13.0 SUMMARY AND CONCLUSIONS

13.1 PROJECT OVERVIEW AND DESIGN

The goal of the Mn/DOT DTP project was to develop a deep test protocol to employ in projects subject to Section 106 of the National Historic Preservation Act and its implementing regulations (36 CFR § 800). The protocol was derived from the direct comparison of the effectiveness of various deep test methods at high-potential alluvial locales as well as at known, buried archaeological sites. The methods compared, which included several geophysical survey techniques, solid-earth coring, and backhoe trenching, are among the most common techniques employed for discovering buried archaeological sites in North America. We evaluated each of the techniques based on cost (the actual money spent applying the method) and outcome (how successful the technique was in discovering buried archaeological deposits and how well it explained the presence/absence of archaeological sites). The results of this cost/benefit provided the basis for the multi-stage deep test protocol presented in Chapter 12.0.

The project design was uncomplicated and intended to answer a very simple question for each method: “What did we learn, and how much did it cost to learn it?” First, six ca. 1.0 ac (0.4 ha) areas that had differing depositional and archaeological characteristics were selected from across Minnesota. These included two areas known to contain buried archaeological deposits (the Fritsche Creek II and Anderson test locales [Chapters 6.0 and 10.0, respectively]) and four others that had never been evaluated but were believed to have a relatively high potential for buried cultural deposits (the Hoff Deep, City Property, Clement, and Root River test locales [Chapters 5.0, 7.0, 8.0, and 9.0, respectively]). The three deep test methods were then applied to each of the test locales by separate, independent field teams. Field work for each method proceeded in the order of each method’s relative impact to the subsurface at each test locale: geophysical survey was first, followed by coring/augering, and lastly backhoe trenching. In general, the results (i.e., whether and where buried archaeological material was found) of each method were kept separate, and each field team independently prepared a report without any knowledge of what was learned by other methods.

Each of these methods has its own unique strengths and weaknesses, which are discussed in detail in Chapters 4.0 and 11.0. The strength of remote sensing is, first and foremost, that it has no impact on buried archaeological resources. Remote sensing procedures included GPR, magnetometry, and resistivity using a multi-probe array. These were selected not only because they are among the most frequently used in archaeology, but also because they map somewhat different soil and archaeological properties. Subsurface maps, pseudo-sections, or three-dimensional reconstructions of buried surfaces can all be created. Additionally, in the proper circumstances, possible buried cultural features can also be identified in the subsurface. Each survey method, however, also has its own specific environmental constraints. Magnetometry is limited to less than about 75 cm (30 in) below the surface, while GPR is nearly useless to employ in areas comprised of silt- and clay-rich alluvium.

The greatest strength of geophysical survey methods, no impact on the buried deposit, is, ironically, also its greatest weakness. The picture of the subsurface that each of the geophysical survey methods provides is virtual. It cannot be seen or directly evaluated. Moreover, we do not even know if it really is present, if it is a methodological artifact, or if it is a natural object or feature. In this study, as has been true on prehistoric sites across North America, geophysical surveys commonly yield false-negative and false-positive results. This clearly creates reliability problems for site discovery and, for this reason, we have not recommended it for site discovery. Geophysical survey, however, does have a place in the buried site evaluation process.

The coring/augering procedure used for the project is a two-step process. First, a set of continuous, 1.75-in (4.5-cm) diameter cores were collected and used to identify target horizons (paleosols or horizons that could include buried archaeological material). Because the probability of actually finding archaeological material in a small diameter core is quite small, these target horizons were then sampled using larger diameter (4 in or 5 in [10 cm to 13 cm]) flight augers. The augers were drilled to the top of a target horizon and spun to clean the hole of overlying sediment. The auger was then run into and through the horizon and the sediment from the horizon was drawn up, then screened for archaeological material. The coring and augering process has little impact on buried archaeological resources, but can still be used to construct maps, cross-sections, and a three-dimensional model of the subsurface. More importantly, it can also yield tangible, not virtual, artifacts and has a generally unlimited depth range. Augering in this study, which is the part of the process likely to yield buried artifacts, is only applied if target horizons are identified.

Although backhoe trenching may seem deceptively simple, in this project we followed a relatively sophisticated procedure. The method focused on collecting data to develop a three-dimensional model of buried landscapes. Several trenches were excavated across the test locales to study specific sedimentary or geomorphological aspects of the landform. The subsurface was mapped based on soils and sedimentary horizons observed in trench profiles, which were also visually inspected for archaeological material. In addition, archaeologically promising horizons were screened to further test for artifacts. Trenching was best at exposing ephemeral soil, sediment and archaeological horizons and also yielded excellent contextual information. Not only are cultural features more likely to be discovered, but sedimentary and pedological details are also more apparent in trench profiles. All of this makes it easy to develop a flood plain depositional history and to understand the relationship between sedimentary units, paleosol development, and archaeological deposits and features. However, this detail comes at a cost. Backhoe trenching is clearly destructive, and backhoe trenches are limited to a maximum of 3 m to 4 m (9.8 ft to 13 ft) depths.

