The Reliability of Using Surface Data for Seriation

Jordan T. Downey

Artifacts found on the surface of an archaeological site are frequently used as a proxy for unexcavated subsurface remains. This is especially true of survey projects or those focused on settlement patterns, where excavating test trenches at a large number of sites is unfeasible and where archaeologists make the assumption—implicit or explicit—that artifacts from the surface of the site are closely related to those that remain belowground. In this way, surface artifacts are used for everything from seriation and determining cultural affiliation to making inferences about the people who lived at a site. Surface artifacts have been widely used in settlement pattern studies and landscape archaeology (e.g., Bevan and Conolly 2009; Billman 1996; Bintliff 2005; Dunnell and Dancey 1983; Wandsnider 2004; Willey 1953; Wilson 1988; Zvelebil et al. 1992), sociopolitical reconstruction (e.g., Sherman et al. 2010; Whalen and Minnis 2012; Winter-Livneh et al. 2010), resource and commodity acquisition (e.g., Barrientos et al. 2015; Wilkinson 1989), population estimates (e.g., Ortman 2016), and mapping of social networks (e.g., Mills et al. 2013), to name just a few examples. Contract archaeology projects also acquire a very large quantity of artifacts

ABSTRACT

Archaeological sites are often dated through seriation analysis of artifacts found on the site's surface. This relative dating method remains common despite the widespread availability of absolute dating methods because it is fast and cost-effective compared to scientific dating methods such as radiocarbon dating. Surface seriation is especially important for regional survey studies that involve a large number of sites and little to no excavation. In this context it is important to ask: are surface assemblages as reliable an indicator of the age of a site as determined through excavation? This unique study addresses this question using data from seven sites in the Virú Valley on the north coast of Peru. Surface assemblages are compared with excavated ones using the G-test statistic. It is found that surface assemblages do not closely resemble excavated ones in a statistically significant sense. Nevertheless, the relative date of surface assemblages typically resembles the relative date of excavated assemblages. Caution is urged when dating surface assemblages purely through seriation because the surface may not actually be representative of excavated assemblages.

Los sitios arqueológicos son a menudo fechados a través del análisis por seriación de los artefactos encontrados en superficie. Este método de datación relativa sigue siendo común a pesar de la amplia disponibilidad de métodos de datación absoluta, ya que es rápido y económico en comparación con los métodos de datación científica tales como la datación por radiocarbono. La seriación de superficie es especialmente importante para los estudios de reconocimiento regional que incluyen un gran número de sitios con poca o ninguna excavación. En este contexto, es importante preguntarse: ¿Es el análisis de las recolecciones de superficie un indicador tan confiable de la edad de un sitio como lo es la excavación? Este es un estudio único que analiza la pregunta utilizando datos de siete sitios en el valle de Virú, en la costa norte de Perú. Se comparan los conjuntos de superficie y los conjuntos excavados utilizando el metodo estadístico de la prueba G. Encontramos que, estadísticamente, los conjuntos de superficie no se parecen mucho a los excavados. Sin embargo, el fechamiento relativo de los conjuntos de superficie normalmente es parecido al fechamiento relativo de los conjuntos excavados. Se recomienda precaución al asignar una fecha a los conjuntos de superficie solamente por medio de seriación ya que es posible que estos no representen los conjuntos excavados.

Advances in Archaeological Practice 5(1), 2017, pp. 26–43

DOI:10.1017/aap.2016.2

from surface contexts, and researchers can use data from these projects to address a bevy of archaeological questions (e.g., Ortman et al. 2007). Projects that rely on surface data typically use a type-proportion model whereby the frequency of one type of artifact of a certain class (e.g., lithics, ceramics) is compared to the frequency of other types in different contexts to make inferences about changing human behaviors across space and through time. Such studies frequently rely on the untested assumption that surface assemblages closely match those that remain buried. This article presents a unique opportunity to test these assumptions using statistical analysis.

Dating sites based on their surface assemblages can be trickier than it may seem at first glance because site formation processes can affect surface and buried deposits in different ways. In general, the surface of an archaeological site is more dynamic than buried levels since an undisturbed stratigraphic level of a site captures a relatively short period of time that is capped by the layers of soil or construction on top of it. As such, the artifacts found in a layer can be expected to date from its period of occupation. Site formation processes must always be considered, however, as a buried level at a site may have been exposed on the surface for a long period of time (Dunnell and Dancey 1983), but at sites with complex architecture, a site surface may be capped very rapidly. Contrast this with the surface of a site, which may represent centuries or even millennia of occupation and reuse and where a slew of site formation processes such as erosion and modern disturbance may bring artifacts from various periods to the surface, resulting in a complex palimpsest where artifacts from several different occupations can potentially be mixed together on the site surface (Barton et al. 2004; Davies et al. 2015; Fanning and Holdaway 2001; Fernández-López de Pablo and Barton 2015; Wandsnider 1992; Zvelebil et al. 1992). This poses a significant, but rarely addressed, problem to the interpretation of archaeological sites based on surface artifacts alone, and its implications are especially important for sites where cultural refuse may be rapidly accumulated and capped while a site is actively occupied, but where the latest layers of the site can sit exposed for a very long time after abandonment.

While surface artifact assemblages have long been used to seriate sites (Lewarch and O'Brien 1981; O'Brien and Lyman 1999), several studies in recent decades have applied statistical techniques to surface assemblages in order to better understand occupational sequences and site formation processes (Christenson 1994; Fanning and Holdaway 2001; Holdaway et al. 2004; Ortman et al. 2007; Roberts et al. 2012; Steponaitis and Kintigh 1993). Seriation methods often assume that a site was occupied for a relatively short time (or for multiple short times), but this is frequently not the case. If formation processes allow a mixture of artifacts from various time periods to appear on the site's surface—not at all an uncommon occurrence—it can be very difficult to date a site through seriation alone. Several authors in recent decades have sought to overcome this problem and sort out mixed or long-term assemblages using various statistical means. Steponaitis and Kintigh (1993) developed and tested two algorithms (type presence and type frequency) to estimate the maximal occupation dates of a site based on ceramic types with known manufacture dates. Their model built upon mean ceramic dating (MCD), a method developed by Stanley South (1977) for seriation in historical archaeology. Steponaitis and Kintigh caution that their method would not work in multicomponent or mixed assemblages. Along similar lines, Christenson (1994) applied MCD to a precontact setting, the Kayenta Anasazi in the U.S. Southwest, and by comparing MCD dates to contexts dated through tree-ring dating he showed that MCD can be useful and accurate in a precontact setting where the precise manufacture dates of any given ceramic type are not known. In a different vein, Fanning and Holdaway (2001; Holdaway et al. 2004) have shown that, even in a setting where soil erosion is high, the type distribution of lithic surface scatters is not significantly impacted, although they caution against viewing surface scatters as intact living floors that represent a specific short-term occupation. Indeed, by obtaining radiocarbon dates from hearth features, they (Holdaway et al. 2004) showed that there was a persistence in the use of place in their study area because individual hearths often dated to decades or centuries apart, and, using this knowledge, they were better able to understand the deposition of lithic scatters as representing millennia of repeated occupation rather than being a single site dated to a well-defined period of time.

Other authors have used advanced statistics to understand the sequence of sites that were occupied for long periods of time, a common shortfall of traditional seriation. Roberts et al. (2012) developed a method that they called "chronological apportioning" and applied it to data from Arizona to better model regional settlement histories using ceramics. Bayesian statistics have also been applied to surface ceramic data. Ortman et al. (2007) use such statistics to examine a very large dataset consisting of more than 3,000 sites in southwestern Colorado. Their dataset was collected for heritage management purposes rather than archaeological research, posing a number of problems. Using Bayesian methods to overcome some of the problems present not only in this dataset but common to surface archaeology in general, the authors find that their model may be of particular use to archaeologists who work with legacy or heritage management data for settlement pattern analysis. Taken together, these studies offer attempts to circumvent some of the inherent problems in surface artifact scatters and are especially relevant to landscape archaeology and settlement patterns.

The purpose of this article is to test the reliability of using surface data as a proxy for the types and quantities of artifacts that would be found in excavations at a site, were they to be done. While the studies discussed above are all focused on obtaining accurate and detailed information from the surface of archaeological sites, none explicitly tests whether the distribution of artifacts found on a site's surface reliably predicts the distribution of artifacts that remain in the ground. The current study is unique in that surface artifact distributions are compared to excavated assemblages from the same site using statistical means, thereby testing the

https://doi.org/10.1017/aap.2016.2 Published online by Cambridge University Press

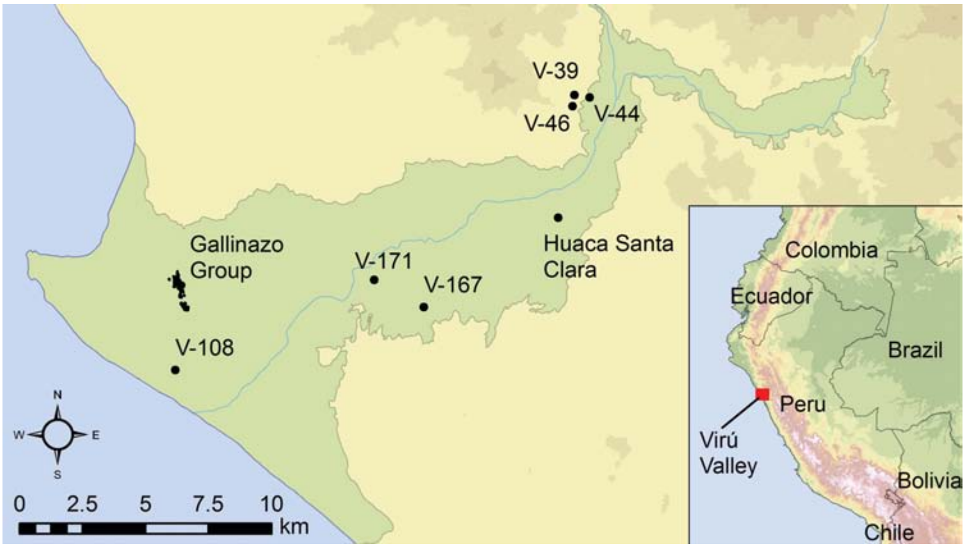

FIGURE 1. Sites discussed in this article. The Gallinazo Group is not used in this analysis but is included here for reference.

implicit assumption that surface artifacts can accurately date archaeological sites using traditional relative dating methods.

