Case Studies in Geoarchaeometry

Jack Arthur Johnson

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2018

Reading Committee:
James K. Feathers, Chair
Donald K. Grayson, Co-Chair
Peter V. Lape

Program Authorized to Offer Degree:
Department of Anthropology

University of Washington

Abstract

Case Studies in Geoarchaeometry

Jack Arthur Johnson

Chair of the Supervisory Committee:
James K. Feathers
Department of Anthropology

This dissertation consists of four standalone papers. Each paper addresses a distinct geoarchaeological challenge through the application of specialized technical methods and experimental data. New approaches to data gathering are developed, and familiar approaches are combined with new archaeological applications and software tools to yield new lines of evidence useful for the examination of significant archaeological questions.

The first paper uses portable X-ray fluorescence (PXRF) to measure elemental concentrations in sediments and ceramics, but identifies several serious issues with common instrumentation and practice, then develops novel protocols and software tools to address these issues. The second paper describes a successful test application of PXRF to the relative dating of rock varnish accumulations atop petroglyphs at Hole-in-the-Ground in southeastern Oregon. The third paper details the use of luminescence dating and Bayesian depositional modeling to create a robust multi-proxy site formation history at Bear Creek in Redmond, Washington. The fourth paper uses luminescence dating of sediments and Bayesian modeling to document over 2000 years of changes in El Niño- and earthquake-driven depositional activity in the Santa and Chao Valleys of Perú, and discusses the implications of these patterns for archaeological research.

Table of Contents

Chapter 1. Foreword

This doctoral dissertation represents the completion of 12 years of research. It feels great to have it written.

It consists of four standalone papers. The first two papers focus on the use of portable X-ray fluorescence (PXRF) technology in geoarchaeological applications, while the second two focus on the use of luminescence dating of sediments and Bayesian depositional modeling to examine culturally relevant ancient environmental change. Each paper seeks to address a distinct geoarchaeological and geochronometric challenge through the application of specialized technical methods and experimental data. Throughout this research, the methods and data employed are rigorously tested, evaluated, and refined, at least to the best of my ability. New approaches to data gathering are developed, and familiar approaches are combined with new archaeological applications and software tools to yield new lines of evidence useful for the examination of significant archaeological questions.

Overall, the intellectual aim of this work is to make contributions to research practice in addition to contributions of novel data and interpretive power. Here, these contributions are made through the use of measurement science as a tool for improving geoarchaeological methods and understanding. In other words, this is a work of geoarchaeometry, as it deploys an archaeometric approach towards geoarchaeological ends.

This work was also guided by ethical considerations. For example, concerns for the preservation of cultural heritage were preeminent throughout. No artifacts were destroyed by this work, and no sites were disturbed electively. Only one study presented here relies on insights

yielded by the excavation of archaeological materials, and in this case excavation was triggered by imminent site disturbance. One study uses pulverized archaeological ceramics as analytes, but pulverization had already been performed by other researchers for other purposes. All other materials were examined non-destructively or collected from non-archaeological contexts. The low-impact nature of this research is by design; I am averse to electively damaging our cultural heritage solely to satisfy my own curiosity.

A collaborative approach was also central to this research. Although all analytical equipment used in this work was provided by the University of Washington, two papers in this volume represent projects conducted under professional contracts, and in these instances collaboration meant working with government agencies, members of affected descendant communities, and private archaeological consultants, while structuring the research itself to meet their concerns and needs. Two papers reflect research which incorporated excavation, and in these instances collaboration meant working with descendant communities and archaeologists during fieldwork, while learning about the past together. Only one paper in this volume was completed in isolation, shaped only by my scholarly interests and objectives. The work detailed here therefore owes much to the aid and influence of others, to whom I am grateful for the many opportunities and lessons provided.

The first paper in this volume examines issues with accuracy, repeatability, and precision in PXRF measurement. Its primary objective is to establish a means of using PXRF to accurately measure elemental concentrations (and uncertainty) in sediments and archaeological ceramics. It documents several barriers to achieving this objective in common practice, and it creates and

tests a complex protocol and statistical code which addresses these barriers. Through extensive experimental data, it also shows that *only* this protocol currently produces verifiably accurate results under observed conditions; other common protocols are tested and shown to be deficient. This paper is of fundamental valuable to my research; all three of the other papers cite its insights on the abilities and limitations of PXRF. It was originally published online in 2012, then in hard copy in 2014 (Johnson 2014).

While the first paper frets about the fundamentals of producing valid PXRF results, the second paper attempts to expand the breadth of PXRF applications to include the relative dating of petroglyphs. It simultaneously constructs and assesses a method of using PXRF in this context, and shows that PXRF sufficiently measures differences in rock varnish manganese content to age sort petroglyphs under observed conditions. At the field site studied—Hole-in-the-Ground within Oregon's Owyhee Canyon—this measurement revealed distinct age groups of petroglyphs, including some which are apparently very ancient, and some which were episodically re-pecked across the centuries. Altogether, then, this study aims to produce both a proof of concept and a glimpse of the types of new archaeological insights available upon its further use. Fieldwork for this study was performed on behalf of SWCA Environmental Consultants under contract to the Vale District Office of the U.S. Bureau of Land Management (BLM). The paper presented in this volume is partly derived from a prior consulting report on file at the BLM, but it also presents new analysis and conclusions.

The third paper in this volume focuses on the application of luminescence dating and Bayesian depositional modeling as part of a multi-proxy effort to firmly establish the antiquity of

the Bear Creek archaeological site in Redmond, Washington, and to provide a basis for interpretation of the depositional and occupational history of that site. Before this effort, the antiquity of the Bear Creek Site was hotly contested. Research presented here makes it clear that the site is around 12,000 calendar years old, making it the oldest known professionally studied and well-accepted archaeological site in the Puget Lowland. The primary value of this study is therefore its contribution to our current understanding of Puget Lowland culture history, but I hope it also provides the Washington cultural resources management (CRM) community an encouraging example of the application of luminescence dating and Bayesian depositional modeling to CRM contexts. At present, such methods are seldom applied in Washington CRM investigations. Work detailed in this paper was performed on behalf of the University of Washington Luminescence Laboratory as part of a collaborative effort led by SWCA Environmental Consultants, whose services were funded by the City of Redmond, Washington.

The fourth paper in this volume uses the same primary methods as the third—luminescence dating and Bayesian depositional modeling—but applies them to a much larger challenge: the late Holocene depositional history of north coastal Perú. This history is intricately linked with El Niño flooding and earthquake activity in the area, and depositional activity stemming from such events has long been argued to have occasionally been a major factor in the lives of ancient Peruvians. But local proxies for that activity remain poorly dated, and therefore their cultural relevance poorly understood. As the first step towards a remedy for this situation, this paper examines Chao Valley flood sediments and Santa Valley beach ridges to build a local chronology of major ancient floods and earthquakes. This chronology is sharpened with

lithologic evidence to link it to independently documented local droughts, allowing the timing and patterning of a wide range ancient local natural disasters—and their cultural implications—to be explored and compared in new detail. As with the preceding papers, its primary objective is explicitly empirically oriented, aiming to supply both methodological (the production of a novel line of chronometric evidence which is accurate, precise, and robust) and documentary (the production of long-term and widely useful data) advances to researchers interested in coastal Perú. As above, I hope this paper encourages greater use of these techniques in coastal Perú, where so much of the material culture is dateable with luminescence, yet so many chronometric challenges await us. Field and laboratory research for this study was funded in part by the National Science Foundation…in 2007. This paper is original to this volume, and it took 12 years to finish. It feels great to have it written.

REFERENCES

Johnson, Jack
2014 Accurate measurements of low *Z* elements in sediments and archaeological ceramics using portable X-ray fluorescence (PRXF). *Journal of Archaeological Method and Theory* 21:563–588.

Chapter 2. Accurate Estimates of Elemental Concentrations in Sediments and Archaeological Ceramics Using Portable X-ray Fluorescence

ABSTRACT

This study seeks to demonstrate the ability of portable X-ray fluorescence (PXRF) to estimate concentrations of potassium (K), calcium (Ca), and iron (Fe) under controlled conditions. After a discussion of the potential confounding factors in PXRF use, a protocol which attempts to address these issues through repeated measurement, calibration, and re-sampling is detailed. Data generated using this protocol are then tested for accuracy and replicability. PXRF is argued to be able to produce accurate estimates of K provided the suggested protocol is used, and able to produce repeatable estimates of K and Ca under these same conditions. Fe results are found to be problematic given the standards used here.

INTRODUCTION

The use of portable X-ray fluorescence (PXRF) is on the rise in archaeology. More papers using PXRF are appearing in journals and edited volumes, PXRF data are featured more often at professional conferences, and these papers and presentations are branching out thematically, as they now encompass everything from the sourcing of lithics (Burley et al. 2011; Craig et al. 2007; Davis 2011; Forster and Grave 2012; Millhauser et al. 2011; Nazaroff et al. 2009; Phillips and Speakman 2009; Sheppard et al. 2011), glasses (Polikreti et al. 2011), and ceramics and clays (Frankel and Webb 2012; Goren et al. 2011; Speakman et al. 2011) to

chemical characterization of sediments (Abrahams et al. 2010; Davis et al. 2011) as well as a range of in situ materials (e.g., Liritzis and Zacharias 2010; Potts and West 2008).

Within this movement lies a body of commentary reminding users to apply the method with conscious, diligent care and a critical eye towards results. The voices of this undercurrent offer a variety of insights ranging from illumination of key issues in the application of PXRF (Shackley 2010) to evaluations of the accuracy of PXRF (Craig et al. 2007; Goodale et al. 2011; Millhauser et al. 2011; Nazaroff et al. 2009). When the latter type of investigation has been conducted, a common result is that PXRF produces source clusters which match those derived from established methods of elemental analysis such as laboratory-based WDXRF, EDXRF, and INAA but fails to produce elemental concentrations which match these independent methods (Craig et al. 2007; Goodale et al. 2011; Nazaroff et al. 2009). In other words, published studies show that PXRF often produces inaccurate elemental concentrations which pattern in accurate ways. This fact should give the rapidly expanding user-base of PXRF pause, especially given 1) growing interest in non-sourcing applications for which elemental concentrations accurate in and of themselves are needed, and 2) the obvious superiority of accurate results for comparison between independent datasets and studies. Going forward, it will be necessary to rigorously test which uses of PXRF yield demonstrably accurate results, to document the conditions under which such results can be obtained, and to publicize these conditions as an aid to future users. Further, as Shackley (2010) points out, it is imperative that work begin now, while the PXRF boom is relatively young and the volume of potentially flawed data relatively small.

The primary goal of the analysis presented here is to address this need through the development of protocols for preparation, measurement, and analysis which consistently yield accurate estimates of elemental concentrations. My specific interest is in deriving estimates of potassium (K), calcium (Ca), and iron (Fe) in sediments from archaeological contexts as an aid to larger investigations into geochemistry and radiometric dosimetry. As regards dosimetric measurement, the particular goal of this study is to evaluate the potential of PXRF for measurement of K in conjunction with luminescence dating, as radioactivity from ^{40}K, which composes about 1.17% of all naturally occurring K, is a major contributor to the generation of luminescence signals (Aitken 1985:62). Given this objective, this study is also concerned with the accurate estimation of errors associated with PXRF use, since these are used in the calculation of errors in final luminescence age estimation.

Because some prior work suggests that elements with low atomic number—such as K and Ca—may pose a significant problem for PXRF measurement (e.g., Goodale et al. 2011), accurate measurement of these elements is particularly likely to be challenging (see also Buhrke et al. 1998:51–54; Lechance and Claisse 1995:204–205). I will attempt to meet this challenge by first diagnosing the potential sources of inaccuracy and imprecision in PXRF measurement, with specific emphasis on the instrument used here, a Bruker Tracer III-V. Once these sources of error are defined and their potential impacts on resultant data are detailed, a protocol which attempts to address these issues through strategic preparation, calibration, and measurement will be set forth. Following this, experimental testing of this protocol's ability to consistently produce accurate results is discussed, along with the broader implications of the test results.

SOURCES OF ERROR IN PXRF

I categorize sources of error in PXRF measurement into two primary conceptual types. These are: 1) errors external to instrumentation, and 2) errors internal to instrumentation. Accounting for both types of error is essential to the derivation of accurate data, and both types will be discussed here.

External Errors and Calibration

External errors are defined here as those errors which are not purely due to fluctuations in machine performance; they are instead due to a range of factors, including sample preparation, instrument calibration, and user error. Of these factors, only instrument calibration needs elaboration here. Issues related to sample preparation, while potentially significant, are relatively easily addressed and are fully detailed elsewhere (e.g., Buhrke et al. 1998).

Calibration

Calibration is a numerical translation from one set of units to another. In the case of XRF, the process translates the intensity of the observed response to X-rays, measured in counts, to elemental concentration, measured in parts per million (ppm) or percent mass. XRF results cannot be reported as elemental concentrations without this translation.

XRF calibration can be undertaken in many ways (Lachance and Claisse 1995:259–275), but in general the process of calibration and estimation used involves a special type of regression analysis known as inverse prediction (see Seber and Lee 2003:145–148). In this analysis, standards—materials of known elemental concentrations—are measured with XRF to observe the counts these concentrations generate, then a best-fit equation is used to describe the

relationship between concentrations (the independent variable) and counts (the dependent variable). In nearly all cases this relationship is imperfect, causing error to be passed to results, which are accompanied by probabilistic confidence intervals to encompass this error. As a general rule, calibrations have less error (and smaller confidence intervals) when more standards are used, more replicate measurements of each standard are used, and standards are homoscedastically distributed throughout the full range of observed values in all dimensions. Once a sound regression is achieved, field samples of unknown concentrations can then have PXRF counts measured and interpolated into the regression equation to infer the concentrations of interest. Error terms in these concentrations are then derived using interpolation into the confidence intervals attached to the regression function.

Importantly, traditional regression analysis relies on assumptions about the nature of the data, and valid estimates of concentrations and error terms can only be obtained if these assumptions are met or if the model is appropriately adjusted to correct for deviations from assumptions. Among these assumptions is the notion that error terms in x-values are both randomly distributed about each x-value and their magnitude is independent of the x-values themselves (Seber and Lee 2003:227). The latter of these conditions is not met by the standards used in this study, as elemental concentration is a strong predictor of reported error for some elements, such as K (Figure 2-1) and Ca ($R^2 = 0.958$). Given this, confidence intervals in calibration for these elements cannot be assumed to be adequately modeled by traditional regression analysis, and results based on uncritical calibration may carry grossly inaccurate error terms. If traditional assumptions are applied to the K values for these standards, for example,

underestimation of error terms will systematically become more pronounced as K concentration increases.

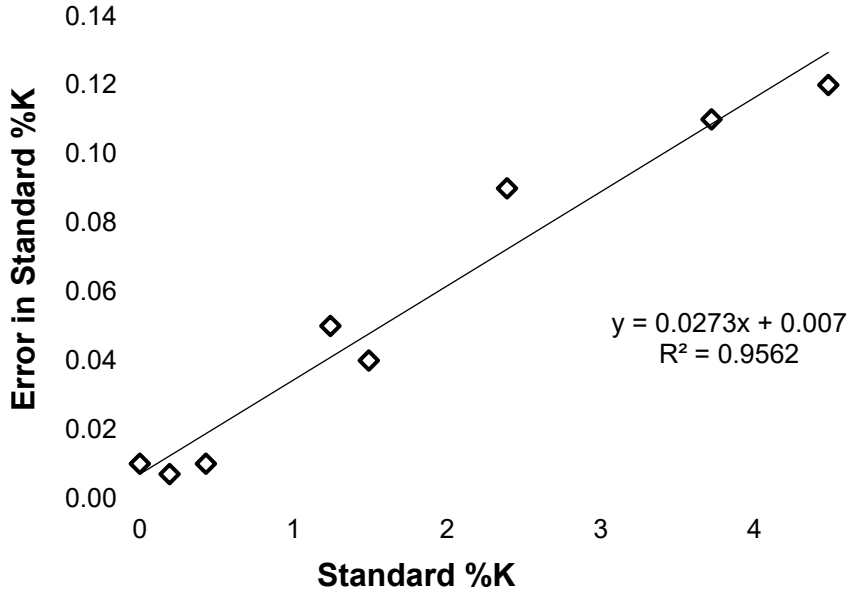

Figure 2-1. Reported 1σ errors in %K concentrations of standards plotted against reported %K concentrations of standards. A value of 0.01 was used for the error in the blank. The two variables are strongly correlated.

