12.0 MINNESOTA DEEP TEST PROTOCOL: A STAGED APPROACH TO SITE DISCOVERY AND EVALUATION

12.1 INTRODUCTION AND OVERVIEW

The fundamental approach to deep testing recommended herein is a multi-step process that does not stop at site discovery divorced from broader cultural resource management issues and Section 106 responsibilities. The protocol first addresses identification. The Section 106 implementing regulation at 36 CFR § 800.4(a)(1), specifies that:

> "The agency official shall make a reasonable and good faith effort to carry out appropriate identification efforts... The agency official shall take into account past planning, research and studies, the magnitude and nature of the undertaking and the degree of Federal involvement…"

The decision to use the deep test protocol takes place during the implementation phase and is evaluated for each undertaking on a case-by-case basis. Hence, the Mn/DOT Cultural Resources Unit (CRU) project manager should review the Mn/Model landscape suitability models within the area of potential effects (APE), if available, to establish the suitability for preservation of buried archaeological resources. Further, if the Mn/DOT CRU project manager determines that the APE exhibits characteristics indicating the possibility of a buried site, the project manager shall assess the "magnitude and nature of the undertaking" to establish if deep testing is prudent. Thus, the Mn/DOT CRU project manager will have to resolve whether to apply the recommended protocol to each Mn/DOT project with the potential for buried sites, even if, for example, the potential impact area is small.

As this project demonstrates, even a relatively small area (i.e., 1.0 ac [0.4 ha]) can be as geologically or archaeologically complicated as much larger locales. For example, the discovery of spatially limited buried archaeological material at the Clement test locale shows that buried archaeological material need not be extensive to be preserved in place within the subsurface. Just how important it is archaeologically, or whether the Clement site is actually National Register eligible, is not known because it was not evaluated. Nonetheless, small sites like Clement have the potential to be important components of the settlement/subsistence system of archaeological cultures that may contribute to our understanding of the past (Monaghan and Hayes 1998, 2000; Monaghan and Lovis 2005). While geological or logistical conditions such as the desire to avoid impacting too large an area or not using heavy equipment at a landowner's request may require some modifications of the preferred testing process, deviation from the protocol or choosing different methods from those recommended by this study are agency decisions that should be made only after careful consideration of all the risks and benefits that may ensue.

The following protocol, thus, is based on the assumption that the Mn/DOT CRU project manager has established, based both on previous research and understanding of the nature and magnitude of the undertaking, that deep testing is an appropriate and necessary method consistent with a reasonable and good faith effort to identify sites within the APE. Further, it is CCRG's contention that the more complete and certain the Phase I identification data are, the better and more efficient and effective the management of archaeological resources will be. This notion justifies the contention that geoarchaeological work is best front-loaded in the archaeological site identification process and underscores the need to maintain a consciously multidisciplinary, geoarchaeological perspective throughout the discovery and evaluation process.

The discovery and evaluation of buried archaeological sites is a multidisciplinary task that focuses on two different aspects of geoarchaeology. The first, discovery, emphasizes the “geology” of geoarchaeology while the second, evaluation, focuses on the “archaeology” of the discipline. This distinction implies that the effective and efficient resolution to first discovery and then evaluation will usually require different approaches and methods. The results of this study demonstrate that backhoe trenching, under most circumstances, is the most effective and efficient method for discovery of buried archaeological deposits. Trenching also is the best method to place the discovered archaeological material and features into their depositional and landform context within the specific sedimentary and pedological stratigraphy revealed in the trenches.

The destructive nature of trenching, however, makes it a poor choice to evaluate the size and significance of the newly discovered site. Coring/augering and remote sensing methods have a minimal impact on the buried archaeological component and are better suited to trace the extent of the buried deposits, discover potential buried features, and help evaluate the National Register eligibility of the site. Backhoe trenches can provide the original, point-source subsurface configuration and chronological placement of sedimentary environments, paleosol configurations, and related buried archeological horizons. These horizons and buried surfaces can then be traced from the point of discovery in the trench and mapped in detail across the site using coring/augering and geophysical methods without significant negative impacts to the archaeological deposits. Such a staged approach to the deep-test process allows the best aspects and strengths of each of the methods studied to be brought to bear on the identification and evaluation of buried archaeological sites.

12.2 DEEP TEST PROTOCOL

12.2.1 Preferred Deep Test Methods

The proposed Phase I deep testing site discovery protocol is illustrated in Figure 12.2.1-1 and the Phase II deep site evaluation protocol is illustrated in Figure 12.2.1-2. These figures illustrate the work flow for the two step process recommended to 1) identify and 2) evaluate deeply buried archaeological sites.

For the recommended protocol, backhoe trenching is the primary method for discovering buried archaeological deposits. Coring/augering, as employed in this study, is secondary or supplemental to trenching and is recommended for circumstances where backhoe access is too difficult or its subsurface impact too great (wet/wooded landscapes, known archaeological sites, lack of landowner permission), where the logistics of trenching are impractical (i.e., more densely wooded terrain, spatially limited urban settings), and/or where sedimentary sequences are suspected to contain archeological deposits at depths exceeding the practical limit of a