The goal of deep testing is to discover whether buried archaeological deposits are present in an area. Consequently, the coring/augering and trenching methods included some standard archaeological practices. These were mentioned in passing in the above discussion and involve mainly screening sediment to search for archaeological material. In the case of the coring and augering procedure, sediment from targeted horizons was passed through a one-quarter inch screen to search for any possible artifacts. This involved only identification of artifacts. No context evaluation, feature studies, or similar archeological skills were employed because the context for discovered artifacts derived from augering is limited to the knowledge that they were retrieved from somewhere within the target horizon. Details of the context of artifacts could be evaluated during backhoe trenching. To determine the presence/absence and the relative density of cultural material, and collect additional contextual data, 50 cm × 50 cm (20 in × 20 in) test units were placed along the wall of backhoe trenches and excavated. Similar to augering, excavations focused on target horizons that either had a high potential for containing archaeological materials or actually included them. Standard archaeological excavation and recordation methods were employed, and sediment was passed through a one-quarter inch screen. Features were identified, and artifact context was evaluated. A primary goal of this procedure was to see if enough information could be collected during Phase I deep testing to make a determination of National Register eligibility.

Given that several states either require or are considering requiring that similar procedures be employed during deep testing, an evaluation of what additional information can be derived from such test units is important. While the screening of augered sediment was essential to the coring/auger process and cannot be eliminated, as it turned out, little was learned through the trench test units that was not already known from carefully sampling and inspecting the trench walls. The test unit procedure, however, complicated the logistics and significantly increased the cost of the trenching procedure. We do not believe that the additional time and expense of excavating test units are worth the minor amount of additional information obtained.

13.2 PROJECT RESULTS AT TEST LOCALES

Trenching yielded the best results for discovering buried archaeological deposits. Sites were "discovered" at both of the test locales that housed known sites and at two of the four test locales where sites were not previously known. This is an impressive 66 percent (four of six) rate of site discovery and, even more significantly, a 50 percent (two of four) new site discovery rate. Coring also yielded good results but discovered only half as many sites (one of two test locales with previously recorded sites and one of four test locales without previously recorded sites). The geophysical surveys yielded mixed results and, as mentioned above, have doubtful utility for efficient buried site discovery. Geophysical survey regularly returned false negative and false positive results, which raise serous questions concerning the reliability of such methods for site discovery. Such methods, however, can play a significant role in the buried site evaluation process, as aids to mapping buried horizons, and searching for buried features at known archaeological sites.

From an outcome standpoint, backhoe trenching proved remarkably successful. The greater success of trenching, compared to the coring/augering procedure, demonstrates how effective this method can be in revealing subtle details of sedimentology and pedology and how these data articulate with the archaeological data. Moreover, the specific successes of buried site discovery for both coring/augering and trenching during this project are also due in no small part to the application of the predictive elements of Mn/Model in selecting areas to investigate. The relatively high new-site discovery rates for both methods clearly indicate that the chances of discovering buried archaeological sites are actually quite good within areas mapped by the Mn/Model Landscape Suitability Models as having a high or moderate potential for buried sites. This observation also substantiates the general utility of the Mn/Model LSRs as a predictive tool for planning purposes and underscores the benefits of expanding Mn/Model LfSA mapping to cover all of Minnesota. Our results, however, also highlight that like all predictive frameworks, the LSRs cannot be a substitute for on the ground testing. For example, at the Hoff Deep test locale archaeological deposits were discovered in relatively thick, Holocene alluvium that occurs in an area mapped by the LfSAs as glaciolacustrine. This is not a criticism of the LfSA mapping, which at the other test locales was generally accurate, but to point out that, like any regional model, it is not always accurate at the site-specific scale. Rather, the LfSAs should be viewed as a "work-in-progress" and should continually be updated as new, more detailed data are collected or new theoretical drainage basin histories are formulated.

The six test locales represent a wide range of geographical situations in Minnesota. The Clement test locale (Chapter 8.0) focused on ridge and swale depositional environments that included relatively coarse-grained sediments deposited within vertical accretion levee systems. The LfSAs indicated that, within such environments, archaeological materials were unlikely to occur deeper than 1 m to 2 m (3.3 ft to 6.6 ft) but were at least moderately probable at more shallow depths. Trenching successfully detected the presence of vertical accretion levee deposits and was able the map the extent of this landform in the subsurface. Coring also determined a similar depositional environment, and both concluded that that high-energy, horizontal accretionary channel or point-bar gravels underlay the sequence at 1 m to 2 m (3.3 ft to 6.6 ft) depths, which, as indicated by the LfSAs, would have destroyed any earlier prehistoric occupations that may have existed on the flood plain.

Coring and trenching diverged in their abilities to discern very subtle details of the depositional and pedogenic structure in the subsurface. Trenching exposed many more details within the series of stacked, ephemeral, and probably short-lived paleosols found in the upper part of the levee. Coring could not resolve these effectively and ultimately considered them unimportant and, with the exception of one core with a buried soil, no augering took place. Regardless, their presence actually marks the beginnings of landform stabilization and, consequently, suggests a higher potential for buried archaeological resources. This potential was fulfilled because low-density prehistoric occupation debris was discovered buried within the paleosol sequence at about 80 cm (32 in) below the surface. Although no diagnostic artifacts were found, a ^{14}C age estimate of 1510 BP indicates a late Middle or early Late Woodland age for the occupation.