This article draws from a large and high-quality ceramic dataset collected by James A. Ford (1949) and his colleagues in the Virú Valley of Peru (Figure 1), a relatively small valley of ca. 185 km^2 that cuts through the arid coastal plain of northern Peru. These data were collected as part of the Virú Valley Project, a multi-institutional research project conducted in the 1940s. This project included archaeologists from the American Museum of Natural History, Columbia University, Yale University, the Field Museum, and the Smithsonian Institution and was designed to offer a major contribution not only to the archaeology of Peru, but also to archaeological method and theory more broadly. The project collected a massive amount of data, which was published in a series of volumes (Bennett 1950; Collier 1955; Ford 1949; Ford and Willey 1949; Strong and Evans 1952; Willey 1953). Surface collections were made at over 250 sites throughout the valley, and deep stratigraphic excavations were carried out at over two dozen sites. Chronology was a major focus of the project and all members worked together to develop a unified ceramic typology.

While all of the scholars involved in the project addressed chronological issues, it was James A. Ford (1949) who was responsible for the bulk of the chronological work. Ford cross-correlated the ceramic assemblages from all excavation trenches to produce a master chronology for the valley and then used this as a basis for seriating surface collections from throughout the valley. Thus, the utility of using surface artifacts to date archaeological sites was central to Ford's (1949) work. Gordon R. Willey (1953) was responsible for summarizing the culture history and the settlement patterns of Virú. Willey worked together with Ford to conduct surface surveys and make artifact collections at some 283 sites throughout the valley.[1] Ford's subsequent seriation then formed the basis for Willey's own dating sequence. Together with his project colleagues, Willey divided the valley's history into nine time periods, each of which coincided with wider north coast cultural expressions (Table 1). Willey's study was and continues to be hugely influential for settlement archaeology, and the dating scheme he proposed for Virú was a significant contribution to the archaeology of the north coast of Peru.

For Ford (1949:51, 1952), artifacts found on the surface of a site represented the mean or average date for that site, except in the case of any site with very deep stratigraphy. Ford (1949:35-37) was primarily concerned with dating the structures at a site but did so by dating the site itself through ceramic seriation, making the assumption that the ceramics at a site closely matched the architecture or, if they dated to different periods, the two could be parsed out. Ford arrived at this conclusion by comparing assemblages collected from the surface of sites with controlled stratigraphic excavations by Strong and Evans (1952), Collier (1955), and some of his own test trenches.[2] Ford (1949:35) saw it as only logical for surface artifacts to represent the mean site date, and he held onto this view strongly (Ford 1952), but he never specified any statistics or method that he used to compare surface artifacts to subsurface ones, and it is difficult to tell whether he systematically compared the surface artifacts from a site to the subsurface ones. To reframe Ford's claim as a hypothesis, the distribution of artifact types from a site surface will be the same as the distribution of artifacts collected from all levels of excavations, except at sites with deep stratigraphy.

As part of an extended critique, Bennyhoff (1952:232) took exception to this hypothesis and argued that surface ceramics from a site generally date the last time period of that site, not an average date for the entire settlement history. Bennyhoff noted that, despite Ford's claims that excavations were conducted at 28 sites, he published data from only nine of these sites, making it difficult to test these claims. Bennyhoff highlighted several cases where Ford's own assessment of a site's date range did not coincide with the excavations made at that site and other cases

https://doi.org/10.1017/aap.2016.2 Published online by Cambridge University Press

TABLE 1. The Nine Time Periods Used by the Virú Valley Project (Willey 1953), Their Approximate Date Range, and the Main Corporate and Domestic Ceramic Types in Use during Each Period.

Period	Date Range (approx.)	Major Corporate Ceramics (Ford 1949)	Major Domestic Ceramics (Downey 2014)
Colonial	A.D. 1532–1824	Colonial	Colonial
Estero (Chimú-Inka)	A.D. 1470–1532	Inca Painted	Late Plainwares Queneto Polished Plain
La Plata (Chimú)	A.D. 1100–1470	San Juan Moulded Polished Black Ware	Late Plainwares Queneto Polished Plain
Tomaval (Lambayeque)	A.D. 750–1100	San Nicolas Moulded Red-White-Black (Coastal Tihuanacoid)	Late Plainwares Queneto Polished Plain
Huancaco (Moche)	A.D. 600–750	Huancaco	Castillo Plain Gloria Polished Plain
[a]Gallinazo/Virú	200 B.C.–A.D. 600	Gallinazo (Virú) Negative Callejón Negative	Castillo Plain Gloria Polished Plain
Puerto Morin (Salinar)	400–200 B.C.	Puerto Morin White-on-Red	Huacapongo Polished Plain
Guañape (Cupisnique)	1200–400 B.C.	Ancón	Guañape Plain
Cerro Prieto (Late Preceramic)	? – 1200 B.C.	N/A	N/A

Note: Corporate ceramic types are from Ford (1949); domestic ceramic types are from Downey (2014) and are a condensed form of Ford's types. Contemporary north coast cultural expressions are listed in brackets where they differ from Virú period names. Dates are derived from Downey and Millaire (2015), Koons and Alex (2014), Quilter (2014), Willey (1953: 37), and Zoubek (1997). The Cerro Prieto and Guañape periods are poorly dated in the valley.

[a]This period was originally called "Gallinazo" by the Virú Valley Project, and this name continues to be widely used. Following Jean-François Millaire (2009a), I prefer to use the term "Virú," the name given by Rafael Larco Hoyle (1945) for the distinctive style of negative-painted corporate ware diagnostic of this period. Both names are listed here for clarity.

where the surface assemblage dated that site's terminal period rather than the mean date. Ford (1952) very briefly rebutted Bennyhoff's detailed critique by emphasizing that the evidence used to date surface assemblages was chosen very carefully by project members and by reiterating that most sites are shallow and, therefore, the surface assemblages date the mean age of the site. However, it remains that Ford never published data to support these claims, making it difficult to accept this at face value.

In this article, I test Ford's hypothesis that the surface ceramics from a site represent the mean or average date for the entire site. In doing so, I address a larger problem in archaeology: when determining a relative date for an archaeological site using seriation, are artifacts found on the site's surface a reliable proxy for those that would be recovered in excavations, were they to be conducted? While Ford never published data to support his hypothesis, the data collected by him[3] and his colleagues in Virú offer a large and high-quality dataset to test this hypothesis. Using these data, as well as data from recent excavations at Huaca Santa Clara, I am able to test Ford's (1949) claims that surface assemblages represent the mean date of a site and, in turn, the hypothesis that surface ceramic assemblages are reliable for dating the prehispanic occupation(s) of sites in the absence of chronometric dating or intensive excavations. I find that Ford's hypothesis is not supported in all cases, but the situation is complex because the relative date indicated by the surface assemblage of most sites is similar to the relative date of assemblages from the upper layers of stratigraphic excavations. Moreover, this study shows the usefulness of revisiting legacy datasets while at the same time highlighting problems that accompany the use of such data.

COMPARING SURFACE AND SUBSURFACE ASSEMBLAGES

The methods used to test Ford's (1949) hypothesis are relatively straightforward. Ford's primary role in the Virú Valley Project was to collect and classify a sample of ceramic artifacts found on the surface of sites throughout the valley.[4] Ford's colleagues William Duncan Strong, Clifford Evans, and Donald Collier (Collier 1955; Strong and Evans 1952) excavated deep stratigraphic trenches at several sites throughout the valley. These trenches were excavated in arbitrary levels, typically 25 cm deep, and were intended to demonstrate how ceramic assemblages changed through time. Ford also made small excavation trenches at a few sites to test the reliability of his surface collections. Ford and his colleagues worked closely together to develop a standardized typology that they used to classify all ceramics, taking measures to minimize inter-observer error. Each scholar described this typology separately with the most detailed version (Strong et al. 1952) being published as an appendix to Strong and Evans (1952). These descriptions were detailed enough to allow other archaeologists in the region to employ the same typology in their own studies.

For this study, I compare the surface and excavated artifacts from the same site using a condensed version of the standardized Virú Valley Project typology described above. Data were collected from Ford's unpublished archives (.F673, Papers of James Ford, Division of Anthropology, American Museum of Natural History; these data were used with permission), from tables published in Collier (1955), and from data collected by Jean-François Millaire's excavations at the site of Huaca Santa Clara.[5] All data on surface

https://doi.org/10.1017/aap.2016.2 Published online by Cambridge University Press

collections come from Ford's archives. Only domestic wares were considered because corporate wares from earlier time periods are very poorly represented in surface collections and because evidence points to corporate and domestic wares developing along different timescales (Downey 2014; Downey and Millaire 2015). Domestic wares were condensed into six types, representing four long-lived traditions in Virú: (1) Guañape Plain, (2) Huacapongo Polished Plain, (3) Castillo Plainwares (including the contemporary Gloria Polished Plain sub-type), and (4) Late Plainwares (including the contemporary Queneto Polished Plain sub-type);[6] see Table 1 for the relative date and period for each type. Detailed descriptions of these traditions and of the reasons why this modified typology was chosen are available elsewhere (Downey 2014). Several of these types were used for very long periods, especially Castillo, which remained quite similar for over one millennium (Donnan 2009). Such long-term trends in ceramic style complicate any seriation of this material because multiple time periods—as defined by corporate ceramics and political cultures—may be represented by the same ceramic type. This has little bearing on the study at present, however, because the ceramic types used are standardized across all contexts and, at present, I am more concerned with comparing these contexts than I am with dating each site.

In total, I am able to compare surface and excavation data from seven sites; these represent four earth-mound sites in the lower and middle valley and three rock-walled sites on the hillsides of the Queneto Quebrada.[7] Although excavations were conducted at over two dozen sites during the course of the Virú Valley Project, most of these cannot be included here due to data incompatibilities. The primary issue is that neither Ford nor his colleagues made surface collections at most sites where excavations took place (or collections at these sites produced a very small sample of artifacts), and therefore, comparable surface and excavation data are available only for a few sites. Fieldwork to make such collections was not presently feasible and was deemed unnecessary considering that several robust contexts were available.

Only three of the sites where Ford excavated test trenches can be used here because most of these excavations were either very limited or were conducted at sites with very small surface assemblages that are not statistically meaningful; I test here Ford's excavations at V-39, V-44, and V-46. Collier's stratigraphic cuts at several midden sites provided good data on long-term ceramic trends, and surface collections were also made at most of these sites. I am able to include only three sites, however, because of various problems encountered by Collier (1955), most notably being the theft of several bags of uncataloged artifacts from the site of V-272. These sites are V-108, V-167, and V-171. Finally, excavations by Jean-François Millaire at the site of Huaca Santa Clara (V-67), a site where Ford collected a large surface assemblage, provide an additional data source. Millaire ensured that ceramic sherds were classified using the same types that were described by Ford and his colleagues. These sites are located throughout the valley and offer a good cross section of the types of site found in Virú.