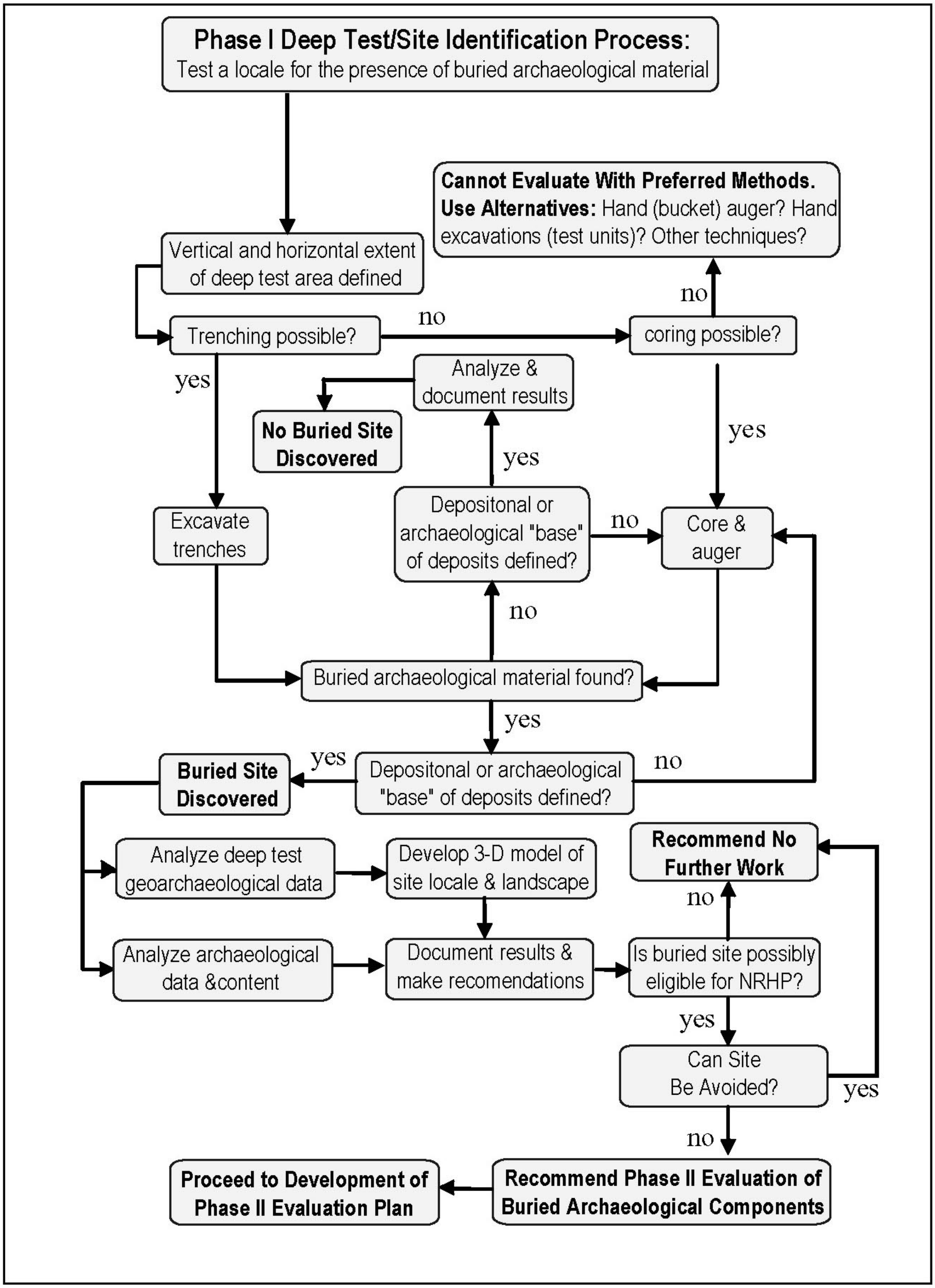

Figure 12.2.1-1. Flow Chart of Phase I Deep Test/Site Identification Process

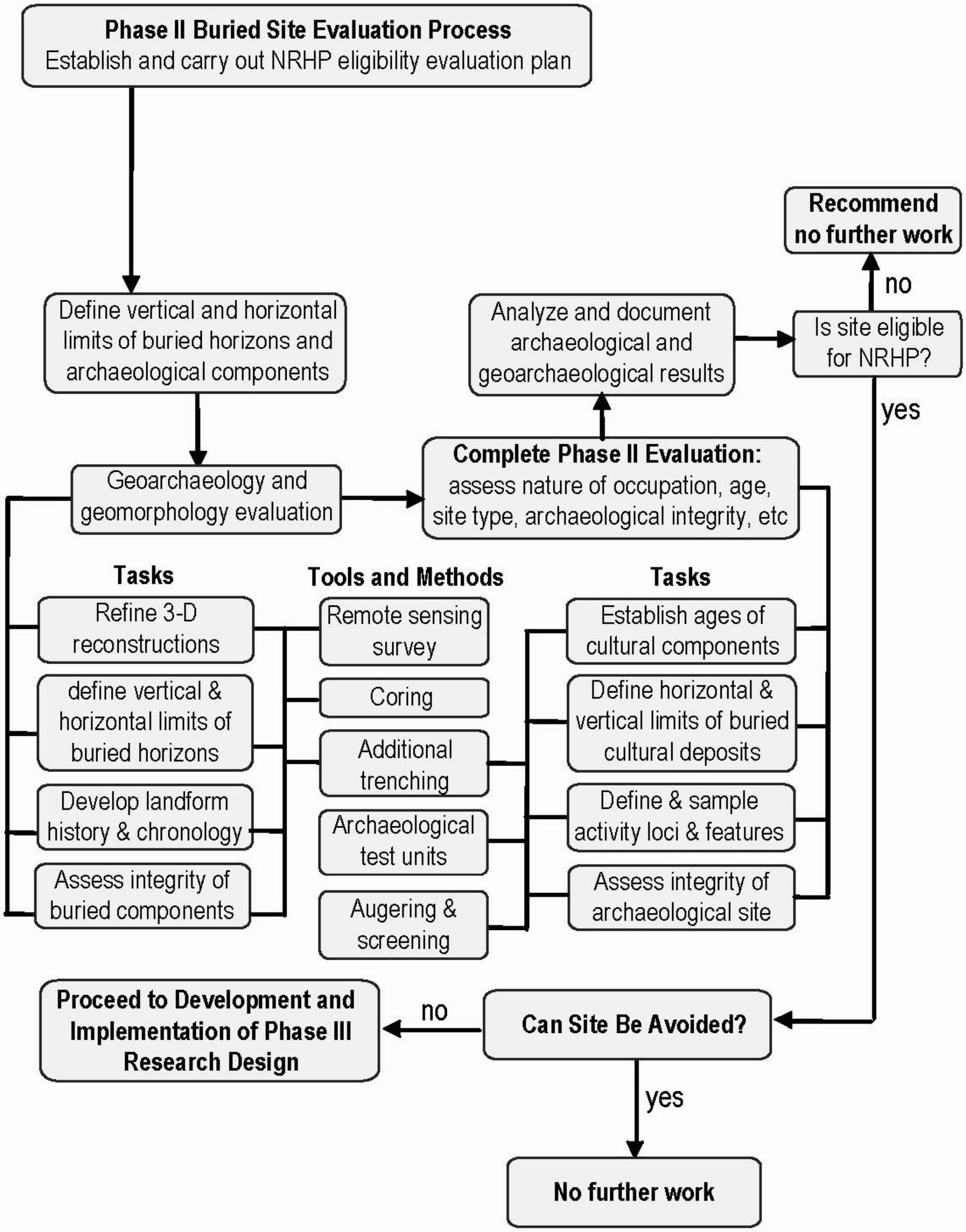

Figure 12.2.1-2. Flow Chart of Phase II Buried Site Evaluation Process

backhoe (>3 m [>10 ft]). Even in the latter case, where deposits extend deeper than can be practically or efficiently reached with backhoe trenching, coring is viewed as supplemental, not primary. In such places, we recommend that the upper parts of the deposits be tested by backhoe trenching.

Sometimes neither trenching nor coring can be employed in a project area if, for example, the area is too wet or too densely wooded for equipment access. In such cases alterative methods, such as hand (bucket) augering or test pit excavation may need to be employed. Such methods, however, are clearly techniques of last resort, as they are neither as effective nor as cost-efficient as the preferred methods of deep testing. They are simply the only alternative to undertake a reasonable and good faith effort to carry out appropriate identification efforts when more effective means are not possible.

