

# Modernizing Spatial Micro-Refuse Analysis: New Methods for Collecting, Analyzing, and Interpreting the Spatial Patterning of Micro-Refuse from House-Floor Contexts

Isaac I. Ullah · Paul R. Duffy · E. B. Banning

Published online: 5 November 2014  
© Springer Science+Business Media New York 2014

**Abstract** Archaeologists can learn a great deal from the distribution of cultural evidence at various scales ranging from large regions, through small communities, down to individual households. Since in many societies a significant proportion of the human experience takes place within and around houses, houses play a prominent role in discussions of *habitus*. Yet archaeologists have also experienced challenges in their attempts to understand this *habitus*, especially when so many archaeological remains pertain to short-term activities that occurred near the end of a house's use life, or even after, and may not even be typical. Focusing on the tiniest debris that accumulates over long periods may help us overcome these challenges, but many archaeologists have been reluctant to employ micro- refuse analysis because of the erroneous perception that the scale of effort it involves must be astronomical. The approach we demonstrate in this paper shows that careful consideration of sampling both in the field and in the lab makes it possible to detect robust patterns from persistent activities with a fraction of the effort that some previous analysts have employed. One of our key findings is that employing large numbers of volunteer counters, in combination with adequate quality assurance protocols, greatly facilitates this type of research.

**Keywords** Micro-refuse analysis · Spatial analysis · Household archaeology · Activity areas

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s10816-014-9223-x) contains supplementary material, which is available to authorized users.

I. I. Ullah (✉)  
Center for Comparative Archaeology, University of Pittsburgh, Pittsburgh, PA, USA  
e-mail: isaaciullah@gmail.com

P. R. Duffy · E. B. Banning  
Department of Anthropology, University of Toronto, Toronto, Canada

I. I. Ullah  
School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, USA

## Introduction

Micro-refuse (also known as micro-artifact, micro-debris, and micro-remains) analysis is the study of very small anthropogenic debris recovered from archaeological contexts. Although micro-refuse may be found in many archaeological features, such as hearths or middens, micro-refuse recovered from ancient living surfaces, such as house floors, can inform on the small-scale activities of individuals and small groups—especially within households or small campsites. This kind of micro-refuse analysis focuses on understanding spatially persistent activity over the life of a household, rather than just the last use, abandonment, or post-abandonment re-purposing of a space. To accomplish these goals, micro-refuse analysis needs to be a kind of spatial analysis, with samples collected, processed, analyzed, and interpreted with spatial research goals in mind. We contend that this “spatial micro-refuse analysis,” when combined with carefully conceived sampling, alleviates much of the perceived difficulty of conducting and interpreting micro-refuse studies in domestic contexts.

Even though researchers ranging from Fladmark (1982) to Hull (1987), Rosen (1986), and Vance (1987) advocated micro-refuse analysis in activity area studies for more than two decades, most archaeologists have found it too daunting (e.g., Tani 1995:245). This is because most of the current methods for micro-refuse analysis were developed for other purposes, such as prospecting for particular micro-artifact types (e.g., charred plant remains), examining individual features (e.g., hearths, storage pits, or middens), or understanding more general diachronic patterns at a site (e.g., analysis of “column” samples). Most of the commonly used methods of micro-refuse analysis were developed by paleoethnobotanists who were looking for very rare ecofacts, and so believed that they should (1) collect *a very large volume* of sediment from features likely to preserve charred plant remains and (2) sort *all* (or nearly all) of that sediment in order to find as many of the very rare preserved plant remains as possible. While this effort may be justified in paleoethnobotanical studies, we contend that it is wholly inappropriate for spatial micro-refuse analysis and this has deterred many researchers from undertaking such studies. Specifically, the concept of counting literally tens of thousands of items in potentially dozens of different bags of sediment is far too slow and costly. For example, just a single 1 l sediment sample in the pioneering Simms and Heath (1990) analysis of micro-refuse at the Orbit Inn site contained more than 10,000 micro-bone fragments and more than 118,000 fragments of shell. Such effort does not justify the result, especially considering that each context in this study only provided a sample size of  $n=1$  (as these are cluster samples). That these early analysts sought to mitigate inter-observer error by having a single individual do all the counting only exacerbated the workload problem and ironically made it impossible to assess the effect of observer bias on the data.

In this paper, we present a set of methods specifically designed for “spatial microarchaeology.” These methods are focused on analyzing micro-refuse samples as spatial records, and we aim to optimize the time and labor associated with sorting and counting micro-refuse, while at the same time controlling for error and providing greater insight into what the data may mean. We propose a work flow that can take a researcher through all stages of analysis, from the field to final interpretations. We present our methods in four sections: (1) field collection, (2) processing and sorting procedures, (3) sampling, outlier detection, and density estimation, and (4) spatial

analysis. To exemplify our method, we will refer to a spatial micro-refuse analysis that we conducted of a house floor from the small Late Neolithic site of Tabaqat al-Bûma in Wadi Ziqlab, northern Jordan. The lower floor of area E33, which we use as our case study, likely dates to the LN4 period of the site (ca. 5350 to 5170 cal. BC at 68 % confidence), an occupation phase that probably lasted between 145 and 400 years (Banning 2007). More complete discussion of spatial micro-refuse analyses from Tabaqat al-Bûma can be found in Ullah (2009, 2012), while more information about the site and its phasing occurs in Banning et al. (2011).

## Towards a Theory of Spatial Micro-Refuse Analysis

Considering that most current archaeological projects are operating under increasingly limited budgets, knowing when and why to include spatial micro-refuse analysis in a research program is not trivial, especially when the time and money devoted to sorting micro-refuse could be used to fund any number of other sophisticated analyses. Justification for spatial analyses of micro-artifact densities relies on its ability to inform about long-term patterns in the use of domestic space in ways that larger artifacts and architectural features cannot. In fact, for some contexts like those in our case study, many of the macro-remains actually pertain to abandonment and post-abandonment use of spaces (e.g., trash disposal) rather than to persistent domestic uses.

Consider several special properties of small debris: (1) The small size of micro-refuse (typically defined as ranging between 0.5 and 5 mm (Ullah 2012)) makes them less vulnerable to some kinds of post-depositional cultural or natural disturbances so that they are more likely to be found as *de facto* refuse in archaeological contexts than are larger artifacts (Hayden and Cannon 1983; Dunnell and Stein 1989). (2) That micro-refuse is likely to be discovered close to where it was deposited means that spatial patterns in the densities of the different micro-refuse classes are directly relevant to the spatial aspects of human behaviors in ancient domestic spaces (or at least to those behaviors likely to create micro-refuse; Fladmark 1982; Rosen 1986, 1993; Sherwood et al. 1995; LaMotta and Schiffer 1997). (3) Most debris-producing activity is likely to produce significantly higher quantities of micro-refuse than it would macro-remains, particularly as fragmentation is a fractal process (Brown et al. 2005). (4) Micro-refuse should accumulate densely in locations where the behaviors that produce them most frequently occur. Thus, spatial micro-refuse actually *benefits* from the fact that most ancient habitation deposits are “time-averaged palimpsests of trash” (Barton et al. 2011: 709). Although we must be mindful that the activities that originally deposited micro-refuse are not the only contributors to micro-refuse distributions—maintenance activities, such as sweeping, and other site-formation processes can be important—we may still exploit these properties to inform about the long-term patterns—the *habitus*—of the use of domestic spaces (Hodder and Cessford 2004).

Following from these four points, we would expect different activities to produce differences in the quantity and “mix” of micro-refuse types. We can draw upon ethnoarchaeological and experimental data to build “middle-range” models that connect specific activities (e.g., butchering animal carcasses or grinding grain) with different abundance profiles of resulting micro-refuse (e.g., abundant bone chips, or abundant basalt chips; Jones 1983; Schiffer 1983, 1987; Gregg et al. 1990; LaMotta and Schiffer 1997; Kamp

2000). Small beads and bead fragments are not present in the Wadi Ziqlab example presented here but can also occur on prehistoric and historic sites, indicating areas of craft production or places where decorated clothing was stored or worn (Guan et al. 2011; Simms and Heath 1990). Interpretation of spatial patterning in the relative density of different micro-refuse classes recovered from ancient floors should then allow us to construct a model of the spatiality of regularly conducted domestic activities.

The preceding discussion of the theory of micro-refuse production, dispersal, and deposition produces tangible methodological implications for the collection and analysis of these small-sized artifacts. If we want to use micro-refuse to understand long-term processes of domestic arrangements, we need to analyze them in a manner appropriate to the particular characteristics of micro-refuse that reveals such long-term patterns. The remainder of this paper lays out our method to achieve this and some example results.

### Collecting Micro-Refuse Samples in the Field

Collection and spatial sampling methods set limits on all consequent analyses, so the research frame of a spatial micro-refuse analysis must begin in the field. Researchers must make trade-offs; they need to obtain enough samples taken at high-enough spatial resolution to produce meaningful results, but should refrain from over-sampling, and from sampling dubious (e.g., recently disturbed or poorly dated) contexts, which would waste time and money. For example, the northern portion of the G34 building at our site of Tabaqat al-Bûma had been disturbed by recent road building, and so we chose not to sample that portion of the floor. This also meant that we could sample the intact portions of this relatively large house floor (the house was approximately 4 m by 6.5 m) at a spatial resolution of 50 cm<sup>2</sup> without unduly increasing the workload associated with analysis.