Importantly, the buried archaeological material within the paleosol sequence at the Clement test locale allowed direct comparison of how effective trenching and the other methods are for detecting low-density occupations within relatively coarse-grained, vertical accretion levee sequences. Because it was the only method to actually confirm the presence of cultural material, trenching was clearly the most successful for discovering such site types and reconstructing the subsurface of similar landforms. The strength of trenching at Clement is not just in site discovery, but also in its ability to collect the kinds of data that allow archaeological deposits to be placed into their appropriate temporal and stratigraphic context within the reconstructed landform depositional framework. This is crucial for designing an effective and efficient site evaluation strategy.

Trenching at the City Property and Root River test locales (Chapters 7.0 and 9.0, respectively) focused on depositional environments associated with smaller tributary systems in upstream (Root River) and stream-mouth (City Property) contexts. Interestingly, these two locales were also the only areas tested that did not include buried archaeological sites, even though sequences of stacked or cumulative paleosols were found within the trenches and cores. The lack of archaeological deposits within the soil and sediment sequences, however, permitted comparisons of how well the methods addressed why occupation was absent and the probability that buried cultural deposits were actually present but were missed. From this standpoint, the reasons for the absence of buried, prehistoric archaeological deposits became clear only after a chronology, based in these two instances on ^{14}C ages of detrital wood/charcoal collected from stratigraphic contexts during trenching, was established for these locales. These data clearly indicated that the sediments are simply too young to likely include prehistoric deposits and/or that, during development, landforms were not composed of environments conducive to human occupation. Without the application of ^{14}C age dating to these stratigraphic sequences, such observations could not have been made, nor could one determine the likelihood that buried resources were present at these locales, but missed. This not only demonstrates the importance of the multidisciplinary approach for buried site discovery, but also highlights that any method employed for deep testing in a compliance framework ultimately must draw defensible conclusions regarding the presence or absence of buried cultural material.

Work at the Fritsche Creek II test locale, which is part of a previously investigated archaeological site (Fritsche Creek II site, 21NL0063) that includes buried Early and Middle Archaic occupations, focused on a stratigraphically and depositionally complex alluvial fan within which archaeological deposits were buried. From the perspective of the Mn/Model LSRs, these valley-margin fan landforms have a high potential to include deeply buried archaeological deposits of nearly all ages. Both coring and trenching successfully reconstructed the alluvial fan depositional environment and discovered relatively shallowly buried artifacts from contexts similar to that described in the original research at the Fritsche Creek II site (Roetzel et al. 1994). However, only small pieces of bone and, significantly, no lithic or ceramic artifacts or fire-cracked rock, were discovered while coring and augering. During trenching, in addition to abundant bone and charcoal, artifacts that included chipped stone tools and debitage and fire-cracked rock were recovered from at least two distinct and temporally separate occupation horizons. In fact, the trenching results showed that the Fritsche Creek II site is clearly different and far more complicated than originally believed. Trenching revealed considerable detail about the character of archaeological deposits and their contextual relationships to sediments and soils that form the site landscape. Regardless of methods, the geoarchaeological work at the site highlights the importance of undertaking complete and detailed geological studies early in the site discovery process. Scientific studies focusing on establishing the stratigraphic and temporal context of archaeological deposits and placing them within their proper place in the landform developmental history are not peripheral to, but rather, are essential to archaeological site discovery and evaluation.

The original plan for work at the Hoff Deep test locale was to compare the effectiveness of the buried site discovery methods in distinguishing between areas underlain by late Wisconsinan varved glaciolacustrine clay with low probability for buried archaeological sites from areas of Holocene clayey alluvium with high archaeological site potential. However, as was true at the Clement test locale, serendipity again played a role when we discovered not only that the entire test grid was underlain by at least 1 m (3.3 ft) of Holocene alluvium, but this alluvial sequence also included stratified archaeological deposits that ranged in age from the Late Archaic through the Woodland periods. These new data show that geomorphic mapping at 1:24,000 scale may not be detailed enough to anticipate specific site specific conditions and boundaries. From the more immediate perspective of this study, low density archaeological material was discovered stratified within the alluvium throughout the Hoff Deep test locale. Unlike at Fritsche Creek II, both coring/augering and trenching discovered a wide range of artifact types including bone (burned and unburned), charcoal, lithics (debitage, tools, and fire-cracked rock), and ceramic material. Trenching was able to provide far more sedimentological and pedological data from the subsurface than coring. Even though comparatively few trenches were excavated, they revealed considerable detail pertaining to the age, landform development and history, and the character of the archeological deposits within the site stratigraphic framework. The long profiles of backhoe trenches were able to effectively clarify the effects that lateritic soils have on site formation, the bank slump process, and details of ephemeral soil horizons.

Nevertheless, the fact that these relatively few trenches had a large impact on the subsurface emphasizes the strength of coring in certain circumstances. The comparative complexity of the Hoff Deep locales demonstrated that coring is far more effective in tracing the extent of subsurface horizons (that it can detect) with a much lower impact than trenching. This is particularly apparent for sites including several distinct and complex depositional settings like the Hoff Deep test locale. Coring's prominent role is during the buried site evaluation process and is discussed in Chapter 12.0. Work at the Hoff site also showed that, through close cooperation among earth scientists and archeologists, the judicial application of ^{14}C chronologies, and an understanding of regional geological history and processes, new ideas integrating long-term, basin wide depositional systems with related archaeological settlement patterns can be formulated. Hopefully, some of these ideas and geoarchaeological data from the Hoff Deep test locale will be incorporated into the Mn/Model LfSAs and refine the Landscape Suitability Ratings (LSRs) for the Red River Valley.