12.2.2 Phase I Deep Testing Site Discovery

Backhoe Trenching

As depicted in Figure 12.2.1-1, backhoe trenching is the first choice for deep testing and should be undertaken following methods similar to those employed in this research (see Chapter 4.0). Although the primary goal is site discovery, the presence or absence of a site should be considered in terms of the observed stratigraphy and landform context as they relate to the Holocene developmental history of the area. This requires reconstruction of a three-dimensional model of the subsurface. When a buried archaeological site is not discovered, such a model can explain why and estimate the probability that undiscovered, buried sites are actually present but were not located during the deep testing. Consequently, more data than just presence or absence of archaeological materials should be collected during the trenching process.

Data collected during backhoe trenching first should be directed toward establishing and exposing the base of Holocene alluvial deposits or depositional units that are unlikely to include archaeological deposits (Figure 12.2.1-1). If basal horizons or depositional units that are unlikely to include archaeological deposits are not found while trenching, whether or not buried archaeological deposits are present within the backhoe trenches, then the coring/augering technique should be employed to test the deeper deposits (Figure 12.2.1-1). The coring/augering procedure is discussed below.

Identification of basal horizons or depositional units that are unlikely to include archaeological deposits is straightforward. Potentially they include non-alluvial, glacial sediments (i.e., outwash, till, etc.), bedrock, or other clearly pre-Holocene deposits, as well as thick sequences of coarse-grained channel or bar deposits indicative of high energy fluvial sedimentation, or low-energy subaqueous lacustrine silt and clay, as well as bog (i.e., peat) deposits. Although environments such as bar or channel deposits could include exposed surfaces for humans to occupy, they also are depositional environments that are not likely to preserve such occupations. Identifying depositional hiatuses that indicate times when no or very limited sedimentation occurred is important because these intervals are most likely to include and preserve archaeological occupation debris. Typically such a hiatus coincides with the top of a paleosol, but it could also be marked by changes in texture, lithology, or even the inclusion of extensive amounts of bone or charcoal. While these latter deposits may not be archaeological in origin, they are certainly strong indictors that the associated horizon could include archaeological material and should be thoroughly investigated for artifacts or cultural features. The need to isolate horizons that have both the geological and archaeological potential to include cultural deposits is one of the key reasons that the buried site discovery process be multidisciplinary.

As this study demonstrates, trenching need not be accompanied by test unit excavation. Instead, each wall of each trench should be carefully inspected for evidence (artifacts or suspected archaeological features) of an archaeological site. This includes visual observation as the trench is excavated, periodic cessation of machine excavation to facilitate closer visual examination and trowel scraping of trench walls, and systematic troweling of all trench walls once the final or maximum OSHA-safe entry depth is reached. Screening sediments to recover archaeological materials is recommended for each stratigraphic unit that appears promising for containing artifacts whether or not they are found through careful inspection of the trench walls. A suite of "grab bag" samples or samples from targeted locations along the profile within the suspect horizon can be screened to recover, or confirm the absence of, cultural material. A minimum sample volume of 20 liters (5.3 gal) is recommended to establish artifact densities within the sampled strata. Collection can be facilitated by marking one or more 5 gal (18.9 liter) buckets (5 gal buckets are readily available and commonly used in archaeological field work) in 5 liter (1.3 gal) increments. They then can be used to collect the desired quantity (a 5 liter [1.3 gal] minimum is recommended for each sub-sample) from each sampling area along the trench profile. The total volume of the sediment sample(s) should be recorded, since circumstances may dictate sample size(s).

While detailed guidelines mandating criteria for trench sizes, types of backhoe bucket, or the maximum distance between trenches could be established, from a practical standpoint these are best selected to accommodate specific environmental, stratigraphic, geomorphological, and/or archaeological conditions at individual project areas. In general, we recommend that trenches be excavated as narrow and short as possible to minimize impact to the archaeological deposits that may be present, but be large enough to meet the testing objectives. Trenches should be long enough to allow a complete disclosure of the stratigraphy and soil profiles and wide enough to allow clear visibility of the profile while allowing safe exit and entry. From a practical standpoint, the trench widths are often based more on the availability of a specific backhoe bucket size for the backhoe used. Buckets are most commonly available in 24 in (61 cm), 36 in (91 cm), and 42 in (107 cm) widths. Although any of these common sizes are adequate, experience indicates that 24-in (61-cm) wide trenches tend to be somewhat narrow, while 42-in (107-cm) wide trenches will cause greater impact to the archaeological deposits that may be present.

In most circumstances, toothless (smooth or ditch) buckets are preferable, particularly for sandy soils, because they allow exposure of features and artifact concentrations in the trench floor with minimal disturbance during trenching. However, toothed buckets may be best for conditions where excavation is difficult, such as in urban fill that includes much rubble, dense clay-rich sediments, or compacted or partly cemented deposits. From a practical standpoint, trenches cannot be excavated as deeply with a toothless bucket as with a toothed bucket. Additionally, except in sandy soil, a toothless bucket results in slower trench excavation, increasing trenching costs.

The trench should be long enough that any variability in composition, sedimentology, and pedology can be confidently determined. This determination frequently depends on the field experience of the researcher. The maximum length of the backhoe arm limits trench length to approximately 4 m (13 ft) long without moving the backhoe. Occasionally trenches need to be extended an additional few meters in length. Regardless of bucket size or trench length and depths, OSHA standards must be followed. For safety purposes, or to allow better exposure of the profile, trenches sometimes need to be widened or stepped back. The amount of stepping depends on the depth and composition of the material excavated. This widening can also help clarify sedimentological and stratigraphic relationship of the subsurface units and contacts. Under no circumstances should a trench be entered for inspection if it is deeper than OSHA dictates. These safety factors constrain the maximum practical depths of backhoe trenches. However, sediment from the backhoe bucket can be inspected without entering the trench if attention is paid to where in the trench the sample was collected. Clearly, the utility of trenching without wall cleaning (i.e., trenches too deep or unsafe to enter) reduces the advantage of the long trench-wall profiles because valuable detail is lost. While stepping, trench boxes, hydraulic shoring, or other similar solutions leaving part of the wall exposed can be used to allow safe entry to deep trenches, these also greatly increase the cost, time, and, in the case of stepping back a trench, site impact.

The placement and number of trenches excavated at each deep test locale depend on the size of the project area, site testing objectives, topography, and stratigraphy. Sometimes accessibility, when constrained environmentally or by property owners, may limit the ultimate placement of trenches. In general, trenches should be placed initially to study the subsurface expression of specific surface depositional features that occur on the landform. Based on the results of these excavations, additional trenches may be placed to trace depositional features, soil horizons, or buried landform expressions in the subsurface. To effectively implement such a testing strategy requires experienced earth scientists applying sound geological principles to formulate an understanding of the developmental history of the landform. Rather than allowing experienced earth scientists to make decisions in the field, many states and federal agencies have begun to dictate the minimum or even standard distances that are allowed between trenches or cores. This is true in Indiana and Pennsylvania and is being considered in Michigan, New York and Virginia. These typically are in the range of 60 m to 100 m (197 ft to 328 ft). While this eases some burden for decision making, and it is certainly possible for Mn/DOT to adopt such a standard, its adoption will not necessarily result in better or more-cost effective deep testing. A more efficient and scientifically effective approach, as recommended in this protocol, is to establish trench distances and numbers on a case-by-case basis, in discussion with the survey team, prior to deep testing.