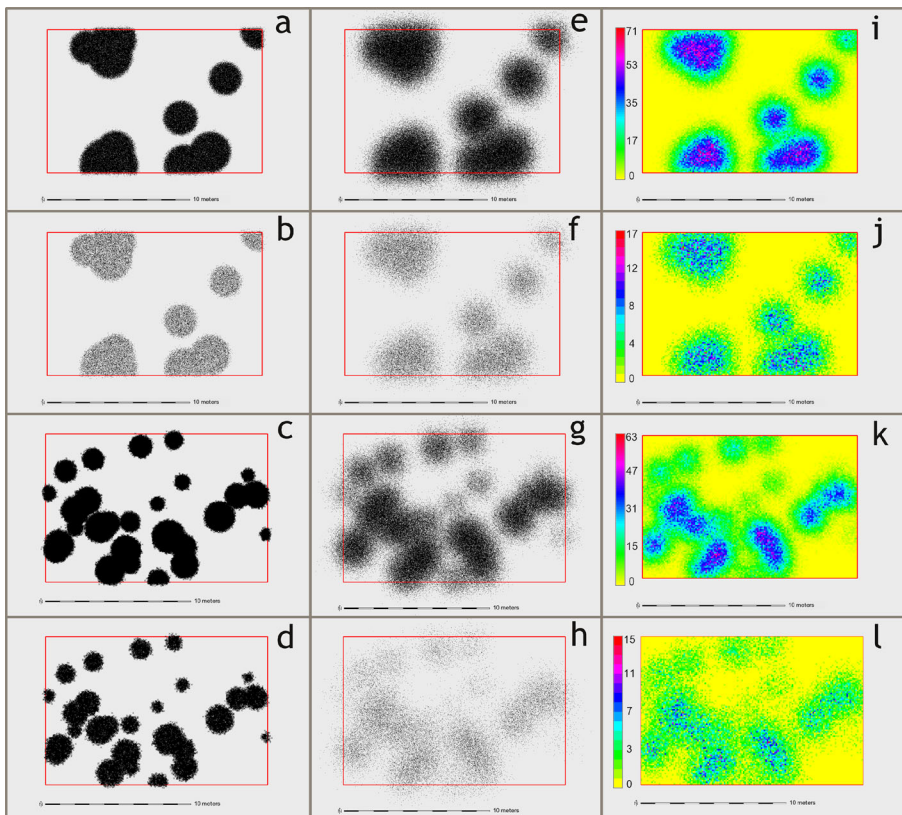
Researchers also need to make practical decisions about what thickness of sediment belongs to a “surface” in the context of their site. As an example of this issue, in our case study in northern Jordan, we were often confronted by buildings that had likely been occupied for long periods (100 or more years (Banning 2007)). The mud and/or loosely plastered floor surfaces that we discovered thus may have been resurfaced many times, but it was not possible for us to confirm or deny this in the field. Therefore, we constrained our samples to a thickness of about 6–10 mm in order to minimize vertical mixing of layers.

The optimum sampling strategy for each site will vary, but it should streamline sample collection and processing, and minimize costs and error. The goals of efficient micro-refuse sampling are as follows: (1) to capture accurately the spatial patterning of micro-refuse from an archaeological surface at a scale appropriate to subsequent analyses, (2) to do so using the smallest number of sample locations, (3) to collect the smallest volume of sediment that will still yield useful results, (4) to provide accurate results for both sparse and abundant artifact types, and (5) to be easy for archaeologists to implement in the field. In this section, we report the results of some simulations of sampling efficiency and accuracy that may guide strategies for collecting micro-refuse samples.

### GIS Simulation of Archaeological Deposits

We used the free and open-source GRASS GIS software suite (GRASS Development Team 2014) to conduct our micro-refuse simulations. We first simulated two possible

arrangements of “activity areas” in a 10 m by 15 m rectilinear domestic floor by randomly distributing different numbers of circular “activity areas” across the floor. We use random distributions of activity areas in our simulations because cultural processes, while not random, are variable enough that there is no *a priori* reason for positioning activity areas in particular places within a structure. The first scenario simulates a small number of large activity areas (Fig. 1a, b; 10 circles with radii of 1–2 m) and the other a larger number of smaller areas (Fig. 1c, d; 30 circles with radii of 0.25–1 m). In our simulation, we “seeded” micro-refuse in these areas at either a high density of  $1/\text{cm}^2$  (Fig. 1a, c) or a lower density of  $0.1/\text{cm}^2$  (Fig. 1b, d). The actual locations of each individual piece of micro-refuse were determined by a random point generator so that point northings and eastings followed a normal distribution. This ensures that artifact deposition is denser in the center of activity areas than at their edges, which conforms to ethnographic data on artifact deposition within activity areas (Kent 1984; Binford 1978). We simulated site-formation processes by “perturbing” the coordinates to distances determined by another random number generator (following a normal distribution of  $\mu=0$  cm and  $\sigma=50$  cm). The results are plausible models of a variety of possible micro-refuse distributions within the two different arrangements of simulated activity areas (Fig. 1e–h). Finally, we created a raster density map for each



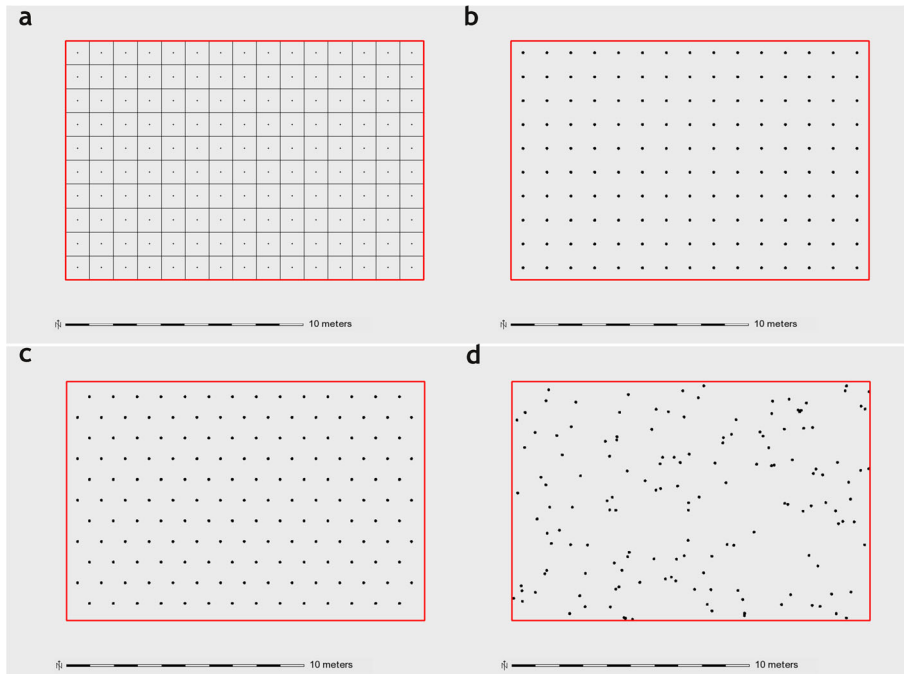
**Fig. 1** The four simulated micro-refuse distributions. **a–d** Initial random assignment of activity area locations and sizes and initial “deposition” of micro-refuse at low and high density. **e–h** Simulation of site formation process by random “perturbation” of micro-refuse points. **i–l** 10 cm resolution kernel density maps of the final micro-refuse density patterns for each simulation

simulation using a kernel density estimator at a 10 cm resolution (Fig. 1j–m) to use as a “control” in the sampling experiments discussed in the next section. We have encapsulated our simulation method in a custom module for GRASS GIS that we call “*r\_floorsim*,” which can be downloaded from the GRASS GIS add-ons repository (see the [supplemental material](#) for the URL).

*Spatial Sampling Frame Experiment Protocol*

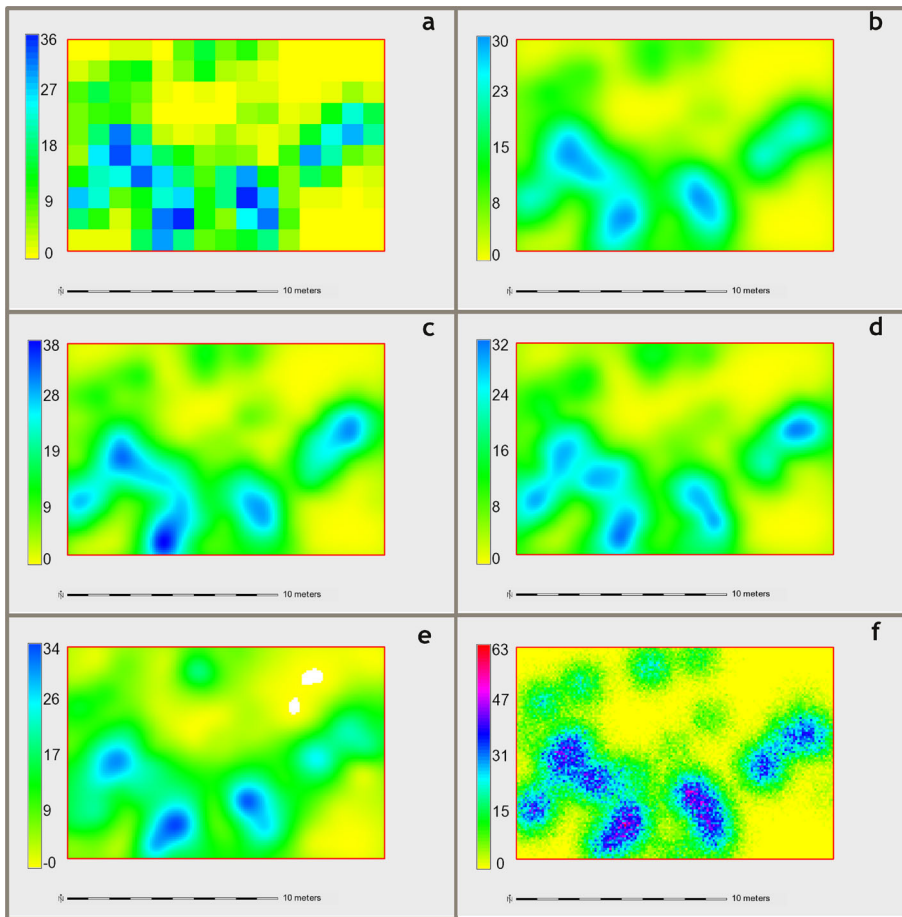
In the field, archaeologists are frequently limited to those collection strategies that are practical to implement with the tools commonly at their disposal, such as carpenter’s tapes, string, and nails. With these constraints in mind, we generated four simple spatial sampling frames from which to draw density estimates for each of the four simulated micro-refuse deposits. The first sampling frame is a standard contiguous square grid of 1 m by 1 m collection cells from which all “sediment” for a sample element would be collected (Fig. 2a). The three remaining sampling strategies employ arrangements of smaller, non-contiguous circular sampling cells of 5 cm radius (i.e., “pinch” samples) from which all “sediment” for a sample element would be collected. The first of these is a 1-m-square lattice of sampling points (Fig. 2b), the second is a 1-m-triangular lattice of sampling points (Fig. 2c), and the final one is composed of 150 sample points randomly located across the surface (Fig. 2d).