There are a number of ways to address this issue (see Seber and Lee 2003:265–292), but they fall into two primary types. The first approach attempts to solve the problem at a more theoretical level and proceeds by modifying the regression model itself as a way to account for the effects of errors in x-values. In practice this approach involves a complex process of identifying trends in variance, estimating the magnitude of these trends, calculating errors in estimation, and then either applying these estimates and errors to the overall regression equation or applying a transformation to data so that they behave normally. This approach complicates matters, especially given the fact that trends in errors in x-values may themselves vary in

strength (i.e., how well the trend fits all x-values and errors), magnitude (i.e., the rate with which error terms grow in proportion to x), and shape (i.e., the mathematical function describing the trend) depending on the set of standards and the element used. For use in PXRF measurement, then, this approach can present a barrier to routine use, as this approach greatly diminishes the ease and possibly the overall precision with which calibration can be performed.

As an alternative, the second approach is to use iterative re-sampling to simulate a distribution of expected outcomes of regression analysis given the reported error in each standard concentration. Here, the x-value assigned to each standard during regression is probabilistically drawn from the Gaussian distribution defined by μ and σ, where μ is the reported elemental concentration of the standard and σ is the reported error in this concentration. Regression can then be performed with these sampled x-values as normal, as these values are empirically reflective of reported errors. This process is repeated a number of times to empirically generate a statistical distribution of outcomes of regression parameters such as slope and y-intercept, as well as confidence intervals about these parameters. In this way, patterned errors in x-value inputs are automatically embedded in regression results, avoiding the need to model them theoretically. This method is relatively easy to carry out with the aid of widely available statistical software, and is therefore the more practical of the two options available.

In addition to errors in x-values, it is probably also necessary to account for errors in y-values used in calibration, since repeated PXRF measurements of a given sample will generate a distribution of counts due to u-drift (see below). Further, as with errors in x, distributions of y-values tend to enlarge as x-values increase, indicating a positive systematic relationship between

x-values and errors in y (see Figure 2-2 for K); this relationship generally also holds for Ca (R^2 = 0.975) and Fe (R^2 = 0.738). This presents a similar problem as described above for errors in x. As above, there is more than one way to incorporate this error in y-values, but here it is sufficient to note that re-sampling will adequately address the issue. To ensure quality results, however, such re-sampling should be used to generate y-values in two phases—first for each standard during calibration, and then for each field sample during interpolation—as both sets are subject to the same source(s) of error in the estimation of their count values. In both cases, repeated PXRF measurement to generate technical replicates should be used to generate y-value distributions before bootstrapping is undertaken. If this re-measurement and re-sampling is not done for both standards and field samples, however, errors in regression and interpolation will not be fully taken into account, and resulting estimates, particularly those of high-concentration field samples, will again be prone to systematic underestimates of error.

Instrumental Errors

Instrumental errors are defined here as those errors due to changes in machine performance which affect the data collected. Such errors are ubiquitous when using complex instruments and cannot be fully prevented, but must rather be taken into account in the process of measurement. In XRF measurement, these fluctuations are often termed "drift," and three major types of drift have been documented for XRF instruments: ultra short-term drift (hereafter u-drift), short-term drift (hereafter s-drift) and long-term drift (hereafter l-drift) (Buhrke et al. 1998). Each has a unique effect on instrument precision and accuracy which is relevant here.

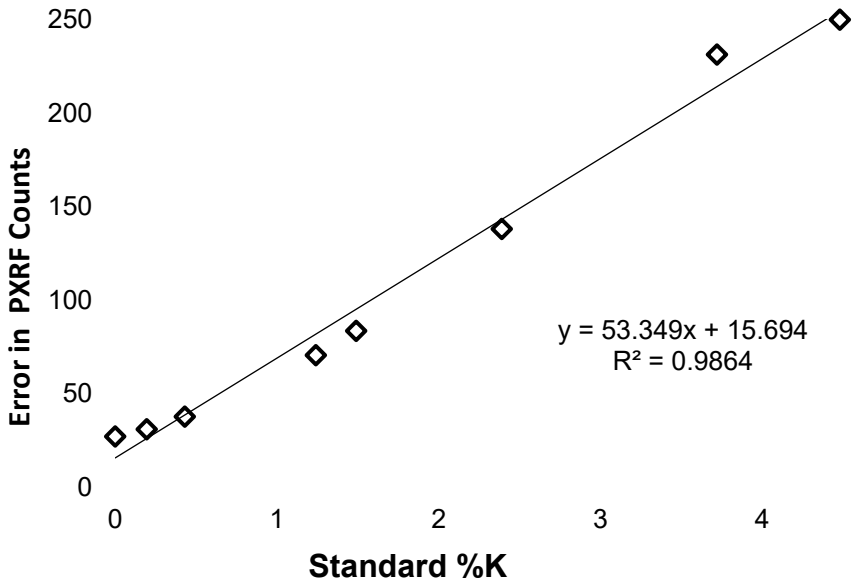

Figure 2-2. 1σ error in observed PXRF counts of K for all standards, averaged over all trials and plotted against reported %K concentrations of standards. The two variables are strongly correlated.

u-Drift

u-drift occurs when the instrument yields different intensity counts for consecutive measurements using identical measurement parameters. Figure 2-3 shows u-drift using a Bruker Tracer III-V at 15 keV and 20-µA excitation with a 25-µm Ti filter and the vacuum attachment to generate a series of 60 one-minute assays using a standard which was unaltered between assays. This type of drift theoretically results in a Gaussian distribution of observations where the mean value represents the best estimate of real PXRF counts and error terms are randomly distributed. In practice, the mean value can be accurately estimated two ways. First, the analyst can use longer assays, as collecting data for more time will average out short-term variation in counts. This approach masks short-term error in favor the obtaining an approximation of the

mean. Alternatively, a series of short, repeated measurements can be used to flesh out the distribution of values caused by u-drift. This method is preferred, since it provides useful additional information relevant to the magnitude of short-term drift, which may itself vary between instruments, instrumental settings, etc. As a general rule, then, a series of short, repeated measurements is probably better than a single, long measurement for a given sample, provided all else is equal.

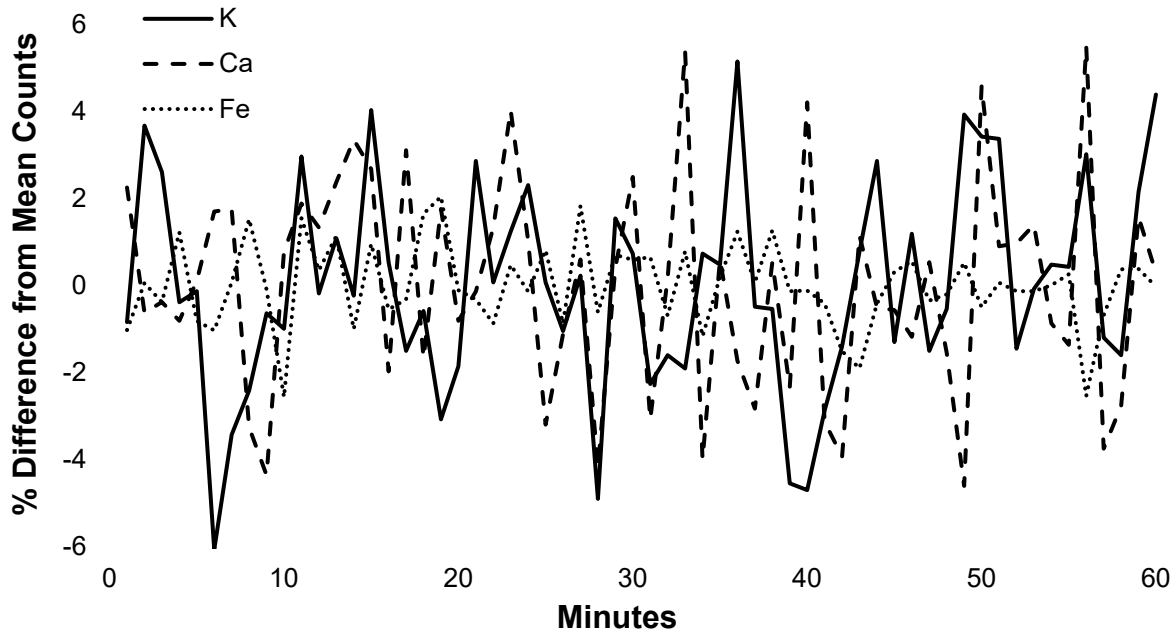

Figure 2-3. u-drift wiggles on a single standard measured with a series of 1-minute assays. The y-axis represents percent difference from each element's 60-minute mean counts.

l-Drift

l-drift occurs when instrument sensitivity is permanently changed over time, typically through the accumulation of parasite lines or tube degradation as a result of routine use (Buhrke

et al. 1998:307). This change in sensitivity results in the gradual accumulation of systematic bias in observed counts. Eventually measurements from a given instrument will no longer be directly comparable with earlier measurements taken from that instrument. Thus the clock is ticking on all PXRF instruments currently in use, as all PXRF instruments are likely to suffer l-drift at some point if they are used long enough.

The solution to this problem is re-calibration of each instrument whenever l-drift causes it to stray far enough from its prior sensitivity, but it is up to the user to monitor the change and determine when re-calibration is necessary, as little information is available for the timescales on which l-drift can affect the variety of instrument models currently in use. As a general rule, then, all PXRF instruments should be regularly checked for sensitivity changes due to l-drift. For the user community at a whole, it would also be helpful for individual users to publicize the timescales on which l-drift is significant for specific instruments and instrumental settings, as this would help set professional standards for the frequency of re-calibration. Publicity would also help users identify which machines are the most stable and durable over the long haul. For the present, however, the degree to which l-drift limits the use-life of a given calibration remains an open question for the instrument used here.

s-Drift

s-drift occurs outside the typical length of a single measurement window (usually 1–5 minutes or so) but within the timeframe of a typical use session of the instrument (say, 3–4 hours). Like u-drift, it tends to occur randomly, it can either increase or decrease instrument sensitivity, and its effects are temporary. Unlike u-drift, however, its effects can last for hours,

essentially causing the mean of the ultra short-term distribution to systematically increase or decrease for a given period of time. Figure 2-4 shows an example of s-drift derived from the same experimental parameters used to document u-drift above, although patterned data of this type has been directly observed for both of the Bruker Tracer III-V instruments tested independent of filter, vacuum, electron voltage, micro-amperage, and counting time used. Further, s-drift occurs after variable durations of use, it lasts for variable amounts of time, and none of the real-time feedback provided by the instrument was predictive of its onset or effects during testing conducted for this study. The effect of a single instance of s-drift can also differ according to the elements analyzed; Figure 2-4 shows a case in which instrumental sensitivity to K and Ca is increased at around the 20th minute, while sensitivity to Fe decreases.

This particular type of drift therefore presents unique challenges to practical use of PXRF, and its effects on instrument performance and resultant data can be profound. If, for example, an analyst were to measure the same sample with the same instrumental parameters 10 times each at the initial and terminal portions of a single 4-hour use session, it is entirely possible that the distributions generated by these two sets of measurements would significantly differ. In the moderate term, then, s-drift produces systematic bias between short-term datasets, which is a problem for the derivation of accurate results. In the long term, however, errors due to s-drift can be seen as effectively random, as s-drift does not appear to consistently affect individual use sessions in the same way.

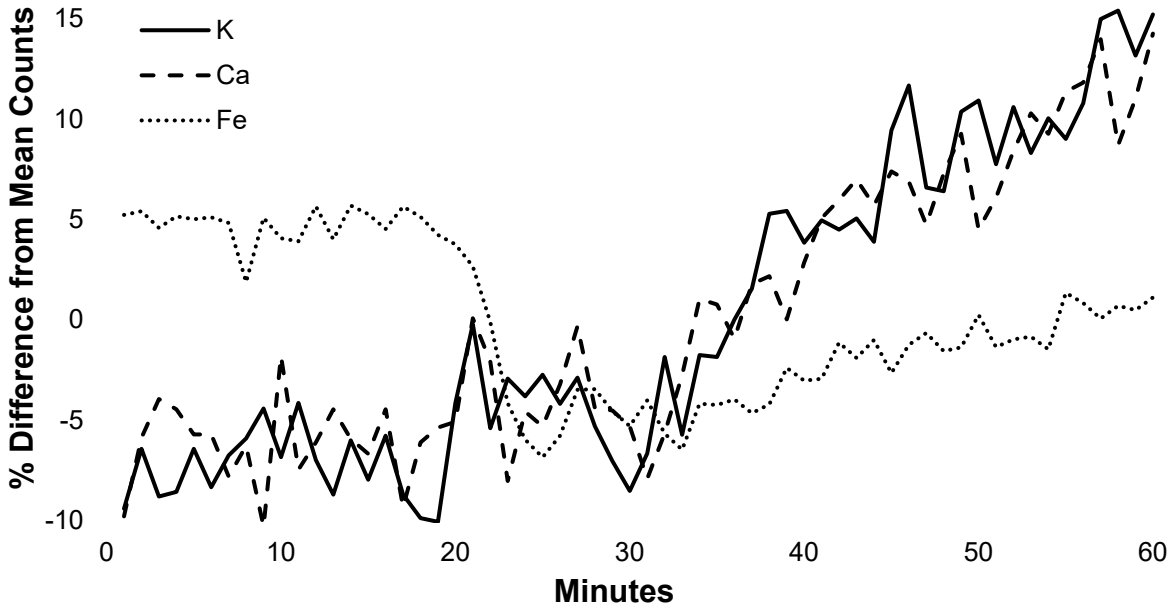

Figure 2-4. s-drift on a single standard measured with a series of 1-minute assays. In this case, K and Ca increased relative to each element's 60-minute mean, while Fe counts decreased. Wiggles are due to u-drift.

In practice, then, s-drift can be addressed through a series of steps designed to randomize the effects of s-drift on PXRF counts data over the long term. First, samples measured within a given use period should be re-measured in such a way that individual measurements of each sample are evenly distributed throughout the use session. This strategy effectively subsumes the effects of s-drift within the distribution of repeated measurements for each sample, allowing the user to attempt to address s-drift and u-drift simultaneously. Second, the median of counts distributions should be used as the measure of central tendency instead of the mean, as the median is more robust to secondary modes and outliers—both possible products of s-drift—than the mean (Schuenemeyer and Drew 2011:10). Third, samples should be measured in multiple

sessions of use, preferably on separate days, so that results of individual sessions can be compared for consistency. This step is critical, as it will allow the user to identify the few cases in which s-drift is potentially responsible for the primary modes of distributions within a given use session, and by extension medians which are heavily contaminated by s-drift. Further, when results from independent use sessions are integrated, any effects of s-drift on each of the individual sessions' results will be randomized, helping to eliminate systematic errors in estimates. Lastly, re-measurements of a set of standards should also be interspersed within each session of use to allow the tracking of s-drift within an individual session, thereby providing diagnostic data which can help inform the user.

As an alternative to the above approach, it is also possible to use data from re-measured standards in an attempt to correct for s-drift by re-scaling s-drift-affected counts values to match counts at time of calibration. Such correction through re-scaling would use a formula such as:

$$\frac{S_m}{S_c} = \frac{F_m}{F_c}$$

where S_m and F_m are observed s-drift-effected counts for a given standard and a field sample, respectively, S_c is the observed counts for standards used in initial calibration, and F_c represents counts for a given field sample on the scale used in instrument calibration. Importantly, however, this correction will be burdened by error from two main sources. First, it will incur errors in estimating the ratio between S_m and S_c, since 1) each of these terms is subject to error in estimation (especially Sc, which embeds errors in the original regression), and 2) any relationship between these terms must hold for all calibrated standards to be effective. In other

words, the definition of S_m:S_c is functionally a mini-calibration—in this case one that translates s-drift-affected counts to counts on the scale used for original calibration—and as such it is subject to all the sample-size effects and errors of any other calibration. Second, it will incur errors in the use of S_m:S_c to infer F_m:F_c, as would interpolation into any other calibration; it will also pass all of these errors on to results. While it is possible to formally account for these errors, it may not be practical to do so each time s-drift asserts itself. The former solution may therefore be more useful in practice, since with this approach 1) the errors s-drift causes don't require extra work to calculate, and 2) its primary inconvenience—the need to measure samples a number of times—is in itself a benefit in generating robust data.