Finally, the Anderson test locale is atypical for this study because it is not located in a fluvial or alluvial depositional environment. Rather, its glacio-fluvial and eolian (dune) setting was chosen because it is composed of coarse-grained sandy sediment rather than fine-grained alluvial silt and clay. The Anderson test locale also was chosen because it includes part of the Anderson site (21AN0008), a large multi-component site that has been investigated since the early twentieth century that includes buried, possibly stratified occupation horizons.

Geophysical survey was able to map cultural features across the landform. In conjunction with pedological and sedimentological information derived from coring and trenching, geophysical survey data also provided details concerning the distribution of subsurface horizons across the landform. While coring revealed no deeply buried soils or target horizons to test, the cores did provide sufficient geomorphological information to conclude that the depositional context is not principally eolian and that the presence of archaeological materials at appreciable depths below the plow zone is probably due to biological and pedological processes. Likewise, no paleosurfaces were uncovered during trenching. Unlike coring, however, relatively deeply buried (>1 m [>3.3 ft]) artifacts and cultural features that extended well below the plow zone were discovering within the backhoe trenches. As was concluded from the coring data, the trenching process also demonstrated that sub-plow zone artifacts at the Anderson site (or at least this portion of the site) were neither associated with clearly defined buried surfaces nor likely buried by eolian deposits or processes.

The results at the Anderson test locale highlight the strengths and weaknesses of the deep test methods discussed throughout this monograph. Geophysical survey, particularly magnetometry, successfully mapped several near-surface cultural features, but also misinterpreted some subsurface horizons as depositional rather than pedogenic in origin. By combining these "misinterpreted" horizons with ground-truthed data derived from other techniques, however, their three-dimensional distribution was successfully reconstructed. As noted for Clement test locale, the coring/augering process is not reliable for deep testing in areas with poorly defined subsurface horizons because buried surfaces are lacking and no augering is undertaken. Furthermore, even at a site like Anderson with its high density of artifacts, the small diameter cores failed to include any artifacts, despite the fact that numerous large and small artifacts were found in every backhoe trench. While the presence of artifacts on the surface at Anderson makes it unlikely that the sub-plow zone component might be missed, we cannot assume that buried sites always have artifacts exposed on the ground surface.

As a result of this work, we concluded that buried artifacts within the Anderson test locale were likely introduced into the subsurface by a combination of cultural- and bioturbation processes (e.g., pit-digging, gopher burrows). Because the Anderson site is so large (i.e., 75 ac [30 ha]) and culturally varied, the conclusions reached from our study may not apply to all areas of the site. Minimally, the previous assumption of site burial by eolian processes should be re-evaluated in light of these results. As recommended in Chapter 12.0, the combination of all the data sets from our investigations together can provide the basis for developing a Phase II research strategy directed toward addressing the relative amount of disturbance and the nature of the interplay between biomantle and cultural turbation processes of sites in this or similar geomorphic settings.

13.3 PROJECT RESULTS: COSTS AND BENEFITS

Per acre costs were calculated for each of the methods and discussed in detail in Chapter 11.0. These were based on the actual time and costs to carry out each deep testing procedure at each test locale. These data are intended to allow direct comparisons of methodological costs rather than serve as a basis for establishing specific costs to carry out deep testing in Minnesota. On the basis of total costs (implementation and logistical costs) per acre (Table 11.3.2-1), geophysical survey was the most expensive, coring/augering (average cost for all test locales) the least expensive, and backhoe trenching without test unit excavation intermediate between geophysical survey and coring/augering. Yet, the costs do not differ appreciably. For example, coring/augering and trenching differ by only $369 per acre, while trenching and geophysical survey differ by only $165 per acre.

Cost is not the only concern in a cost and benefit analysis, particularly in light of the negligible differences summarized above. It must be considered in relationship to 1) methodological success in finding buried archaeological deposits and 2) effectiveness in providing independent lines of evidence to support the presence or absence of an archaeological site on the prehistoric landscape. In terms of a method's success in locating buried archaeological deposits, backhoe trenching was most effective with a 66 percent overall detection rate and an even more impressive 50 percent new-site discovery rate. Coring/augering was also effective, but not nearly as successful as backhoe trenching. This process had a 33 percent overall detection rate and a 25 percent new-site discovery rate. Geophysical survey discovered possible buried archaeological phenomena at all sites except one, but left many uncertainties as to what was actually present. Determination of exactly what these possible archaeological phenomena are can only be obtained by applying ground disturbing activities (i.e., trenching, coring, augering, test unit excavation). This, plus the fact that remote sensing identified possible cultural features so frequently, forced us to reach the conclusion that remote sensing is not well suited for use in buried site discovery. Rather, it is better suited for judicious use in the evaluation of archaeological sites.