Using each of these sampling frames, we first extracted counts of “micro-refuse” for each sample element from each of the simulated micro-refuse distributions. We then



**Fig. 2** Four potential spatial sampling frames for collecting micro-refuse. **a** Cellular grid-based method (central dot shows location of interpolation points), **b** rectangular grid of “pinch” sampling points, **c** triangular grid of “pinch” sampling points, and **d** a random distribution of “pinch” sampling points

converted the counts into density by dividing them by the area of each sample element (e.g., multiplying by  $10,000 \text{ cm}^2$  for the gridded square sampling elements, and by  $78.54 \text{ cm}^2$  for the 5 cm circular “pinch” sampling elements). Finally, we used these data to interpolate raster density maps for each simulated micro-refuse distribution at 10 cm resolution (see [Interpolation of Density Maps](#) for a discussion of interpolation techniques). For the grid-based approach, we created both a “coarse” raster density map at the resolution of the sampling grid (1 m) and an interpolated 10 cm resolution density map, created by assigning the density estimate for the entire cell to a point in the center of that cell (Fig. 2a) from which to run the interpolation. In this way, we achieved five different maps of estimated density for each of the four simulated activity area distributions, resulting in four experiments with 20 maps in total (Fig. 3a–e).



**Fig. 3** The interpolated density surfaces from each of the different sampling frames for the high-density simulation of many small activity areas. **a** “coarse” 1 m resolution grid-based sampling, **b** “fine” 10 cm resolution interpolated grid-based sampling, **c** 10 cm resolution interpolated rectilinear point-based “pinch” sampling, **d** 10 cm resolution interpolated triangular point-based “pinch” sampling, **e** 10 cm resolution interpolated random point-based “pinch” sampling, and **f** 10 cm kernel density “control” map

*Simulation Experiment Results*

We compared the density values in the interpolated maps with the kernel-density “control” maps by calculating the correlation coefficient between them, and also to 3×3 cell and 5×5 cell mean-smoothed versions of each control model (Table 1). The smoothed control models remove the influence of small-scale variations in density, and thus, provide an understanding of how well the interpolated maps correspond to larger-scale trends. Each of the 20 density maps correlated reasonably well with the actual density of micro-refuse in the simulated deposits (the minimum correlation was 0.687 for triangular “pinch” sampling for the low-density simulation with many activity areas, which is still significant at  $p < 0.1$ ), although every collection method produced better correlations in the high-density experiment than in the low-density one. Interpolating from the cellular grid count provided the most highly correlated result in all experiments, likely because the larger “sampling radius” of the 1 m grid cell was less susceptible to small perturbations in micro-refuse density (although this may also mean that this method potentially under-estimates very high densities due to spatial averaging). The triangular or rectilinear “pinch” sampling frames provide the next most accurate estimates for all but the low-density experiment of many small activity areas. The coarse-resolution 1 m grid data was typically little better than a random sampling pattern, and both usually provided the worst density correlations of the tested sampling frames, except in the low-density simulation of few large activity areas.

Interestingly, all of the experimental maps had higher correlation with the smoothed control datasets than with the original 10 cm kernel density control map. In all but the high-density experiment of a few large activity areas, furthermore, correlation *increased* with the size of the smoothing neighborhood. This suggests that all of the sampling and interpolation procedures we tested here better predict broad-scale trends

**Table 1** Table of correlation statistics showing the correlation between the density surfaces interpolated from the different sampling frames with the kernel density map for each simulated micro-refuse distribution

	Control	Control 3×3 smoothed	Control 5×5 smoothed	Control	Control 3×3 smoothed	Control 5×5 smoothed
	10 large areas			30 small areas		
High density simulations						
Grid cells	0.924938	0.947849	0.9372	0.892586	0.927697	0.932699
Interpolated grid	0.955905	0.979535	0.966663	0.919693	0.955951	0.962663
Rectilinear points	0.942822	0.966167	0.95916	0.907355	0.943007	0.949496
Triangular points	0.944964	0.968307	0.956301	0.913275	0.948922	0.955027
Random points	0.93062	0.953614	0.937951	0.884719	0.919556	0.92605
Low density simulations						
Grid cells	0.821658	0.921419	0.932433	0.718998	0.884836	0.904214
Interpolated grid	0.852584	0.954296	0.966663	0.747251	0.918137	0.940607
Rectilinear points	0.826407	0.924877	0.936748	0.700751	0.859377	0.879895
Triangular points	0.828716	0.927589	0.938313	0.687665	0.843687	0.863098
Random points	0.820205	0.917944	0.929395	0.703816	0.864561	0.884801



in the density data than they do patterning at scales smaller than the original collection grid. Our simulations suggest, however, that very small-scale details are more likely the result of post-depositional site-formation processes than of long-term human activities that initially produced the deposit. For example, our technique would accurately delineate the flintknapping corner of the house (a region bigger than the resolution of the collection grid), but would only hint at small pockets of flakes introduced to other parts of the floor by bioturbation (e.g., animal burrows smaller than the collection grid). As very small-scale spatial variations are not likely to provide insight into long-term trends of domestic behavior, a collection scheme focused on small-scale patterning may actually obscure the larger patterns that are more relevant to persistent household practice. Interpolation essentially removes fine-scale “noise,” thereby, providing a clearer picture of the general trend of micro-refuse densities at the scales that are more relevant to repeated human behavior, and analogous to the “activity areas” of Susan Kent’s (1984) ethnoarchaeological studies.

### Field Sampling and Collection Methods

The results of the brief simulation experiments reported in [GIS Simulation of Archaeological Deposits](#) provide empirical evidence that can better guide the choice of a spatial micro-refuse sampling frame for a research plan and knowledge of current field conditions, an expectation of the possible micro-refuse densities at the site, and the budget for field and laboratory work. If time and budget allow, the traditional grid-based methods of spatial micro-refuse sample collection can yield very accurate estimates of the spatial patterning of micro-refuse densities if combined with an interpolation procedure, but will take the most time to collect and process. Grid-based collection schemes should not be used without interpolation, however, as the coarse, uninterpolated results of a grid-based sampling strategy will not justify the added expense. Although point-based “pinch” samples cannot capture smaller-scale patterning, they do capture the general trends well, and may also provide better estimates of density in the densest parts of the distribution. “Pinch” samples may therefore be preferable if minimization of total collected sediment volume is desirable. For example, in the broad horizontal exposure style of excavation at our Tabaqat al-Bûma case study, we conducted grid-based sampling in most buildings (with sample elements of about 1–2 l each), but, at more recent, smaller-scale excavations at other Late Neolithic sites in the region, we have opted for lattice-based “pinch” sampling frames (with sample elements of about 100–200 ml each), which are more practical to deploy in smaller excavation units that obliquely bisect portions of living surfaces (Fig. 4).

Another aspect of field collection is the choice of surfaces to sample. In this paper, our case study is limited to variation across floors inside buildings. However, it can also be useful to study micro-refuse in exterior surfaces and it is useful to take samples from comparable deposits off-site, especially in contexts where the cultural identification of some classes of material is not clear. For example, perhaps, it is difficult to be sure that basalt chips result from the use of grinding stones (in our context, however, basalt does not occur naturally in Wadi Ziqlab’s drainage). Regardless of field collection methods, our simulations show that our ability to produce reliable data is closely related to the density of micro-refuse in the ancient living surface and the spatial configuration and size of the activity areas that may be present. Any interpretation of the spatial patterning



**Fig. 4** The larger exposures at Tabaqat al Būma allowed easy implementation of grid-based sampling frames aligned to the orientation of the house-floors (a). Smaller excavation units at sites like al-Basafin (Gibbs et al. 2006) exposed only portions of living surfaces, which were more efficiently sampled with lattice-based “pinch” sampling frames (b)

of micro-refuse should take this into account, and we will return to the issue of micro-refuse data in the “[Sampling Procedure and Optimum Sampling Volume](#)” section.

### Separating, Identifying, and Counting Micro-Refuse

After collecting samples in the field, the next phase of spatial micro-refuse analysis is to estimate the density of different micro-refuse types in those samples. To do so, one must first find the micro-refuse, which means it must be separated from the sediment matrix.

#### Mechanical Separation of Micro-Refuse

Although simple wet screening (or even dry sieving in depositional contexts such as sandy or silty layers) through two or more nested screens can be employed to remove the “fines” (silt and clay-sized particles) and large clasts (larger gravels, pebbles, and cobbles) of sediment to be sampled for micro-refuse, most micro-refuse analysts rely on flotation to separate delicate charcoal and botanicals from the heavier clasts. Screening is simpler, very inexpensive, and commonly used (e.g., in CRM, where a sample sorted by 1/8” screen is often required in important contexts), but we advocate the use of flotation in spatial micro-refuse studies whenever it is feasible. Although typically more time-consuming, inexpensive methods of flotation do exist (e.g., “bucket” flotation (Fairbairn 2005) or “hand pump” flotation (Shelton and White 2010)), and flotation offers significant advantages for later analysis because it results in cleaner clasts and micro-refuse that is easier to sort and identify, especially if surfactants or other safe reagents (such as a 5 % solution of distilled white vinegar or other mild organic acid) were added to the flotation liquid.