Summary of Issues and Experimental Goals

Each of the above sources of error has the potential to impede accurate PXRF results. Given this, the most reliable way to derive accurate PXRF results is to control as many of these sources as possible, while accounting for the others. Errors due to sample preparation are easily controlled if simple procedures are followed. Errors in calibration and interpolation are inevitable, but they can be fully taken into account through statistical regression models which make use of re-sampling from distributions of known concentrations and observed PXRF counts. Instrumental errors are inherent and pose a significant challenge, but their effects on PXRF data can be minimized through strategic use of the same redundant measurements needed for calibration and interpolation. This strategy should involve re-measurements of individual samples which are evenly distributed throughout a single session of use and repeated over multiple sessions of use in order to account for s-drift. It should also involve repeated

measurements of standards as a part of standard practice to help diagnose the onset of s-drift and l-drift and provide feedback about instrument performance. If these steps are taken, it should be possible to produce PXRF results which are both consistent and reproducible by independent methods.

The experiment described below is designed to test this assertion through intensive analysis of a set of calibration standards and field samples. It seeks to do so by creating relatively ideal conditions in sample preparation, calibration, and measurement in an attempt to verify the accuracy of the instrument under these conditions. The protocol used in preparation, measurement, and analysis used to create these conditions is detailed first. Next, the degree to which calibrations accurately described the variation in data derived from standards is discussed. Then, the instrument's ability to accurately derive K concentrations is evaluated, along with the statistical consistency with which the instrument is able to derive results for other elements of interest. Lastly, an evaluation of instrumental performance over the period of experimentation is provided, and its implications discussed.

EXPERIMENTAL METHODS

Field samples

In total, 71 field samples were selected for analysis. Fifty-one of these were sediment samples collected from a series of sites in north coastal Perú. These are samples for which elemental compositions have not yet been independently measured. Their primary value to this study is in providing data relevant to the evaluation of the consistency of PXRF results throughout repeated measurements. The other 20 samples were archaeological ceramics from a

series of sites in North Carolina loaned from the University of Washington (UW) Luminescence Laboratory. K concentrations for these ceramics were independently measured by ICP-MS for use in luminescence measurements carried out by Jean-Luc Schwenninger at Oxford (Feathers 2009), and these samples are therefore useful for the evaluation of instrument accuracy.

Sample Preparation

Bulk samples were first dried in a 50°C oven for 24 hours to remove moisture content. Next, samples were pulverized using a sterilized tungsten-carbide orbital rock crusher until the consistency of fine flour was achieved for each sample. This step homogenized each sample and put each sample in a similar form as the standards used, each of which is critical to ensuring XRF measurements would be maximally comparable between assays and between analytes (see Buhrke et al. 1998:35–49). Archaeologists intending to take advantage of the non-destructive potential of PXRF may want to avoid this step, but in cases where analytes are compositionally heterogeneous (such as sediments and ceramics) it is highly recommended when possible. Cost considerations prevented use of a glass-forming agent or a pellet press in sample preparation, and samples were left in powdered form for analysis. Although not ideal, this shortcut did not appear to have any damaging effect on results.

After pulverization, samples and standards were each placed in sterile Teflon bottles and bottle apertures were covered with 4-µm Ultralene film to create measurement windows. This film was secured in place to seal samples inside bottles for the duration of experimentation, helping minimize potential contamination caused by repeated handling. Sample and standard bottles were stored in a desiccation cabinet when not undergoing measurement, preventing re-

hydration of samples over the period of experimental work, as such re-hydration could affect apparent elemental concentrations.

Calibration Standards

Standards, defined here as materials of known elemental values for use in calibration, are detailed in Table 2-1. To ensure accuracy, only standards with elemental values verified by multiple independent methods were used. Specific standards were chosen to provide a broad range of values for relevant elements, thereby allowing for more robust regression fitting. Importantly, one of the standards was a "blank" composed of the same type of Teflon bottle and Ultralene film described above, but with no sample inside. This blank provides a zero value for calibration and a means of measuring background counts values. In total, eight standards were used. Preliminary testing indicated this number produced sufficient resolution in calibration while minimizing measurement time over the multiple cycles of repeated measurement described below.

Table 2-1. Standards used in Calibration, with Relevant Elemental Concentrations and 1σ Errors

Standard	%K	+/-	%Ca	+/-	%Fe	+/-
AGV-2 Andesite	2.39	0.09	3.72	0.09	4.68	0.09
BCR-2 Basalt	1.49	0.04	5.09	0.08	9.65	0.15
BHVO-2 Basalt	0.43	0.01	8.17	0.12	8.63	0.14
Blank*	0.00	0.01	0.00	0.01	0.00	0.01
DNC-1 Dunite	0.19	0.01	8.21	0.05	6.97	0.10
G-2 Granite	3.72	0.11	1.40	0.06	1.86	0.12
GSP-2 Granodiorite	4.48	0.12	1.50	0.04	3.43	0.11
PACS-2 Marine Sed*	1.24	0.05	1.96	0.18	4.09	0.06

* Denotes non–U.S. Geological Survey standards. PACS-2 was made available by the National Research Council Canada.

Measurement Parameters

Measurements were carried out using a Bruker Tracer III-V with a 25-μm Ti filter and the external vacuum pump attachment. The machine was selected for its flexibility in configuration, the filter was selected after some experimentation as the best way to reduce noise in the keV range of interest to K, Ca, and Fe, and the vacuum pump was used to help prevent attenuation of the extremely low-energy X-rays which characterize this range. To prevent heterogeneity across the measured surface, all samples were presented to the instrument in such a way that the sample completely covered the beam profile. Additionally, all samples measured were at least 6 mm thick across the beam profile to fully ensure "infinite thickness" and thus reliable comparison between samples. Samples were also gently tapped prior to measurement to remove any air pockets which might have affected the penetration and escape of X-rays. The instrument was always turned on at least 0.5 hour prior to initial measurement to allow the detector temperature to fully cool and stabilize. Likewise, the instrument's X-ray tube was always turned on for at least 1 minute prior to initial measurement to help eliminate the effects of initial fluctuation in electrical current—and therefore spikes in the data—as exploratory analysis often yielded such spikes within the first few seconds of tube activation. Given this, instrument hardware was never manipulated once measurements had begun, as measurements themselves were initiated solely with the use of software. All measurements used a 15-keV, 20-μA stimulation with a 1-minute counting window, and Bruker's S1PXRF software (Version 3.8.22). Extensive preliminary experimentation with these samples showed these parameters to generate sufficient signals for all elements of interest to this study.

Measurement Sequences

Each measurement cycle was composed of two tiers. The lower tier, the *assay*, was a single 1-minute interval in which a single sample was exposed to X-ray stimulation and the resultant intensities counted. This is the basic temporal unit of measurement in this study. The higher tier, the *trial*, was always a sequence of 180 individual assays; 10 assays each for 10 different samples, and 10 assays each for the 8 standards. Assays were interwoven in such a way that two samples would be measured once each, followed by two standards, then two samples, etc. This process was repeated until each sample and standard had been measured 10 times, with each of the 10 measurements being interspersed throughout the trial to help account for s-drift. As used here, the trial was the basic unit of analysis; all trials carried their own self-sufficient local calibration data, and all trials contained enough sample data to provide a robust estimate of elemental concentrations with error terms. These repeated local calibrations allowed statistical comparison of the relative stability of the instrument over repeated trials, as well as the degree to which l-drift affected calibration parameters over the course of experimentation.

Each of the 71 samples examined here underwent no less than two individual trials. The majority (n = 50) underwent at least three trials. Again, this provided redundant measurements to help eliminate potential errors due to user oversight or s-drift, while allowing for testing of the replicability of results and the stability of the instrument over time. Repeated trials also helped improve precision in results due to increased sample size, as most samples were measured 30 times total. Each of the 20 ceramics used for the evaluation of instrument accuracy was measured

30 times. Altogether, results reported here make use of data from around 2,100 sample assays and exactly 1,680 standard assays derived from exactly 21 individual trials.

Data Selection

Raw data used for analysis were selected by a process designed to incorporate only those data most directly representative of the elemental concentrations of interest. Bruker provides software which will do this task, but this software is a bit of a "black box" in its operation and is also not compatible with the calibration program in R software described below. To circumvent this black box in a way useful to this study, a custom process for identifying and incorporating raw count data was used.

The first step in this process was to examine the visible spectrum for each assay in S1PXRF to identify the most pronounced peak for each element of interest. For purposes of clarity, Ca will be used as an example here, although for all of the elements of interest to this study the relevant peak is always the more intense K-orbital peak, also known as the K_α peak. Next, the range of electrical channels potentially relevant to the Ca K_α peak were identified using S1PXRF. Relevant channels were defined as those which indicated Ca counts distinctly exceeding background noise. Raw data for each trial were then compiled into a table, and a total of within-trial counts was derived for each of the electrical channels previously identified as potentially relevant to the K_α peak for Ca. The relevant channel with the highest total was taken to be the apex of the Ca peak for that trial. This channel, hereafter referred to as the primary channel, was assumed to be the channel most sensitive to Ca, as this channel by definition exhibited the most pronounced XRF response to samples' Ca concentrations. Next, the

correlation of each channel within the identified range with the primary channel was calculated to yield R^2 values describing the strength of the relationship between the primary channel and each of the other channels under consideration. Channels highly correlated with the primary channel were also taken to be directly indicative of Ca; channels less correlated with the primary channel were likely responding to other factors, including background noise. Channels which were highly correlated with the primary channel, hereafter referred to as secondary channels, were included in analysis. In practice it was easy to identify secondary channels, as there was in all cases a large drop-off in R^2 between highly correlated channels and less-correlated channels. Additionally, in all cases the secondary channels identified were adjacent to the primary channel and contiguous with one another, making the breaking point between secondary channels and noisy channels easily identifiable. Once secondary channels were identified, counts from the primary and secondary channels were totaled together for each assay to calculate total counts for Ca for each assay. This process was repeated for all elements of interest, and channels 80–85 were used for K, 90–96 for Ca, and 155–164 for Fe for all trials to ensure totaled counts and calibrations would be directly comparable from one trial to the next. These channels are likely to vary according to instrumental settings, and users replicating this method are advised to define channels appropriate to their own uses.

Calibration and interpolation

Calibration was performed for each trial in R Version 2.13.1 using a three-stage re-sampling regime (Table 2-2) to help account for error in the following: standard concentration estimates, standard PXRF counts, regression residuals, and field sample PXRF counts.

Table 2-2. Detail of the Process used to Calibrate, Interpolate, and Estimate Errors for Each Trial

Stages of Calibration and Interpolation
Stage 1: Re-sampling x and y in standards
a) Take 10 samples (with replacement) from the distribution of reported errors in a standard.
b) Use the mean of this distribution as the standard x-value.
c) Take 10 samples (with replacement) from within-trial repeated measurements of PXRF counts of standards.
d) Take the median of this distribution as the standard y value.
e) Repeat 1a to 1d for each within-trial standard.
f) Fit a regression line to (x,y) values for standards.
Stage 2: Re-sampling residuals
a) Sample (with replacement) observed residuals. Take one sample for each x-value.
b) Modify sampled residuals to reflect variance between standards.
c) Randomly replace all observed residuals with modified residuals.
d) Re-fit the regression to modified residuals.
Stage 3: Re-sampling y in field samples/interpolation
a) Take 10 samples (with replacement) from within-trial repeated measurements of PXRF counts of standards.
b) Take the median of this distribution as the field sample y value.
c) Interpolate field sample y value into regression fitted in 2c.
d) Repeat 3a to 3c for each within-trial field sample.
Stage 4: Reiteration
a) Repeat steps 1 through 3 5,000 times to generate a distribution of outcomes for each field sample.
b) Repeat all steps for each element of interest.
c) Repeat all steps for each trial.

The first stage empirically fit a regression function which fully incorporated errors in standard concentrations and observed standard XRF counts. To do this, 10 elemental concentration values were sampled (with replacement) for each standard from the Gaussian distribution described by the mean and standard deviation reported for each standard. Next, 10 replicates of standard intensities (counts) were sampled (with replacement) from the 10 assays for each standard to probabilistically generate 10 y-values for each standard. Then a regression function was fitted to eight points, each corresponding to a single standard, and each derived

using the mean of the corresponding standard's simulated distribution of x-values (as this distribution is reported as Gaussian) and the median of its y-values (as the median is more robust to s-drift).

The second stage involved the modification of first-stage residuals to help ensure variation in residuals was adequately modeled, as the assumption of constant variance about the regression line is another fundamental assumption in regression (Seber and Lee 2003:227). This is therefore an important step in creating a regression which is statistically robust to differences between standards and in particular to the potential leveraging effects of a small number of standards used for regression fitting. Modification was accomplished by first sampling the residuals from the regression fitted as above to generate bootstrapped residuals. Next, bootstrapped residuals were modified using the formula:

$$Y_b = \hat{Y} + e_b\sqrt{\left(\frac{n}{n-2}\right)}$$

where e_b is the bootstrapped residual, $\hat{Y}$ is the predicted count value, n is the number of observations, and Y_b is the modified residual. Observed residuals are then randomly replaced with modified residuals. Residuals at this point have a constant variance that emulates the observed variation in residuals and models the effects of this variation across the full range of x-values used in the regression. The regression line is then re-fit to these modified residuals.

The third stage interpolated observed counts of measured field samples into the regression using realistic representations of counts distributions. This was accomplished by first sampling (with replacement) 10 replicates from the 10 assays for each field sample. The median

of each distribution was then interpolated into the regression function to derive an estimate of the elemental concentration for each field sample measured in the trial.

Each of these three stages was then iteratively repeated 5,000 times per trial to generate a distribution of estimates of elemental concentrations for each element and each field sample. From these distributions, mean elemental concentrations and confidence intervals were derived for each trial. Means were used as best estimates of true field sample concentrations, and are reported with 1σ error terms in this study.

Calculation of Final Estimates

Results of individual trials were combined by deriving a single mean value and error terms weighted by precision. As a result, trials with less scatter due to lower u-drift and/or s-drift had greater influence on final estimates than those with higher scatter. In cases where single-trial error terms were asymmetrical as a result of the effects of s-drift, error terms were made symmetrical by inflating errors on the smaller "side" of the concentration to match the higher. Final estimates are not listed in entirety here due to space considerations.

RESULTS AND DISCUSSION

Accuracy of Calibrations

Table 2-3 provides a summary of data derived from regression fitting for each of the 21 trials, including within-trial R^2 values for each element. All values listed are derived from linear regression.