Survey design will vary for each project area, but should be based on the specific testing objectives, site conditions, and physical or environmental constraints at particular test locales. Consequently, the number of trenches required to complete deep testing will necessarily depend on both the physiographical, expected sedimentological, and anticipated stratigraphic complexity of the subsurface. Moreover, depending on the subsurface conditions actually encountered, the actual number of trenches may change once work begins at the site. In developing a work plan and formulating a budget, the survey team should provide either a geomorphological and/or sedimentological map (based either on Mn/Model LfSAs or other criteria), and anticipate the number of trenches and trench placement strategy that will be required to test the landform components. The basic rational for trench numbers and placement should be outlined by the survey team as part of the work plan supplied to Mn/DOT. The work plan should place the subsurface, based on the data available, into the type of three-dimensional, geoarchaeological framework outlined throughout this report. Personnel at Mn/DOT should then review this work plan for adequacy. The formulation of this work plan by the survey team and the agency review process are important because they provide a clear understanding for all parties concerning the testing goal and the minimum expectation for deep testing. The work plan is also the only way to provide reasonable budgeting and planning. Clearly, as stated in Chapter 11, the more trenches excavated, the less likely the deep testing will lead to a false-negative outcome (i.e., a buried site missed by deep testing). More trenching, however, will also mean that the deep testing will be more costly. Ultimately, judging the balance between deep test cost and the risk of false-negative outcome must be an agency decision based on consultation with the selected contractor.

Once deep testing commences, if the subsurface is actually more complicated or simpler than expected, the numbers of trenches and the budget, can be adjusted accordingly. Such adjustment should be done in consultation with Mn/DOT archaeologists and should be based on a clear and succinct explanation by the survey team of why such changes (both in scope and budget) are necessary. Given trenching's destructive nature, if archaeological deposits are encountered, then excavation should be halted within the immediate area of discovery. As discussed below, trenching is not akin to shovel testing and should not aim to determine the site boundaries during Phase I. However, it may be continued elsewhere on the landform and test locale.

Data recordation for profile descriptions by the survey team's geoarchaeologist should follow generally accepted professional standards. A standard form or set of forms could be developed by Mn/DOT. Because the focus of trenching is geoarchaeological, critical information concerning the general lithology, sedimentology, pedology (soil formation), and extent of each identified stratigraphic unit should be recorded. The depth below surface of the top and bottom of each of the stratigraphic units encountered, as well as their general thicknesses, should be measured. Lithological information such as general color, texture, and lithology of the units, as well as that of any minor interbeds or interstratified deposits included within the units, should be noted according to currently accepted standards (e.g., Munsell soil color charts, Wentworth scale). Major sedimentological information should include, but not be limited to, bedding (e.g., cross-bedded, tabular beds), sorting, grading, and any deformation observed in the unit. Additional pedological information, such as evidence of post-depositional inclusions, transferrals, development of structure, consistency, and bioturbation, should be noted. The contact between lithologic units should also be described. Importantly, when encountered, samples of organic material (both floral and faunal) should be collected for possible analysis at a later date (i.e., ^{14}C and flotation samples). The processing of organic samples (i.e., ^{14}C and flotation samples) would depend on the type(s) of information needs for each project and should be undertaken in consultation with and at the discretion of the CRU project manager.

Though working in tandem with the geoarchaeologist, the survey team's archaeologist should independently take standard archaeological field notes, possibly using a standardized form, and draw profiles of trench walls, including the locations of artifacts, possible features, and stratigraphy. When there is no evidence of buried sites in a trench, a profile of one wall is sufficient. Profile drawings should be coordinated with those of the geoarchaeologist for consistency and any divergence of opinion noted and, hopefully, rectified in the field.

The maximum depth of trenches at a deep test location depends principally on the stratigraphy observed within the trench while excavating. Because the major objective of the deep test excavation is to locate areas of buried archaeological deposits, the recognition of soils and sediments that are of Holocene age is critical. This is particularly true for alluvial settings in the glaciated terrain of the upper Midwest. With this in mind, we recommend that trenches be excavated to a depth defined by the maximum extension of the backhoe boom (typically 3 m to 4 m [10 ft to 13 ft]), until the water table is penetrated (i.e., trench fills with groundwater), or until the geoarchaeologist is reasonably certain that the trench has penetrated late Wisconsinan or older age stratigraphic units (i.e., glacial, glaciolacustrine, glacio-fluvial, bedrock). If reaching greater depths is deemed necessary because the depositional or archaeological base of the deposit was not penetrated, coring/augering procedures should be considered (Figure 12.2.1-1).

Coring/Augering

The role of coring as a secondary option for deep testing is depicted on Figure 12.2.1-1. Coring is employed when trenching is not possible. Trenching, for example, might not be possible because the surface or subsurface impact is deemed too great; the deep test parcel is densely wooded, marshy, or steep; access is spatially limited, as it often is in urban settings; or the depth of the deposits needing testing exceeds the reach of the backhoe arm. High groundwater conditions also limit the maximum depths that trenches can penetrate. However, "dual-tube" type setups, which are available with most coring devices (including GeoProbe), do allow cores to be collected from below the water table and are recommended for areas with high water table conditions.

When used during the site discovery phase of deep testing, the coring/augering procedure should be undertaken following methods similar to those employed in this project (i.e., see Chapter 4.0) and in a two-step process. First, solid-earth cores are collected using a GeoProbe or similar continuous coring device. As with trenching, these cores should establish the base of Holocene alluvial deposits or at least penetrate depositional units that are unlikely to include *in situ* archaeological deposits. During the coring process, depositional hiatuses, which mark times when the landform stabilized and are likely to include archaeological occupation, should be sought and identified when present. These typically represent the tops of paleosols but could also include changes in texture, lithology, or even the inclusion of extensive amounts of bone or charcoal. From these data, specific target horizons should be selected to test for the presence of buried archaeological material using augering.

Initially, the coring process is best carried out using a systematic grid. In this study, core data were collected every 20 m (66 ft) within the testing area, and from this information target horizons were identified for coring. While this sample pattern was generally adequate, the fact that no target horizons were identified near the buried archaeological deposits at the Clement test locale suggests that the sample interval may have been too coarse to intersect low-density archaeological deposits (i.e., Kintigh 1988; Krakker et al. 1983; McManamon 1984; Nance and Ball 1986, 1989; Shott 1985, 1989). Closer interval sampling, however, will also greatly increase the coring costs. For example, a 10 m (32.8 ft) coring interval will essentially double the costs of coring (i.e., without augering) but would also increase the confidence that low-density buried deposits are discovered. Leaving changes in the coring interval to the discretion of the deep test team in the fields might be a more cost effective means of increasing confidence. Because coring often will be performed in the event that trenching did not find the base of the deposits, some information concerning the basic subsurface configuration will be known. These data can provide guides to the most effective coring interval at the test locale. As was noted for trench intervals, beyond providing guidelines for a maximum coring interval, mandating core spacing may be more restrictive than helpful because specific environmental, stratigraphic, geomorphological, and/or archaeological conditions encountered at individual test locales may provide better parameters by which to adjust the sampling interval.