The relative differences in site discovery outcomes between trenching and coring/augering is due to the fact that 5-m to 6-m (16.4-ft to 19.7-ft) long trench profiles are simply better able to reveal subtle aspects of the subsurface than a 1.75-in (4.5-cm) diameter core. Predictably, therefore, the sites missed by coring/augering were those of low density and associated with ephemeral soil horizons (Clement test locale) or where the subsurface archaeology was not associated with clear subsurface depositional or soil horizons (Anderson site). As numerous statistical site location studies demonstrate, the probability of discovering a site is directly related to the size of the sample area (Kintigh 1988; Krakker et al. 1983; Nance 1983; Nance and Ball 1986; Shott 1985, 1989). This factor alone suggests that the discovery of buried archaeological deposits is more likely to occur by using trenches rather than coring/augering. Moreover, the fact that coring/augering may be biased against finding sites such as Clements and Anderson not only places its effectiveness at discovering sites into question, but may also lead to the development of incomplete settlement and subsistence models because whole classes of sites are missed.

The survey procedures selected for this project provide stratigraphic, pedological, and landform history data that can be used to place the archaeological site into a three dimensional model of the subsurface. These data and the model can then serve as independent lines of evidence to explain why a site is present or likely to be present and, more importantly, why a site is not present. In the later case, scientific evidence is used to support management decisions in compliance projects rather than relying on the mere presence or absence of an archaeological site. Using this approach will be more effective than simple hand methods in the long run as it minimizes the occurrence of false negative results. The success of a deep testing method and confidence that the deep testing process can reliably determine the presence or absence of buried archaeological deposits, with some certainty in most environments, are as critical to developing a deep test protocol as cost. The cost of failure in the deep testing process is simply too high to rely on any but the most reliable methods for buried site discovery.

While alternative techniques to mechanized deep test methods were not a part of our field investigations, we also considered their costs in addition to their strengths and weaknesses . These alternatives are purportedly cheaper to implement because they do not employ the multidisplinary, geoarchaeology framework that requires highly trained personnel. The techniques that are often suggested include hand or power augering, hand excavation of test units, or sometimes even just “deep” shovel testing using bucket augers with extensions. To estimate their costs, we extrapolated the costs for some of our procedures to achieve a reasonable per acre costs for hand and power augering (see Chapter 11.0). Our analysis suggests that these techniques are neither as effective as nor necessarily less expensive than the three primary techniques used for this study.

In the final analysis, the deep test protocol proposed herein for Mn/DOT promotes effective and efficient methods and provides guidelines that will reliably discover buried archaeological resources. The geoarchaeological approach also facilitates an understanding of the integrity of any archaeological site(s) that may be present and provides feedback to Mn/Model LfSAs and refinement of LSRs. Perhaps more importantly, in the absence of such buried deposits, the protocol provides the means to assess the probability of false-negative results and understand why archaeological deposits are not buried in the project area, so that management decisions are informed and justifiable. Such assessments, however, can be made only if both geological development and archaeological site formation and their interrelationships are understood. This multidisciplinary approach, which we believe was critical to the success of this project, is particularly important for efficient and effective deep testing because of the complexity and expense of assessing the significance and nature of buried archaeological deposits.

13.4 DEEP TEST PROTOCOL: A STAGED APPROACH

This deep test protocol is based on the premise that the discovery of buried archaeological sites is primarily a geoarchaeological process and should be undertaken as a multidisciplinary study by earth and archaeological scientists. We have proposed a multi-step procedure that first focuses on the recommended methods and procedures for buried site discovery phase and then outlines the recommended methods and procedures for evaluating the National Register eligibility of discovered or rediscovered sites. The protocol assumes that Mn/DOT's management process for projects subject to Section 106 review and other federal and state legal mandates includes analysis by Mn/DOT CRU project managers of whether deep testing is appropriate and necessary for the specific area under investigation. Such a decision should be based both on previous research as well as an understanding of the nature and magnitude of the local and regional depositional history (i.e., site burial potential) and archaeological conditions. The Mn/Model LfSAs, for example, already provide an excellent framework for making these decisions in many places in the state and should be regularly employed for assessing the deep test needs of an area. In places where the LfSAs have not been mapped and a decision has been made to initiate an archaeological survey, decisions about the need for deep testing should be undertaken in consultation with geoarchaeological or other earth science specialists using an array of maps, soils information, and other pertinent environmental data. Furthermore, we strongly recommend that Mn/DOT continue to update and expand LfSA mapping, particularly with respect to the buried site analyses, with the goal of extending coverage throughout the state. Presently, Mn/DOT plans to expand mapping coverage as funds become available.

Given the results presented in this report and summarized above, we believe that backhoe trenching is the best method for discovering buried archaeological sites and recommend that it be the preferred method for deep testing. If trenching is not possible, coring/augering should be employed. Additionally, this procedure may also be appropriate in other instances, such as where deposits with archaeological potential lie deeper than can be reached using a backhoe. When neither trenching nor coring/augering can be used, alterative methods, such as hand (bucket) augering or test pit excavation, may need to be employed as a last resort.