Table 2-3. K, Ca, and Fe Linear Regression Parameters and Limits of Detection (LOD) for all Trials (LOD values in % concentration)

Trial	K Calibration						Ca Calibration						Fe Calibration					
	R^2	*a*	+/-	*b*	+/-	LOD	R^2	*a*	+/-	*b*	+/-	LOD	R^2	*A*	+/-	*b*	+/-	LOD
1	0.998	636.7	11.9	636.7	11.9	0.13	0.998	974.7	16.9	260.1	80.6	0.25	0.982	5329.0	295.8	3907.2	1718.7	0.97
2	0.989	652.4	28.1	652.4	28.1	0.30	0.994	926.5	30.4	295.0	145.0	0.47	0.976	5043.8	321.9	4169.8	1870.1	1.11
3	0.995	587.9	17.6	587.9	17.6	0.21	0.996	929.0	24.0	214.0	114.5	0.37	0.973	5057.6	344.3	4088.7	2000.1	1.19
4	0.998	612.6	11.9	612.6	11.9	0.14	0.996	927.4	25.5	246.1	121.4	0.39	0.974	4994.8	333.9	4415.2	1939.7	1.17
5	0.996	621.8	15.7	621.8	15.7	0.18	0.989	901.5	39.6	338.0	188.7	0.63	0.969	4949.9	362.5	4326.3	2106.0	1.28
6	0.997	625.8	14.9	625.8	14.9	0.17	0.996	1032.6	28.0	178.0	133.7	0.39	0.982	5368.0	295.4	3719.9	1716.3	0.96
7	0.988	638.9	28.8	638.9	28.8	0.32	0.991	1051.9	40.0	116.6	190.7	0.54	0.986	5456.8	262.3	3611.6	1524.2	0.84
8	0.998	589.4	10.1	589.4	10.1	0.12	0.991	1011.2	38.6	104.9	183.9	0.55	0.969	5232.5	380.3	4400.2	2209.6	1.27
9	0.996	621.5	15.7	621.5	15.7	0.18	0.996	1033.3	26.8	169.9	127.8	0.37	0.983	5528.3	295.6	3796.6	1717.2	0.93
10	0.996	609.8	16.5	609.8	16.5	0.19	0.998	1001.6	19.9	244.5	94.7	0.28	0.984	5601.3	294.7	3751.7	1712.3	0.92
11	0.995	596.3	18.0	596.3	18.0	0.21	0.998	1004.3	20.2	258.8	96.5	0.29	0.979	5617.5	339.1	3633.6	1970.0	1.05
12	0.997	536.8	11.3	536.8	11.3	0.15	0.996	940.3	23.5	207.0	111.9	0.36	0.974	5084.8	338.9	3764.5	1969.2	1.16
13	0.997	571.1	13.7	571.1	13.7	0.17	0.998	962.8	17.3	215.7	82.4	0.26	0.975	5176.4	336.6	3801.5	1955.7	1.13
14	0.997	501.0	12.2	501.0	12.2	0.17	0.995	876.9	26.7	172.8	127.5	0.44	0.963	5491.4	437.8	4510.1	2543.5	1.39
15	0.997	512.4	10.8	512.4	10.8	0.15	0.999	895.2	11.9	207.7	56.5	0.19	0.971	5532.0	390.3	4568.3	2267.7	1.23
16	0.995	535.6	16.3	535.6	16.3	0.21	0.998	961.3	16.1	150.2	76.8	0.24	0.973	5205.8	357.6	3608.0	2077.4	1.20
17	0.996	608.5	15.8	608.5	15.8	0.18	0.993	1079.7	36.4	173.5	173.7	0.48	0.979	5585.6	332.8	3296.4	1933.2	1.04
18	0.997	542.7	12.2	542.7	12.2	0.16	0.998	959.2	18.4	159.8	87.8	0.27	0.973	5172.0	349.0	3638.0	2028.0	1.18
19	0.996	541.1	13.5	541.1	13.5	0.17	0.999	945.1	12.8	194.4	61.0	0.19	0.977	5162.5	523.6	3443.9	1880.2	1.09
20	0.991	526.8	20.6	526.8	20.6	0.27	0.998	959.7	18.6	135.8	88.5	0.28	0.970	5119.6	368.8	3764.1	2142.6	1.26
21	0.993	536.4	19.1	536.4	19.1	0.25	0.999	970.2	9.8	189.4	46.8	0.14	0.974	5289.6	351.1	3334.8	2040.1	1.16
Mean	0.995	581.2	15.9	581.2	15.9	0.19	0.996	968.8	23.9	201.5	113.8	0.35	0.976	5285.7	348.2	3883.4	1967.7	1.12

These data suggest excellent linear regression fits for K for all standards, as R^2 values indicate nearly perfect fits (mean $R^2 = 0.995$) between known standard concentrations and observed intensities. Examination of R output for residuals (Figure 2-5) consistently shows homoscedasticity about the regression line, indicating the relationship between variables is, in fact, linear. Overall, local calibrations for K were highly accurate for all trials, since even the lowest observed R^2 value (0.989, Trial 2) provides an excellent predictive relationship and homoscedastic residuals. The average limit of detection (where LOD = b+3σ when b is the

predicted y-intercept and σ is the standard error in the estimation of the intercept (Lachance 1995:272) for K is about 0.19%, meaning K concentrations below this value are on average considered statistically indistinguishable from background noise for these data. Linear regression fits for Ca are likewise consistently excellent (mean $R^2 = 0.996$, homoscedastic residuals, minimum R^2 value of 0.989). Mean LOD for Ca is about 0.35%, which is slightly higher than K but still below any of the predicted values for field samples examined here. The methods used have therefore produced good results for these two elements, at least through the calibration stage of analysis.

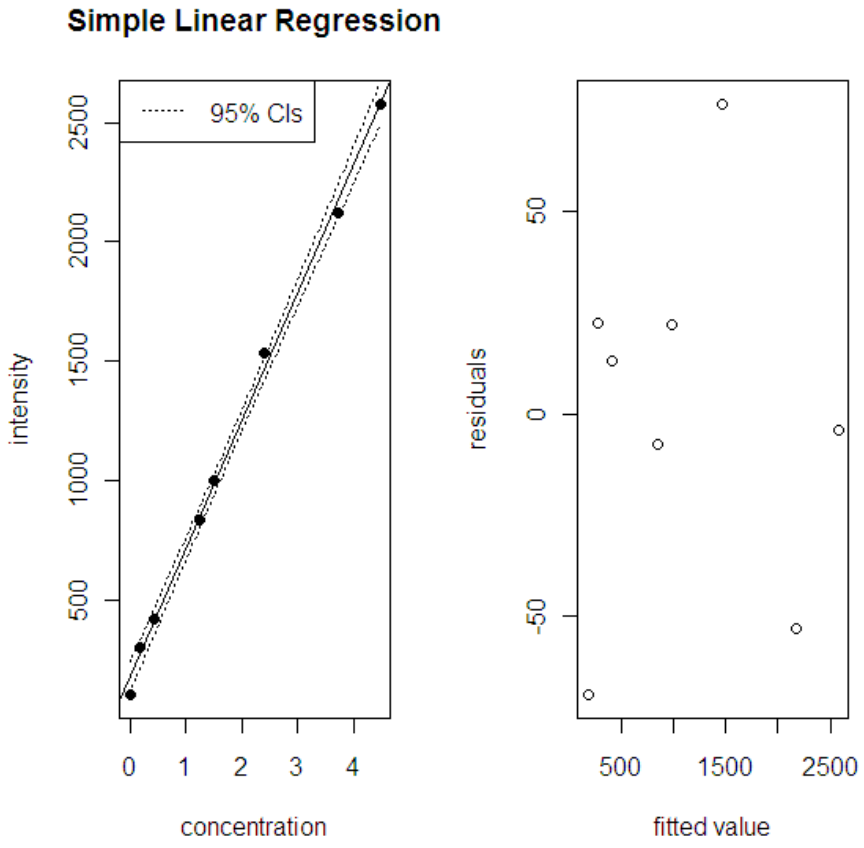

Figure 2-5. Linear regression and residuals from R output for K values of Trial 12.

Fe results are a slightly different story. While average regression correlation for Fe was generally quite high (mean $R^2 = 0.976$), residuals consistently show evidence of heteroscedasticity (Figure 2-6). The relationship between standard concentrations and counts is therefore probably not linear for Fe. Given this, if a linear regression is used to calibrate for Fe in this case, systematic bias will be introduced in resulting estimates. Here, this bias would result in overestimation of concentrations at low and high counts and underestimation of concentrations for moderate counts. This deviation from linearity is likely due to a degree of saturation on the part of the instrument; this is caused by the instrument becoming marginally less sensitive as Fe concentrations increase. Importantly, the particular pattern of the residuals here indicates a quadratic relationship between variables (Scheunmeyer and Drew 2011:121), and indeed the application of a quadratic regression to Fe data yields homoscedastic residuals as well as higher R^2 value on a trial-by-trial basis. Unfortunately, however, robust quadratic regression requires the estimation of an additional parameter when compared to linear regression, and therefore requires a greater number of standards to resolve. Fe results based on quadratic regression are therefore not reported here, as they are possibly reflective of bias due to an "over-fitted" statistical model (see Seber and Lee 2003:230). Additionally, a number of predicted Fe concentrations in field samples were higher than reported values for standards, and extrapolated values do not merit detailed reporting or analysis here. Fe results based on these data and linear regression are likely still useful (sans error terms) on an ordinal scale, although ratio-scale results for these samples will require further testing with additional standards.

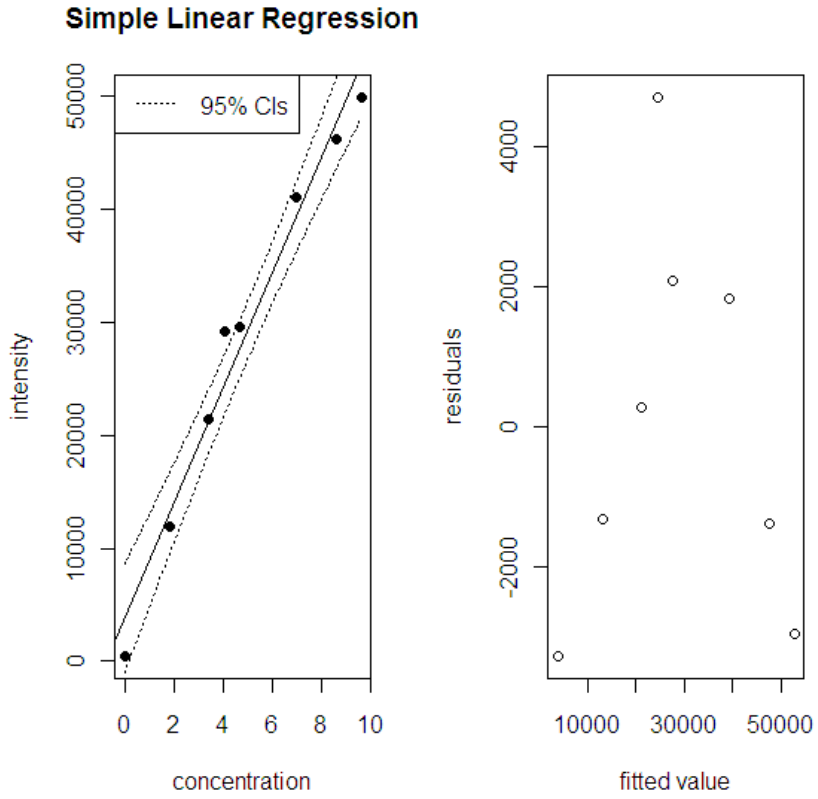

Figure 2-6. Linear regression and residuals from R output for Fe values of Trial 12.

The fact that PXRF calibration significantly varies by element provides a powerful example of the need for thorough attention to detail in the calibration and use of PXRF instruments. Similar experiments with the calibration of manganese (Mn) using the same standards, settings, and samples, for example, found that for this element the best fit is probably to be found by applying logarithmic transformation to each axis (with additional adjustments to accommodate the blank standard, as zero values are nonsensical on a log scale), suggesting that there may be multiple formulae applicable to proper PXRF calibration. Perhaps appropriate regression functions are patterned as a product of relative elemental concentration, element

atomic mass, or some other variable, although such patterning is not within the scope of this study. At present, it is therefore probably best to rigorously check the applicability of each calibration for each element of interest to avoid biased results. In many cases, this will require users to undertake manual calibration rather than rely on manufacturer software, which often does not provide much information relevant to the overall performance of a given calibration.

Repeatability of Calibrations

Given the overall accuracy of the calibrations generated for K and Ca, evaluation of the degree to which these calibrations remain consistent over repeated uses provides a good measure of the overall effects of drift on instrument sensitivity. To this end, regression parameters from each trial's internal calibration were statistically compared to examine variation in these parameters, and therefore long-term consistency in instrument sensitivity. These parameters include fitted regression slope (*a*) and y-intercept (*b*) from the linear equation and were compared using independent samples t-testing (two-tailed, n = infinite, 95% confidence) to examine the compatibility of one calibration with each of the others. The ability to perform such comparison is another benefit of manual calibration, as regression parameters and accompanying error terms are not always provided by manufacturer-produced calibration software.

On the whole, K calibrations vary significantly between trials. The slope of a given calibration line is on average statistically compatible with only 0.85 calibrations from other trials; y-intercepts are on average compatible with 2.19 other trials. Moreover, incompatible calibrations are very statistically divergent, as the average p value in t-testing of K slopes is < 0.0001. Interestingly, the few regression lines which are compatible are typically not derived

from consecutive trials, as only one regression slope is compatible between consecutive trials (Trials 17 and 18), and only six intercepts are compatible between consecutive trials; this indicates that the observed bimodal distribution of slope values (Figure 2-7) for K regression is at least partly the product of s-drift differentially affecting data between trials. On the other hand, use time is somewhat predictive of slope ($R^2 = 0.604$, negative predictive relationship) and y-intercept ($R^2 = 0.776$, positive predictive relationship), indicating that systematic error due to l-drift is also a meaningful factor over the period of experimentation (about 80 hours total) for K measurement. Specifically, the instrument seems to have more background noise and less sensitivity to K the longer it is used, even within two work-weeks of relatively continuous measurement. By contrast, instrument use-time exhibits no significant power to predict other indices of calibration performance such as R^2 of regression fit ($R^2 < 0.001$) and LOD ($R^2 = 0.005$), so the instrument has not lost any baseline ability to measure K over this time period, provided local calibration is used.

Calcium results show similar instability between individual trials, as slopes are compatible with an average of 1.33 other trials and y-intercepts with an average of 4.67. Here, three trials have slopes compatible with the following trial, and six have y-intercepts compatible with the following trial. Usage time is a poor predictor of calibration slope ($R^2 < 0.001$) and y-intercept ($R^2 = 0.225$), so long-term drift is not a major factor in Ca calibration error over the timescale observed. This leaves random fluctuations in sensitivity as s-drift as the primary causes of the pronounced differences between one trial's Ca calibration and the next. As with K, within-trial calibration fit bears no relationship to use time for Ca.

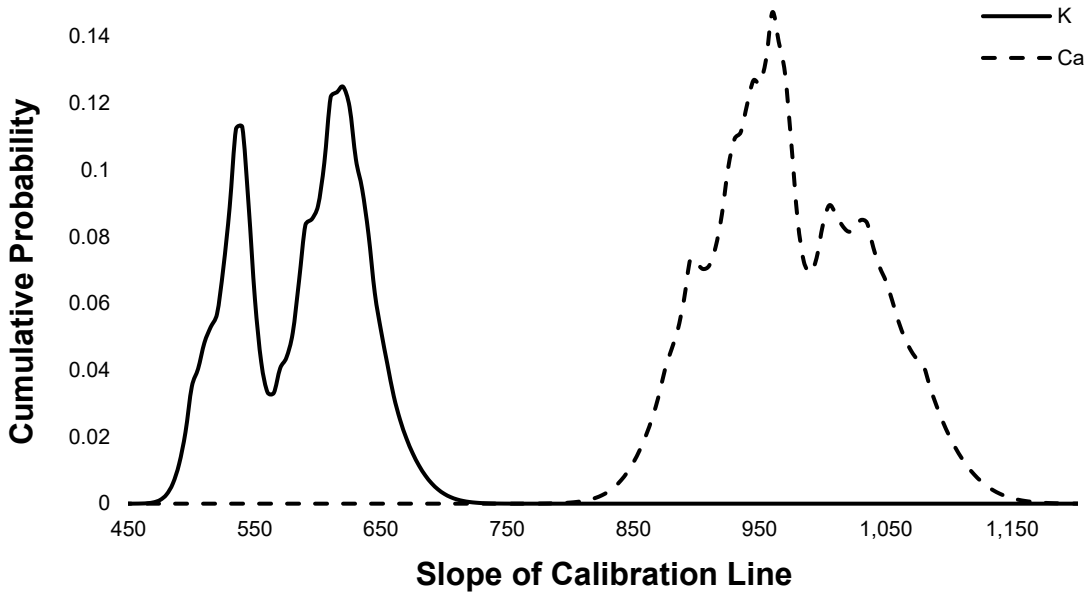

Figure 2-7. Probability density function of observed slope values for all K and Ca calibration lines.

A few implications of these results are significant here. First, even though each of the calibrations generated was quite accurate, there is a very low chance that any PXRF calibration for K or Ca will be valid outside the use period in which it is generated. In fact, in these experiments there was no observed instance where consecutive trials produced compatible K calibrations *and* compatible Ca calibrations. Functionally, then, it is almost as if each trial were performed on a different instrument, providing consistent incentive to make use of a calibration local to each trial. In light of this, many users may want to re-think the use of a large number of standards to build a single, permanent calibration for PXRF measurement. In addition, it is also noteworthy that the chance a given calibration will be applicable to future measurements may vary somewhat by element, and this chance is a product of a combination of random and systematic fluctuations in instrument performance over moderate and long timescales of use. The

effects of l-drift in particular appear to vary by element, even if instrumental settings are held constant. This variation fits with previous observations of differential performance between elements (e.g., Goodale et al. 2011) and means that a one-size-fits-all remedy may not suffice, and given this a great deal of experimentation may be needed in order to test instrument performance for the range of elements of interest to researchers.