To the extent possible, data recordation recommended for the trenching procedure above should be followed. Although difficult in small diameter cores, sedimentological information such as bedding (e.g., cross-bedded, tabular beds), sorting, grading, and any deformation within units should be recorded.

The principal goal of the coring process is to identify stratigraphic horizons that represent stable surfaces of an age compatible with human occupation. These are identified based on their stratigraphical, pedological, and sedimentological characteristics. The depths to the top and base of these horizons are defined based on the core data, and then these target horizons are sampled with augers for the buried archaeological materials. We recommend using at least six 4-in (10-cm) or four 5-in (13-cm) diameter flight augers. The total area augered would then be 486 cm^2 (75 in^2) and 507 cm^2 (79 in^2) or the equivalent of a 25-cm (10-in) diameter shovel test. Because they are more efficient, 5-in (13-cm) or larger augers are preferable.

To recover a standard sample of the sediment, enough length must be added to the auger to reach the target horizon. The auger is spun to clear the overlying spoil from the hole and then augered into the layer that potentially contains archaeological materials. The soil and sediment from the target horizon should be screened to locate any artifacts that may be present. We recommend one-quarter inch mesh for site discovery. While a smaller mesh would allow for recovery of microdebitage, most archaeological sites will contain a broad range of debitage sizes that can be collected through standard recovery techniques. In addition, the cost comparisons shown for coring and augering in this study are based on the use of a one-quarter inch screen. Costs would significantly increase if finer mesh were employed during the augering process. Moreover, numerous logistical issues associated with screening matrix through fine mesh screen or floting soil samples make these procedures either impractical or, at best, not cost effective for Phase I site discovery.

12.2.3 Phase II National Register Evaluation of Deeply Buried Archaeological Sites

Evaluation Methods

If evidence for buried archaeological deposits is not found during either trenching or coring/augering, then the research team can confidently argue that the proposed undertaking within the deep test locale (APE) will have no effect on buried archaeological resources. If buried archaeological materials are found and further archaeological investigations are recommended, however, the proposed deep testing protocol recommends that additional geoarchaeological investigations accompany the Phase II archaeological evaluation. Because they require very different tools, and in the end have different goals, site evaluation is separated from site discovery. This second stage of deep testing, however, is an extension or refinement of the deep test process. It should aim to provide enough information to facilitate the evaluation of the NRHP eligibility of the site, as well as fully reconstruct the geologic context of the site and its implication(s) for understanding the archaeological structure and function of the site. While the goals and rationale of Phase II evaluation of a buried site are the same to those of surface sites, the process is vastly more complicated because it is buried. We believe that the evaluation should be undertaken as a two-step process, and, like the Phase I site discovery process, it should proceed within a multidisciplinary, geoarchaeological framework. The first step of this process should focus largely on defining and/or refining the geological parameters of the buried deposits and providing greater detail of the actual and/or potential extent of buried archaeological materials. The second step, on the other hand, concentrates more on an archaeological evaluation of the site. Details of how these steps should be implemented on the ground, their design, and the information that should be collected are described below (Figure 12.2.1-2).

The goals of site evaluation are five-fold: 1) to refine the horizontal and vertical limits of the site within the APE; 2) to define the nature of the site, including the density and distribution of artifacts, the presence of cultural features, and any post-depositional impacts to the site; 3) to place the site components in their proper regional and local cultural context; 4) to provide chronometric dates if possible; and 5) to gain a sufficient understanding of the nature of the occupation relative to other sites in the region so that its National Register eligibility can be assessed (Mn/DOT 2004:5).

We propose, in contrast to buried site discovery, that the evaluation of the newly discovered buried site integrity, horizontal and vertical extents, and significance be conducted using a combination of coring, remote sensing, and/or more traditional archaeological excavation methods (Figure 12.2.1-2). The destructive nature of backhoe trenching means that once a site is discovered, further trenching should proceed only with great caution. In fact, as noted above, trenching should probably cease in the general area of site discovery. If the site area represents the only part of the landform undergoing deep testing, the initial discovery phase of deep testing should be ended. If other, not yet tested elements of the landform require investigation, however, these can be trenched. Further trenching to discover site limits or for site assessment is not appropriate unless part of a clearly defined, evaluation work plan. Although backhoe trenching can be an effective part of an archaeological evaluation testing plan, it should be undertaken only when the risks and relative gains of trenching a site are clearly recognized. A clearly articulated and multidisplinary testing plan must be developed before the buried site can be effectively and efficiently evaluated.

The staged approach to discovery and evaluation of buried archeological deposits suggested here is neither new nor particularly innovative and has been successfully employed elsewhere. For example, the authors applied it to a buried site (Converse site [20KT2]) near Grand Rapids, Michigan (Hambacher et al. 2003). Here, a thick multi-component occupation midden was discovered under about 2 m (6.6 ft) of historic fill within a deep-test backhoe trench. Subsequent investigations minimized impact to the buried midden by coring on a 10-m (33-ft) grid pattern to trace the vertical and horizontal extent of the archaeological horizon. A set of isopach and other contour maps based on these data was used to map the paleo-land surface. The real strength of this staged approach is apparent from the Converse site experience. Because the burial depths, thicknesses, and spatial extent of the archeological horizon were known and clearly mapped upon completion of the Phase II analysis and report, the data recovery work plan could include parameters defining the areas that would be impacted by the proposed bridge piers; the amount of stripping necessary to expose the archaeological horizon; and the sediment volumes per test unit. This is in spite the fact that the target horizon was only observed in detail from one backhoe trench excavated during the site discovery process.

Coring and geophysical survey methods are well suited to trace and map details of subsurface horizons across the buried archaeological site with little impact on the buried component(s). This information can be used to map the three-dimensional configuration of a buried archaeological site, which aims to refine the initial reconstructions obtained during the site discovery process. In addition, the step-1 Phase II efforts should also assist in identifying features and potential post-depositional impacts to the site as reflected in variations in the stratigraphic sequence. Large historic cultural disturbances can also be revealed through this work. The appropriateness of one method versus the other is based on several factors. These include: the nature of the soils and sediment, lithological contrasts between stratigraphic layers, groundwater issues, and potential interference of the geophysical signals from features such as electrical power lines or metal on and near the site. Even such aspects of the site as amount and type of vegetation can dictate what survey methods can be realistically undertaken.

