and Todd, L.C. 2006. "Grain, Extent, and Intensity: The Components of Scale in Archaeological Survey." In: Lock, G., Molyneaux, B.L. (eds.) *Confronting scale in archaeology*. Springer, Boston, 235-255.
- Caraher, W., Moore, R., Pettegrew, D., Andriotti, M., Kardulias, P., Nakassis, D., and Olson, B. 2014. *Pyla-Koutsopetria I: Archaeological Survey of an Ancient Coastal Town*. Boston, MA: The American Schools of Oriental Research.
- Chadwick, A. M., and Evans, H. 2000. "Reading Roystone's rocks: Landscape survey and lithic analysis from test pitting at Roystone grange, Ballidon, Derbyshire, and its implications for previous interpretations of the region." *Derbyshire Archaeological Journal* 120, 101- 122.
- D'Amico, V. 1953. *Jelsi e il Suo Territorio Dall'antichità Remota ad Oggi*. Campobasso.
- Dench, E. 1995. *From barbarians to new men: Greek, Roman, and modern perceptions of peoples of the central Apennines*. Oxford.
- García-Sánchez, J., and M. Cisneros. 2013. "An Offsite Approach to Late Iron Age and Roman Landscapes on the Northern Plateau, Spain." *European Journal of Archaeology* 16-2, 289-313.
- Hodder, I., and C. Orton. 1976. *Spatial Analyses in Archaeology*. Cambridge University Press.
- Jones, H. (ed.) 2004. *Samnium: Settlement and Cultural Change. The Proceedings of the Third E. Togo Salmon Conference on Roman Studies*. Providence.
- Kintigh, K. W. 1988. "The effectiveness of subsurface testing: A simulation approach." *American Antiquity* 53, 686- 707.
- Kraker, J. J., Shott, M. J. and Welch, P. D. 1983. "Design and evaluation of shovel-test sampling in regional archaeological survey." *Journal of Field Archaeology* 10-4, 469- 480.
- La Regina, A. 1989. "I Sanniti." In: Pugliese Carratelli, G. (ed.) *Italia Omnium Terrarum Parens*. Milan, 301-432. Milan.
- Lewarch, D. E., and O'Brien, M. J. 1981. "The Expanding Role of Surface Assemblages in Archaeological Research." *Advances in Archaeological Method and Theory* 4, 297-342.
- Lightfoot, K. G. 1989. "A defense of shovel-test sampling: A reply to Shott." *American Antiquity* 54, 413-416.
- Marchi, M. L., 2016. "Carta Archeologica d'Italia - Forma Italiae. Project: Research Method." In: Burgers, G. L. M., Kluiving, S.J., and Hermans, R. A. E. (eds.) *Multi-, Inter- and Transdisciplinary Research in Landscape Archaeology. Proceedings of the 3rd International Landscape Archaeology Conference LAC2014*, 1-9.
- Mateo Corredor, D. and Molina Vidal, J. 2016. "Archaeological quantification of pottery: The rims count adjusted using the modulus of rupture (MR)." *Archaeometry* 58, 333- 346.
- Millett, M. 2000. "Dating, Quantifying and Utilizing Pottery Assemblages from Surface Survey." In: Francovich, R. and Patterson, H., eds. *Extracting Meaning from Ploughsoil Assemblages. The Archaeology of Mediterranean Landscapes* 5. Oxbow Books, 53-59.
- Molina Vidal, J. 1997. *La dinámica comercial romana entre Italia e Hispania Citerior (siglos II a. C. - II d. C.)*. Universidad de Alicante, Instituto de Cultura Juan Gil-Albert, Alicante.
- Mueller, J. W. 1975. *Sampling in Archaeology*. The University of Arizona Press.
- Nance, J. D. 1983. "Regional Sampling in Archaeological Survey, The Statistical Perspective." *Advances in Archaeological Method and Theory* 6, 289-356.