Accuracy of Elemental Estimates

Table 2-4 provides an outlay of K concentrations for 20 ceramic samples as predicted by seven different analytical methods. The first six methods are variations on PXRF analysis which make use of different approaches to calibration; differences between these approaches will provide additional insight into the degree to which accuracy varies as a result of either 1) the software used, or 2) the temporal proximity of calibration data to field sample data. The first method used was PXRF analysis following the protocol described above, including the creation of a calibration for each trial using R software. Results from this method are referred to as "R Local" results here. The second method is a variation on the first, as it adheres to the above protocol with one key exception: it uses a single calibration derived from data generated from the first of the six trials in which the 20 ceramic samples were measured (Trial 16). These data are referred to as "R Single" results here. The third method is much like the second, except it uses data from only the first overall trial (Trial 1) to calibrate and derive estimates. This method is called "R Global" here. The fourth method used Bruker's S1CalProcess (Version 2.2.29), which was provided with purchase of the instrument, instead of R to create a local calibration for each trial. This method is referred to here as "S1Cal Local." The fifth method also makes use of

S1CalProcess, but (like the second method) makes use of a single calibration based on Trial 16 to derive all estimates. This method is referred to as "S1Cal Single" here. The sixth method used a single S1CalProcess calibration based on Trial 1 to derive all estimates, and this method most closely represents the way most users currently undertake PXRF analysis. This method is referred to as "S1Cal Global" here. Lastly, ICP-MS results are displayed (no error terms were provided with analysis). These ICP-MS results were assumed to be "true" concentration values, thereby allowing the quantification of the accuracy of the PXRF methods by comparison. Ca estimates for these samples were not provided by available ICP-MS data, so only potassium results will be discussed here.

Table 2-4. Display of %K Estimates Derived for 20 Pulverized Ceramic Samples from Seven Different Methods (listed errors are at 1σ)

Sample	R-Calibrated PXRF						S1CalProcess-Calibrated PXRF						ICP-MS
	Local		Single		Global		Local		Single		Global		
	%K	+/-	%K	+/-	%K	+/-	%K	+/-	%K	+/-	%K	+/-	%K
1776	1.15	0.09	1.21	0.09	1.12	0.05	1.27	0.03	1.28	0.03	1.48	0.07	1.00
1777	0.31	0.06	0.30	0.11	0.36	0.05	0.92	0.02	0.92	0.02	0.68	0.05	0.36
1778	1.03	0.09	1.08	0.09	1.00	0.05	1.25	0.03	1.26	0.04	1.45	0.08	1.07
1779	1.34	0.10	1.38	0.10	1.26	0.05	1.33	0.03	1.35	0.03	1.62	0.06	1.22
1781	1.27	0.10	1.34	0.11	1.23	0.07	1.60	0.05	1.61	0.06	2.14	0.11	1.08
1782	1.23	0.09	1.29	0.10	1.19	0.05	1.07	0.02	1.08	0.03	1.04	0.06	1.17
1783	1.15	0.07	1.14	0.08	1.07	0.05	1.03	0.02	1.04	0.01	0.94	0.03	1.18
1785	1.22	0.07	1.22	0.09	1.13	0.05	1.51	0.04	1.51	0.04	1.95	0.07	1.15
1786	0.62	0.06	0.62	0.10	0.63	0.05	1.02	0.02	1.02	0.02	0.90	0.04	0.59
1787	0.76	0.08	0.76	0.09	0.74	0.05	1.49	0.04	1.49	0.04	1.91	0.08	0.81
1788	1.11	0.08	1.11	0.09	1.04	0.05	1.09	0.02	1.10	0.02	1.08	0.05	1.06
1792	1.30	0.08	1.30	0.09	1.21	0.05	1.29	0.03	1.29	0.03	1.50	0.06	1.28
1793	1.04	0.09	1.10	0.09	1.03	0.05	1.09	0.03	1.10	0.04	1.08	0.09	0.90
1794	0.44	0.08	0.43	0.10	0.46	0.05	0.86	0.02	0.87	0.01	0.53	0.03	0.53
1795	1.30	0.09	1.35	0.10	1.23	0.05	1.88	0.06	1.91	0.07	2.65	0.11	1.27

Sample	R-Calibrated PXRF						S1CalProcess-Calibrated PXRF						ICP-MS
	Local		Single		Global		Local		Single		Global		
	%K	+/-	%K	+/-	%K	+/-	%K	+/-	%K	+/-	%K	+/-	%K
1796	1.27	0.09	1.32	0.10	1.20	0.05	1.67	0.04	1.69	0.06	2.28	0.11	1.17
1797	1.02	0.07	1.01	0.09	0.96	0.05	0.98	0.02	0.98	0.01	0.81	0.03	1.20
1798	0.32	0.08	0.31	0.11	0.36	0.05	0.81	0.02	0.81	0.01	0.40	0.03	0.49
1799	1.20	0.06	1.20	0.10	1.11	0.06	1.49	0.03	1.51	0.04	1.94	0.08	1.04
1800	1.14	0.07	1.13	0.09	1.06	0.05	1.31	0.03	1.31	0.03	1.56	0.07	1.27

Where S1CalProcess was used, data selection also deviates from the R protocol, as S1CalProcess automatically selects data for inclusion using its own criteria. Also, when using this software, estimates represent the median and simple standard deviation of all assays instead of the weighted mean, as above. This was done because weighted mean calculation for these data resulted in mean errors of about +/- 0.01% potassium, and such low errors are not reasonable given average reported errors in standards of about +/- 0.06%. In part this is due to the fact that S1CalProcess neither considers error terms in x-values nor reports error terms in calibration and estimation for each trial, and given this the program probably has a tendency to underestimate real errors in results. Using simple standard deviation helps inflate these error terms to a slightly more realistic +/- 0.04%, while encompassing scatter across all trials.

Accuracy testing used two statistical measures. First, a paired samples t-test (two-tailed, n = 20 pairs, 19 df) was used to examine whether independent methods produced sets of results which were statistically identical overall. Second, each of the six methods of deriving PXRF estimates for each sample was compared with ICP-MS estimates by calculating z-scores on a sample-by-sample basis. This second test was performed to highlight the proportion of individual

sample estimates which were statistically compatible with ICP-MS data. The probability that the null hypothesis—in all cases, that observed values are statistically identical—is met are listed in Table 2-5 for paired t-tests and in Table 2-6 for z tests of each sample. Tests significant at the 95% level of confidence are shaded in these tables. As a whole, t-test p values show that only PXRF results generated with the R Local and R Single calibrations produced results which are statistically compatible with the ICP-MS estimates overall. All other methods produced results which statistically differed from the ICP-MS data, including R Global.

Table 2-5. Paired t-test Results Comparing %K Estimates Derived from Various Methods

		R			S1CalProcess			ICP-MS
		Local	Single	Global	Local	Single	Global	
R	Local		0.019	<0.001	0.001	<0.001	0.001	0.445
	Single	0.019		<0.001	0.002	0.001	0.002	0.222
	Global	0.043	0.062		<0.001	<0.001	0.001	<0.001
S1Cal	Local	0.237	0.219	0.280		<0.001	0.057	<0.001
	Single	0.245	0.226	0.288	0.008		0.067	<0.001
	Global	0.385	0.367	0.428	0.148	0.140		0.002
ICP-MS		0.019	0.038	0.024	0.256	0.264	0.404	

Note: Test p values are high and right. Low and left shows the $\bar{X}_D$ value for test; this provides an index of average difference between methods' estimates of %K.

Table 2-6. p Values from z-tests Comparing Each Sample's PXRF Estimates of %K as Derived from a Variety of Calibrations with ICP-MS Estimates of %K

Sample	R Calibration			S1CalProcess Calibration		
	Local p value	Single p value	Global p value	Local p value	Single p value	Global p value
1776	0.086	0.019	<0.001	<0.001	<0.001	<0.001
1777	0.410	0.613	<0.001	<0.001	<0.001	<0.001
1778	0.704	0.938	<0.001	<0.001	<0.001	<0.001
1779	0.234	0.117	0.201	<0.001	<0.001	<0.001
1781	0.070	0.016	<0.001	<0.001	<0.001	<0.001

Sample	R Calibration			S1CalProcess Calibration		
	Local p value	Single p value	Global p value	Local p value	Single p value	Global p value
1782	0.552	0.195	<0.001	<0.001	<0.001	0.020
1783	0.653	0.633	<0.001	<0.001	<0.001	<0.001
1785	0.350	0.441	<0.001	<0.001	<0.001	<0.001
1786	0.564	0.770	0.335	<0.001	<0.001	<0.001
1787	0.562	0.570	0.002	<0.001	<0.001	<0.001
1788	0.532	0.570	<0.001	0.177	0.100	0.721
1792	0.804	0.856	<0.001	0.770	0.775	<0.001
1793	0.148	0.032	<0.001	<0.001	<0.001	0.045
1794	0.278	0.329	<0.001	<0.001	<0.001	0.920
1795	0.774	0.428	<0.001	<0.001	<0.001	<0.001
1796	0.266	0.114	0.081	<0.001	<0.001	<0.001
1797	0.013	0.024	<0.001	<0.001	<0.001	<0.001
1798	0.043	0.092	<0.001	<0.001	<0.001	0.008
1799	0.006	0.087	<0.001	<0.001	<0.001	<0.001
1800	0.078	0.097	<0.001	0.182	0.184	<0.001

The two types of R calibration which produced results compatible with ICP-MS estimates differed significantly from each other ($p = 0.019$), and are therefore not completely interchangeable. Of these two methods, the R Local was superior, for two reasons. First, it produced a slightly higher rate of compatible individual estimates (17 out of 20 versus 16 out of 20 for R Single) in z-testing. Second, in paired sample t-testing, the $\overline{X}_D$ value of R Local results was half that of R Single results (0.019 versus 0.038) when each was compared to ICP-MS estimates, indicating that the R Single results were on average twice as divergent from ICP-MS data as R Local results. Of all the PXRF methods used, then, the R Local produced the best results, the R Single produced results which are technically accurate as a dataset but less

harmonious on a sample-by-sample basis, and the R Global and the S1CalProcess software produced inaccurate results. Figure 2-8 shows the contrast between plots of the best and worst methods used here.

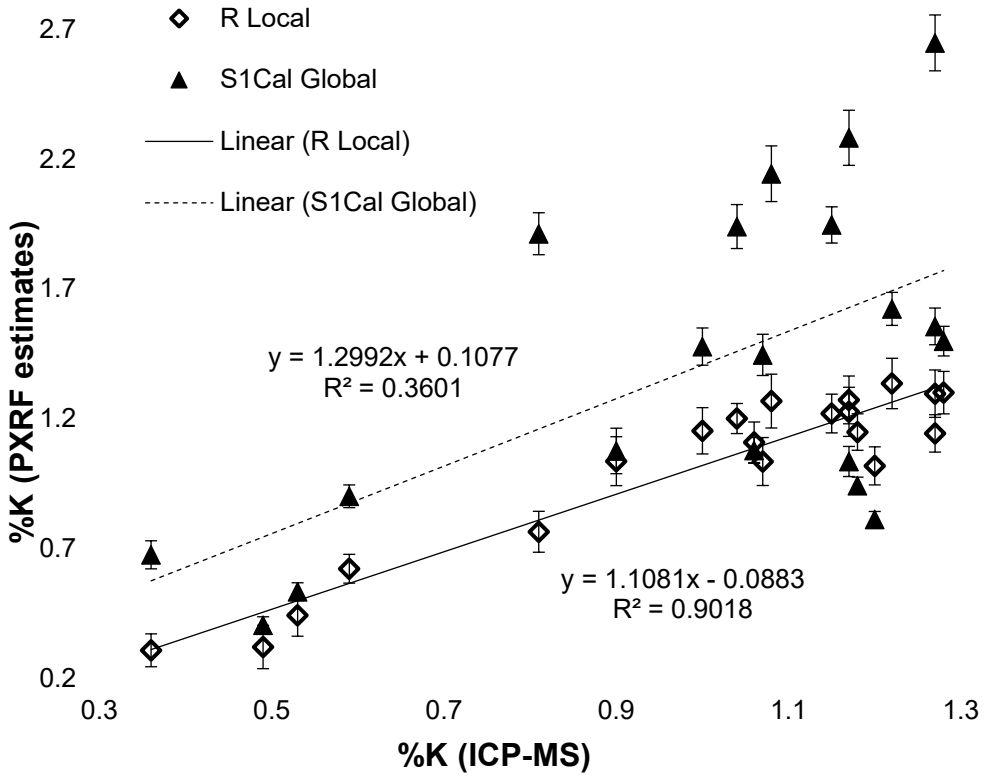

Figure 2-8. %K estimates using two PXRF calibrations compared with ICP-MS %K estimates. Error terms are 1σ.

The relatively poor performance of the S1CalProcess software could be due in part to the fact that it is probably not designed to work well with such a small number of standards. Typically, an analyst using this software would want to include more standards to build a robust regression. On the other hand, these data generally support the idea that local calibration is

better, as temporal proximity of calibration data to field sample data yields marginally better correspondence with ICP-MS data regardless of software used. For example, not only are R Local results better than R Single results and R Single results better than R Global results (particularly in the number of significant z scores), but S1Cal Local results have a marginally better $\overline{X}_D$ value than ICP-MS data than S1Cal Single results, which in turn have a marginally better $\overline{X}_D$ value than S1Cal Global results. Since frequent re-calibration is apparently necessary for the derivation of accurate K estimates, the incentive to use calibration software which needs a large number of standards to function is reduced, since the use of such software will require quite a bit of measurement time. Further, because S1CalProcess does not take error in standard concentrations into account, it is reasonable to question whether its inaccuracy is also in part due to embedded problems with its regression fitting and interpolation. For the present it is therefore probably best to avoid S1CalProcess in favor of local calibrations which make use of a small number of standards while effectively modeling errors throughout the process. If this is done, however, demonstrably accurate results can be achieved, at least given the experimental parameters used here.

Repeatability of Elemental Estimates

As with calibration data, examination of the consistency with which elemental estimates could be derived provided important information about instrument stability, as well as the repeatability and the reliability of results. Evaluation of this consistency made use of statistical comparison of elemental concentrations from one trial to the next. This comparison was carried out two ways. First, paired sample t-testing (n = 63) was carried out as above to compare the first

trial with the second trial for all samples. Second, an independent two-sample t-test (two-tailed, n = 10 for each sample to reflect the number of replicate measures per sample per trial) was used to compare estimates of elemental concentration on a sample-by-sample basis. For this test, an f-test was first conducted to assess compatibility of variances; Student's t-test (18 df) was used in cases of compatible variances, Welch's t-test (variable df) in cases of incompatible variances. These analyses were performed (at 95% confidence) for all samples for both K and Ca results; Fe results needed to be based on quadratic calibration to constitute a valid test of accuracy here and were therefore omitted.

Paired-sample t-tests of K results show that K estimates are significantly different as a whole from the first trial to the second (p = 0.004). Individual sample estimates are statistically compatible at a rate of about 70% (44 out of 63) from the first trial to the next, however, showing that statistical redundancy is achieved for the majority of individual samples after only two trials. S-drift was again the likely factor when results were not repeated between trials; each of the 21 trials contained samples and/or standards which showed a degree of bimodality in counts distributions relevant to K, indicating s-drift was pervasive throughout experimentation. Of the 19 samples which statistically differed between trials 1 and 2, however, 15 (79%) produced a Trial 3 estimate which was statistically compatible with estimates from either Trials 1 or 2. Thus after three trials, statistical redundancy was achieved in K estimates for 59 out of 63 field samples (94%), indicating at least three trials are needed to generate a large enough sample of measurements to average out s-drift and achieve repeatable concentrations of K using these experimental conditions.

Paired-sample t-tests of Ca results show that Trials 1 and 2 are statistically compatible (p = 0.06), although just barely. Individual sample estimates are statistically compatible at a rate of about 90% (57 out of 63) from Trial 1 to Trial 2. Here, s-drift was just as pervasive, but its resultant error was not as disruptive to overall consistency in results as it was for K, possibly as a result of the fact that larger average errors in Ca estimates (+/- 0.14 versus +/- 0.08 for K estimates) drove greater overall statistical compatibility. Of the six samples for which Trials 1 and 2 produced incompatible results, all but one produced a Trial 3 measurement which was statistically compatible with one of the previous trials, meaning statistical redundancy was achieved for over 98% of the Ca estimates after three trials. Again, three trials seems like the magic number in deriving repeatable estimates in spite of s-drift.

CONCLUSIONS

The protocol proposed here addressed significant issues in calibration and instrument performance to successfully produce PXRF estimates of K and Ca which are consistent over repeated measurements. In the case of K measurement, this protocol also produced PXRF estimates of K which agree with independent estimates derived from established laboratory methods, despite known difficulties in measuring elements of low atomic number. In fact, these estimates agree on 85% of the samples independently measured for K, which is a high rate considering the fact that even well-established methods such as ICP-MS often produce estimates which diverge from one another (Liritzis and Zacharias 2010), especially for K and Ca (Murphy et al. 2002). This indicates that the accuracy of K estimates from PXRF can be roughly on par with more established methods provided measured samples exceed the relatively high minimum

concentrations detectable by PXRF. Further, accurate results were accompanied by error terms which fully reflect true errors in PXRF analysis while achieving precision roughly comparable to average reported errors in standards themselves, indicating PXRF using this protocol is a viable alternative for measurement of K and Ca. Measurement of Fe was less successful, but this was largely an issue in calibration, and the inclusion of additional standards with higher known Fe concentrations seems likely to fix the problem.