The Virú Valley is typical of the coastal plain in northern Peru. The coast of Peru is an arid desert strip that runs between the Andean foothills to the east and the Pacific Ocean to the west. The area receives essentially no rainfall in an average year, although periodic El Niño Southern Oscillation (ENSO) events can bring catastrophic torrential rain every seven to 20 years (Moseley 2001). This environment is thought to be stable throughout the Holocene. The Virú River is one of several east–west river systems that run from the Andes to the ocean, and irrigation canals have expanded each river to a wide valley system. The middle and upper Virú Valley is bordered by steep-sided rocky slopes and *quebradas* that form alluvial fans full of sub-angular boulders. Erosion through direct, though rare, rainfall likely affects all sites throughout the valley. Looting activity is also common at all sites in the valley.

Site formation processes are little studied in Virú. Anthropogenic formation processes can be expected to be similar at all sites: structures may consist of simple walled residential houses or adobe pyramid mounds but, in all cases, floors will be built up using fill, plaster, and adobes, and some structures can attain a significant height either through intentional building-up or through a tell-like formation of architecture and refuse. Many sites were reused as prehispanic (and in some cases modern) cemeteries after the site was no longer occupied and burials within the floor of houses were common in some periods. The sites of V-39, V-44, and V-46 are all rock-walled residential settlements located on the alluvial fan of the Queneto Quebrada (Figure 2; Willey 1953). Each site is built directly on the desert floor, and there is little vegetation and no soil formation. ENSO rains have cut channels through each site, but outside of these channels, there is likely to have been little water movement. The site of Huaca Santa Clara (V-67) is a large adobe structure on top of a steep rocky outcrop in the middle valley (Figure 3; Millaire 2004). There is no water source at the site, and fluvial transformation of artifacts is unlikely. While windblown sand covers some parts of the site there is no soil formation to speak of. The sites of V-108, V-167, and V-171 are all adobe structures that rise several meters above the valley floor in lower Virú (Willey 1953). Such sites in the valley are generally dry and unaffected by irrigation canals or aquifers that are on the valley floor. Soil probing at a nearby site suggests that adobe structures in the lower valley were originally built on sand dunes that have been built up and expanded by construction (Millaire and Eastaugh 2011, 2014). Windblown sand typically accumulates at such sites, sometimes to a depth of several meters. These site formation processes offer context to the discussion below.

Statistics and Data Used

There are several ways to compare the surface and subsurface ceramic remains from any given site. One method is to compare the frequency or percentage of ceramic types between various contexts. This is essentially what Ford did. Ford (1949:44–47, Figure 4) compiled data from stratigraphic excavations conducted by Strong and Evans (1952), Collier (1955), and from Cut A at the site of V-60, which he excavated himself. He converted all data to the percentage of types from each strata and graphed these systematically using bars whose width represented the percentage of each type. He created a master chart by interdigitating each stratum from every site so as to represent all excavation cuts and the valley's entire ceramic sequence on a single graph.[8] He subsequently used this master graph to seriate all surface collections. All percentages, both from surface and excavated assemblages, were based on sherd count rather than minimum number of vessels. Ford (1949:37) recognized that the

https://doi.org/10.1017/aap.2016.2 Published online by Cambridge University Press

FIGURE 2. A large room at V-39, a typical site in the Queneto Quebrada.

FIGURE 3. Main platform of Huaca Santa Clara (V-67).

https://doi.org/10.1017/aap.2016.2 Published online by Cambridge University Press

quantity of sherds does not necessarily accurately represent the quantity of vessels since larger vessels will produce more sherds and thinner vessels will break more easily. He dismissed this issue because he thought that vessel size and thickness changed through time and that the percentage of types therefore remained valid for dating, and because he did not consider this potential inaccuracy to affect his primary goal of dating assemblages. Ford (1949:34–35) noted that efforts were made to ensure that surface ceramics were collected from a large area of each site so as to be representative of the total assemblage and not just of a small number of large pots. Some of the Virú Valley Project ceramic material is available at the project members' respective institutions, but an unspecified quantity of this material is effectively lost,[9] and it would therefore be difficult to convert the sherd count to another more reliable quantifier, such as minimum number of vessels (MNV). That said, excavated and surface ceramics alike tend to be highly fragmentary in the Virú Valley and, at present, there is no reason to think that there is a systematic difference in the sherd counts between excavated and surface collections.

From what I am able to tell, Ford never systematically compared the surface and excavated data from any single site and did not employ statistical tests to confirm the goodness of fit between different contexts. Rather, he felt that the surface was a reliable indicator of subsurface artifacts at shallow sites because these had very little accumulated refuse and because site formation processes acted differently at different parts of each site, such that a living floor representing an occupational period may be variously buried and exposed at different parts of the site (Ford 1949:35). He indicated that stratigraphic excavations were made at most sites with deep cultural stratigraphy. Moreover, although he does not say so explicitly, Ford appears to have come to the conclusion that surface ceramics reliably dated sites without deep stratigraphy by visually comparing the master excavation graph described above with percentage bar graphs created for each surface assemblage. Such a method can be productive, although it is more qualitative than strictly quantitative. Indeed, in the analysis presented below, I consider the percentage of types in various surface and excavated contexts but do so on a site-by-site basis, rather than in the valley-wide way that Ford appears to have done.

In addition to this more qualitative method, I statistically test the association between surface and subsurface assemblages at each site using the G-test, a statistic that uses the chi-square distribution (McDonald 2014a:54, 68; Sokal and Rohlf 1981). Originally developed for biostatistical applications, the G-test compares the observed frequencies of count data to expected frequencies and serves as a more robust form of the chi-squared test of independence, and is applicable to archaeological data. There are two variants of the G-test, the goodness-of-fit test and the G-test of independence. Both tests compare observed frequencies in an $R \times C$ table to expected frequencies and operate using the same equation, but the goodness-of-fit test uses theoretically derived expected frequencies whereas the test of independence derives expected frequencies from the dataset itself. The formula for the G-test is

$$G = 2 \sum_i O_i \cdot \ln\left(\frac{O_i}{E_i}\right)$$

The G-test of independence is used for all statistics throughout this article (Table 2). Calculations were made using a spreadsheet developed by McDonald (2014b). The test calculation produces a statistic, G, with degrees of freedom equal to the number of rows minus one, multiplied by the number of columns minus one.[10] The spreadsheet automatically compares this statistic to the chi-square distribution to produce a p-value; any p-value of less than .05 was considered significant. Significant p-values indicate that the observed and expected frequencies do not match making it likely that the contexts are independent and that the surface assemblage is not simply a random sample of the subsurface artifacts. Ford's hypothesis cannot be upheld in such cases.

In all cases, I compare the surface ceramic assemblages to the entire subsurface assemblage (i.e., all levels added together from a single stratigraphic cut). If Ford's (1949) hypothesis that a surface assemblage represents the mean cultural date of the entire site is to be upheld, there will be no statistically significant differences between the surface and excavated assemblages. In cases where data is available for multiple excavation levels, I test the surface against both the upper and lower levels with the hypothesis that, if the surface assemblage is skewed toward a later date in time, then it will be more similar to upper levels than to lower. These tests are performed for each site individually below. As will be seen below, the results from the G-test are inconsistent, resulting in varying acceptance and rejection of Ford's hypothesis. However, qualitative comparison of frequency bar graphs between surface and excavated contexts shows that surface types do line up with excavation levels moderately well, except at sites with very deep stratigraphy, and this means that an archaeologist would likely arrive at a similar relative date for both contexts in a ceramic seriation, despite the lack of statistically significant correlations between the differing contexts.

V-39

V-39 is a small rock-walled village in the Queneto Quebrada that Willey (1953) dated to the Middle and Late Virú periods. Ford excavated a small trench 20 cm into Room 1 at this site; excavation data are from Ford's archival notes for the site. Ford did not describe the geological or archaeological context of this trench. See Figure 4a for the domestic ware types present at the site and for their distribution. The subsurface and surface assemblages are significantly different ($G = 80.604$; $df = 3$; $p < .001$; Guañape and HPP eliminated due to small numbers); in this case, the surface distribution actually skews toward the earlier Castillo type, rather than the Late Plainware that is abundant in excavations. That the surface dates earlier than the upper layers of excavation runs contrary to what should be expected, and it is not clear why this is so. Ford's hypothesis that the surface assemblage of a site represents its mean cultural date cannot be supported at V-39 because earlier types are actually overrepresented on the surface.

V-44

V-44 is a rock-walled site at the base of the large, rocky Queneto Quebrada that Willey (1953:317) dated to the La Plata period and interpreted as being a community center with a residential settlement. Ford excavated a small trench at this site, which he

https://doi.org/10.1017/aap.2016.2 Published online by Cambridge University Press

TABLE 2. Summary of Site Context and Results of G-test Comparison between Surface and Successive Excavation Units.

Site	Site Type	Geomorphic Context	Surface vs.	Level Depth	G =	df =	p =
V – 39	Rock Wall Houses	Alluvial Plain	Room 1	0–20 cm	80.604	3	< .001
V – 44	Rock Wall Houses	Alluvial Plain	Cut 1	Undefined	4.027	3	.259
V – 46	Rock Wall House	Alluvial Plain	Levels 1–2	0–20 cm	61.261	3	< .001
			Levels 3–5	20–50 cm	134.198	3	< .001
			Excavation Total	0–50 cm	1.215	1	.27
V – 67	Adobe Pyramid/Castillo Fortification Complex	Hilltop	Excavation Total	Various	597.76	4	< .001
V – 108 Cut A	Adobe Pyramid Mound	Valley Floor	Levels 1–2	0–50 cm	52.362	3	< .001
			Levels 3–5	50–125 cm	35.413	2	< .001
			Excavation Total	0–125 cm	69.282	3	< .001
V – 167	Adobe Pyramid/Earth Refuse	Valley Floor	Levels 1–2	0–50 cm	9.906	3	.01938
			Levels 3–4	50–100 cm	16.627	3	< .001
			Levels 5–6	100–150 cm	29.203	3	< .001
			Levels 7–8	150–200 cm	19.652	3	< .001
			Levels 9–10	200–250 cm	16.707	3	< .001
			Levels 11–12	250–300 cm	34.139	3	< .001
			Levels 13–14	300–350 cm	Eliminated due to small sample size		
			Excavated Total	0–350 cm	16.576	3	< .001
V – 171 Cut A	Rectangular Adobe Compound	Valley Floor	Levels 1–3	0–75 cm	47.376	2	< .001
			Levels 4–6	75–150 cm	54.834	2	< .001
			Levels 7–10	150–250 cm	204.504	2	< .001
			Cut A Total	0–250 cm	59.907	2	< .001
V – 171 Cut B			Level 1	0–40 cm	2.808	2	.246
			Levels 2–3	40–90 cm	13.023	2	.0015
			Levels 4–6	90–165 cm	140.897	2	< .001
			Levels 6–16	140–415 cm	Eliminated due to small sample size		
			Cut B Total	0–415 cm	90.100	4	< .001
V – 171 Cut C			Levels 1–3	0–100 cm	.259	2	< .001
			Levels 4–6	100–175 cm	53.294	3	< .001
			Levels 7–12	175–325 cm	Eliminated due to small sample size		
			Cut C Total	0–325 cm	257.399	5	< .001