Whether floated or not, we advocate separation of multiple size classes of micro-refuse (and associated sediment clasts). Although considerable debate exists about the size range for micro-refuse, most researchers have focused on artifacts that are smaller

than 5 mm but larger than 0.5 mm, typically analyzing all artifacts in their size-definition as a single sample. Because there are considerable data on the effects of fluvial, colluvial, and aeolian transport process on sedimentary grains (and therefore, artifacts) of differing sizes (e.g., Sahu 1964; Visher 1969; Tucker 1980; McLaren and Bowles 1985; Pye and Tsoar 2009; Bertran et al. 2012), we strongly suggest separation of floated micro-refuse by nested geological sieves with mesh sizes aligned to standard sedimentary clast size divisions for sands and small gravels (either the “Udden-Wentworth” scale used in North America (Wentworth 1922), or the ISO 14688-1 scale used internationally (ISO-BSEN 2002; Table 2). It is clear that formation processes differentially affect artifacts of different sizes (Shott 2010), but barring additional ethnoarchaeological or experimental justification, our initial experience from Tabaqat al-Bûma suggests that the “Very Coarse Sand” and “Very Fine Gravels” of the Udden-Wentworth scale (i.e., clasts between 1 and 4 mm ) are reasonable “default” size fractions with which to conduct initial analyses, and the results reported in the “[Sampling Procedure and Optimum Sampling Volume](#)” section focus on the 1.4–2.0 mm size fraction. These size classes are easily viewed under low magnification with binocular microscopes and are sufficiently large to resist traction or transport by moderately powered natural processes (e.g., normal winds, rain splash, etc.) while still small enough to evade removal during the use and maintenance of floors.

### Sorting Procedure, Training, and Sorting Ergonomics

Previous attempts at micro-refuse analysis have often attempted to escape inter-observer variation by having a single analyst count all the micro-refuse in all the sample elements. This does indeed mitigate inter-observer error but causes serious bottlenecks in analysis and, ironically, makes evaluation of observer error difficult, if not impossible. We therefore endorse the use of multiple “amateur” student volunteers to sort the samples, which can be done in the laboratory well after excavations have been completed (see the “[Sampling Theory, Outlier Detection, and Density Estimation](#)” section for justification). It is vital to have a standardized training program for these volunteers, however, so as to minimize inter-observer biases. We suggest this be done in a group session to minimize differences in training, and that a standardized reference collection with 10–30 examples of each type of micro-refuse from the site is always

**Table 2** Udden-Wentworth (North American) and ISO 14688-1 (International) clast-size divisions relevant to micro-refuse studies (Wentworth 1922; ISO-BSEN, 2002)

Udden-Wentworth scale		International scale (ISO 14688-1)	
Clast size	Size fraction name	Clast size	Size fraction name
8–16 mm	Medium gravel	6.3–20 mm	Medium gravel
4–8 mm	Fine gravel	2.0–6.3 mm	Fine gravel
2–4 mm	Very fine gravel	0.63–2.0 mm	Coarse sand
1–2 mm	Very coarse sand	0.2–0.63 mm	Medium sand
1/2–1 mm	Coarse sand		
1/4–1/2 mm	Medium sand		

available. Training should include detailed description of the distinguishing features of each type of micro-refuse, such as the sharp edges and ventral surface of a micro-flake, the different colors of flint to expect, the apparent “sponginess” and inclusions in tiny chips of pottery, the cellular structure of a charcoal fragment, the sheen of archaeological shell, or the fibrous aspect of coprolite fragments. Our recommended training protocol also requires volunteer analysts to re-count the first few samples that they sorted under the supervision of a more experienced sorter, discarding the data that they recorded while they were still inexperienced at identifying micro-refuse.

Each volunteer analyst should maintain a log of each sediment context as it was sampled, noting the size fraction and the date of counting (see the [supplemental materials](#)). This ensures that each counter will examine a subsample from each context at least once. As discussed in the “[Sampling Procedure and Optimum Sampling Volume](#)” section, sampling should be done with replacement so analysts should be prevented from working on the same context at the same time. Samples should be drawn from the larger bag after mixing its contents gently, using a graduated cylinder to determine the correct sampling volume (3 ml in our case; the “[Sampling Procedure and Optimum Sampling Volume](#)” section). Samples should be sorted into small piles by category on a Petrie dish under binocular optical microscopes (Fig. 5). After the initial sorting, analysts should re-examine the piles and ensure identification of micro-materials by checking against the reference collection or requesting supervisory confirmation. Counts of each micro-refuse type should then be entered on a central form, along with the date and the analyst (see the [supplemental materials](#)).

The ergonomics of sorting should be ensured by adjusting the height of microscopes and lab furniture for a variety of body sizes, reducing fatigue error. Since this is eye-straining work under any conditions, we suggest that analysts only work for sessions of about one hour. We find that analysts can count one or two subsamples during this time and thus the subsamples for a particular floor and size fraction can be sorted in just a



**Fig. 5** An example of a sorted micro-refuse sample element from our Tabaqat al-Būma case study

few weeks of part-time work. Finally, we also recommend setting up the microscope near a window so that counters can periodically look out and change their focus, which also relieves eye strain (Fig. 6).

### Sampling Theory, Outlier Detection, and Density Estimation

Sampling theory has long demonstrated that it is much better to have a large number of observations from which we can measure central tendency (e.g., mean or median) and its associated error, than a single observation, no matter how extensive or careful that single observation might be. In other words, it is always advantageous to have a large  $n$  in order to have a small standard error on an estimate of the mean. Consequently, it follows that a solution to the workload dilemma is to employ a larger number of observers (e.g., student volunteers) and to subdivide the work. We must be confident, however, that differences in apparent micro-refuse abundance or density (abundance adjusted for unit volume) from context to context are “real” and not simply due to inter-observer error. The first solution to mitigate inter-observer error is to ensure that sediment from every spatial sampling location is examined by every observer. Our analyst training program (see the “[Sorting Procedure, Training, and Sorting Ergonomics](#)” and the “[Determining the Optimum Sample Size \(Number of Sample Elements\)](#)” sections) and supervisory assistance also help but, even after training, some observers will inevitably be incapable of accurately identifying particular classes of material, and will consistently over- or under-report counts for those classes. Typically, over-counting is the greater problem; there is theoretically no upper limit to the possible counts, but the lowest count any observer can have is zero (a Poisson distribution). Our method for dealing with these consistently poor counters is statistical, but varies depending on the density of a particular micro-refuse type. In any case, we attach a unique observer ID tag and date stamp to each count so that we can track trends in observer differences (see the [Supplemental Materials](#) for templates of the laboratory forms that we use to accomplish this).

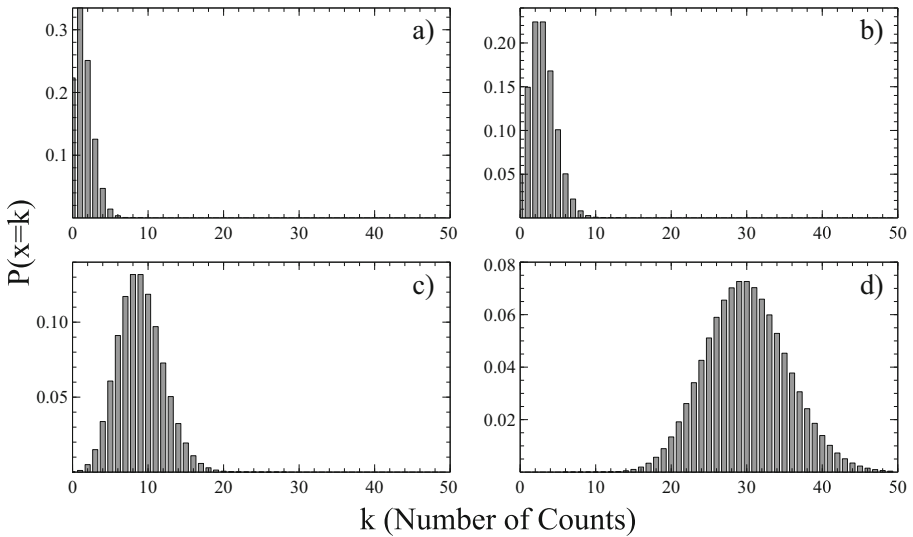


**Fig. 6** Volunteer student analysts sorting micro-refuse in our laboratory on the University of Toronto campus. Note the abundant natural light, variety of seating/working positions, and the height of the lab tables