Depending on the how much information was initially obtained regarding the horizontal and vertical limits of horizons and strata that may include cultural deposits, this may require considerable or minimal effort. For example, if a great deal of relatively detailed stratigraphic data were obtained during the Phase I process, if the configuration of the subsurface was relatively simple, or if the site was not buried very deeply, then the geological characterization of the of site during the first step of Phase II should be simple. Conversely, if the site was more deeply buried or the stratigraphy and depositional framework was particularly complex, then greater effort and expense might be required to adequately understand and characterize the subsurface distribution and relationships of the cultural deposits and associated soil/sediment horizons. The level of effort needed cannot be predicted and should be defined by the agency archaeologists based on the initial geoarchaeological data report by the Phase I research team. In addition, the best tools and techniques to address these issues (Figure 12.2.1-2) are also highly variable and will depend on the environmental conditions at a specific locale.

The Anderson site is a case in point. While many facets of the sedimentology and pedology were apparent in sediment cores, they actually added limited data, beyond that already known by trenching, regarding the nature and distribution of the archaeological components at the site. In part, this was due to the generally homogenous nature of the deposits and the absence of clear and visually traceable archaeological or soil horizons. These factors, on the other hand, along with the coarse-grained nature of sediment at the site, allowed nearly all geophysical survey methods to reveal considerable detail about the subsurface. Likely surface and sub-plow zone cultural features were clearly apparent and traceable with both resistivity and magnetometry. Additionally, although at first inferred to be buried surfaces, once the trench data showing the stratigraphy and soil horizons at Anderson were made available, the GPR survey could be reinterpreted to discern and map the upper meter of bioturbated or culturally mixed sediment across the site and distinguish the beginnings of Fe-rich lamella and B-horizons. Clearly, the application of all of the geophysical surveys could greatly enhance the site evaluation process at Anderson by allowing archaeological test units to focus on examining specific cultural indications in the subsurface.

At other locations, such as the Hoff Deep test locale, where sediments were very fine-grained and included only limited lithological contrasts between layers, coring yielded far more useful information than geophysical survey methods. For example, because the GPR signal attenuates quickly in silt, clay, or saturated sediment, it could barely penetrate beyond the upper meter of deposits at Hoff Deep. Even when observed, the radar signal was often too weak to reveal much in the way of subsurface detail. Considerable sedimentological and pedological details, however, were noted within the sediment cores. These were also useful for mapping the bottom of alluvial deposits by tracing the paleosol that developed within the top of the glaciolacustrine deposits in the southern and eastern part of the at Hoff Deep test grid. As was the case at sites with a strong surface expression of the archaeological component (e.g., Anderson), the magnetometry survey of the Hoff Deep test locale, nevertheless, did show several potential near-surface features as well as indications for some type of historic component. Magnetometry generally should be employed during Phase II evaluations that include archaeological surface expressions to aid in assessing the distribution of near-surface features. Surprisingly, none of the geophysical survey methods yielded particularly useful results at the Fritsche Creek II test locale, even though the subsurface conditions should have allowed adequate resolution of both GPR and resistivity methods. Unfortunately, methodological ineffectiveness is often only apparent after the fact.

Occasionally, both coring and remote sensing will be unsuccessful in revealing the precise nature of a complex subsurface configuration or, at best, yield results that cannot be sensibly interpreted. For example, regardless of how extensive the coring and geophysical surveys were at the Hoff Deep test locale, the complex stratigraphy and depositional associations related to sediment slumping found in Trench 3 are neither resolvable nor interpretable. Arguably, they were not even seen through these methods and would be extremely confusing in a 1 m × 1 m (3.3 ft × 3.3 ft) archeological test unit. The long profiles exposed by backhoe trenches are essential to understand such relationships. This indicates that sometimes additional trenching may be a good choice or even requirement as part of the Phase II site evaluation process.

In the final analysis, exactly which method(s) is chosen to map an archaeological component in detail or trace subsurface horizons in three dimensions is dependent on the specific conditions at the site being evaluated. The methods used should be selected as part of the deep test Phase II work plan based on the information needs outlined in the research design. Moreover, the method should be chosen because it is able to solve specific problems, answer specific questions, or address specific limiting conditions at a particular site. Additionally, sometimes the best method is actually a combination of methods discussed in this report.

Once the vertical and horizontal limits of the site are defined, the second step for Phase II site characterization can be undertaken to evaluate the archaeological component of the site. First, augers or hand excavated test units should be used to collect additional archaeological data relevant to defining the density and distribution of artifacts and the presence of cultural features. These will be located based on the three-dimensional mapping of the subsurface achieved during the initial, geological step of the Phase II process. The choice of which of these techniques to employ will largely depend on the nature of the deposit, as defined during site identification, and information derived during the Phase II evaluation regarding the horizontal and vertical extents of the site. Despite the relatively high labor costs associated with the excavation of test units, they may be more appropriate than augers if the site occurs at depths less than 1.5 m [4.9 ft]) or if heavy machinery can be used to strip off sterile layers to reach greater depths safely. Augers, on the other hand, have less impact and are a reasonable and cost effective alternative for testing the deepest of buried sites. Because test units provide a larger sample than augers, augers may be appropriate only for evaluating relatively artifact-rich sites. In addition, test units, perhaps complemented with backhoe trenches in situations in which the stratigraphy is particularly complex, may be appropriate to provide more controlled recovery of materials. In contrast, the fine scale vertical control needed to discern complex or subtle stratigraphy is not provided by augering. Finally, test units may be required to sample features and to collect materials for chronometric dating, if the latter were not collected during the Phase I trenching.

In sum, earth scientists can provide information regarding the spatial extent, integrity, and geoarchaeological context(s) of the deposits critical to fully evaluating the National Register eligibility of a site. Archaeologists will have to use their discretion to decide the best means of obtaining a sufficient sample of the archaeological deposits to define the function (i.e., site type) and age of the occupation(s) and determine if the site is eligible for the National Register.

Discussion

The approach to buried site evaluation proposed here maximizes the amount and quality of geoarchaeological information available to contextualize the archaeological deposits with minimal impact to the site. The evaluation of the archaeological deposits, thus, is directly tied to the accessibility of sound and complete geological, pedological, and stratigraphic contextual data. This approach and, consequently, the front-loading of geoarchaeological data into a consciously multidisciplinary Phase I/Phase II process, assures that vital geoarchaeological background data will be available for input into the archaeological evaluation process.

Although the process of buried site discovery we have recommended focuses on geoarchaeological processes, the overall needs of site evaluation are essentially archeological.

Moreover, if the assessment of site significance and National Register eligibility involve principally archaeological practices, the results of archaeological excavations must also be incorporated into the geological and depositional framework to be useful in developing a site evaluation work plan. Clearly, the archaeological materials and features at the site and the site's context and cultural meaning are affected by geoarchaeological processes related to site formation during occupation as well as post-occupation sedimentation and weathering (pedological) events. These relationships underscore the fact that archaeological investigations of stratified and/or buried sites require a conscious and overtly multidisplinary focus to successfully deconstruct the history of site usage and abandonment as well as the cultural meaning of site components.