By contrast, deviations from this protocol failed to produce accurate results. Experimentation showed that inaccuracies in this case were likely due to a number of factors, including: 1) both random and systematic changes to instrument sensitivity over multiple timescales, 2) calibration software used, and 3) calibration data used. Portions of the protocol essential to success therefore appear to be 1) repeated measurements of samples and standards within trials, 2) local calibration, 3) empirical simulation of true errors in calibration, and 4) redundant trials.

Unfortunately, at a practical level this protocol is quite rigorous, and it strips away much of the appeal of PXRF for users interested in results which are both rapid and non-destructive. Further experimentation will reveal to what degree the experimental controls used here can be relaxed without compromising results, as well as the degree to which results here are applicable to other materials, elements, instruments, settings, etc. For the present, however, it seems that the accuracy of PXRF is limited primarily by the time investment required for re-sampling, even in cases where samples are homogenous. The method may therefore be a valid alternative for bulk ^{40}K dosimetry in a devoted laboratory setting, provided access, cost, and productivity are

commensurate with larger laboratory resources and needs. In cases where heterogeneous samples are measured, however, more re-measurement will be necessary to account for additional scatter within observed counts due to micro-compositional differences between sampled locations. Given this, in situ applications which require accurate estimates of K, Ca, and possibly other elements will probably not be served well by PXRF in many cases. In situ measurements of heterogeneous materials may still have some use, however, provided users are interested in ordinal-scale discrimination between samples.

ACKNOWLEDGEMENTS

This paper was originally published in the *Journal of Archaeological Method and Theory*. The final publication is available at https://link.springer.com/article/10.1007/s10816-012-9162-3. Significant assistance in statistics, calibration, and R software was provided by UW statisticians Soyoung Ryu, Paul Sampson, and Jonathan Gruhl. Sediment samples used for this study were collected during fieldwork supported by National Science Foundation DDIG No. 0731529. PXRF units and equipment used for this study were purchased with funds awarded by UW STF Award No. 2008-068-1. Access to field sites in Perú was graciously given by Dr. Santiago Uceda (Universidad Nacional de Trujillo) and by Mr. Francisco Burga (Agroindustrias San Simon S.A.). Ceramic samples and ICP-MS data were made available by Dr. James Feathers (UW). Jim also deserves thanks along with Dr. Donald Grayson (UW) for providing insightful comments on drafts. Rhiannon Held provided assistance with copy editing.

REFERENCES

Abrahams, P.W., J.A. Entwistle, and R.A. Dodgshon
2010 The Ben Lawers Historic Landscape Project: Simultaneous Multi-element Analysis of Former Settlement and Arable Soils by X-ray Fluorescence Spectrometry. *Journal of Archaeological Method and Theory* 17:231–248.

Aitken, M.J.
1985 *Thermoluminescence Dating*. Academic Press, London.

Buhrke, V.E., R. Jenkins, and D.K. Smith, eds.
1998 *A Practical Guide for the Preparation of Specimens for X-ray Fluorescence and X-ray Diffraction Analysis*. John Wiley & Sons, New York.

Burley, D.V., P.J. Sheppard, and M. Simonin
2011 Tongan and Samoan volcanic glass: pXRF analysis and implications for constructs of ancestral Polynesian society. *Journal of Archaeological Science* 38(10):2625–2632.

Craig, N., R.J. Speakman, R.S. Popelka-Filcoff, M.D. Glascock, J.D. Robertson, S.M. Shackley, and M.S. Aldenderfer
2007 Comparison of XRF and PXRF for analysis of archaeological obsidian from southern Perú. *Journal of Archaeological Science* 34(12):2012–2024.

Davis, L.G.
2011 Return to Cooper's Ferry Site: Studying Cultural Chronology, Geoecology, and Foragers in Context. *Idaho Archaeologist* 34:1–4.

Davis, L.G., S.J. Macfarlan, and C.N. Henrickson
2012 A PXRF-based chemostratigraphy and provenience system for the Cooper's Ferry site, Idaho. *Journal of Archaeological Science* 39(3):663–671.

Feathers, J.K.
2009 *Luminescence Dating of Ceramics from Three Sites in Cumberland County, North Carolina, Including Interlaboratory Comparisons*. Prepared for New South Associates. University of Washington Luminescence Dating Laboratory.

Forster, N., and P. Grave
2012 Non-destructive PXRF analysis of museum-curated obsidian from the Near East. *Journal of Archaeological Science* 39(3):728–736.

Frankel, D., and J.M. Webb
2012 Pottery production and distribution in prehistoric Bronze Age Cyprus. An application of pXRF analysis. *Journal of Archaeological Science* 39(5):1380–1387.

Goodale, N., D.G. Bailey, G.T. Jones, C. Prescott, E. Scholz, N. Stagliano, and C. Lewis
2012 pXRF: a study of inter-instrument performance. *Journal of Archaeological Science* 39(4):875–883.

Goren, Y., H. Mommsen, and J. Klinger
2011 Non-destructive provenance study of cuneiform tablets using portable X-ray fluorescence (pXRF). *Journal of Archaeological Science* 38(3):684–696.

Lachance, G.R., and F. Claisse
1995 *Quantitative X-ray Fluorescence Analysis*. John Wiley & Sons, Chichester, UK.

Liritzis, I., and N. Zacharias
2010 Portable XRF of Archaeological Artifacts: Current Research, Potentials and Limitations. In *X-Ray Fluorescence Spectrometry (XRF) in Geoarchaeology* edited by M. S. Shackley, pp. 109–142. Springer, New York.

Millhauser, J.K., Rodriguez-Alegria, E., and Glascock, M.D.
2011 Testing the accuracy of portable X-ray fluorescence to study Aztec and Colonial obsidian supply at Xaltocan, Mexico. *Journal of Archaeological Science* 38(11):3141–3152.

Murphy, K.E., S.E. Long, M.S. Rearick, and O.S. Ertas,
2002 The accurate determination of potassium and calcium using isotope dilution inductively coupled "cold" plasma mass spectrometry. *Journal of Analytical Atomic Spectrometry* 17(5):469–477.

Nazaroff, A.J., K.M. Prufer, and B.L. Drake
2009 Assessing the applicability of portable X-ray fluorescence spectrometry for obsidian provenance research in the Maya lowlands. *Journal of Archaeological Science* 37(4):885–895.

Phillips, S.C., and R.J. Speakman
2009 Initial source evaluation of archaeological obsidian from the Kuril Islands of the Russian Far East using portable XRF. *Journal of Archaeological Science* 36(6):1256–1263.

Polikreti, K., J.M.A. Murphy, V. Kantarelou, and A. Germanos Karydas
2011 XRF analysis of glass beads from the Mycenaean palace of Nestor at Pylos, Peloponnesus, Greece: new insight into the LBA glass trade. *Journal of Archaeological Science* 38(11):2889–2896.

Potts, P.J., and M. West
2008 *Portable X-ray Fluorescence Spectrometry: Capabilities for In Situ Analysis*. The Royal Society of Chemistry, Cambridge, UK.

Schuenemeyer, J.H., and L.J. Drew
2011 *Statistics for Earth and Environmental Scientists*. John Wiley & Sons, Hoboken, NJ.

Seber, G.A.F., and A.J. Lee
2003 *Linear Regression Analysis*. John Wiley & Sons, Hoboken, NJ.

Shackley, M.S.
2010 Is there reliability and validity in portable X-ray fluorescence spectrometry (PXRF)? *The SAA Archaeological Record* 10(5):17–20.

Sheppard, P.J., G.J. Irwin, S.C. Lin, and C.P. McCaffrey
2011 Characterization of New Zealand obsidian using PXRF. *Journal of Archaeological Science* 38(1):45–56.

Speakman, R.J., N.C. Little, D. Creel, M.R. Miller, and J.G. Iñañez
2011 Sourcing ceramics with portable XRF spectrometers? A comparison with INAA using Mimbres pottery from the American Southwest. *Journal of Archaeological Science* 38(12):3483–3496.

Chapter 3. Experimental Use of Portable X-Ray Fluorescence for the Relative Dating of Petroglyphs at Hole-in-the-Ground, Malheur County, Oregon

ABSTRACT

Portable X-ray fluorescence (PXRF) technology offers a potential means of measuring elemental deposits resulting from ongoing weathering of stone surfaces. By extension, PXRF has the potential to enable measurement and comparison of the relative degree of rock varnish accumulation atop petroglyphs carved into rock surfaces, and therefore the relative ages of petroglyphs themselves. Research detailed here attempted to apply PXRF to the assessment of relative ages of petroglyphs at the Hole-in-the-Ground site in Malheur County, Oregon, on behalf of the Bureau of Land Management. After a brief discussion of previous relevant archaeological research, methodological considerations for this application of PXRF are discussed in detail. Data are then presented and used to evaluate the viability of the application of this technique at Hole-in-the-Ground. Results show that manganese (Mn) accumulations on weathered rock surfaces provide a sufficient mechanism for assessing relative petroglyph age in this context, allowing for novel insights about the history of petroglyph carving at Hole-in-the-Ground. Suggestions for future work are then briefly discussed.

INTRODUCTION

Background

The Vale District Office of the Bureau of Land Management (BLM) was interested in applying portable X-ray fluorescence (PXRF) to the measurement of rock varnish accumulation

on the numerous petroglyphs at the Hole-in-the-Ground area (sites 35ML168 and 35ML169) along the Owyhee River in Malheur County, Oregon. The primary goal of this application was to use non-destructive methods to obtain a relative chronology for a sizeable subset of petroglyphs at these sites. Through this effort, the BLM was also interested in furthering interpretation of the area's occupation and use history, assessing and documenting ongoing preservation issues at these sites, and advancing recommendations for further research and preservation efforts at these sites.

The Vale District Office of the BLM retained the services of SWCA Environmental Consultants (SWCA) to conduct this work. SWCA, in turn, retained the services of the author who used analytical equipment on loan from the University of Washington Department of Anthropology to conduct fieldwork and analysis. All parties had the understanding that the application of PXRF to the measurement of petroglyphs remains highly experimental, and that successful application of PXRF is highly dependent on a number of geologic, environmental, and logistical conditions which could not be fully anticipated prior to actual fieldwork. These caveats were acceptable to all parties, and the project proceeded in November 2014.

This report gives a detailed description of all work executed for this project. It begins with a brief description of the natural and cultural setting in the project area. This is followed by a detailed description of the methods used to apply PXRF to the Hole-in-the-Ground petroglyphs, then by a detailed discussion of the results of this analysis. This report then concludes with recommendations for future work at Hole-in-the-Ground.

Natural and Cultural Setting

The broad natural and cultural setting of the project area has been recently, extensively, and cogently discussed elsewhere (Andrefsky 2010; Fisher 2010; Noll 2009; Sudman 2005). These references should be considered the definitive resources on the subject, and therefore extensive review and discussion is not presented here. Instead, a brief review of salient background particularly relevant to the execution of PXRF analysis at Hole-in-the-Ground will be the focus of this discussion.

The Hole-in-the-Ground petroglyph site (35ML168 for the west portion, 35ML169 for the east portion) sits on the northern bank (river left) of the Owyhee River near the confluence between the river and Juniper Canyon (Figure 3-1). The site is located between the river and a large basaltic boulder field which sits atop a topographic bench well above the flowing waters of the modern river (Figure 3-2). This boulder field was probably exposed and/or deposited during a single fluvial scouring event caused by a massive outburst flood originating from pluvial Lake Alvord between 13,000 and 14,000 radiocarbon years before present (BP) (see Carter et al. 2006); boulders have been subjected to subaerial weathering since this time.

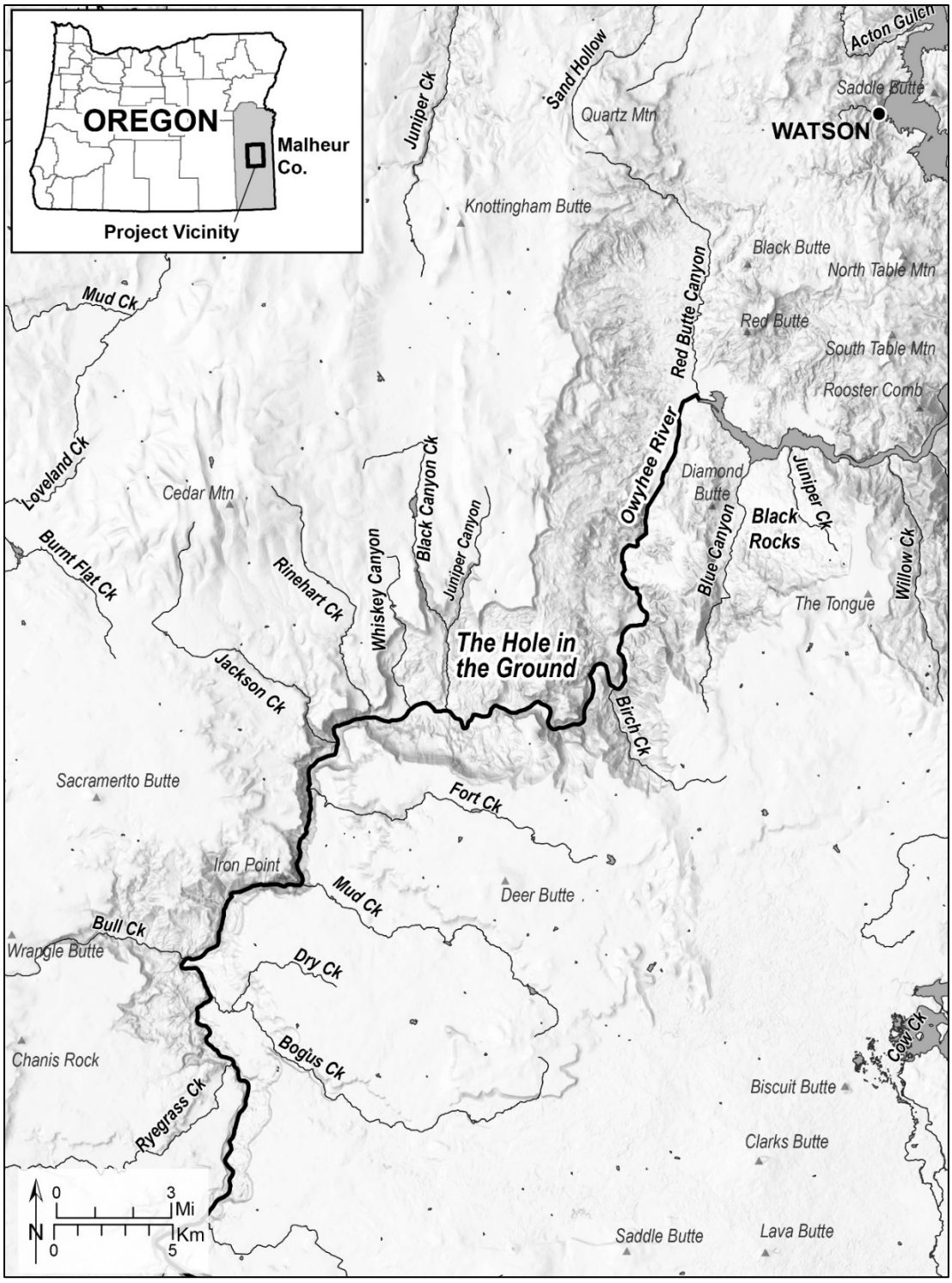

Figure 3-1. Overview map showing the location of Hole-in-the-Ground.