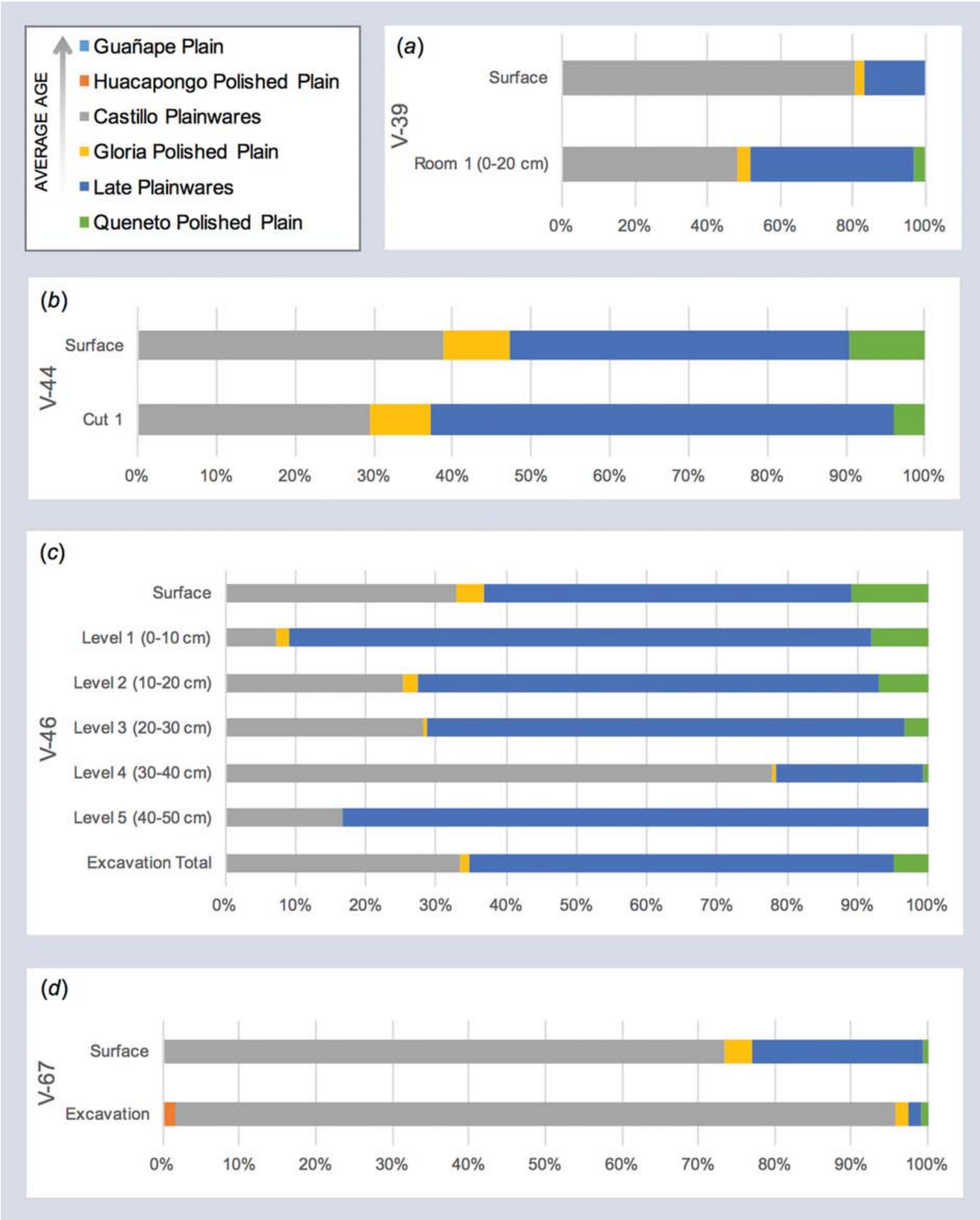

FIGURE 4. Surface and excavated ceramic distribution for sites V-39, V-44, V-46, and V-67 (Huaca Santa Clara). All graphs are 100 percent stacked bar graphs, with the percentage of each type represented by the size of its respective bar. Types are standardized across all graphs and are arranged from Guañape Plain (leftmost bar, when present) to Queneto Polished Plain (rightmost bar, when present) following the approximate sequence by which types were introduced into the Virú Valley sequence.

https://doi.org/10.1017/aap.2016.2 Published online by Cambridge University Press

called Cut 1; excavation data are available in Ford's field notes. This trench was cut across the south wall of the building and cleared to floor level, but Ford did not provide additional details about excavation contexts (.F673, Papers of James Ford, Division of Anthropology, AMNH). See Figure 4b for the domestic ware types present at the site and for their distribution. For V-44, the surface and subsurface distributions are quite similar; indeed, there is no significant difference between them ($G = 4.027$; $df = 3$; $p = .259$). In this case, Ford's hypothesis is supported, and surface ceramics are a good proxy for subsurface ceramics; however, the excavated artifact quantity is quite small and may be from only one level.

V-46

This small site in the Queneto Quebrada consists of rock-walled enclosures built on terraces. Willey (1953:300) dated this site to the La Plata Period. Ford excavated a small test trench extending 50 cm into this site (Figure 4c); excavation data are from Ford's field notes. Ford excavated five arbitrary 10-cm-thick levels at this site (.F673, Papers of James Ford, Division of Anthropology, AMNH). Level 0–10 cm was in an ash-filled room just below two looted burials. Level 10–20 cm cut through rubbish that largely consisted of ash, and Ford noted that some material from looted graves may be present at this level. Level 20–30 cm consisted of ash and a large quantity of sherd material, and a fire pit with charcoal was uncovered. The ash fill layer ended in level 30–40 cm and was followed by rock and soil fill to create a level room since this site was built on a hillside. Finally, excavation was stopped at level 40–50 cm, which was primarily soil fill; Ford noted that an ash deposit near the center of the room was chosen to extend the cut, but I cannot find any additional mention of this or any data associated with it. For the current analysis, I compare the surface assemblage to the entire excavated assemblage, to the top 20 cm of excavation, and to the lower 30 cm of excavation. For Ford's hypothesis to be supported, the surface assemblage needs to be similar to the entire excavated assemblage.

The surface is significantly different from the aggregate total of all excavation units ($G = 63.632$; $df = 3$; $p < .001$; Guañape and Huacapongo Polished Plain types eliminated due to small numbers). This indicates that Ford's hypothesis that surface ceramics represent the mean cultural date of the site cannot be supported. The surface ceramic assemblage is also significantly different from the upper levels (Levels 1–2 combined; $G = 61.261$; $df = 3$; $p < .001$) and from the lower levels (Levels 3–5 combined; $G = 134.198$; $df = 3$; $p < .001$). That these differences are significant is a little surprising, however, considering the distribution of ceramic types (Figure 4c). In overall distribution, the surface appears to be quite similar to the total excavation distribution; this likely led Ford to conclude that the surface was a good proxy for the excavated levels. The statistically significant differences cannot be overlooked, but they are likely the result of two minor types, Gloria and Queneto, appearing in higher numbers than expected; when these types are added to their parent type (Castillo and Late Plainware, respectively), the distribution between the surface and the aggregate total of all excavations is not significant ($G = 1.215$; $df = 1$; $p = .27$). This result indicates that Ford's hypothesis can be supported for V-46.

Huaca Santa Clara (V-67)

Huaca Santa Clara is one of six *castillos*, natural hills topped with large ceremonial, administrative, and military structures located in the middle valley (Downey 2014; Millaire 2004, 2009b). Huaca Santa Clara lies on a large isolated hill on the valley floor in a strategic location and consists of four large adobe platforms, one on top of the hill and the other three on natural terraces. Willey (1953:225–226) dated the site to the Huancaco period based on Ford's surface collection, and this site was thought to be a key location of the supposed Moche takeover of the Virú Valley. Millaire (2004, 2009b) conducted extensive excavations within various architectural contexts at the site and found that it was actually a large Virú administrative center and town with minor occupations in the Huacaco and later periods. Millaire found evidence not of a violent Moche conquest at this site but, rather, of a more gradual transition from Virú to Huancaco. Huaca Santa Clara is an ideal location to test Ford's hypothesis because it has large ceramic samples from both surface and excavated contexts.

Excavation strategies at Huaca Santa Clara were designed to obtain samples from across the site rather than focus on deep stratigraphic trenches and, for this reason, I consider it most useful to analyze all excavation data together. There is a much larger proportion of Late Plainwares on the surface of Huaca Santa Clara compared to the aggregate total of all excavated contexts (Figure 4d).[11] The two contexts are significantly different ($G = 597.76$; $df = 4$; $p < .001$). This result indicates that, in the case of Huaca Santa Clara, the surface ceramic assemblage is not a good proxy for the subsurface ceramics. Instead, Late Plainwares are overrepresented on the surface; this is logical, as later ceramics should be more common than earlier ones on the surface of an undisturbed site, but this pattern is not observed at all sites. Ford's hypothesis is not upheld for Huaca Santa Clara because the surface assemblage represents late types in far greater quantity than excavated frequencies predict.

V-108

V-108 is a rectangular adobe-brick structure located in the sandy flats of the lower Virú Valley, 1 km from the coast. Collier (1955:30, Table 1) excavated two cuts (A and B) at this site and dated it to the La Plata Period. I only include Cut A, a 1.5-m-×-3-m trench cut into a refuse mound. The top 30 cm consisted of shell and faunal remains in a sand and ash matrix, followed by 60 cm of a sandy fill with adobe rubble and soil with less shell and faunal remains than above (Collier 1955:30–31). This sandy fill layer continued from 60–120 cm but with fewer inclusions, and a sterile windblown sand layer extended from a depth of 120 cm to the limit of excavations at a depth unspecified by Collier.