- Nance, J. D., and B. F. Ball. 1986. "No Surprises? The Reliability and Validity of Test Pit Sampling." *American Antiquity* 51-3, 457-483.
- Nance, J. D., and B. F. Ball. 1989. "A Shot in the Dark: Shott's Comments on Nance and Ball." *American Antiquity* 54 -2, 405-412.
- Odell, G. H., and F. Cowan. 1987. "Estimating Tillage Effects on Artefact Distributions." *American Antiquity* 52-3, 456-484.
- Orton, C. 2000. *Sampling in Archaeology*. Cambridge University Press.
- Pelgrom, J. and Stek, T.D. 2010. "A landscape archaeological perspective on the functioning of a rural cult place in Samnium: field surveys around the sanctuary of S. Giovanni in Galdo (Molise)." *Journal of Antique Topography* XX, 41-102.
- Plog, S. 1976. "Relative Efficiencies of Sampling Techniques for Archaeological Surveys." In: Flannery, K.V. (ed.) *The Early Mesoamerican Village*, 136-158.
- Plog, S. 1978. "Sampling in Archaeological Surveys: A Critique." *American Antiquity* 43-2, 280-285.
- Salmon, E. T. 1967. *Samnium and the Samnites*. Cambridge.
- Schon, R. 2000. "On a Site and Out of Sight: Where Have Our Data Gone?" *Journal of Mediterranean Archaeology* 13, 107-111.
- Schon, R. 2002. *Seeding the Landscape: Experimental Contributions to Regional Survey Methodology*. Unpublished Ph.D. dissertation, Bryn Mawr College.
- Scopacasa, R. 2015. *Ancient Samnium: Settlement, Culture, and Identity Between History and Archaeology*. Oxford.
- Shott, M. J. 1989. "Shovel-Test Sampling in Archaeological Survey. Comments on Nance and Ball, and Lightfoot." *American Antiquity* 54-2, 396-404.
- Stek, T.D. 2009. *Cult Places and Cultural Change in Republican Italy: A Contextual Approach to Religious Aspects of Rural Society After the Roman Conquest*. Amsterdam Archaeological Studies 14.
- Stek, T.D. 2018. "Exploring non-urban society in the Mediterranean: hill-forts, villages and sanctuary sites in ancient Samnium, Italy." *Antiquity* 92, 1-7.
- Stek, T.D., Modrall, E.B., Kalkers, R.A.A., van Otterloo, R.H. and Sevink, J. 2015. "An early Roman colonial landscape in the Apennine mountains: landscape archaeological research in the territory of Aesernia (Central-Southern Italy)." *Analysis Archaeologica: An International Journal of Western Mediterranean Archaeology* 1, 229-291.
- Stek, T.D. and Pelgrom, J. 2005. "Samnite Sanctuaries Surveyed: preliminary report of the Sacred Landscape Project 2004." *BABesch* 80, 65-71.
- Stek, T. D. and Waagen, J. 2022. "Scratching the Surface: Integrating Low-Visibility Zones and Large Rural Sites in Landscape Archaeology Using Point Sampling." *Journal of Field Archaeology*, 47-4, 228-244.
- Tagliamonte, G. 1997. *I Sanniti. Caudini, Irpini, Pentri, Carricini, Frentani*. Milan, Longanesi.
- Tol, G. W. 2012. *A fragmented history. A methodological and artefactual approach to the study of ancient settlement in the territories of Satricum and Antium*. Barkhuis and Groningen University Library.
- Van de Velde, P. 2001. "An Extensive Alternative to Intensive Survey: Point Sampling in the Riu Mannu Survey Project, Sardinia". *Journal of Mediterranean Archaeology* 14-1, 24-52.
- Van Leusen, P.M., Pizziolo, G. and Sarti, L. 2011. *Hidden Landscapes of Mediterranean Europe Cultural and methodological biases in pre- and protohistoric landscape studies*. Proceedings of the international meeting Siena, Italy, May 25-27.
- Verhagen, P. 2013. "Site Discovery and Evaluation Through Minimal Interventions: Core Sampling, Test Pits and Trial Trenches." In: Corsi, C., Slapšak, B., and Vermeulen, F. (eds.) *Good Practice in Archaeological Diagnostics, Natural Science in Archaeology*, 209-225.
- Verhagen, P., Rensink, E., Bats, M., and Crombe, P. 2012. "Establishing Discovery Probabilities of Lithic Artefacts in Palaeolithic and Mesolithic Sites with Core Sampling." *Journal of Archaeological Science* 40-1, 240-247.
- Verhagen, P., and Tol, A. 2004. "Establishing Optimal Core Sampling Strategies: Theory, Simulation and Practical Implications." In: Fischer Ausserer, A., Börner, W., Goriany, M., and Karlhuber-Vöckl, L. (eds.) *Enter the Past. The E-Way Into the Four Dimensions of Cultural Heritage. CAA 2003. Computer Applications and Quantitative Methods in Archaeology*, 416-419.
- Voorrips, A., Loving, S. H., and Kamermaans, H. 1991. *The Agro Pontino Survey Project. Methods and Preliminary Results*. Studies in Prae- en Protohistorie 6. Amsterdam.
- Waagen, J. 2014. "Evaluating Background Noise: Assessing Offsite Data from Field Surveys Around the Italic Sanctuary of S. Giovanni in Galdo, Località Colle Rimontato, Molise, Italy." *Journal of Field Archaeology*

39-4, 417–429.

Waagen, J. 2021. “Breakage, Bias and the Archaeological Surface Record. Assessing the Quantification Problem in Archaeological Field Survey.” *Archaeometry* 64-2, 1–16.

Wandsnider, L. and Camilli, E. L. 1992. “The character of surface archaeological deposits and its influence on survey accuracy.” *Journal of Field Archaeology* 19, 169–188.

Yorston, R. M., Gaffney, V. L., and Reynolds, P. J. 1990. “Simulation of Artefact Movement Due to Cultivation.” *Journal of Archaeological Science* 17-1, 67–83.