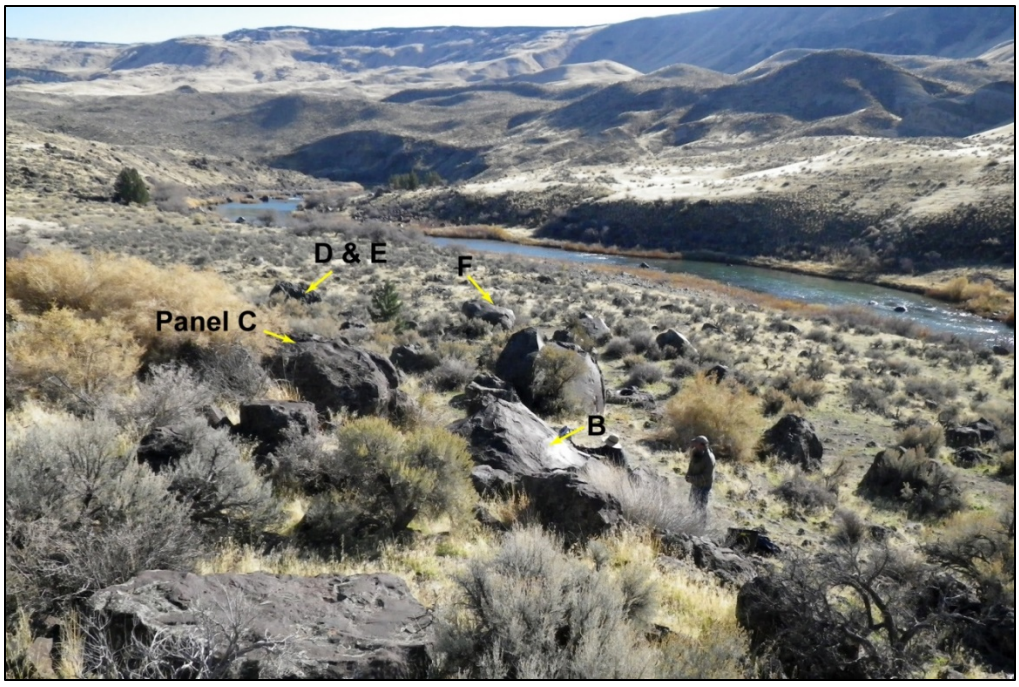

Figure 3-2. Overview of boulders at 35ML169, view south across the Owyhee River.

There are at least 55 boulders with petroglyphs documented at the site: one at 35ML168 and 54 at 35ML169 (Sudman 1993). The petroglyphs were typically pecked or incised from rock faces to remove the very dark, almost black manganese-rich patina on the weathered surfaces of the boulders. These petroglyphs number in the hundreds and consist of an array of motifs, including rows of dots, abstract curvilinear and geometric shapes, and representational forms such as human figures and canoes. Detailed systematic stylistic comparisons between these motifs and the other known petroglyph sites in the region (see Sudman 2005) have not been conducted. Many of the Hole-in-the-Ground glyph-bearing boulders are actively exfoliating, though no glyph-bearing spalls are to be found at the site, suggesting many have been carried away by site visitors. Some of the boulders at or near the site exhibit apparent recent vandalism,

including a boulder carved with the year 2006 (Figure 3-3), and another carved with the year 1957 (Figure 3-4).

In addition to the glyph-bearing basaltic boulders, other archaeological materials are visible on the surface at the site, including lithic debitage and retouched tools made from a range of materials including rhyolite, chert, and obsidian, as well as a number of mortars and pestles, including bedrock mortars with pestles remaining in situ. Sediments are generally thinly deposited at the site, although lower-lying areas near 35ML168 and east of 35ML169 may contain some stratified deposits. The site itself was likely used for a range of purposes, including tool-making, food preparation, and perhaps as a temporary camp.

Figure 3-3. Boulder vandalism dated to 2006.

Figure 3-4. Boulder vandalism dated to 1957.

Previous Work at Hole-in-the-Ground

Initial documentation of the Hole-in-the-Ground site and petroglyphs was conducted as early as the 1930s by Luther Cressman (Sudman 1993). The site was visited again by Reg Pullen in 1976, at which time the original site form was created (Sudman 1993). Numerous BLM employees and archaeologists visited the site in the 1980s (Sudman 1993), including James Huntley, who conducted a 1986 survey of known archaeological sites on the lower Owyhee (Sudman 2005). In 1991, Natalie Sudman visited the site and updated the site form, at which time she noted that some of the petroglyphs evident in 1976 had since been obliterated by vandalism (Sudman 1993). Sudman and colleagues again visited the site in 1992 and 1993 to conduct comprehensive cataloging and documentation of glyphs at the site. Although these

efforts were apparently abandoned prematurely due to an unexpected budgetary shortage, they produced a detailed inventory of glyph-bearing boulders, including several sketches and a management plan for the site itself (Sudman 1993). These materials comprise the majority of the current reference materials available for the site. Between 1994 and 2005, periodic visits were made to Hole-in-the-Ground as part of an ongoing survey and monitoring of the known archaeological sites along the Owyhee River and between Rome, Oregon, and the Birch Creek Site (Sudman 2005), although no updates to the site form were made during this time. Little if any documented work has been conducted at Hole-in-the-Ground between 2005 and the time of writing, although BLM river ranger Dan Thomas informed SWCA that BLM staff periodically visit and monitor the site to assess the potential impacts of visits from river rafters, especially during the spring rafting season (D. Thomas, pers. comm.).

Previous Work at Nearby Sites

Sudman (2005:10) notes that around 200 sites have been recorded for the Lower Section of the Owyhee, and around 140 more have been described but not documented. Discussion of these sites is beyond the scope of this study. Instead, a brief summary of nearby sites will be confined to those sites which 1) were inventoried in 2005 by Sudman or extensively examined by William Andrefsky and his students at Washington State University (e.g., Andrefsky 2008; Noll 2009) and 2) lie along the Owyhee River within a linear distance of roughly 5 km from Hole-in-the-Ground itself.

These sites represent a mix of glyph sites (35ML168, 35ML169, 35ML1168), small lithic scatters (MO-136 and MO-97), and occupation sites (35ML181, and perhaps 35ML168 and

35ML169). Work at these sites has been primarily confined to occasional monitoring and informal surface survey, with the exception of the Birch Creek Site (35ML181). Here, extensive research and excavation has been conducted in conjunction with a series of field schools held by Washington State University under the direction of Dr. Andrefsky; this research has produced a number of theses and published papers examining various aspects of the site's material record (e.g., Andrefsky 2008, 2010; Fisher 2010; Noll 2009).

As a result, much of the prehistory of the immediate vicinity of Hole-in-the-Ground is reasonably well documented, especially the lithic assemblages recovered from the Birch Creek Site. While the region has probably been occupied for as long as 14,000 years (Aikens et al. 2011), current direct archaeological evidence for local occupation only stretches back around 5,300 years at the Birch Creek Site (Fisher 2010). At this site, occupation is argued to be representative of groups of highly mobile people who may have been making seasonal use of resources in the area, and this is thought to have been the general use-pattern of both the site and the Lower Section of the Owyhee for millennia (Andrefsky 2010; Fisher 2010). In this light, Hole-in-the-Ground and its neighboring sites likely represent palimpsests of relatively short-term use episodes which led to long-term accumulations of both material culture and petroglyphs.

METHODS

Methodological Background: Portable X-ray Fluorescence

X-ray Fluorescence (XRF) measurement is a method of elemental analysis that makes use of a sample's characteristic secondary X-ray emissions as a means of estimating its elemental composition. Broadly, the process of measurement proceeds through the following steps: 1) the

sample is exposed to a controlled spectrum of high-intensity radiation, 2) this radiation stimulates the emission of secondary X-rays characteristic of the elements comprising the sample, 3) these secondary X-ray emissions are detected by specialized instrumentation, and 4) the presence and intensity of the detected characteristic X-rays is converted into estimates of the sample's elemental composition. Often, results are expressed in percentage composition (for common elements) or in parts per million (for trace elements), although such expressions require careful calibration of the instrument using standards of known elemental composition, and calibration is itself a complex process (see Lachance and Claisse 1995 for a detailed discussion). When calibration is absent, sample elemental concentrations can be expressed in terms of counts or peak intensities detected by the instrument; this approach allows analysis sufficient for many applications to proceed, but can make results more difficult to compare with independent estimates derived from other techniques. XRF has long been used for compositional studies of a range of archaeological materials, including lithics, ceramics, metals, and sediments (see Shackley 2011 for a review), and is favored as a low-cost and non-destructive method of elemental analysis.

Portable X-ray fluorescence (PXRF) measurement is a subtype of XRF that makes use of a smaller, more portable radiation source and detector, allowing for in situ measurements of archaeological materials. Over the last decade PXRF has become popular amongst archaeologists for its logistical flexibility, low cost, and non-destructive nature. Despite its surge in popularity, the application of PXRF remains highly experimental, as a number of published studies have shown that PXRF is generally less precise than laboratory-based XRF and that estimates of

absolute elemental concentrations derived from PXRF instruments typically differ significantly from those derived from established independent methods (Craig et al. 2007; Goodale et al. 2012; Nazaroff et al. 2009), especially when used to measure elements of relatively low atomic number such as calcium (Ca), potassium (K), manganese (Mn), and iron (Fe) (Goodale et al. 2012). Exceptions to this situation have been documented (Johnson 2014a), but only in cases where protocols used in preparation, measurement, calibration, and analysis were extremely strict. These protocols unfortunately cannot be applied non-destructively to samples measured in situ. To date, then, no published study has shown that PXRF measurements conducted in situ produce absolute elemental concentrations that match those derived independently from established laboratory methods. Given this, *absolute* elemental concentrations derived from in situ PXRF measurements should not be considered accurate until proven otherwise, nor should they be used as a basis for analysis.

On the other hand, *relative* concentrations derived from PXRF remain generally viable for most elements of interest, even when ideal measurement conditions are relaxed to accommodate in situ applications. Thus if given two hypothetical samples—A and B—of differing elemental concentrations measured in situ, PXRF can be expected to reliably and accurately discern the following:

> Sample A has roughly twice as much Mn (or Fe, Ca, K, etc.) as B.

But not the following:

> Sample A has exactly 3.57% (+/- 0.09%) Mn (or Fe, Ca, etc.).

Given this, the application of PXRF to in situ samples is most reliable when using analysis that relies on a comparison of characteristic X-ray peak intensities only. Such a conservative approach may someday be obviated by research demonstrating that absolute elemental concentrations derived from in situ PXRF measurements match independent data, but for the present this is not the case.

The Application of PXRF to Petroglyph Weathering

Because of its portability, non-destructive nature, and its ability to yield data on elemental composition, PXRF is an ideal method for examining the progressive weathering of petroglyphs, which should be reflected by similarly progressive changes in the elemental composition of weathered surfaces.

Previous researchers (Lytle et al. 2008; McNeil and Cecil n.d.) have explored the potential of PXRF for this application, although these researchers have focused their efforts primarily on the use of PXRF to derive absolute (calendrical) dates for petroglyphs. This is problematic, as no one has yet demonstrated that such absolute dating of petroglyphs is reliable, despite claims by Lytle and colleagues. In fact, McNeil and Cecil correctly critique the claims of Lytle and colleagues by demonstrating in part why these claims—which include the proposition by Lytle and colleagues that a "universal" conversion from Mn concentration to petroglyph age has been established and is linear in nature—are simplistic and not supported by experimental data. To this critique should be added two points, each of which further undermines Lytle and colleagues' suggestion that one can simply measure Mn and apply a linear conversion to determine the age of a glyph. These are:

- Nowhere in the peer-reviewed literature of any discipline is the idea of a universal conversion (linear or otherwise) from Mn to glyph age set forth or supported. Abundant published research indicates weathering and varnish accumulation rates differ between locations and substrates (e.g., Colman 1981; Hunt 2015; Liu and Broecker 2000).
- Rock weathering rates—and their chemical signatures—vary widely between environmental contexts, making a universal conversion from weathering-related Mn to absolute age impossible on its face. Rock weathering is instead a complex suite of processes that are highly dependent on an array of factors (see Colman 1981) which can vary at small spatial scales as a result of a variety of micro-environmental factors such as slope, aspect, rock face micro-topography, and microorganisms, in addition to variation over time as a result of climate change.

Despite these objections, however, it is generally true that rock weathering in the form of desert varnish generally proceeds (in part) through the progressive accumulation of elements such as Mn on exposed rock surfaces. These accumulations form as a result of the interaction of biota such as fungus and/or oxidizing bacteria with eolian dust to gradually create sedimentary accumulations on rock surfaces (Cremaschi 1996; Hungate et al. 1987; Nagy et al. 1991). The composition of such accumulations can be measured in situ with XRF, provided the following assumptions hold:

- **Assumption 1:** Chemical change induced by varnish sediment is sufficiently pronounced to be detectable by the instrument. This assumption can be tested in the field by

comparing measurements performed on unweathered surfaces with measurements on varnished surfaces to determine whether a detectable gradient in elemental composition exists.

- **Assumption 2:** Baseline geology of the petroglyph rocks is sufficiently uniform. If glyphs are on rocks which are compositionally heterogeneous, elemental differences within the unweathered rock may obscure compositional changes due to the accumulation of varnish alone. This assumption can be tested in the field by measuring a number of unweathered rock surfaces from across the site to determine the degree to which they are compositionally homogenous.
- **Assumption 3:** Rock weathering in the form of varnish accumulation is sufficiently uniform across the surfaces of measured rocks. If this assumption does not hold, differences in glyph varnish accumulation cannot be directly attributable solely to differences in age. Careful sample selection to isolate glyphs on rock faces of similar aspect and slope, both of which measurably affect varnish accumulation (Lytle et al. 2008; McNeil and Cecil n.d.), can help ensure this assumption is met. Further, this assumption can be tested in the field by measuring a number of "fully weathered" rock surfaces—portions of boulder faces where no patina has been removed by the etching of glyphs—to determine the degree to which they are compositionally homogenous. If heterogeneity in the composition of weathered surfaces is evident, this heterogeneity should be quantified and attached to results as error terms, since the differential

weathering of rocks obfuscates the measurement and comparison of weathered glyphs adorning them.

- **Assumption 4:** Petroglyph pecking fully removed the rock varnish in antiquity. If this assumption is not met, residual varnish persisting through the glyph-making process may make petroglyphs appear more weathered—and therefore older—than they actually are. Further, if some glyphs were pecked more thoroughly than others, an artificial gradient in age may emerge due to differential pecking alone. As above, careful sampling by selecting for measurement only glyphs exhibiting thorough pecking (wide and deep glyphs) may help ensure this assumption is met. Further, performing multiple measurements on a single glyph, then making use of the measurement indicative of the lowest degree of varnish accumulation should help ensure the most well-pecked portion of a given glyph is used for analysis and interpretation. Finally, testing well-pecked recent glyphs to determine how closely these glyphs match the composition of unweathered rock surfaces can help evaluate the strength of this assumption.

Overview of Project Fieldwork at Hole-in-the-Ground

Field application of PXRF to the Hole-in-the-Ground petroglyphs was designed and conducted with the preceding discussion in mind, and is discussed in detail in subsequent sections of this report. Fieldwork was conducted by J. Johnson and C. Undem from SWCA's Seattle Office on November 11 to 12, 2014. River ranger D. Thomas of the BLM's Vale office was present to provide guidance to and from the site location, an orientation to site layout, and

camping advice and transportation across the Owyhee River between the project campsite and the study area itself.

Fieldwork began as scheduled at daybreak on November 11 with a brief reconnaissance of site petroglyphs, followed by intensive measurement and documentation of petroglyphs for the remainder of daylight hours, as well as the entire duration of daylight on November 12.

Conditions in the field were cold and clear, with daytime temperatures in the mid-30s Fahrenheit and nighttime lows much colder. A portable generator was used to power a space heater to maintain high enough temperatures to protect the sensitive electronics of the PXRF from extreme nighttime cold, and also to supply a means of charging instrumental batteries in the field. The looming onset of Winter Storm Bozeman on November 12 prompted fieldwork to be cut one day short to avoid being stranded by snow accumulation on the steep canyon roads providing access to the site; this change resulted in an exit from the site at dusk on November 12 and return to Seattle on November 13. As a result, measurement time in the field was about one-third shorter than anticipated.

Sampling Strategy

Given restrictions on time available for fieldwork and the large number of boulders with petroglyphs at Hole-in-the-Ground, comprehensive PXRF measurement at the site was not feasible. Instead, it was necessary to limit analysis to a subset of the site's boulders and glyphs.

Ideally, our sampling would have made use of earlier documentation efforts detailed on the site form, but prior documentation of the Hole-in-the-Ground glyphs was simply not useful in

the field, especially given the short window for fieldwork in this project. Issues with these documents which prevented their use include, but are not limited to:

- The absence of a map showing the spatial relationships between numbered rocks or rock faces.
- Recorded dimensions of glyph-bearing rocks which are incomplete, ambiguous, or uncertain. For example, available documents contain multiple question marks attached to the listed dimensions of rocks, making use of these measurements cumbersome at best. Further, the notes on these documents openly question the validity of the north-south designations of these dimensions, as well as the base numbering system used to identify the rocks themselves, again limiting their utility in identifying stones in the field.
- The absence of field-portable photographs that comprehensively display labeled boulders, sets of glyphs, or individual glyphs which would allow the recorded numbering system to be visually matched with boulders or glyphs in the field.
- Boulder sketches which are difficult to relate to actual boulders at best and which inconsistently label boulders at worst. Many boulders, for example, are given different numbers in different photos and sketches.