12.2.4 Deep Testing for Buried Sites in Urban Settings

The proposed deep testing protocol process outlined in Figures 12.2.1-1 and 12.2.1.2 has been formulated mainly to address archaeological site discovery within natural, geological depositional contexts (e.g., alluvial, colluvial, and eolian processes). Although sites can be buried in many different ecological and cultural settings, in reality most are associated with alluvial landforms and, accordingly, such settings are the main focus of the project. While other natural depositional landforms, such as eolian, colluvial, or littoral/lacustrine, may have their own set of problems, the two-stage deep test protocol described above is designed to be applied to any depositional setting. The steps outlined can be followed in nearly any setting because the basic principle guiding the protocol centers on placing archaeological deposits within a processual framework and treating the archaeological site as another component or horizon within an evolving landscape. The discovery and evaluation of buried archeological resources, regardless of depositional environments, relies on a systematic reconstruction of the depositional history and processes revealed by the stratigraphy of the associated landform deposits rather than on specific idiosyncratic aspects of cultural or geological processes. How this principle applies to other natural depositional settings, such as dunes, shorelines, or wetland deposits, is clear. However, by maintaining the geoarchaeological perspective and focus on sediment depositional processes for buried site discovery, the same basic procedures that apply to deep testing in natural settings can also be applied to contexts where sites can be buried by non-geological processes. This is true even for deep testing areas underlain by recent fill deposits in urban settings.

Overall, the process of discovery and reconstruction of depositional sequences in urban or other historically developed settings is generally similar to that employed within more natural contexts. Deep testing in urban environments, however, also offers some unique challenges for understanding site formation processes and depositional sequences. This is particularly true for discovering and characterizing prehistoric sites buried within urban contexts. In fact, because many of the same cultural and environmental factors apply to human selection of settlement locations during both the historic and prehistoric periods, such as resource availability and transportation along rivers, urban settings are likely to include prehistoric as well as historic-period archaeological sites in both mixed (disturbed) and stratigraphic in situ contexts. For example, the modern land-use pattern of Euro-American settlers often results in the preservation of deeply buried surfaces and landform components that were covered during the historic development of the present urban ground surface. This process is particularly significant in riverbank and floodplain environments, not only because such areas are often the first developed, but are usually “wet” or poorly-drained and require filling and reshaping for effective building or industrialization (Hambacher et al. 2003; Lovis 2004; Monaghan and Lovis 2005). Consequently, geomorphological features that can be used to guide deep testing in alluvial settings are often buried within relatively thick urban fill deposits, which then effectively mask any evidence of alluvial landform development. This phenomenon has been noted in both large and small sites throughout Great Lakes region where intact prehistoric deposits, which often include archaeological sites, have been buried under extensive urban fill sequences (Demeter et al. 1994; Demeter and Weir 1983; Hambacher et al. 2003; Larsen and Demeter 1979; Lovis [ed.] 1993, 2002; Monaghan and Lovis 2005). Additionally, during the formation of urban landscapes, components of important early historic sites are often buried by fill within areas that are do not include natural burial mechanisms and would not normally be selected for deep testing (Demeter and Weir 1983; Lovis 2004). These are particularly significant in metropolitan areas that include early, long-term historic-period settlement along the eastern seaboard (Blakey 1998; Hayes and Monaghan 1998, 1999; Medford [ed.] 2004; Parrington and Wideman 1986) as well as the Midwest and Great Lakes regions (Demeter and Monaghan 1997; Demeter and Weir 1983; Kolb 2003a, 2004).

Prehistoric deposits in urban settings are not necessarily deeply buried or, as mentioned above, associated with settings that are otherwise conducive to archaeological site burial and preservation. For example, Hambacher et al. (1995) described Late Woodland features preserved directly under sidewalks, while Lovis [ed.] (2002) noted and investigated intact and undisturbed Middle to Late Archaic deposits, which included human burials, directly under middle to late nineteenth century road beds in Bay City, Michigan. Clearly, the presence of human remains preserved in urban landscapes is particularly problematic from a compliance and cultural heritage perspective. Their discovery during construction phases of projects has led to significant construction delays and heated socio-political conflicts. The unearthing and subsequent mitigation of a nineteenth century African-American cemetery in Manhattan in 1993 (New York Times 1993), for example, presented major technical and ethical challenges for the preservation community (Blakey 1998), but also ultimately led to an increased awareness of the importance of deep testing in urban settings (Monaghan and Lovis 2005; Roberts et al. 1993).

Although urbanization is increasingly recognized as an important site burial mechanism (Kolb 2004; Lovis 2004; Lovis and O’Shea 1994; Monaghan and Lovis 2005), the processes that form urban landscapes result in relatively few predictors for the presence or absence of archaeological deposits or buried site potential. This is true for both historic and prehistoric sites and presents significant challenges to deep testing. Just as challenging, the fills themselves often include concrete, brick, or even cemented sequences that are difficult to effectively penetrate during testing and always require power equipment (Hambacher et al. 2003; Monaghan and Hayes 2002). Except for their occurrence, urban fills seldom possess morphological, sedimentological, or environmental markers that indicate whether intact Holocene surfaces or prehistoric archaeological sites might be buried beneath them. Urban fill sequences, however, can include important historic archaeological deposits that may be NRHP eligible themselves. Such deposits should be scrutinized by appropriate personnel trained in historical archaeology. Additionally, fill sequences may also include “prehistoric” deposits within them. Clearly, such prehistoric archaeology material is not in primary context. Its presence, however, may hint at the occurrence of *in situ* prehistoric materials below the fill or elsewhere in the project area.

While detailed predictors in the fills themselves may be lacking, the fact that these derived during the historic period, however, means that historic maps of variable quality and detail sometimes do exist. These can reveal the land use history of urban areas as well as show the presence and chronology of buildings and other construction episodes (i.e., foundations, basement excavation, etc.) that might have resulted in destruction or preservation of the pre-settlement ground surface. As discussed below, determining the age and construction methods of such structures is important because construction techniques and standards that may have acted to enhance or diminish preservation varied greatly during the nineteenth and twentieth century.

Even when masked by fill sequences, the present geomorphological configurations can provide some broad guidance for deep testing in urban environments. Because prehistoric archaeological sites are often concentrated along river or lake waterfronts, which are also topographic lows in the landscape, they are likely to be covered with fill deposits and buried during the early phases of urbanization (see Demeter et al. 1994; Hambacher et al. 2003; Lovis [ed.] 2002). For example, Demeter et al. (1994) describe stratified Middle to Late Woodland deposits preserved in primarily alluvial context under more than 3 m (9.8 ft) of nineteenth century fill in Bay City, Michigan. Even though historic settlement was first concentrated along waterfronts, such deposits may be more commonly preserved than destroyed because eighteenth and nineteenth century excavations were mainly by hand and, therefore, considerably more difficult than filling. Consequently, filling rather than cutting probably was more characteristic of early urbanization (Lovis 2004) and, if so, preservation of waterfront prehistoric sites may be more common than not (Demeter et al. 1994; Hambacher et al. 2003; Monaghan and Lovis 2005). See Lovis (2004) for a more detailed and in-depth discussion of these and other urban taphonomic processes.