The protocol stresses that data collected during backhoe trenching should be first directed toward establishing the stratigraphy, identifying horizons that have archaeological potential, and exposing the base of deposits that are likely to include archaeological material. The key to the site discovery process is the identification of depositional hiatuses that indicate times when no or very limited sedimentation occurred, as these intervals are most likely to include and preserve archaeological occupation debris. Identification and definition of the deepest limits of archaeological horizons is also straightforward and usually corresponds to either the base of Holocene deposits or, in fluvial systems, the top of high-energy channel or bar gravels.

Our research demonstrated that trenching need not be accompanied by archaeological test unit excavation. Test units yielded little, if any, information that altered or enhanced the data and interpretations that were produced by the trenches, and were relatively costly. Instead, each wall of each trench should be carefully examined by the project archaeologist for evidence (artifacts or suspected archaeological features) of an archaeological site through regular inspection and troweling at regular intervals during the excavation process. Though working in tandem with the geomorphologist, the archaeologist should independently record the locations of artifacts, possible features, and stratigraphy. When encountered, artifacts and samples of organic material should be collected for analysis and dating. Finally, the archaeologist's profile drawings should be coordinated with those of the geomorphologist for consistency, and any divergence of opinion rectified in the field. Although formal archaeological test units are deemed unnecessary, screening a consistent volume of sediments to recover archaeological materials is recommended for specific stratigraphic units that appear promising for containing archaeological materials. A 20 liter (5.3 gal) sample collected from several places along the profile within the suspect horizon(s) that is passed through one-quarter inch screen is recommended. The goals of this process are to confirm the absence of cultural material and/or to establish the density of artifacts. Consequently, the total volume of the sediment sample should be recorded, since circumstances may dictate sample size(s).

While detailed guidelines mandating criteria for trench sizes, types of backhoe bucket (i.e., smooth or toothed), or the maximum distance between trenches could be established, from a practical standpoint these are best selected to accommodate specific conditions encountered at individual project areas. 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, as well as meeting OSHA and other safety guidelines. The placement and number of trenches excavated at each deep test locale are necessarily variable and depend on the size of the project area, site testing objectives, topography, and stratigraphy. Sometimes accessibility, either due to environmental conditions or logistical considerations (e.g., landowner restrictions), is a limiting factor and can determine the ultimate placement of trenches. In general, trenches should be placed to study the subsurface expression of specific surface depositional features. Based on the results of these excavations, the placement of additional trenches should trace subsurface depositional features, soil horizons, and buried landform expressions.

The coring/augering procedure is recommended as the first option for deep testing if trenching is neither feasible nor possible. Such situations include the need to reach depths greater than is practical or safe with trenching; in densely wooded, marshy, or steep places; or in some urban settings when site conditions render trenching impractical. When used during the site discovery phase of deep testing, the coring/augering procedure should follow the two-step method used in this study. First, a set of solid-earth cores are collected using a continuous coring device such as a GeoProbe or Giddings Rig. The goals of the coring phase are similar to those for trenching and should aim to develop a stratigraphy, assess the archaeological potential of the subsurface, and at least penetrate depositional units that are unlikely to include in situ archaeological deposits (e.g., base of the Holocene). During the coring process depositional hiatuses, representing places in the depositional sequence that are likely to include archaeological occupation, are sought and identified when present. From these data, specific target horizons are selected to test for the presence of buried archaeological material using augering.

We recommend that augering be undertaken using a set of at least six 4-in (10-cm) or four 5-in (13-cm) diameter flight augers, which produce the equivalent area of a 25-cm (10-in) diameter shovel test. The soil and sediment from the target horizon should then be screened to recover any artifacts that may be present. We believe that screening the sediments through one-quarter inch mesh is sufficient for site discovery. Although a smaller mesh would allow for recovery of microdebitage and may be more appropriate for addressing certain research questions, most archaeological sites will contain a broad range of debitage sizes, as well as other categories of culturally derived materials, that can be collected through standard recovery techniques. Moreover, numerous logistical issues associated with screening matrix through fine mesh screen or floting soil samples make these procedures impractical and/or dramatically increase the cost of deep testing.

During site discovery, coring is best carried out initially using a systematic grid, which in this study was every 20 m (66 ft). While this sampling interval was generally adequate for locales that had well-developed paleosols and/or continuous/dense archaeological horizons (i.e., Fritsche Creek II and Hoff Deep test locales), the fact that no target horizons were identified during coring at the Clement test locale suggests that the sample interval may have been too coarse to detect ephemeral paleosols and low density archaeological horizons. This same phenomenon has been noted in studies addressing shovel test intervals (i.e., Kintigh 1988; Krakker et al. 1983; McManamon 1984; Nance and Ball 1986, 1989; Shott 1985, 1989). Closer interval sampling will greatly increase the coring/augering costs 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 field, however, might be a more cost effective means of increasing confidence. Similar discretionary decisions concerning trench placement were also suggested for backhoe trenching.