Based on the distribution graph (Figure 5a), it appears that the surface of V-108 is not very similar to any level of the excavation. Indeed, the surface is significantly different from the excavation total ($G = 69.282$; $df = 3$; $p < .001$). The surface is also significantly different from levels 1–2 combined ($G = 52.362$; $df = 3$; $p < .001$) and from levels 3–5 combined ($G = 35.413$; $df = 2$; $p < .001$). Although the earlier ceramic Castillo type is present at this site in much lower quantities than Late Plainwares, it is more abundant on the surface than it is in any excavation level, an unexpected pattern and the opposite of what was observed at Huaca Santa Clara. It is possible that there is a slightly earlier

https://doi.org/10.1017/aap.2016.2 Published online by Cambridge University Press

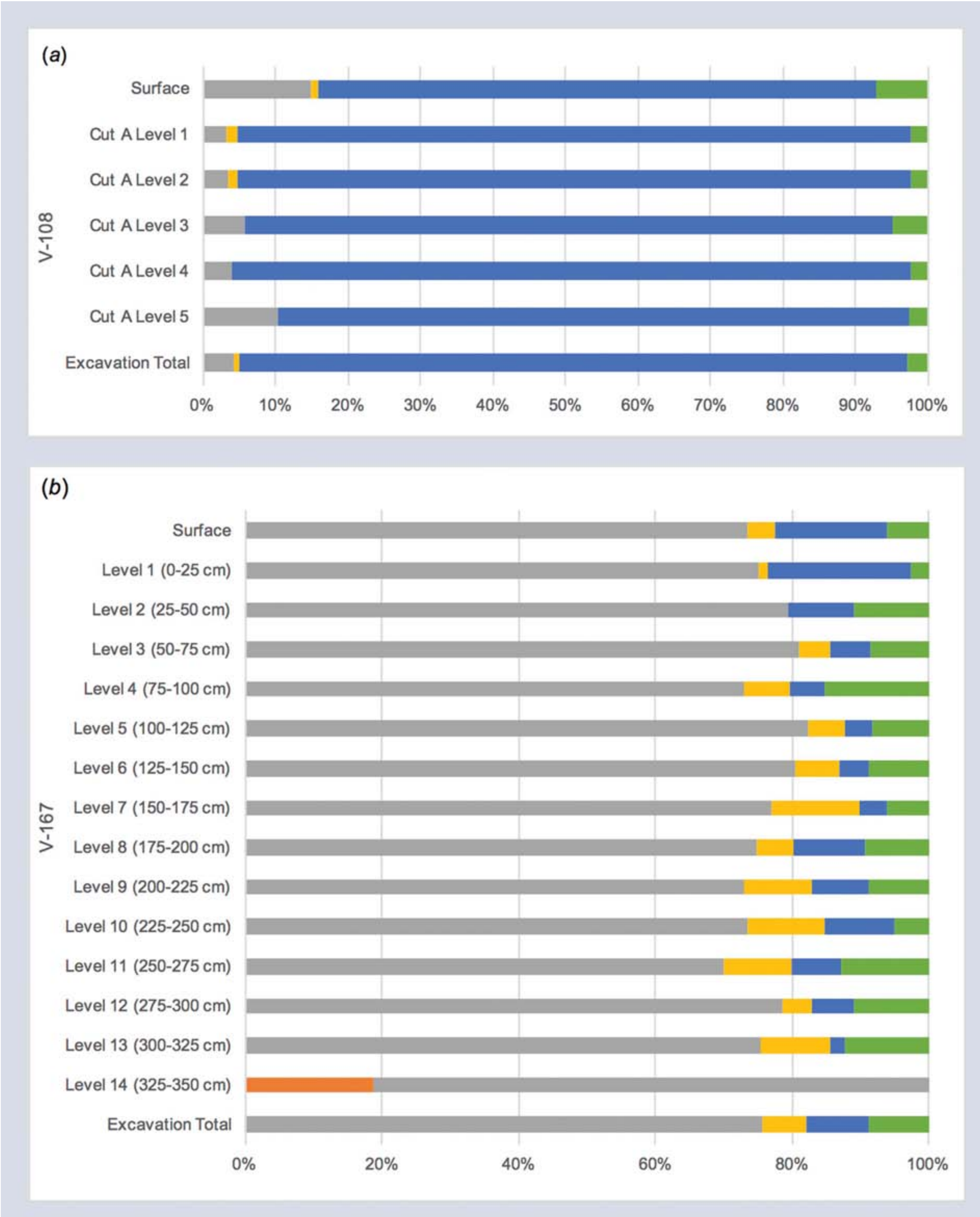

FIGURE 5. One hundred percent stacked bar graphs showing the surface and excavated ceramic distributions for sites V-108 and V-167. All graphs use same legend and ordering as Figure 4.

https://doi.org/10.1017/aap.2016.2 Published online by Cambridge University Press

component at this site that was detected in surface collections but was not detected by the single excavation trench compared here; Ford (1949) made two surface collections at this site—which I have combined here—and presumably covered its full extent. Nevertheless, while the differences observed here are statistically significant, they are minor, and the site primarily contains Late Plainwares and the surface and all excavation levels date securely to late periods.

V-167

V-167 is a large, low earthen mound located in the lower valley on the south side of the river, near the large site of Huancaco. Collier (1955:55, Table 8) described the site as consisting of two small pyramid mounds and one refuse mound. Collier excavated one 3-m-×-4-m trench (Cut A) in the center of the refuse mound and dated all but the lowest layers to the Tomaval Period[12] (Figure 5b). Collier did not detail the archaeological or geological contexts of this cut but noted that midden refuse and ceramic sherds were abundant to very abundant in levels 1–12 and rare in levels 13–14, while a 75-cm-deep test pit below layer 14 exposed only a compact and mottled brown clay layer that was sterile (Collier 1955:55-57). Collier did note that an adobe wall with no apparent floor was encountered between a depth 50 cm and 150 cm and that a thin (2–3 cm) layer of shell and charcoal was encountered at a depth of 265 cm in layer 11, a layer that contained a large number of sherds.

While the surface distribution appears to be similar to the aggregate total of all excavated contexts, they are significantly different ($G = 16.576$; $df = 3$; $p < .001$), and Ford's hypothesis is not supported. This site is ideal for comparing the surface distribution against the distribution from specific excavation levels to test whether the surface is more similar to upper levels and less similar to lower ones (Table 2); indeed, this is the case, but the surface is significantly different from all excavated levels, although the difference is not large between the surface and the uppermost levels. In the case of V-167, the surface assemblage is a relatively good proxy for the upper levels of the site but is not a good proxy for the lower levels. Once again, although the surface and excavated distributions are significantly different when measured statistically, they are broadly similar; Castillo sherds dominate both the surface and subsurface assemblages.

V-171

V-171 is a large rectangular compound in the lower valley, south of the river. Collier (1955:49–52, Tables 5–7) excavated three trenches here and found evidence that the site had been in continual use from Guañape times through to the end of the Estero Period. Following Ford (1949), Collier considered the surface ceramic assemblage at the site to date to the Estero Period. Collier's three stratigraphic excavation cuts offer perhaps the best opportunity in Virú to examine the relationship between surface and subsurface ceramic assemblages because they follow the pattern that is expected for an undisturbed, long-term occupation, where earlier types become gradually more prominent in the lower levels of excavation. Still, it is necessary to test this pattern against Ford's surface collection from the site. I test the surface assemblage against each of Collier's three cuts separately because there is no reliable way to merge them into one.

Collier described the archaeological and geological strata at this site in greater detail than at the other sites discussed in this article. He also provided accurate profile drawings for each cut, which was not done for the other sites; these should be consulted for a thorough understanding of the contexts at V-171. Cut A is a 2-m-×-4-m trench in the center of a small mound that rose 1.4 m above the valley floor and was cut adjacent to a tapia wall (Collier 1955:49). The upper 140 cm of excavation was fill but contained three strata of dark earth, with a layer of adobe rubble encountered at a depth of 75 cm. The tapia wall ended at a depth of 140 cm—level with the valley floor—with no evidence of a floor. Cultural material continued to be encountered in a matrix of mottled yellow-brown clay fill that transitioned to a gray-brown sand until a depth of 250 cm. The excavation was continued to a depth of 325 cm, but no further cultural strata were encountered. Cuts B and C were carried out in the central mound of the site, which was a 15-m-×-40-m mound that rose 1.1 m above the valley floor (Collier 1955:52). Both had similar stratigraphy that consisted of alternating layers of brown earth (not further described but likely a loam), yellow-brown clay, and sandier soils (Collier 1955:Figure 23). These layers were of varying compaction and sometimes mottled together. Cut B was 2 m × 4 m and extended to a depth of 415 cm. Ceramics were abundant to a depth of 140 cm and became sparser in deeper levels until almost none were encountered in levels 15–16 (365–415 cm). Collier considered the lower 250 cm of fill to have accumulated very slowly over several major periods of occupation, leading to the overlap of ceramic styles from different periods. A small looter's pit cut into levels 1–3 (to a depth of about 75 cm), but Collier felt that it did not displace ceramics within these levels. Cut C, a 3-m-×-6-m trench that was excavated to a depth of 450 cm, was located just south of cut B in order to confirm the cultural stratigraphy of cut B (Collier 1955:52–54). The top 125 cm contained abundant ceramic material, but ceramics were sparser below with the exception of level 8 (200–225 cm), which contained abundant sherds. No ceramics were encountered below 300 cm, although evidence of human occupation in the form of shells, ash, and burned clay continued to level 15 at a depth of 400 cm. The soils became very wet below 375, and Collier estimated that the water table was no more than a meter below the limit of excavation.

Collier's Cut A at V-171 is, at first appearance, the most similar to Ford's surface collection, but the surface is significantly different from the subsurface total (Figure 6a; see Table 2 for all statistics). The surface is also significantly different from each of the levels of Cut A. Cut B shows a considerably different ceramic distribution in the lower levels than in the upper levels (Figure 6b), and it is therefore not surprising that the surface is significantly different from the excavation total for the cut ($G = 90.100$; $df = 4$; $p < .001$). Viewing the frequency distribution, the surface appears most similar to the upper levels of the cut; indeed, the surface distribution is not significantly different from the level 1 distribution ($G = 2.808$; $df = 2$; $p = .246$; Guañape, HPP, and Gloria excluded due to low numbers), but is significantly different from levels 2–3 combined ($G = 13.023$; $df = 2$, $p = .0015$) and from levels 4–6 combined ($G = 140.897$; $df = 2$; $p < .001$). The surface distribution could not be tested against the lower excavation levels due to small sample numbers, but it is readily apparent that the ceramic distribution in the lower levels are remarkably different from the surface distribution (with a much higher percentage of earlier types in the lower levels). The distribution of Cut C is similar to Cut B (Figure 7). The aggregate total for Cut C