## Sampling Procedure and Optimum Sampling Volume

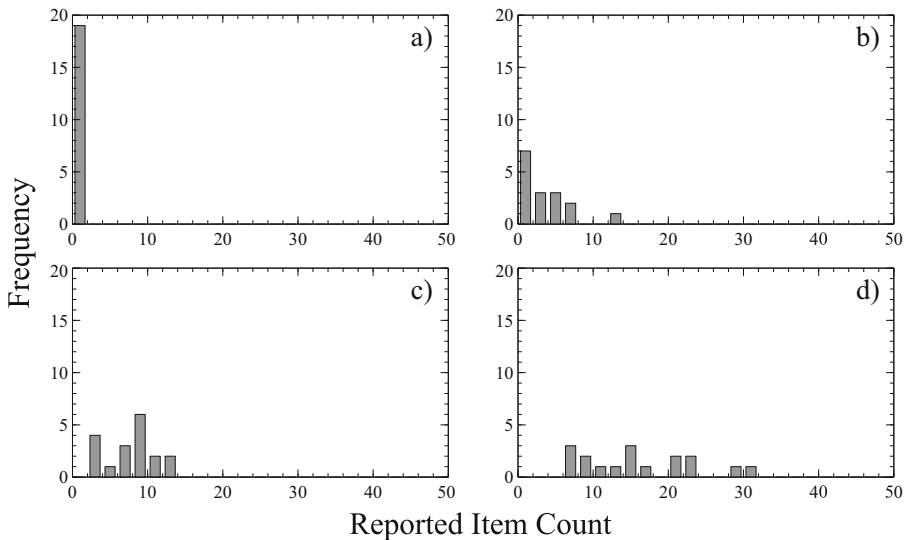
The simulation experiments in the “[Collecting Micro-Refuse Samples in the Field](#)” section showed how the actual density of micro-refuse in the sampled deposit and the style of collection can affect the accuracy of the resultant density maps. This is exacerbated by the volume of the collected sample elements; all else being equal, a smaller volume of sediment will contain fewer micro-refuse than a larger volume. These factors also affect the laboratory procedure of sorting, counting, and outlier removal. On the one hand, having sampling elements that are too small can result in too many counts of zero for all micro-refuse types, leading to biased estimates of density (Fig. 7a). On the other hand, sorting large volumes improves density estimates (Fig. 7d) but at the cost of increasing the workload and thus the potential for analyst fatigue and miscounts, or of decreasing sample size ( $n$ ). There is an inevitable trade-off between the number of sample elements (i.e., sample size,  $n$ ) and the size of the individual sample elements. Assuming that the total budget of analytical effort is fixed, if we increase the size of the sample elements, we must decrease the total number of samples that can be studied. The “pinch” sampling strategies discussed in the “[Collecting Micro-Refuse Samples in the Field](#)” section are one way to mitigate this (i.e., by limiting the initial sample size during field collection). However, as we showed in the “[Collecting Micro-Refuse Samples in the Field](#)” section, the larger collection volumes that grid-cell sampling yields may produce better results, and so may be the preferred collection method in many cases.

We have therefore developed a sequential sampling procedure that reduces the workload associated with sorting large sample elements. In this procedure, multiple small volumes for a given size class are drawn—with replacement—from the total volume that has been collected from each grid cell. Each of these small volumes is analyzed by a different analyst. The actual volume of sediment to be included in the sample should be determined by experimentation, using the density of the most common or most important types of micro-refuse (e.g., bone, basalt fragments, lithics, shell, and pottery) at the very coarse sand through very fine gravel size ranges as the indicator of adequate sample size. It is important to adapt the specifics of the experimentation procedure to the conditions of each site (i.e., the specific types, densities, and sizes of micro-refuse), and we detail the procedure we used at Tabaqat al-Būma only as an example of how this may be achieved. According to the logic in the “[Mechanical Separation of Micro-Refuse](#)” section, we chose to initially focus on the very coarse sand sized particles (1–2 mm, Table 2) of our size-sorted heavy fractions. In order to increase sampling efficiency, we also chose to break that category up into smaller (1–1.4 mm) and larger (1.4–2 mm) size fractions to see if there were differences in micro-refuse density or identifiability within the very coarse sand size fraction. For each of our two size classes, we initially drew samples using an element of 0.5 ml, but found far too many counts of 0 and 1 to be acceptable. We also experimented with elements of 1.0 ml, finding only little improvement, before settling on 3.0 ml as fairly satisfactory for most micro-refuse types, in that the mean or density was sufficiently high and our estimates of relative error leveled off (see the “[Determining the Optimum Sample Size \(Number of Sample Elements\)](#)” section). This sample element size produced adequate amounts of these artifacts for statistical rigor, while still being small enough for most analysts to sort and count in about an hour (Fig. 8 and compare Fig. 7), and the 1.4–



**Fig. 7** Ideal Poisson distribution functions for sample element sizes of **a** 0.5, **b** 1, **c** 3, and **d** 10 ml, with an overall density of 3,000 items per liter. As sample element size increases, the distribution approaches normality and the accuracy of estimates of mean and standard deviation increases, but sorting effort drastically increases

2 mm size class was much easier for analysts to sort accurately into micro-refuse categories. We note that this procedure to select an adequate element volume takes only a few days, and can also be used to produce the comparative collections that we advocate in the “[Sorting Procedure, Training, and Sorting Ergonomics](#)” section. Finally, we note that if “pinch” collection methods are used, this type of experimentation is not necessary because the sample element sizes are predetermined in the field.



**Fig. 8** Histograms of reported counts of **a** charcoal, **b** bone, **c** shell, and **d** flakes for the grid “G” context of our Tabaqat al-Bûma case study with a sample element size of 3 ml. These types occur in increasing abundance, with charcoal being very rare, and flakes being the most abundant. Compare these distributions to those in Fig. 4

For pinch samples, we instead advocate sorting the entire sample for each size fraction of interest by multiple analysts. The outlier detection, stopping rule, and density estimation routines discussed in the next section (“[The Effect of Abundance and Outlier Detection](#)” section) apply to both collection strategies.

### The Effect of Abundance and Outlier Detection

For the more abundant classes of micro-refuse, a variety of statistical analyses can highlight cases where the identity of an observer is a better explanation of the variation in counts than is the identity of the sediment context. The approach that we advocate is one that is analytically and computationally simple; we transform the raw counts into standard deviation units away from the mean ( $Z$ -scores), and then identify “bad counters” for a particular class of micro-refuse as those consistently above some confidence threshold (e.g., the 90 % confidence interval or  $\pm 1.65$  SD). We then exclude data from those observers from the subsequent analysis of counts for that artifact type, but it is not necessary to remove them from all aspects of the analysis; someone poor at counting lithics might be perfectly competent to count basalt, for example. With the most consistent outliers removed, the impact of remaining outliers is largely mitigated by using either trimmed means or medians (keeping in mind that micro-refuse studies entail cluster samples rather than element samples) to estimate the actual density of each micro-refuse type in counts per unit volume of clasts in the studied size fraction. Trimmed means and medians are reasonable methods of assessing central tendency for skewed distributions that contain many outliers, including Poisson-distributed phenomena like our micro-refuse count data (Hoaglin et al. 2000).

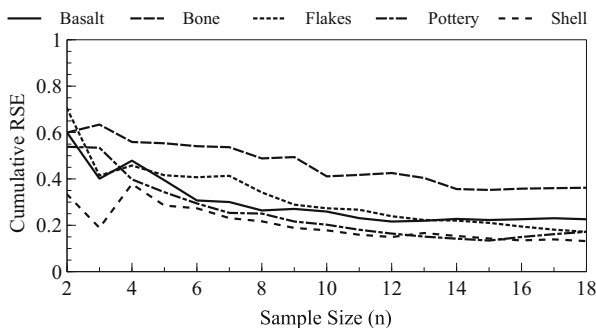
Some classes of material (e.g., burned bone, burned lithics, most micro-ecofacts) occur at much lower density than the more abundant ones, with the result that our sample element volume (3.0 ml), even after the procedure discussed in the “[Sampling Procedure and Optimum Sampling Volume](#)” section, is highly likely to yield only counts of 1 or 0, with only rare instances of higher counts. For example, in our samples from Wadi Ziqlab, micro-ecofacts such as coprolites may only occur in one out of ten samples, and then at very low densities (one or two observations). Rather than increasing the volume of the sample element (which feeds back into the workload problem), we recommend the simpler solution of treating the rare items with decreased precision. Analysts still count these rare items, but we then convert these counts to simple presence or absence. Outlier detection then becomes a matter of identifying observers who either never see anything other than the most abundant artifact types or who consistently count rare items that no other counters noticed. We stress that micro-refuse densities will vary from site to site, and even within contexts at the same site, so this conversion should only be undertaken when rare items occur at densities that are significantly (e.g., an order of magnitude) sparser than the more commonly discovered micro-refuse types *in the same context*. Although density cannot be calculated for these rare items, their ubiquity (the relative frequency of non-zero observations) in each grid cell or sample point provides a quantitative measure with which to distinguish spatial patterns. It is important to keep in mind that ubiquity is not directly comparable to density or abundance, but it is also reasonable that contexts that tend to include at least one observation of a certain micro-refuse class a significant proportion of the time are different in some way from contexts that almost never have such observations. For



example, the presence of a few pieces of burned bone in one corner of a house could mark the corner as an area of interest if no other burned bones were discovered in other portions of the house. This is especially useful for types that are so rare as to be difficult for many observers to identify correctly even when they are present. The number of possible values that ubiquity can have depends entirely on the total number of observations—with only four observations ubiquity can only be one of 0, .25, .5, .75, or 1.0, for example—but the sample sizes produced by our stopping rule (see the “[Determining the Optimum Sample Size \(Number of Sample Elements\)](#)” section) are sufficiently large for ubiquity to take on many possible values so that it almost approaches a continuous measurement.