If no depositional units that include archaeological deposits are found during either trenching or coring/augering, then the research team can argue that the proposed undertaking within the deep test locale will not affect buried archaeological resources. While the absence of buried archaeological material is not certain (and can actually never be certain), such an argument is bolstered by the deep testing process and protocol proposed in this report. The philosophical axiom of science, *the absence of evidence is not evidence of absence*, clearly must be kept in mind and dictates that only a general probabilistic statement concerning the absence of an archaeological site can be ventured. In fact, the strength of the multidisplinary, geoarchaeological approach we advocate for deep testing is its ability to provide a rationale for reasonably assessing false-negative deep testing outcomes. The process we propose explicitly requires that the sedimentary horizons and paleosols noted in the subsurface be placed into a three-dimensional stratigraphic framework that aims to reconstruct the development of the landform as well as map the types and distribution of depositional environments. This reconstruction provides that basis for understanding why buried archaeological material was not found, determining places in the stratigraphic system where an archaeological site might occur, and assessing the chances that it actually exists in the subsurface but was missed.

The above discussion begs the question of what to do if buried archaeological deposits are absent, but the environmental and depositional conditions at the site nevertheless suggest that a reasonable chance exists that they could be present. Unfortunately, the answer to the question "at what point CAN we confidently assume there are no resources present?" is we never can be 100 percent sure. While the presence of archaeological material can be verified, its absence within an area (i.e., a negative assertion) can never be proven unquestionably. Again, this is why the proposed deep test protocol stresses the importance of developing a model of the age, stratigraphy, and depositional history of the subsurface within the project area so that probabilities for sites can be objectively assessed. In the rare case that no archaeological materials are found in an area with undeniably high potential, the options are limited to excavating additional trenches, which only reduces the area that a site may occur within, monitoring during construction, or avoidance. The resolution of this dilemma by choosing one of the above options is best made through discussions between the deep test survey team and Mn/DOT and subsequent cost/benefit analysis by Mn/DOT given the parameters of the proposed construction project.

If no depositional units that include archaeological deposits are found during either trenching or coring/augering, then the research team can confidently argue that the proposed undertaking within the deep test locale will not affect buried archaeological resources. The sedimentary horizons and paleosols noted in the subsurface should be placed into a three-dimensional stratigraphic framework that aims to reconstruct the development of the landform as well as map the types and distribution of depositional environments. This reconstruction provides that basis for understanding why buried archaeological material was not found, determining places in the stratigraphic system where an archaeological site might occur, and assessing the chances that it actually exists in the subsurface but was missed.

If buried archaeological materials are found and further archaeological investigations are recommended, the proposed deep testing protocol recommends that additional geoarchaeological investigations accompany the archaeological evaluation. As was true for site discovery, this process is also geoarchaeological and multidisciplinary. The role of earth scientists is to 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. The role of archaeologists in this process is to focus on understanding the significance of cultural material and place it into its proper cultural, temporal, and environmental contexts. The goal of this work is to determine if the site is eligible for the National Register. The process, however, is usually idiosyncratic to the specific geological conditions at a site and, by necessity, must involve discretion as to the best means of obtaining a sufficient sample of the geoarchaeological deposits to define the site function, type, and age of the occupation(s).

The first stage of the evaluation is an extension or refinement of the deep test process and should occur prior to the archaeological evaluation (Complete Phase II Evaluation in Figure 12.2.1-2). While the entire evaluation process should aim to provide enough information to facilitate the determination of the National Register eligibility of the site, the first stage focuses more on fully reconstructing the geologic context of the site and its implications for understanding the archaeological structure and function of the site. This stage focuses on the evaluation of site integrity, definition of the site's horizontal and vertical extents, and site significance. A combination of coring, remote sensing, and/or more traditional archaeological excavation methods may be appropriate. The destructive nature of backhoe trenching means that, once a site is discovered, further trenching should proceed only with great caution. While trenching could be part of the process, it should be employed only to address specific and essential parts of the archaeological assessment.

Coring, supplemented by limited augering and geophysical survey methods, is well suited to trace and map details of buried horizons across the buried archaeological site with little impact on the buried components. The core logs or geophysical data are used to map the subsurface of a buried archaeological site, refining the initial Phase I information regarding the horizontal and vertical limits of the site. As noted for the Hoff Deep test locale, coring is very effective at tracing and mapping soil and archeological horizons. In addition, it may also assist in identifying features and potential post-depositional impacts to the site as reflected in variations in the stratigraphic sequence. The appropriate geophysical technique is selected based on the nature of the soils and sediments, lithological contrasts between stratigraphic layers, groundwater issues, and potential interference of the geophysical signals related to factors such as electrical power lines or metal on and near the site. In the final analysis, exactly which method or combination of methods is chosen to map an archaeological component in detail or trace subsurface horizons in three dimensions depends on the specific conditions at the site being evaluated. The methods used should be selected based on the information needs as indicated by the Phase I results.

Once the geoarchaeological properties of the subsurface are well understood and the vertical and horizontal limits of the site have been defined, archaeological evaluation should proceed. Work undertaken during this second phase is similar to the more traditional evaluation process in Section 106 review projects and aims to evaluate the archaeological characteristics of the sites. Augers or hand excavated test units can be employed to collect additional archaeological data relevant to defining the density and distribution of artifacts and the presence of cultural features. One might also consider using remote sensing methods to assist in finding possible cultural features to test. The choice of technique largely depends on the nature of the deposit, as defined during site identification, and information derived during the geoarchaeological phase of the evaluation regarding the horizontal and vertical extents of the site and the general distribution of materials. Despite the relatively high labor costs associated with the excavation of test units, they may be more appropriate than augering if the site occurs at depths less than 1.5 m (4.9 ft) and/or if heavy machinery can be used to strip off sterile layers to reach greater depths quickly and safely. Test units placed on a grid can easily be tied in with data recovery excavations both vertically and horizontally. In addition, test units, perhaps complimented with backhoe trenches in situations in which the stratigraphy is particularly complex, may be appropriate to provide more controlled recovery of materials. 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. The fine scale vertical control needed to discern complex or subtle stratigraphy is not provided by augering, and the results would be more difficult to integrate into data recovery excavations. 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.