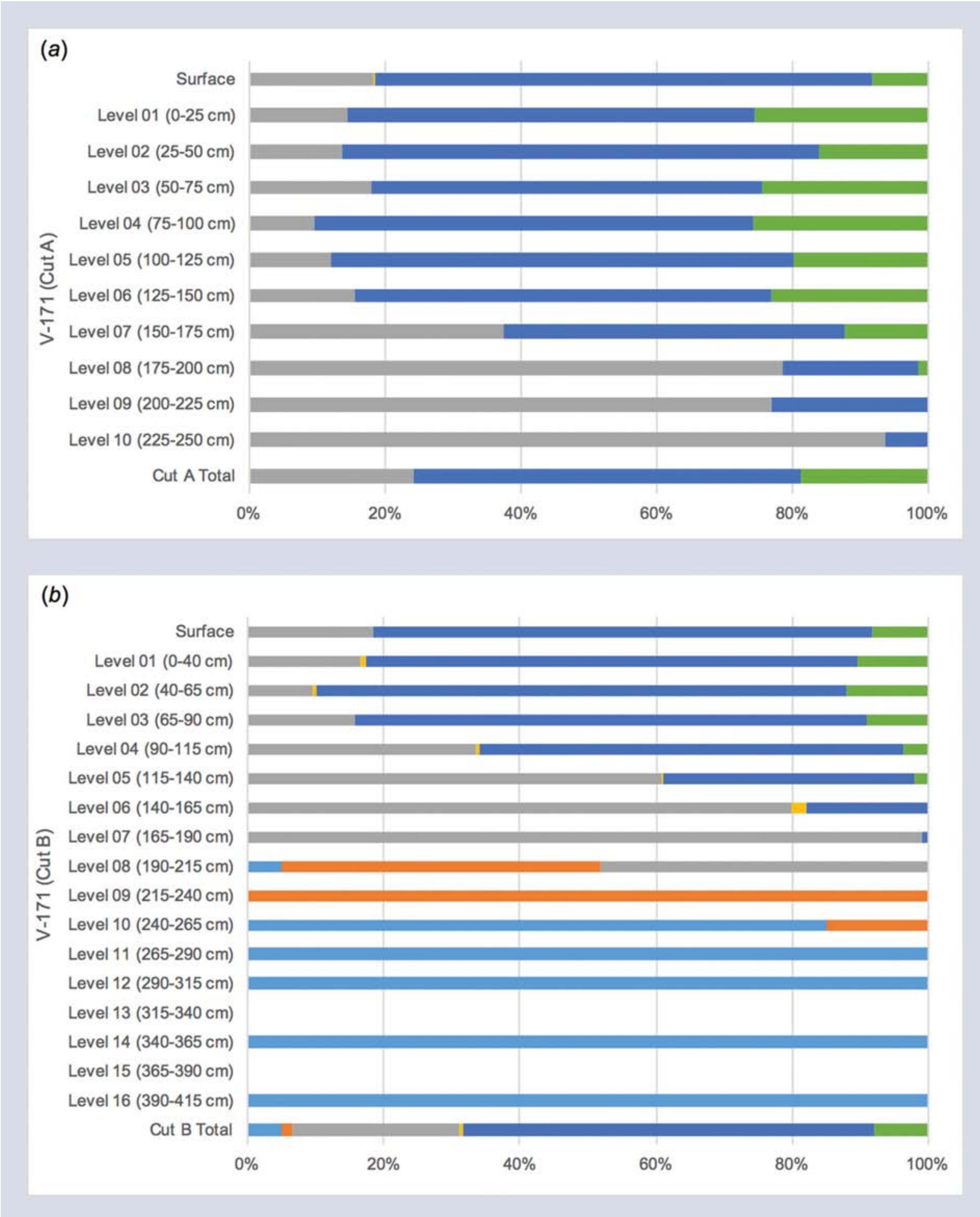

FIGURE 6. One hundred percent stacked bar graphs showing the surface and excavated ceramic distributions for cuts A and B at site V-171. All graphs use same legend and ordering as Figure 4.

https://doi.org/10.1017/aap.2016.2 Published online by Cambridge University Press

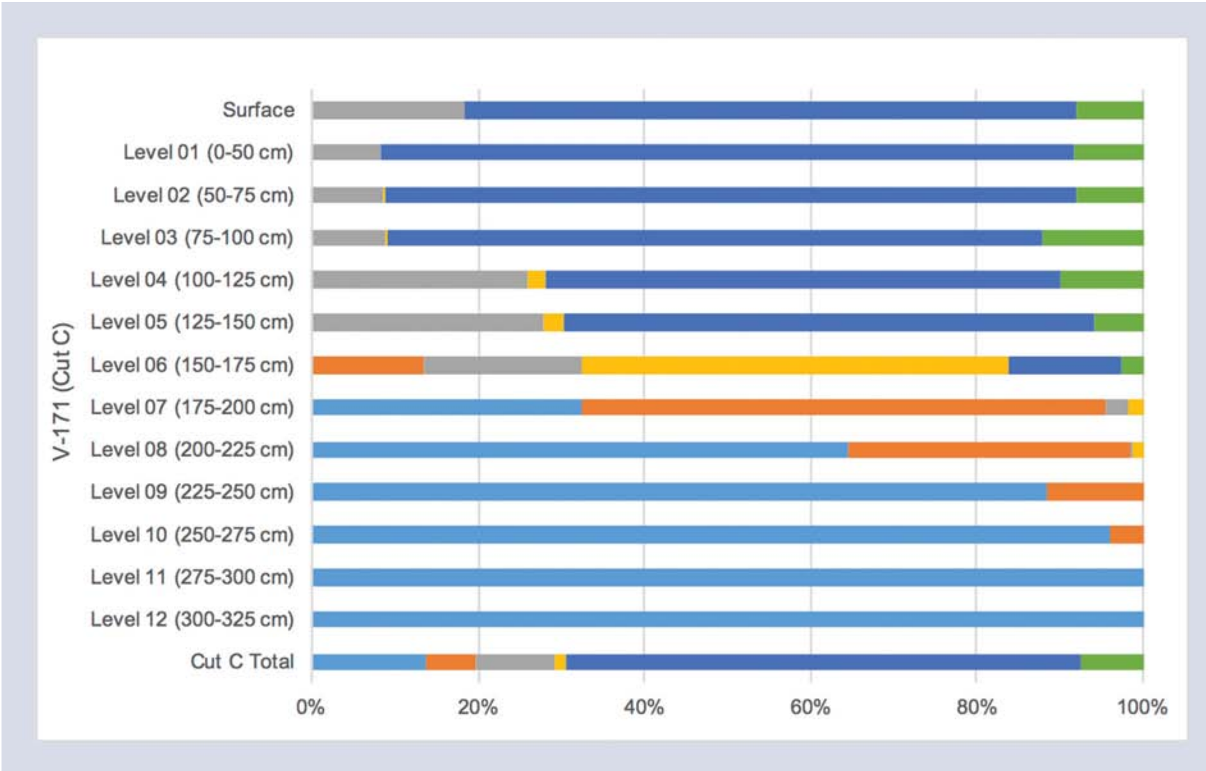

FIGURE 7. One hundred percent stacked bar graphs showing the surface and excavated ceramic distributions for cut C at V-171. Graph uses same legend and ordering as Figure 4.

is significantly different from the surface collection for the site ($G = 257.399$; $df = 5$; $p < .001$), from levels 1–3 combined ($G = .259$; $df = 2$; $p < .001$; Guañape, HPP, and Gloria excluded), and from levels 4–6 combined ($G = 53.294$, $df = 3$; $p < .001$; Guañape and HPP excluded). As with Cut B, the lower levels of Cut C cannot be tested using the G-test or the Chi-Square test, but these levels contain much higher proportions of earlier ceramics than the surface does.

In summary, in nearly all cases, the surface distribution of V-171 is significantly different from the entire subsurface distribution of any excavation cut, therefore refuting Ford's (1949) hypothesis that the surface assemblage represents a mean measure of the subsurface assemblage. However, the surface assemblage is also significantly different from nearly *any* level of excavation. While this would suggest that the surface assemblage is not a good proxy for the subsurface assemblage at this site, when considering the actual distribution of each cut, the surface assemblage dates to roughly the same time as the upper levels of excavation. A large site like V-171 is complex and was surely built and occupied in stages; Cut A is excavated in a different mound from Cuts B and C, and the former appears to date much later and to have been used for a shorter period of time than the latter two (Collier 1955). The surface assemblage cannot detect these nuances in the occupational history of a site; nor does it appear to detect in any way the fact that there were much earlier (Guañape and Early Virú) occupations at the site.

DISCUSSION AND CONCLUSIONS

Ford (1949) developed the hypothesis that surface ceramic assemblages represent the mean ceramic date of the site. This hypothesis is not entirely supported; in almost all cases, the surface of any site was significantly different both from the average subsurface distribution (calculated by summing together all excavated levels) and from any individual level of excavation. Still, it is not possible to reject this hypothesis outright because significant correlations do exist in several cases. Moreover, despite the lack of statistically significant differences in most cases, the frequency distribution of types across the different contexts follows an expected pattern: the distribution of types on a site's surface is generally similar to the distribution of types from all excavated contexts at shallow sites and similar to the uppermost layers of deeper sites.

This finding has significant implications for archaeological dating and particularly for the archaeology of regions. Surface artifacts are commonly used to date archaeological sites through seriation (Lyman and O'Brien 2006). Seriation methods rely on the ratio of artifact types or attributes in an archaeological unit of interest (e.g., site, excavation block, etc.) to establish relative dates for the unit. Typological sequences and regional seriation sequences have typically been established for many decades such that an archaeologist can be expected to date a site based on a few diagnostic types and their a priori knowledge of the local artifact

https://doi.org/10.1017/aap.2016.2 Published online by Cambridge University Press

sequence. For regional survey, this often means surface data. The results presented here highlight the need for caution. An archaeologist may well uncritically date a surface site based on the proportion of types he or she finds on the surface, and a comparison with the proportion of types in excavated contexts would confirm this result: based on the frequency of types alone, the surface seems to be a reasonably good proxy for the upper ~100 cm of subsurface remains, at least at sites whose formation processes match those described here. At the sites presented here the surface and near-surface excavation levels would be dated to the same time period if dated by their type proportions. In other words, qualitatively the surface and subsurface match, except at sites with very deeply buried levels.

Nevertheless, the results here show that the surface distribution of sites does not closely match subsurface distributions in a quantitative or statistically significant sense. Typically, the surface distribution at the sites examined here is different from all excavation levels combined (which can be considered the overall site average) but is also different from most individual excavation levels except for the uppermost levels at a few sites. Can we trust a seriation that is based entirely on the ratio of different types of ceramic found on a site's surface? While not often acknowledged, several authors have recognized the problems inherent in dating ceramic assemblages and have devised methods for obtaining more accurate results from surface ceramic artifacts (Christenson 1994; Ortman et al. 2007; Roberts et al. 2012; Steponaitis and Kintigh 1993). Bayesian analysis (Ortman et al. 2007) offers the greatest promise for deriving accurate dating results from surface ceramic assemblages. Bayesian analysis uses prior knowledge of a distribution to create a probabilistic model that, in its archaeological application, can determine the probability that an observed assemblage fits into the prior known sequence. The Virú Valley is the ideal location to seriate sites using Bayesian analysis because of the large and standardized dataset collected by the members of the Virú Valley Project: the known distribution of types can be derived from stratigraphic trenches, and Bayesian analysis can then be used accurately to seriate surface assemblages. The present article highlights the need for such an analysis because it shows that type-proportion seriation methods are not preferable because surface assemblages fail to accurately represent subsurface assemblages and therefore fail to accurately date the site's occupation period.