The process of prehistoric site preservation in urban settings is illustrated by the recent discovery and subsequent excavation of the Converse site (20KT2) related to a Michigan Department of Transportation (MDOT) project to reconstruct the US-131 bridge across the Grand River in Grand Rapids, Michigan (Hambacher et al. 2003). Here, as mentioned above, an up to 1 m (3.3-ft) thick, Late-Archaic through Contact period midden located near a presumably destroyed Middle Woodland mound group (Converse Mounds) was discovered under ca. 2 m (6.6 ft) of nineteenth-century fill along the Grand River in downtown Grand Rapids (Hambacher et al. 2003). The mound group and several nearby sites, described as kitchen middens and village sites in nineteenth century documents (Hambacher et al. 2003), were believed to have been destroyed when Euroamerican settlers and subsequent urban development removed the mounds and industrialized the waterfront. Remarkably, considerable portions of an occupation midden in the APE of the bridge project were undisturbed, even though a foundry, numerous utility trenches, a parking lot, a flood-control wall, and a six-lane expressway bridge (US-131) were all constructed within site boundaries during the past 150 years. The occupation midden was discovered using the type of deep-testing procedure outlined above (i.e., backhoe trenching), and its preservation confirms several important factors that control site preservation in urban areas.

First, as is true elsewhere, filling rather than cutting was apparently the first and primary urbanization activity that occurred along the waterfront at the Grand River. This preserved much of the ground surface at the time of nineteenth century settlement and effectively sealed it from general urban disturbance relatively early in urbanization (Hambacher et al. 2003). Thus, when more extensive structures were built during the latter half of the nineteenth century, this fill actually protected the prehistoric archaeological deposits. Moreover, because hand excavation predominated during the nineteenth century, footing and basement preparations were often minimal and tended to be only as large as absolutely required. Consequently, only minimal lateral disturbance occurred at the Converse Site throughout most of the nineteenth century. Indeed, because they are placed on generally filled wetland, basement excavations for buildings located on the sites were avoided from a structural stability standpoint. Thus, once built, these buildings, with their shallow, minimal footings and simple floors, tended to further protect the site. In fact, much of the cobblestone floor of what originally was a nineteenth century foundry was found intact and preserved the occupation midden. The same minimal ground disturbances that occurred during early urbanization were not true during most of the twentieth century when the use of heavy equipment became more common. As a result, most the actual disturbance of the Converse site occurred during the construction of the US-131 bridge piers during the 1960s (Monaghan and Hayes 2002).

The mitigation plan for the Converse Site also serves to illustrate how the protocol outlined above (Figures 12.2.1-1 and 12.2.1-2) can be directly applied to urban settings. The processes and procedures for deep test in urban contexts, however, do have some important differences from those used for more typical natural, geological depositional contexts. This is true for specific equipment needs as well as background information necessary and available to undertake effective deep testing. Regardless of setting, the goal of trenching, as indicated in Figure 12.2.1-1 for deep testing in general, is site discovery, and the choice of equipment is controlled mainly by the testing goals and subsurface conditions. For example, deep testing at the Converse site area in Michigan was undertaken using a large track-backhoe (excavator), but because the test area was under a parking lot, it first involved cutting pavement. A comparatively large and powerful track-excavator was selected for trenching because of the potential for encountering building foundations, buried pavements, and/or layers of brick or concrete below the surface (Monaghan and Hayes 2002) and, from an efficiency stand point, this type of machine should probably always be the first choice in urban setting. However, surface obstacles may dictate that other means be used. That continuous solid-earth coring may be a better or more efficient choice in some circumstances has been demonstrated in urban settings near St. Louis, Missouri (Kolb 2003a, 2004).

Regardless of equipment used for the site discovery phase of deep testing (see Figure 12.2.1-1), however, goals of the deep test in urban contexts should focus on addressing the following concerns: 1) mapping the depth and extent of urban fill, 2) mapping the depths, extent, and depositional environments of Holocene deposits present below the fills, 3) determining the extent and magnitude of disturbance of natural, pre-urbanization sediments and soils below the fill sequence, and 4) determining the presence and/or potential for preserved archaeological deposits within the sequence. Except for the first concern above (extent of urban fill deposits), the steps for assessing an area for the presence of archaeological materials are identical to those shown in Figure 12.2.1-1. Questions such as: “Are there archaeological deposits within the sequences?”

and “Has the base of deposits been defined?” (Figure 12.2.1-1) must be addressed. If archaeological deposits are encountered, the geoarchaeological context and archaeological content must be analyzed to determine if the site is National Register-eligible, just as the protocol suggests for natural depositional settings (Figure 12.2.1-1). Importantly from the standpoint of a deep test protocol, by considering urban fill as another type of “sediment” that is deposited during a specific interval (i.e., historic period) and conditions (rapid accumulation related to urbanization), then the same philosophical approach outlined throughout this volume for natural sediments can also be applied. As is true for naturally deposited sediment that could include many different types of archaeological sites, the fill sequence, particularly the lower part, may preserve important historic-period sites (see Demeter and Weir 1983; Monaghan and Hayes 2002; Lovis 2004; Medford 2004).

Once discovered, sites that have been buried in urban contexts should be evaluated using the multi-step procedure similar to that outlined in Figure 12.2.1-2. At the Converse site (20KT2), described in Section 12.2.3 above, the Geoarchaeology and Geomorphology Evaluation step in Figure 12.2.1-2 was completed after the agency archaeologists at MDOT and the Michigan SHPO had concurred with CCRG’s recommendation that the site was National Register eligible during an onsite meeting during the initial backhoe trenching. More typically, the second stage of buried site evaluation (i.e., “Complete Phase II Evaluation”; Figure 12.2.1-2) will be applied to historic or prehistoric sites buried beneath urban fill deposits and follow the Geoarchaeology and Geomorphology Evaluation step.

In some cases coring may either be impractical or not effective, and geophysical survey techniques will be more appropriate for mapping the subsurface. In fact, some remote sensing procedures are quite effective for large urban prehistoric or historic sites, or within urban landscapes where relatively extensive cultural features (such as building foundations and earthworks) and large metal artifacts occur in abundance (Chávez et al. 2001; Hargrave et al. 2002). As was suggested for evaluation of sites in natural depositional settings, the best method(s) for mapping the subsurface depends on the specific conditions at the site and the needs of the Phase II work plan. Whether geophysical survey, coring, or additional trenching is deemed necessary for tracing, mapping, or otherwise evaluating a buried site, the particular method(s) should be selected to solve specific problems, answer specific questions, and/or address specific limiting conditions at a particular site. In fact, the best method may actually be a mix of these procedures.