13.5 CONCLUSIONS

The Mn/DOT DTP project goals were to test various methods of identifying buried archaeological sites, conduct a cost/benefit analysis, and develop a protocol of methods and procedures to follow in identifying buried archaeological sites. Though the project researchers expected the project to be a success, the results of the cost/benefit analysis were particularly enlightening because 1) they objectified the choice of methods for the protocol and rendered any personal biases moot and 2) cost turned out to be a less important factor than we anticipated. The small difference in per-acre cost for each of the three methods tested allowed us to focus on the efficacy of the methods. While each of the tested methods had their own successes, the results of the project demonstrate that trenching unquestionably had the highest success rate of identifying buried archaeological sites.

The choice of trenching as the primary method employed in the deep test protocol was based on its success at discovering buried archaeological sites in the greatest number of geomorphological settings rather than its utility to advance the multidisciplinary approach advocated by the research team. However, the success of trenching (as well as the coring/augering procedure) underscores the effectiveness of a multidisplinary team of archaeologists and earth scientists over a team comprised on only archaeologists or only earth scientists. We reiterate that finding buried archaeological sites is not a daunting task if undertaken within its proper geoarchaeological context. For example, the explanation of why a site is not present, which is required to justify a management decision to move forward with a project, is critical and requires the multi-disciplinary approach we advocate if management decisions are to have a high success rate and are scientifically valid. Explanations of why a site is not present require a geoarchaeological explanation that uses geological, geomorphological, pedological, and archaeological knowledge and methods and theory. This is, in our opinion, the most important contribution this project makes to the management of deeply buried archaeological sites.

Likewise, the evaluation of the National Register eligibility of a site is best approached from a multidisciplinary perspective. Certainly, the evaluation of sites can be straightforward and require only minimal non-archaeological input. Nevertheless, when the archaeological data is ambiguous, a geoarchaeological approach is critical for developing a sound, scientifically based explanation of why further work is or is not recommended. Unlike plow zone sites where large areas can be exposed by mechanical stripping to determine the presence and density of subsurface features in a cost effective manner during a Phase II evaluation, the evaluation of deeply buried sites is not only a more complex and difficult task, but also has the potential to be extraordinarily expensive. The approach we advocate employs the methods and tools of earth and archaeological sciences when they are most effective and cost efficient during the site identification and site assessment phases. The types of broad areal exposure possible at the surface can be both logistically problematic and prohibitively costly in evaluating a deeply buried site. This is particularly true for sites with complicated stratigraphic and depositional frameworks that are more than just layer-cake in structure. With buried sites we simply do not have the luxury of knowing *a priori* what the site contains archaeologically or depositionally. However, as demonstrated throughout this project, we believe that a multidisciplinary approach has a much better chance of be efficiently developing and undertaking the assessment of a site's eligibility than either discipline alone.

From an archaeological viewpoint, the most unexpected result of the project is that test unit excavation did not produce information that superseded the archaeological information obtained from examining the test trench walls. This observation is also probably surprising to most archaeologists. The excavation of test units adjacent to the test trenches was made part of the trench methodology because 1) several states where CCRG research team members have conducted deep testing have such provisions in their guidelines for projects subject to Section 106 review and 2) the **perceived** experience of certain members of the CCRG research team was that test unit investigations conducted in the past produced significant archaeological data that negated the significant additional cost to the project. That test unit excavation failed to provide additional significant information beyond that acquired investigating the trench walls at all four test locales that included a buried archaeological sites plus the magnitude of the added cost of hand excavations compelled us to drop this procedure from the protocol. This is not to say that no useful information was provided, but rather that most of the information derived from test units merely duplicated what was known from just trenches alone. The advantages of deep test trenching and test unit excavation are clearly different. The former (trenching) is best at site discovery and placing any buried archaeological material within its proper depositional, stratigraphical, and pedological contexts while the latter is best used to assess archaeological content. Ultimately, too few archaeological test units were excavated to realistically address questions of archaeological content, which is more appropriate during the evaluation stage. Moreover, the long trench profiles afforded more cost effective information concerning archaeological context than the 50 cm × 50 cm (20 in × 20 in) test units.

The fact that traditional archaeological test units were dropped from the Phase I deep test protocol, however, did not compel us to also leave archaeologists out of the process. Rather, we believe that archaeologists must be present during trenching and merely suggest that their time and effort are better expended on working directly with project earth scientists on understanding the stratigraphic and geoarchaeological context of sediments and cultural deposits. Again, archaeological knowledge of prehistoric and historic cultural sequences and artifacts, as well as how the discovered archaeological materials fit into archaeological site formation processes and depositional contexts, are critical to understanding what was found and formulating management decisions.