Interestingly, at shallow rock-walled sites in the Queneto Quebrada (V-39, V-44, V-46), the surface distribution actually dates slightly earlier than the subsurface distribution, and the surface of V-108 and V-171—both deeply stratified sites—dates somewhat earlier than the upper levels of excavation. The surface distribution at V-167 is roughly similar to all subsurface levels, while the lower levels of V-171 contain large quantities of earlier wares that do not appear on the surface of that site. The Huaca Santa Clara (V-67) surface assemblage is skewed toward later periods, as can be expected because the ceramics from later periods should be more common on the surface than the ceramics from earlier periods. These discrepancies highlight an important point: site context must always be taken into account. Not only are Virú sites subjected to varying site formation processes, looting is also extremely common in the valley (Contreras 2010). Looting was common even when Ford conducted his survey, as seen in his field notes and in site descriptions published by Willey (1953), and looting has continued unabated in the decades since. Among other problems, looting impacts surface collections because buried artifacts from looter's pits are frequently scattered on a site's surface such that the area around a pit may have a distribution of artifacts that dates earlier than undisturbed parts of a site, even if the different parts of the site were abandoned at the same time. This may account for the cases in Virú where a site's surface appears to date earlier than excavated levels. It is difficult to suggest generalized solutions to this problem other than to emphasize that site context can never be neglected when an archaeologist decides whether to accept or reject the relative date of a site based on the seriation of artifacts collected from its surface. This highlights another problem with using surface data for traditional seriation techniques because much of the data may be residual, and types from different levels and occupations may be comingled on the surface.

Beyond the merits of this analysis as a case study of a single coastal Andean valley, it should also serve as a cautionary tale of the limits of using the seriation of surface data to date sites without excavation. In the cases presented here, there is no consistent skew in the dates indicated by ceramic types, and there is no way to account for the variability. Although I contend that surface ceramic assemblages are a reasonable proxy for subsurface contexts in the Virú Valley, because the approximate date indicated by surface ceramics is generally similar to the approximate date indicated by excavated contexts, this is not necessarily the case in other regions, and the surface of a site cannot be reliably trusted to indicate the true age of that site without test excavations. There is greater reliability if one limits their dating of a site to very general time periods, such as the centuries-long traditions used here, than if one attempts to date a site to specific periods or sub-periods based solely on surface finds.

This study tested the hypothesis developed by Ford (1949) that ceramic collections made from the surface of sites in the Virú Valley of Peru represented the mean or average date of occupation at the site where the collection was made, except in rare cases of sites with very deeply buried layers. Employing a standardized typology and using both the G-test of independence and a visual comparison of the frequency of types, the distribution of ceramic artifacts between surface and subsurface or excavated contexts was compared from seven sites from the valley. Ford's hypothesis was not upheld in all cases, but there was no consistent skew or pattern in the data, making it difficult to reject outright Ford's hypothesis. In most cases, the distribution of ceramic types on the surface of a site showed statistically significant differences to the distribution of sites from the excavated contexts. Despite these differences, the frequency of types between the varying contexts was generally similar except at sites with deep stratigraphy, where the deepest layers dated earlier than the shallow layers and the surface. Because of this, were the surface and subsurface assemblages dated through seriation, they would date to the same general time period, with the exception of deeply buried assemblages. This study cautions against regional studies simply accepting a site date obtained through the seriation of artifacts collected from the site's surface without testing the reliability of such dates through excavation. This study also highlights the importance of considering site formation processes when dating sites through the seriation of surface artifacts.

Acknowledgments

I would first of all like to acknowledge the help, support, and ideas of Jean-François Millaire from the conception to the completion of this project. I am also grateful to Charles Spencer and the American Museum of Natural History for providing access to the papers of James Ford, which form the basis of this article. I also appreciate Dr. Spencer's comments on an earlier version of this article. The research discussed here was undertaken while I was a doctoral candidate at the University of Western Ontario and portions of this article were written while I was under the support of the Ontario Graduate Scholarship and the Western Graduate Research Scholarship. I would like to thank James Keron for his help and advice with statistics, as well as Theresa Topic, Lisa Hodgetts, Chris Ellis, and Micha Pazner for their guidance throughout this project and their comments on an earlier version of this article. Jose Gutierrez saved me much trouble by kindly translating the abstract into Spanish. I would like to thank three anonymous reviewers for their input, which significantly improved this article. Finally, I must acknowledge the support and friendship of Flannery Surette, who gave me many ideas throughout this project, and the love of Stacey Guy, whose support helped me to finish it. Permits were not required for the work conducted in this analysis.

Data Availability Statement

Excavation data from sites V-39, V-44, and V-46, as well as all data from surface contexts, are reproduced in a modified form from the archives of James Ford and are available on file at the American Museum of Natural History (.F673, Papers of James Ford, Division of Anthropology, AMNH). Excavation data from Huaca Santa Clara (V-67) is provided courtesy of Jean-François Millaire and is on file at the University of Western Ontario. Excavation data from sites V-108, V-167, and V-171 are reproduced from a published volume by Donald Collier (1955). In all cases these data have been modified using a condensed typology and are provided in supplemental materials attached to this article. These data, as well as the details on the condensed typology and additional data on corporate and domestic ware ceramics in the Virú Valley, are also available as supplemental files at http://ir.lib.uwo.ca/etd/2687/.

Supplementary Material

To view supplementary material for this article, please visit http://doi.org/10.1017/aap.2016.2.

REFERENCES CITED

Barrientos, Gustavo, Luciana Catella, and Fernando Oliva
2015 The Spatial Structure of Lithic Landscapes: The Late Holocene Record of East-Central Argentina as a Case Study. *Journal of Archaeological Method and Theory* 22:1151–1192.

Barton, C. Michael, Joan Bernabeu, J. Emili Aura, Oreto Garcia, Steven Schmich, and Lluis Molina
2004 Long-Term Socioecology and Contingent Landscapes. *Journal of Archaeological Method and Theory* 11(3):253–295.

Bennett, Wendell C.
1939 *Archaeology of the North Coast of Peru: An Account of Exploration and Excavation in Virú and Lambayeque Valleys.* Anthropological Papers of the American Museum of Natural History Vol. 37, No. 1. American Museum of Natural History, New York.
1950 *The Gallinazo Group, Virú Valley, Peru.* Yale University Press, New Haven, Connecticut.

Bennyhoff, James A.
1952 The Virú Valley Sequence: A Critical Review. *American Antiquity* 17:231–249.

Bevan, Andrew, and James Conolly
2009 Modelling Spatial Heterogeneity and Nonstationarity in Artifact-Rich Landscapes. *Journal of Archaeological Science* 36:956–964.

Billman, Brian R.
1996 *The Evolution of Prehistoric Political Organizations in the Moche Valley, Peru.* Ph.D. dissertation, Department of Anthropology, University of California, Santa Barbara.

Bintliff, John
2005 Human Impact, Land-Use History, and the Surface Archaeological Record: A Case Study from Greece. *Geoarchaeology* 20(2):135–147.

Christenson, Andrew L.
1994 A Test of Mean Ceramic Dating Using Well-Dated Kayenta Anasazi Sites. *Kiva* 59(3):297–317.

Collier, Donald A.
1955 *Cultural Chronology and Change as Reflected in the Ceramics of the Virú Valley, Peru.* Fieldiana: Anthropology Vol. 43. Chicago Natural History Museum, Chicago, Illinois.

Contreras, Daniel A.
2010 *Huaqueros* and Remote Sensing Imagery: Assessing Looting Damage in the Virú Valley, Peru. *Antiquity* 84:554–555.

Davies, Benjamin, Simon J. Holdaway, and Patricia C. Fanning
2015 Modelling the Palimpsest: An Exploratory Agent-Based Model of Surface Archaeological Deposit Formation in a Fluvial Arid Australian Landscape. *Holocene* 26(3):450–463.

Donnan, Christopher B.
2009 The Gallinazo Illusion. In *Gallinazo: An Early Cultural Tradition on the Peruvian North Coast*, edited by Jean-Françcois Millaire and Magali Morlion, pp. 17–32. Cotsen Institute of Archaeology Press, University of California, Los Angeles.

Downey, Jordan T.
2014 *Statecraft in the Virú Valley, Peru, in the First Millennium A.D.* Ph.D. dissertation, Department of Anthropology, University of Western Ontario, London, Canada.

Downey, Jordan T., and Jean-François Millaire
2015 Anchoring the Absolute to the Relative: Recent Chronological Research in the Virú Valley, Peru. Manuscript on file, Department of Anthropology, University of Western Ontario, London, Canada.

Dunnell, Robert C., and William S. Dancey
1983 The Siteless Survey: A Regional Scale Data Collection Strategy. *Advances in Archaeological Method and Theory* 6:267–287.

Fanning, Patricia, and Simon Holdaway
2001 Stone Artifact Scatters in Western NSW, Australia: Geomorphic Controls on Artifact Size and Distribution. *Geoarchaeology* 16(6):667–686.

Fernández-López de Pablo, Javier, and C. Michael Barton
2016 Bayesian Estimation Dating of Lithic Surface Collections. *Journal of Archaeological Method and Theory* 22:559–583.

Fogel, Heidy
1993 *Settlements in Time: A Study of Social and Political Development during the Gallinazo Occupation of the North Coast of Perú.* Ph.D. dissertation, Department of Anthropology, Yale University, New Haven, Connecticut. University Microfilms, Ann Arbor.

Ford, James A.
1949 Cultural Dating of Prehistoric Sites in Virú Valley, Peru. In *Surface Survey of the Virú Valley, Peru*, edited by James A. Ford and Gordon R. Willey, pp. 29–89. Anthropological Papers of the American Museum of Natural History Vol. 43, No. 1. American Museum of Natural History, New York.
1952 Reply to "The Virú Valley Sequence: A Critical Review." *American Antiquity* 17:250.

Ford, James A., and Gordon R. Willey
1949 Virú Valley: Background and Problems. In *Surface Survey of the Virú Valley, Peru*, edited by James A. Ford and Gordon R. Willey, pp. 11–28.