### Determining the Optimum Sample Size (Number of Sample Elements)

Although it is possible to estimate the ideal size of  $n$  for a particular budgeted effort, estimated density, and desired error and confidence level (Thompson et al. 1992; Drennan 1996; Shennan 1997; Banning 2000), we recommend the alternative approach of sequential sampling. Sequential sampling involves adding sample elements until some stopping rule is satisfied. Here, we use a very simple stopping rule that identifies when the decrease in the cumulative relative standard error (RSE) begins to level off as we increase  $n$ . A heuristic plot with  $n$  on the  $x$  axis and the estimate of RSE along the  $y$  axis (Fig. 9) show how RSE changes with increases in  $n$ . When  $n$  is small, RSE is not stable from one size of  $n$  to the next. However, as  $n$  increases, RSE usually begins to fall and then level out, so that the RSE of, say, 15 observations is not that different from that for 16 observations. We conducted exploratory data analyses with our Tabaqat al-Bûma data sets, in which we purposely over-sampled grid cells (e.g., Fig. 9) to test the validity of this method, and to determine a reasonable general stopping rule. Our results suggest that sampling can be suspended after three consecutive estimates that are not significantly different. Because the graphical representation of the cumulative RSE curve (e.g., Fig. 9) is an easily interpretable heuristic image, we recommend using it first to confirm visually that the last three estimates appear to make a “reasonably flat” line (i.e., the RSE curve “flattens out” over 3 or more consecutive data points). We then confirm this by calculating the “local slope” of the plot line for each micro-refuse type (the slope of a regression line drawn through three consecutive data points). We have found that a 3-point “slope” of less than 0.03 for three consecutive measures of this

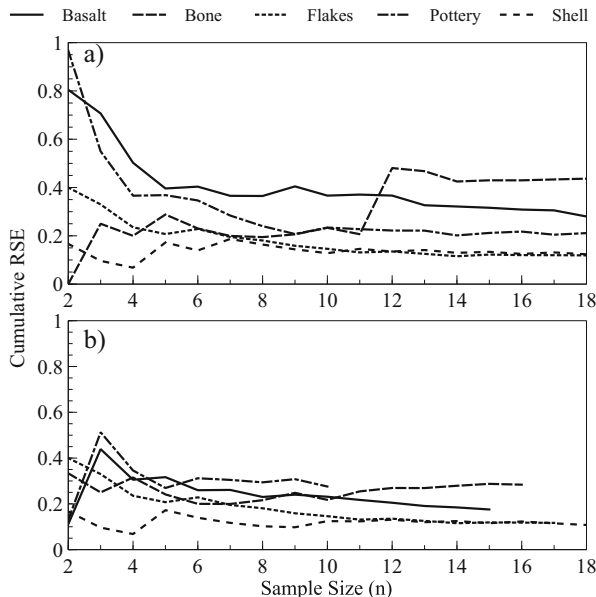


**Fig. 9** Plot of cumulative relative standard error versus number of observations ( $n$ ) for grid “B” of the E33 floor. After the trend lines level out, further increases in sample size cease to add precision to our estimate of the density of micro-refuse

*slope* indicates that there will be no significant benefit from further sampling. It is easy to program a check on this slope into a spreadsheet with a simple formula.

It is important to note that significant outliers in the data can affect this stopping rule. As an example, notice how the Cumulative RSE curve for Bone in Fig. 10a “spikes” at 12 samples. Examination of the raw count data from this context show that sample element 12 returned a count of “36,” which is much higher than most of the other counts (including the immediately preceding count of “4”). In the case of “true” outliers (e.g., due to miscounts or other human error), we can trim them (see “The Effect of Abundance and Outlier Detection” section), and this will likely alleviate the problem (e.g., Fig. 10b). However, in the case of rare micro-refuse with many 0 counts, such “spikes” are caused even by counts of 1. In these cases, we advocate applying the stopping rules to the more abundant types in the same collection unit, unless one of the rare classes is particularly critical to the anticipated analysis. Finally, if the data appear to suffer from neither significant outliers nor rarity, then there may simply be significant error inherent in the data. This is not likely to occur, however, so should be the last conclusion drawn for a curve that refuses to flatten out quickly.

This approach provides some grounds for confidence that the estimates of density or ubiquity for the most important common and rare micro-refuse are reasonably accurate, and it also generates statistical errors and confidence limits for those results. Importantly, this approach most often accomplishes this with sample sizes that require much less effort than has usually been associated with micro-refuse. We estimate that it took approximately 200 h to sort the entire E33 floor in our case study, with the workload spread across the two dozen volunteer and work-study students that we employed over two separate semesters in our lab at the University of Toronto. This estimate includes the time allocated to the training procedure outlined in the



**Fig. 10** Plots of cumulative RSE versus  $n$  for grid “G” of the lower E33 floor **a** before outlier trimming and **b** after trimming

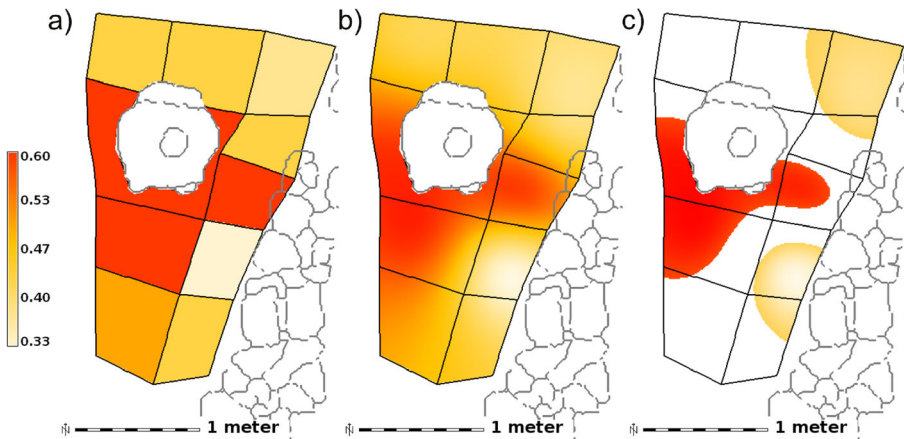
“Mechanical Separation of Micro-Refuse” section, some experimentation with sample element size and micro-refuse size classes, and deliberately “over sampling” half of the grid squares to confirm our stopping rules, yet still yields a total analysis time of about 20 h per context. Although it is difficult to precisely assess all the incidental costs, we estimate that the final cost of the analysis for this one floor context was less than CAN\$1,000 (this includes the costs associated with gathering, packing, shipping, and analyzing the samples for this floor). This can be contrasted with the famous “Heartbreak Hotel” spatial micro-refuse study by Metcalfe and Heath (1990), who reported that “literally hundreds of person-days were required over the four years devoted to isolating the microrefuse [sic] from the floor and hearth samples” (p. 793). That was for 42 contexts, and presuming at least 200 8-h days, yields a minimum estimate of 1,600 total sorting hours. It, thus, took Metcalfe and Heath an average of at least 38 h to sort a single context, and all of these hours were accrued by a single analyst. We do not know the financial cost of Metcalfe and Heath’s study, but we believe our method to have cost significantly less. Our methods not only saved sorting time and cost, but provided a vastly improved understanding of the error in our data. To further encourage a renewed interest in spatial micro-refuse studies, we have created a spreadsheet template that facilitates our counting, outlier detection, and stopping rules routines, which can be freely downloaded from [microcommons.org](http://microcommons.org), which is an online repository of micro-refuse analysis tools, data, and publications (see the [Supplemental Material](#) for the URL).

### Analyzing Spatial Patterns in Micro-Refuse Density

The final phase of a spatial micro-refuse analysis is the actual spatial analysis that will inform our interpretation of the habitual use of space on the ancient surface. Tables of micro-refuse density data are not easily interpretable, and most of the common grid-based spatial statistics, such as local density analysis (Johnson 1984) and grid-based unconstrained clustering (Kintigh 1990), are unsatisfactory for two main reasons. They only produce grid statistics at the (often very coarse) resolution of the original collection grid, and they are often difficult to interpret and must be mentally cross-correlated with other spatial data to produce an understanding of the spatial patterning in question. We instead advocate the use of quantitative graphical methods such as interpolation and unsupervised multiband-image classification.

#### Interpolation of Density Maps

The simplest quantitative graphics are thematic color maps at the resolution of the original collection grid (Fig. 11a), but these are often too coarse (grid-based collection) or spread (“pinch” samples) to be heuristically meaningful. We therefore advocate the use of GIS software to interpolate visually pleasing density plots that also serve as inputs for mathematically complex statistical analyses. Many interpolation methods are available, but we endorse regularized spline-tension interpolation as a tunable and highly accurate method, although any higher-order interpolation procedure (e.g., kriging or bicubic spline interpolation) will produce better results than lower-order interpolation procedures (such as the inverse distance-weighted method; Mitasova and



**Fig. 11** Micro-basalt density on the lower E33 floor **a** represented at the original scale of the collection grid, **b** interpolated to 1 cm resolution, and **c** converted to Z-score and clipped to the  $\pm 1\sigma$  threshold

Mitas 1993; Mitas and Mitasova 1999). To do this with grid-based data, the densities of each grid square are assigned to a point in the center of that grid square. Because the interpolation algorithm statistically “fills the gaps” between each input point, it produces output maps that predict the spatial pattern of micro-refuse densities at a precision finer than that of the original collection grid (Fig. 11b). We must reiterate, however, that these interpolated density maps are only as good as the input data, and as such are still subject to all the potential problems associated with grid-based data. That being said, the simulation experiments discussed in the “Collecting Micro-Refuse Samples in the Field” section suggest that grid-based methods produce reasonable results when combined with interpolation. The density plots provide excellent heuristic power with which to visualize concentrations and sparse areas in each micro-refuse type (Fig. 11b), but if the densities are relatively normally distributed across the map area, then the interpretive power of the density maps can be increased by converting them to Z-score units, which, with a threshold at the  $\pm 1\sigma$  level, will better highlight areas of significant density and sparseness (Fig. 11c).

### Unsupervised Classification of Activity Areas

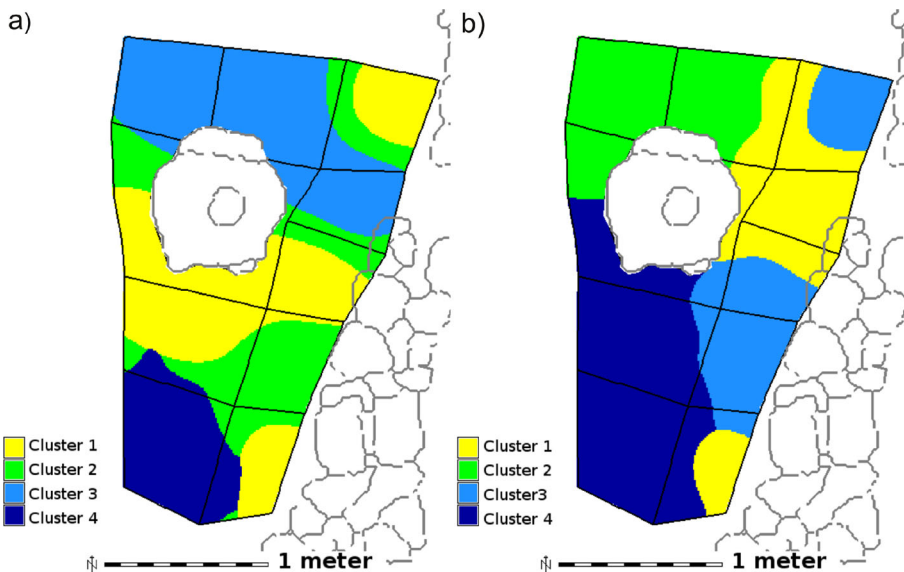
Our discussion of theory in the “Towards a Theory of Spatial Micro-Refuse Analysis” section suggests that activity areas should be defined by a unique signature of densities for all of the various micro-refuse rather than by examining just one type of micro-refuse at a time. Visually, it is difficult to compare the distributions of more than two artifact types at a time, so we advocate the use of “unsupervised classification”, a type of cluster analysis commonly used to classify multi-band satellite imagery (Lillesand et al. 2004), to partition the data from living surfaces into unique groupings of density values across many or all micro-refuse types simultaneously. This technique, which can be conducted in a variety of open-source and commercial imagery-analysis and GIS software, first spatially registers each density plot into an image “stack,” and then compares the multi-spectral “signature” at each pixel of the map. Pixels are assigned into a predetermined number of clusters based on a user-defined measure of similarity (typically the Maximum Likelihood statistic) in an iterative process so that the similarity within each cluster is greater than that between

clusters. Cluster separation occurs above a user-specified confidence interval (we use 99 %). The resulting cluster maps (Fig. 12) are produced as visual heuristic products at a scale that is easily interpretable by a human analyst. We stress that this process is not a “black box”—the analyst must control and interpret the procedure in order for it to produce a meaningful result.

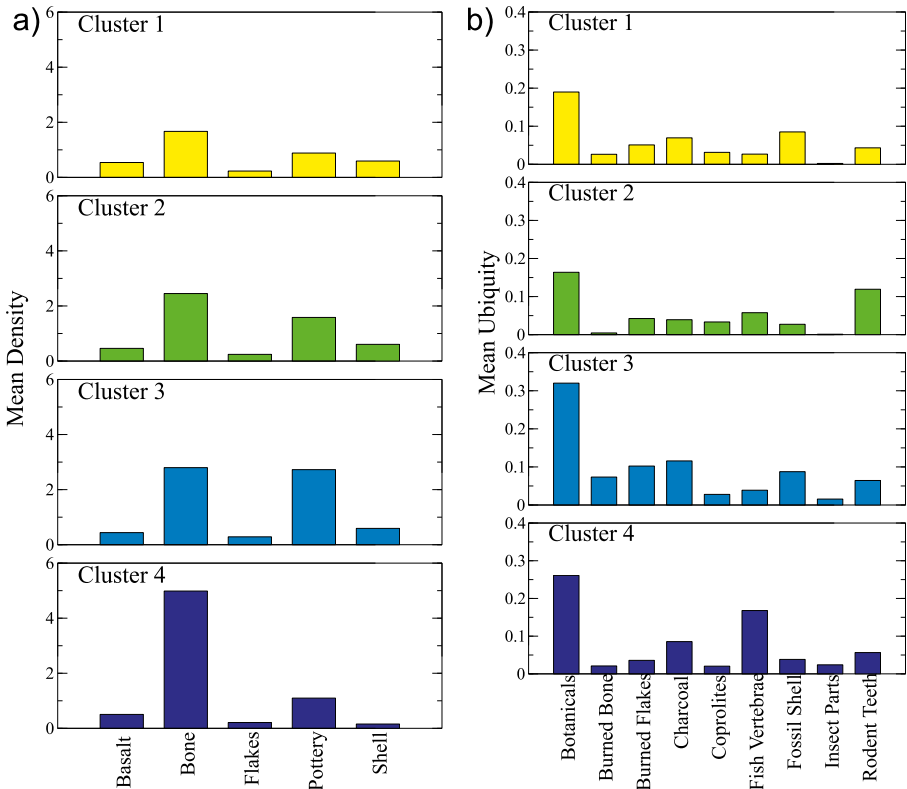
In our use of this technique, each “spectrum” is actually the density or ubiquity map of a different micro-refuse type, so the frequency of each of the spectra (micro-refuse types) within each resulting cluster indicates the relative importance of each artifact class in the areas defined by that cluster’s boundaries. A useful way to visualize this for abundant data types is to display the average density of each micro-refuse type within the spatial boundaries of each cluster as bar plots (Fig. 13 and see Table 3 for raw data). Comparing these plots not only shows which micro-refuse types are more frequent *within* a cluster, but also the relative importance of a particular micro-refuse type across *all* clusters. It should be noted, however, that while density and ubiquity measures *can* be used together in the cluster classification, ubiquity and density *cannot* be plotted on the same *y* axis in the cluster constituent chart. It may therefore be more useful to undertake separate classifications with each type of data, which is what we have done with our case study data (Fig. 13a, b).

### Results in a Case Study from Tabaqat al-Bûma, Jordan

Although space precludes detailed discussion of our particular case study here, we draw attention to a few of the results already illustrated as examples of the kinds of persistent activities that we can plausibly reconstruct through micro-refuse analysis.



**Fig. 12** Map of clustering results for the lower E33 floor for **a** abundant micro-refuse types using density values and **b** rare micro-refuse types using ubiquity value. See Fig. 8 for types included in each analysis and for cluster constituents



**Fig. 13** Cluster compositions for the cluster configurations shown in Fig. 7 for the lower E33 floor **a** for abundant micro-refuse using density values and **b** for rare micro-ecofacts using ubiquity values

Even the simple smoothed density plots provide interesting insights. For example, the density of micro-basalt, which can only occur at the site anthropogenically, and most likely originates from the use of basalt grinding stones imported to the site, is slightly higher in an area close to a large limestone mortar (Fig. 11b, c). It is plausible to interpret this either as evidence for the use of a basalt pestle or for the use of a portable basalt grinding slab that was removed from the structure before its abandonment, while the very heavy mortar was left in place. Similarly, bone chips that might result from some aspects of food preparation are concentrated to one side of a possible plaster hearth feature and micro-lithics cluster fairly tightly near the likely entrance, which would have had better daylight lighting conditions than other parts of the room.

The cluster results for relatively abundant micro-refuse (Figs. 12a and 13a) also appear to identify several distinct activity areas, each with a different suite of micro-refuse densities. In the corner near the possible hearth, cluster 4 is highly influenced by that concentration of bone fragments already mentioned, and shows fairly average levels of micro-pottery and low levels of other types, with shell especially low. Cluster 1, which occurs mainly in the likely grinding area, has above-average basalt, below-average bone and pottery and about average lithics and shell. The highest levels of

**Table 3** Densities of the five most common micro-refuse types for each cluster in the supervised classification results

Cluster	Basalt	Bone	Flakes	Pottery	Shell
1	0.540045	1.66965	0.230215	0.883101	0.593715
2	0.460488	2.45048	0.242947	1.58377	0.606337
3	0.439374	2.79819	0.286754	2.72552	0.593319
4	0.505159	4.98588	0.211477	1.09817	0.156426

lithics and pottery occur in cluster 3, which could have been a locus for tool retouch and maintenance and pot cleaning or storage in pots. Cluster 2 is probably the “background” that would include swept and walking or sitting areas, and perhaps, the micro-refuse that was included in the floor-construction material, such as the earth excavated nearby.

However, we should interpret these results in light of cluster analysis of the ubiquity data as well (Figs. 12b and 10b). Especially notable is that clusters 3 and 4, with the greatest ubiquity of botanical remains, are mainly in the southern part of the room, also consistent with food preparation there, and fish vertebrae, also likely associated with food preparation, only reach high ubiquity in cluster 4. Burned bone and burned flakes are only common in cluster 3, where we also find relatively high ubiquity of charcoal in close association with the hypothesized hearth.

## Conclusions

Problems of time investment, precision and error enter micro-refuse studies at many points from the initial collection of sediment samples, through sampling and counting in the lab, through quality assurance of the data, to analysis and interpretation. Through experimentation over many years, we have found that all these problems are surmountable. Attention to the spatial collection frame guides the scale of the rest of the analysis, but even coarse, grid-based collection can result in accurate, high-resolution density and ubiquity plots. Use of many analysts sorting multiple small samples significantly decreases the cost and processing time over traditional methods of sorting micro-refuse, yet still yields reliable results with an enhanced understanding of the error of the final density estimates. Moreover, by ensuring adequate training, ergonomics, and quality assurance protocols, we minimize the effects of human error. Our preferred method for identifying clusters from interpolated density maps is an intuitive graphical approach that allows simultaneous analysis of many categories of micro-refuse. All together, the workflow outlined in this paper leverages modern digital technology and sampling theory to yield easily interpretable results that, we think, help us better understand the daily lives of ancient people. That our method does so while maintaining budgetary efficiency and with attention to error control will, we hope, encourage the expansion of spatial micro-refuse studies in household archaeology.