Anthropological Papers of the American Museum of Natural History Vol. 43, No. 1. American Museum of Natural History, New York.
Holdaway, Simon, Justin Shiner, and Patricia Fanning
2004 Hunter-Gatherers and the Archaeology of Discard Behavior: An Analysis of Surface Stone Artifacts from Sturt National Part, Western New South Wales, Australia. *Asian Perspectives* 43(1):34–72.
Koons, Michele, and Bridget Alex
2014 Revised Moche Chronology Based on Bayesian Models of Reliable Radiocarbon Dates. *Radiocarbon* 56(3):1039–1055.
Larco Hoyle, Rafael
1945 *La cultura Virú*. Museo Rafael Larco Herrera, Chicliń, Peru.
Lewarch, Dennis E., and Michael J. O'Brien
1981 The Expanding Role of Surface Assemblages in Archaeological Research. *Advances in Archaeological Method and Theory* 4:297–342.
Lyman, R. Lee, and Michael J. O'Brien
2006 *Measuring Time with Artifacts: A History of Methods in American Archaeology*. University of Nebraska Press, Lincoln.
McDonald, John H.
2014a *Handbook of Biological Statistics*. 3rd ed. Sparky House Publishing, Baltimore.
2014b G-test of Independence. Electronic document, http://www.biostathandbook.com/gtestind.html, accessed May 1, 2016.
Millaire, Jean-François
2004 Gallinazo - Moche Interactions at Huaca Santa Clara, Virú Valley (North Coast of Peru). Paper presented to the 69th Annual Meeting of the Society for American Archaeology, Montreal.
2009a Gallinazo and the *Tradición Norcosteña*. In *Gallinazo: An Early Cultural Tradition on the Peruvian North Coast*, edited by Jean-Françcois Millaire and Magali Morlion, pp. 1–16. Cotsen Institute of Archaeology Press, University of California, Los Angeles.
2009b Woven Identities in the Virú Valley. In *Gallinazo: An Early Cultural Tradition on the Peruvian North Coast*, edited by Jean-François Millaire and Magali Morlion, pp. 149–165. Cotsen Institute of Archaeology Press, University of California, Los Angeles.
2010 Moche Political Expansionism as Viewed from Virú: Recent Archaeological Work in the Close Periphery of a Hegemonic City-State System. In *New Perspectives on Moche Political Organization*, edited by Jeffrey Quilter and Luis Jaime Castillo, pp. 223–251. Dumbarton Oaks, Washington, D.C.
Millaire, Jean-François, and Edward Eastaugh
2011 Ancient Urban Morphology in the Virú Valley, Peru: Remote Sensing Work at the Gallinazo Group (100 B.C.–A.D. 700). *Journal of Field Archaeology* 36(4):289–297.
2014 Geophysical Survey on the Coast of Peru: The Early Prehispanic City of Gallinazo Group in the Virú Valley. *Latin American Antiquity* 25:239–255.
Mills, Barbara J, Jeffery J. Clark, Matthew A. Peeples, W. Randall Haas Jr., John M. Roberts Jr., J. Brett Hill, Deborah L. Huntley, Lewis Borck, Ronald L. Breiger, Aaron Clauset, and M. Steven Shackley
2013 Transformation of Social Networks in the Late Pre-Hispanic US Southwest. *Proceedings of the National Academy of Sciences* 110(15):5785–5790.
Moseley, Michael E.
2001 *The Incas and Their Ancestors: The Archaeology of Peru*. 2nd ed. Thames and Hudson, New York.
O'Brien, Michael J., and R. Lee Lyman
1999 *Seriation, Stratigraphy, and Index Fossils: The Backbone of Archaeological Dating*. Kluwer Academic/Plenum Publishers, New York.
Ortman, Scott G.
2016 Uniform Probability Density Analysis and Population History in the Northern Rio Grande. *Journal of Archaeological Method and Theory* 23:95–126.
Ortman, Scott G., Mark D. Varien, and T. Lee Gripp
2007 Empirical Bayesian Methods for Archaeological Survey Data: An Application from the Mesa Verde Region. *American Antiquity* 72:241–272.
Quilter, Jeffrey
2014 *The Ancient Central Andes*. Routledge, London.
Roberts, John. M., Jr., Barbara J. Mills, Jeffery J. Clark, W. Randall Haas Jr., Deborah L. Huntley, and Meaghan Trowbridge
2012 A Method for Chronological Apportioning of Ceramic Assemblages. *Journal of Archaeological Science* 39:1513–1520.
Sherman, R. Jason, Andrew K. Balkansky, Charles S. Spencer, and Brian D. Nicholls
2010 Expansionary Dynamics of the Nascent Monte Albán State. *Journal of Anthropological Archaeology* 29:278–301.
Sokal, Robert R., and F. James Rohlf
1981 *Biometry: The Principles and Practice of Statistics in Biological Research*. 2nd ed. W.H. Freeman, New York.
South, Stanley
1977 *Method and Theory in Historical Archaeology*. Academic Press, New York.
Steponaitis, Vincas P., and Keith W. Kintigh
1993 Estimating Site Occupation Spans from Dated Artifact Types: Some New Approaches. In *Archaeology of Eastern North America: Papers in Honor of Stephen Williams*, edited by Stephen Williams and James B. Stoltman, pp. 349–361. Archaeological Report 25. Mississippi Department of Archives and History, Jackson, Mississippi.
Strong, William D., and Clifford Evans Jr.
1952 *Cultural Stratigraphy in the Virú Valley Northern Peru: The Formative and Florescent Epochs*. Columbia University Press, New York.
Strong, William D, Clifford Evans Jr., and Rose Lilien
1952 Appendix I: Description of Pottery Types. In *Cultural Stratigraphy in the Virú Valley Northern Peru: The Formative and Florescent Epochs*, by William D. Strong and Clifford Evans Jr., pp. 253–351. Columbia University Press, New York.
Wandsnider, LuAnn
1992 Archaeological Landscape Studies. In *Space, Time, and Archaeological Landscapes*, edited by Jacqueline Rossignol and LuAnn Wandsnider, pp. 285–292. Springer Science+Business Media, New York.
2004 Solving the Puzzle of the Archaeological Labyrinth: Time Perspectivism in Mediterranean Surface Archaeology. In *Side-by-Side Survey: Comparative Regional Studies in the Mediterranean World*, edited by Susan Alcock and John F. Cherry, pp. 49–62. Oxbow, Oxford.
Whalen, Michael E., and Paul E. Minnis
2012 Ceramics and Polity in the Casas Grandes Area, Chihuahua, Mexico. *American Antiquity* 77:403–423.
Wilkinson, Tony J.
1989 Extensive Sherd Scatters and Land-Use Intensity: Some Recent Results. *Journal of Field Archaeology* 16(1):31–46.
Willey, Gordon R.
1953 *Prehistoric Settlement Patterns in the Virú Valley, Perú*. Smithsonian Institution Bureau of American Ethnology Bulletin Vol. 155. Smithsonian Institution, Washington, D.C.
Wilson, David J.
1988 *Prehispanic Settlement Patterns in the Lower Santa Valley Peru: A Regional Perspective on the Origins and Development of Complex North Coast Society*. Smithsonian Institution Press, Washington, D.C.
Winter-Livneh, Rona, Tal Svoray, and Isaac Gilead
2010 Settlement Patterns, Social Complexity and Agricultural Strategies during the Chalcolithic Period in the Northern Negev, Israel. *Journal of Archaeological Science* 37:285–294.
Zoubek, Thomas Andrew
1997 *The Initial Period Occupation of Huaca El Gallo/Huaca La Gallina, Virú Valley, Peru and Its Implications for Guañape Phase Social Complexity*. Ph.D. dissertation, Department of Anthropology, Yale University, New Haven, Connecticut. University Microfilms, Ann Arbor.
Zvelebil, Marek, Stanton W. Green, and Mark G. Macklin
1992 Archaeological Landscapes, Lithic Scatters, and Human Behavior. In *Space, Time, and Archaeological Landscapes*, edited by Jacqueline Rossignol and LuAnn Wandsnider, pp. 193–226. Springer Science+Business Media, New York.

https://doi.org/10.1017/aap.2016.2 Published online by Cambridge University Press

NOTES

1. Willey ultimately identified 315 sites in Virú. Several of these sites were excavated and dated by other project members, and no ceramic collections were made from their surface.
2. The individual members of the Virú Valley Project used a unified typology to classify all ceramic sherds collected from both excavations and surface assemblages and worked together to minimize inter-observer differences (Ford 1949:42–43).
3. Ford did not publish any typological data for the Virú Valley material but made these data available in his archives, housed at the American Museum of Natural History. I was granted access to these data and digitized them for the purposes of this study.
4. This sample was not randomly selected. Ford, along with Gordon R. Willey, made collections at 254 sites with a bias toward larger sites. These represented the majority of the 315 sites that Willey (1953) described in his summary of Virú settlement patterns; the remaining sites were investigated by other members of the Virú Valley Project, and surface collections were not made at most of those sites. Willey estimated that, in total, approximately one-quarter of the valley's archaeological sites were investigated and described by all project members.
5. Although the excavations by Strong and Evans (1952) at several important sites provided a wealth of high-quality data on ceramic sequences, none of these sites can be included in this study. This is because surface collections from these sites either were of a limited scope providing a very small sample or else were not conducted at all.
6. The Late Plainware tradition is left intentionally broad for the purposes of this study and could be further subdivided into two or three types.
7. Although the Gallinazo Group is perhaps the best-investigated settlement in Virú (Bennett 1939, 1950; Fogel 1993; Downey and Millaire 2015; Millaire 2010; Millaire and Eastaugh 2011, 2014; Strong and Evans 1952), it cannot be included here because Ford did not make a surface collection at the site, and Bennett classified ceramics using a different typology than his colleagues making it difficult to reconcile data from surface surveys and excavations at the site.
8. Note that no excavation covered the entirety of the Puerto Morin Period, when the Huacapongo Polished Plain type was most popular. Ford (1949:47) filled this gap on the master graph by including surface collections from eighteen sites that had very high percentages of Huacapongo Polished Plain. Thus the master graph is essentially a hybrid of excavation and surface data.
9. In a letter dated February 1977 and contained within Ford's archives at the AMNH, Junius Bird commented that an unspecified number of artifact bags were taken from the collection by the Museo Nacional in Lima and were therefore not curated at the AMNH. It is not known how or whether these artifacts were curated, or whether they are still stored at any museum in Peru.
10. As with the chi-square test, G-tests do not operate well when expected frequencies are low. The spreadsheet used included a "minimum expected" calculation, which was used to determine whether the sample size was sufficiently large. In cases where this condition was not met, poorly represented ceramic types were removed from analysis and/or separate excavation levels were combined so that numbers would be sufficient. Insufficient sample sizes at many sites permitted only the seven sites identified above to be used in this analysis.
11. Excavation data for this site were kindly provided by Jean-François Millaire (personal communication, 2012) and are used here with permission.
12. The lowest levels of Cut A contained three sherds of HPP and 13 of Castillo. While this is the only HPP at the site, these quantities are too low to assign a time period to this layer.

AUTHOR INFORMATION

Jordan T. Downey ■ Department of Anthropology, University of Western Ontario (alumni), 748 Platt's Lane, London, Ontario, Canada, N6G 5E5 (jtdowney@gmail.com)

https://doi.org/10.1017/aap.2016.2 Published online by Cambridge University Press