**Acknowledgments** This research was funded by several grants from the Social Sciences and Humanities Research Council of Canada, and we acknowledge their generous support. We would like to extend our profound thanks to the dozens of University of Toronto student volunteers and work-study students who participated in microrefuse sorting in the Department of Anthropology's Wadi Ziqlab laboratory, and to Matthew Betts, Paul Racher, and Danielle MacDonald who supervised them in earlier phases of this project. The Department of Antiquities of Jordan and its Director at the time, Dr. Ghazi Bisheh, greatly facilitated the fieldwork at Tabaqat al-Bûma that provides the case study for this paper. We also wish to acknowledge the GRASS GIS development team who, together, has created an excellent and highly extensible open-source GIS environment, without which this research could not have taken place. Finally, we would also like to thank microcommons.org for hosting our spreadsheet template, and for helping the cause of spatial micro-refuse analysis in general.

## References

- Banning, E. B. (2000). *The archaeologist's laboratory: the analysis of archaeological data*. New York: Springer.
- Banning, E. B. (2007). Wadi Rabah and related assemblages in the southern Levant: interpreting the radiocarbon evidence. *Paléorient*, 33, 77–101.
- Banning, E. B., Gibbs, K., & Kadowaki, S. (2011). Changes in material culture at Late Neolithic Tabaqat al Buma, in Wadi Ziqlab, northern Jordan. In J. Lovell & R. M. Yorke (Eds.), *Culture, chronology and the chalcolithic: theory and transition* (pp. 36–60). Oxford: Oxbow Books.
- Barton, C. M., Riel-Salvatore, J., Anderies, J. M., & Popescu, G. (2011). Modeling human ecodynamics and biocultural interactions in the late Pleistocene of western Eurasia. *Human Ecology*, 39, 705–725. doi:10.1007/s10745-011-9433-8.
- Bertran, P., Lenoble, A., Todisco, D., et al. (2012). Particle size distribution of lithic assemblages and taphonomy of Palaeolithic sites. *Journal of Archaeological Science*, 39, 3148–3166. doi:10.1016/j.jas.2012.04.055.
- Binford, L. R. (1978). Dimensional analysis of behavior and site structure: learning from an Eskimo hunting stand. *American Antiquity*, 43, 330–361. doi:10.2307/279390.
- Brown, C. T., Witschey, W. R., & Liebovitch, L. S. (2005). The broken past: fractals in archaeology. *Journal of Archaeological Method and Theory*, 12, 37–78.
- Drennan, R. D. (1996). *Statistics for archaeologists: a common sense approach*. New York: Springer.
- Dunnell, R. C., & Stein, J. K. (1989). Theoretical issues in the interpretation of microartifacts. *Geoarchaeology*, 4, 31–42.
- Fairbairn, A. (2005). *Simple bucket flotation and wet sieving in the wet tropics*. Canberra: RSPAS, Australian National University.
- Fladmark, K. R. (1982). Microdebitage analysis: Initial considerations. *Journal of Archaeological Science*, 9, 205–220. doi:10.1016/0305-4403(82)90050-4.
- Gibbs, K., Kadowaki, S., Banning, E. B. (2006). The Late Neolithic at al-Basatîn in Wadi Ziqlab, northern Jordan. *Antiquity*, 80(310), <http://antiquity.ac.uk/projgall/gibbs/index.html>.
- GRASS Development Team (2014) Geographic Resources Analysis Support System. <http://grass.osgeo.org>.
- Gregg, S. A., Kintigh, K. W., Whallon, R. (1990). Linking ethnoarchaeological interpretation and archaeological data: the sensitivity of spatial analytical methods to postdepositional disturbance. *The Interpretation of Archaeological Spatial Patterning*.
- Guan, Y., Gao, X., Wang, H., Chen, F., Pei, S., Zhang, X., & Zhou, Z. (2011). Spatial analysis of intra-site use at a Late Paleolithic site at Shuidonggou, Northwest China. *Chinese Science Bulletin*, 56(32), 3457–3463. Accessed 12 September 2014.
- Hayden, B., & Cannon, A. (1983). Where the garbage goes: refuse disposal in the Maya Highlands. *Journal of Anthropological Archaeology*, 2, 117–163. doi:10.1016/0278-4165(83)90010-7.
- Hoaglin, D. C., Mosteller, F., & Tukey, J. W. (2000). *Understanding robust and exploratory data analysis* (2nd ed.). New York: Wiley.
- Hodder, I., & Cessford, C. (2004). Daily practice and social memory at Catalhoyuk. *American Antiquity*, 69, 17–40.
- Hull, K. L. (1987). Identification of cultural site formation processes through microdebitage analysis. *American Antiquity*, 52, 772–783. doi:10.2307/281385.
- ISO-BSEN (2002). 14688–1. 2002. Geotechnical investigation and testing. Identification and classification of soil. Part 1: identification and description. *British Standards Institution*.



- Johnson, I. (1984). Cell frequency recording and analysis of artifact distributions. In H. Hietala (Ed.), *Intrasite spatial analysis in archaeology* (pp. 75–96). Cambridge: Cambridge University Press.
- Jones, G. E. M. (1983). *The use of ethnographic and ecological models in the interpretation of archaeological plant remains: case studies from Greece*. Cambridge: University of Cambridge.
- Kamp, K. A. (2000). From village to tell: household ethnoarchaeology in Syria. *Near Eastern Archaeology*, 63, 84–93. doi:10.2307/3210745.
- Kent, S. (1984). *Analyzing activity areas: an ethnoarchaeological study of the use of space*. Albuquerque: University of New Mexico Press.
- Kintigh, K. W. (1990). Intrasite spatial analysis: A commentary on major methods. In A. Voorrips (Ed.), *Mathematics and information science in archaeology: a flexible framework* (pp. 165–200). Bonn: Helos.
- LaMotta, V., & Schiffer, M. B. (1997). Formation processes of house floor assemblages. In P. M. Allison (Ed.), *The archaeology of household activities* (pp. 19–29). New York: Routledge.
- Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (2004). *Remote sensing and image interpretation* (5th ed.). Hoboken: Wiley.
- McLaren, P., Bowles, D. (1985). The effects of sediment transport on grain-size distributions. *Journal of Sedimentary Research* 55.
- Metcalfé, D., & Heath, K. M. (1990). Microrefuse and site structure: the hearths and floors of the Heartbreak Hotel. *American Antiquity*, 55, 781–796. doi:10.2307/281250.
- Mitas, L., & Mitasova, H. (1999) Spatial interpolation. In: P. A. Longley, M. F. Goodchild, D. J. Maguire, & D. W. Rhind (Eds), *Geographical information systems: principles, techniques, management and applications* (pp 481–492). New York: Wiley.
- Mitasova, H., & Mitas, L. (1993). Interpolation by regularized spline with tension: I. Theory and implementation. *Mathematical Geology*, 25, 641–655.
- Pye, K., & Tsoar, H. (2009). *Mechanics of aeolian sand transport. Aeolian Sand and Sand Dunes* (pp. 99–139). Berlin: Springer.
- Rosen, A. M. (1986). *Cities of clay: the geoarchaeology of tells*. Chicago: University of Chicago Press.
- Rosen, A. M. (1993). Microartifacts as a reflection of cultural factors in site formation. In P. Goldberg, D. T. Nash, & M. D. Petraglia (Eds.), *Formation processes in archaeological context* (pp. 141–148). Madison: Prehistory Press.
- Sahu, B. K. (1964). Depositional mechanisms from the size analysis of clastic sediments. *Journal of Sedimentary Research*, 34.
- Schiffer, M. B. (1983). Toward the identification of formation processes. *American Antiquity*, 48, 675–706.
- Schiffer, M. B. (1987). *Formation processes of the archaeological record*. Albuquerque: University of New Mexico Press.
- Shelton, C. P., & White, C. E. (2010). The hand-pump flotation system: a new method for archaeobotanical recovery. *Journal of Field Archaeology*, 35, 316–326. doi:10.1179/009346910X12707321358838.
- Shennan, S. (1997). *Quantifying archaeology* (2nd ed.). Edinburgh: Edinburgh University Press.
- Sherwood, S. C., Simek, J. F., & Polhemus, R. R. (1995). Artifact size and spatial process: macro- and microartifacts in a mississippian house. *Geoarchaeology*, 10, 429–455. doi:10.1002/gea.3340100603.
- Shott, M. J. (2010). Size dependence in assemblage measures: essentialism, materialism, and “SHE” analysis in archaeology. *American Antiquity*, 75(4), 886–906.
- Simms, S. R., & Heath, K. M. (1990). Site structure of the Orbit Inn: an application of ethnoarchaeology. *American Antiquity*, 55, 797–813. doi:10.2307/281251.
- Tani, M. (1995). Beyond the identification of formation processes: behavioral inference based on traces left by cultural formation processes. *Journal of Archaeological Method and Theory*, 2(3), 231–252. doi:10.1007/BF02229008.
- Thompson, S. K., Ramsey, F. L., & Seber, G. A. F. (1992). An adaptive procedure for sampling animal populations. *Biometrics*, 48, 1195. doi:10.2307/2532710.
- Tucker, P. (1980). A grain mobility model of post-depositional realignment. *Geophysical Journal International*, 63, 149–163. doi:10.1111/j.1365-246X.1980.tb02614.x.
- Ullah, I. I. T. (2009). Within-room spatial analysis of activity areas at Late Neolithic Tabaqat Al-Buma, Wadi Ziqlab, Al Koura, Jordan. *Studies in the History and Archaeology of Jordan*, X. The Department of Antiquities of Jordan, Amman, pp 87–95
- Ullah, I. I. T. (2012). Particles from the past: Microarchaeological spatial analysis of ancient house floors. In B. J. Parker & C. P. Foster (Eds.), *New perspectives in household archaeology* (pp. 123–138). Winowna Lake: Eisenbrauns.
- Vance, E. D. (1987). Microdebitage analysis in activity analysis: an application. *Northwest Anthropological Research Notes*, 20, 179–189.
- Visher, G. S. (1969). Grain size distributions and depositional processes. *Journal of Sedimentary Research*, 39, 1074–1106.
- Wentworth, C. K. (1922). A scale of grade and class terms for clastic sediments. *The Journal of Geology*, 30, 377–392.