Given these issues, sampling for this project proceeded independent from prior systems of numbering/cataloguing the Hole-in-the-Ground boulders or glyphs as a matter of necessity. This was accomplished by first selecting “panels” of glyphs to include in this study. A panel as defined here is a set of glyphs occurring together on a portion of a boulder with relatively

constant slope and aspect. This is roughly equivalent in meaning to the term “face” used in the original site report, but we have selected the new term “panel” to avoid confusion with previous nomenclature.

Panels were selected for a variety of reasons. First, panels of slope between roughly 30 and 45 degrees and a generally southerly aspect were chosen to help ensure weathering, including varnish accumulation, would be comparable across multiple panels. Second, panels exhibiting multiple glyphs with a range of degrees of visual patination were generally preferred in order to facilitate within-panel comparison of the broadest possible array of probable glyph ages. Third, some panels were chosen for special attributes such as obvious signs of panel endangerment due to active rock exfoliation (Panel A) and/or vandalism (Panels D and F), including vandalism of known age, since documentation of panels of this type was important to both analytical and overall project goals.

In total, six panels were examined using PXRF. Basic panel details are given in Table 3-1; where possible we have included our best guess as to how our panel designations match the boulder or face numbers given by Sudman (1993). Each of the studied panels was extensively photographed and sketched to allow for easy field identification in the future, and the georeferenced location of each was recorded using a professional-grade GPS unit to allow detailed mapping of panel locations on the landscape (Figures 3-5 and 3-6).

Table 3-1. Panels Examined During Fieldwork

Panel Name	Site	Site Report Rock/Face No.	Qty. Glyphs Measured	UTM Zone 11T, NAD83		Notes
				Northing	Easting	
Panel A	35ML168	N/A	7	4783944.757	452318.423	Active Exfoliation
Panel B	35ML169	Unknown	3	4783712.882	452563.520	
Panel C	35ML169	Face 46	20	4783712.728	452568.780	
Panel D	35ML169	Face 6	1	4783692.101	452584.399	Modern Vandalism
Panel E	35ML169	Face 11	10	4783690.856	452586.891	
Panel F	35ML169	Face 1	6	4783683.370	452578.634	Modern Vandalism

Within panels, glyphs were selected for measurement based on two main criteria. First, glyphs which appeared both sizeable and thoroughly pecked were chosen for analysis, as these glyphs were most likely to yield measurements which were free of residual patination prior to glyph creation. In practice, this precaution excluded many glyphs from analysis, including two glyphs representing recent vandalism of known age—the 1957 glyph and the 2006 glyph—as these glyphs were not sufficiently pecked to remove the weathered patina over the roughly 4 × 4–mm swath observed by the instrument's detector. Second, glyphs for which pecking was too deep to allow the instrument to directly abut the stone were excluded, since direct contact between instrument detector and stone face is mandatory for accurate measurement. In practice, this precaution excluded few glyphs from analysis. In sum, most glyphs from selected panels were included in the measurements, providing good overall coverage for panels selected for analysis. As with panels, all measured glyphs were carefully sketched and photographed to allow future in situ identification and additional analyses moving forward.

Figure 3-5. Aerial photograph showing location of Panel A at 35ML168.

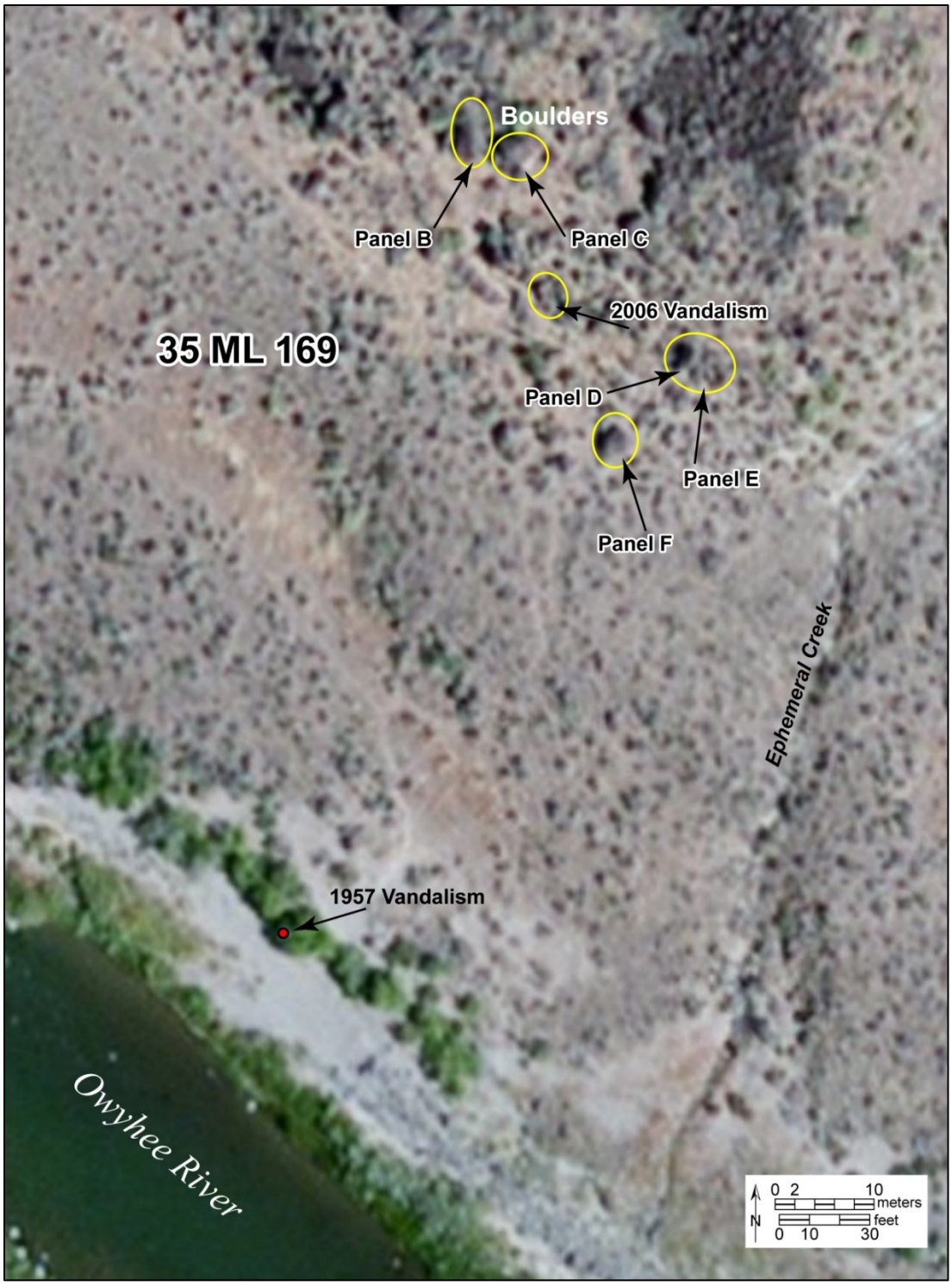

Figure 3-6. Aerial photograph showing locations of Panel B to F and vandalism at 35ML169.

Measurement

Measurement of selected panels and glyphs was conducted using a Bruker Tracer III-V PXRF instrument with a 1-mil titanium (Ti) filter and the PDA interface running BrukerS1 software in Analyze mode. Prior to measurement, the instrument's "library" was updated with a broad array of basalt, obsidian, and clay reference standards used in previous PRXF analyses conducted with the instrument (see Gjesfjeld 2014; Johnson 2014a; Phillips and Speakman 2009) to provide instrument software with a reference database of materials roughly similar in composition to the expected makeup of the basaltic Hole-in-the-Ground petroglyph boulders. Full calibration of the instrument was not performed due to the fact that effective calibration protocols (see Johnson 2014a) were impossible given field circumstances and time constraints. The absence of a calibration for these data was not a limiting factor in interpretation, however, since it in no way hampers the ability of this study to quantitatively compare Hole-in-the-Ground glyph weathering using X-ray peak intensities.

At least three 2-minute assays were conducted for each measured glyph. Mn and Fe values resulting from each assay were manually recorded in a measurement log along with error terms associated with these values. Instrument output represents these values as percentages by element and they are therefore recorded as such, but it is again important to remember that—given the accuracy issues of PXRF discussed above—these values are better thought of as relative measurements akin to "scores" rather than as absolute concentrations. Glyphs which showed signs of re-pecking in the form of lighter-colored areas were measured more than three times to generate data for both the "older" and "younger" parts of the pecked glyph.

Multiple reference measurements of fully weathered rock surfaces were also taken for each panel, including each portion of panels which have been extensively affected by exfoliation, resulting in baseline rock surfaces exhibiting various degrees of weathering within a given panel (e.g., Panel A). Additionally, reference measurements of newly broken rock faces were collected from a number of basaltic spalls found near glyph boulders. Again, these reference measurements were necessary to help evaluate assumptions of homogenous rock weathering and homogenous baseline geology.

RESULTS

Raw Results

Raw results are given by Table 3-2, and represent relative concentrations of Mn and Fe with error terms reported by the instrument. These raw results are useful as a basis for evaluating whether the assumptions discussed above were met, and also as a basis for derivation of the "scaled weathering scores" for each glyph detailed below.

Table 3-2. Raw Results of PXRF Measurement of Glyphs and Reference Locations

Panel	Glyph or Reference?	Glyph/ Ref. No.*	Msmt. No.*	%Mn	+/-	%Fe	+/-	Unweathered (U) or Weathered (W) Reference?
A	Glyph	AG1	AG1.1	10.66	0.19	70.78	0.36	
A	Glyph	AG1	AG1.2	14.35	0.17	75.73	0.22	
A	Glyph	AG1	AG1.3	15.50	0.16	72.61	0.22	
A	Glyph	AG2	AG2.1	21.50	0.15	68.80	0.21	
A	Glyph	AG2	AG2.2	23.23	0.13	66.43	0.18	
A	Glyph	AG2	AG2.3	18.35	0.20	71.56	0.30	
A	Glyph	AG3	AG3.1	16.64	0.16	72.93	0.21	
A	Glyph	AG3	AG3.2	23.07	0.16	66.69	0.22	

Panel	Glyph or Reference?	Glyph/ Ref. No.*	Msmt. No.*	%Mn	+/-	%Fe	+/-	Unweathered (U) or Weathered (W) Reference?
A	Glyph	AG3	AG3.3	16.81	0.18	72.08	0.24	
A	Glyph	AG4	AG4.1	9.66	0.15	74.21	0.28	
A	Glyph	AG4	AG4.2	11.24	0.18	78.61	0.25	
A	Glyph	AG4	AG4.3	8.03	0.16	76.86	0.33	
A	Glyph	AG5	AG5.1	9.80	0.16	78.19	0.13	
A	Glyph	AG5	AG5.2	15.42	0.15	72.96	0.21	
A	Glyph	AG5	AG5.3	9.19	0.08	71.49	0.15	
A	Glyph	AG6	AG6.1	4.38	0.12	76.95	0.31	
A	Glyph	AG6	AG6.2	4.71	0.76	76.96	0.90	
A	Glyph	AG6	AG6.3	4.23	0.04	78.49	0.11	
A	Glyph	AG7	AG7.1	26.86	0.13	62.23	0.18	
A	Glyph	AG7	AG7.2	24.71	0.08	58.20	0.12	
A	Glyph	AG7	AG7.3	17.93	0.08	71.65	0.11	
A	Reference	AR1	AR1.1	23.29	0.13	66.60	0.18	W
A	Reference	AR2	AR2.1	19.79	0.15	70.66	0.20	W
A	Reference	AR3	AR3.1	20.98	0.08	59.62	0.12	W
A	Reference	AR4	AR4.1	17.69	0.14	73.00	0.18	W
A	Reference	AR5	AR5.1	19.12	0.16	70.84	0.21	W
A	Reference	AR6	AR6.1	22.54	0.17	66.46	0.24	W
A	Reference	AR7	AR7.1	17.14	0.17	73.24	0.23	W
A	Reference	AR8	AR8.1	21.15	0.19	68.66	0.26	W
A	Reference	AR9	AR9.1	19.82	0.22	69.77	0.30	W
A	Reference	AR10	AR10.1	2.33	0.13	82.38	0.44	U
A	Reference	AR11	AR11.1	2.35	0.09	81.20	0.45	U
A	Reference	AR12	AR12.1	2.32	0.11	81.11	0.36	U
B	Glyph	BG1	BG1.1	18.10	0.15	71.22	0.21	
B	Glyph	BG1	BG1.2	8.95	0.18	74.40	0.35	
B	Glyph	BG1	BG1.3	10.17	0.12	78.09	0.18	

Panel	Glyph or Reference?	Glyph/ Ref. No.*	Msmt. No.*	%Mn	+/-	%Fe	+/-	Unweathered (U) or Weathered (W) Reference?
B	Glyph	BG2	BG2.1	9.47	0.14	76.37	0.21	
B	Glyph	BG2	BG2.2	9.36	0.14	77.90	0.21	
B	Glyph	BG2	BG2.3	12.00	0.15	77.55	0.20	
B	Glyph	BG3	BG3.1	12.32	0.16	76.56	0.22	
B	Glyph	BG3	BG3.2	15.89	0.16	72.82	0.21	
B	Glyph	BG3	BG3.3	11.33	0.22	77.85	0.30	
B	Reference	BR1	BR1.1	22.13	0.14	68.66	0.19	W
B	Reference	BR2	BR2.1	24.54	0.16	64.92	0.23	W
B	Reference	BR3	BR3.1	17.61	0.13	72.69	0.18	W
B	Reference	BR4	BR4.1	15.44	0.14	75.44	0.18	W
B	Reference	BR5	BR5.1	18.35	0.17	72.22	0.23	W
B	Reference	BR6	BR6.1	16.99	0.21	72.91	0.28	W
B	Reference	BR7	BR7.1	2.38	0.11	80.13	0.38	U
B	Reference	BR8	BR8.1	2.57	0.10	79.44	0.32	U
B	Reference	BR9	BR9.1	2.56	0.10	81.44	0.33	U
C	Glyph	CG1	GC1.1	12.71	0.08	74.06	0.12	
C	Glyph	CG1	GC1.2	21.01	0.22	68.91	0.29	
C	Glyph	CG1	GC1.3	10.05	0.16	72.73	0.30	
C	Glyph	CG2	GC2.1	5.62	0.15	78.16	0.35	
C	Glyph	CG2	GC2.2	7.49	0.18	73.77	0.40	
C	Glyph	CG2	GC2.3	13.92	0.19	73.93	0.15	
C	Glyph	CG3	GC3.1	8.74	0.18	73.03	0.38	
C	Glyph	CG3	GC3.2	10.09	0.24	78.28	0.35	
C	Glyph	CG3	GC3.3	11.81	0.09	75.57	0.12	
C	Glyph	CG4	GC4.1	7.87	0.17	74.24	0.35	
C	Glyph	CG4	GC4.2	6.39	0.15	77.54	0.35	
C	Glyph	CG4	GC4.3	8.62	0.21	73.99	0.43	
C	Glyph	CG5	GC5.1	11.07	0.21	78.83	0.30	

Panel	Glyph or Reference?	Glyph/ Ref. No.*	Msmt. No.*	%Mn	+/-	%Fe	+/-	Unweathered (U) or Weathered (W) Reference?
C	Glyph	CG5	GC5.2	8.02	0.17	75.87	0.36	
C	Glyph	CG5	GC5.3	7.16	0.07	74.34	0.15	
C	Glyph	CG6	GC6.1	9.83	0.23	73.25	0.45	
C	Glyph	CG6	GC6.2	6.77	0.14	78.20	0.31	
C	Glyph	CG6	GC6.3	8.97	0.07	72.76	0.15	
C	Glyph	CG7	CG7.1	9.72	0.15	78.87	0.22	
C	Glyph	CG7	CG7.2	8.14	0.16	74.24	0.32	
C	Glyph	CG7	CG7.3	8.42	0.77	75.04	1.55	
C	Glyph	CG8	CG8.1	9.29	0.30	73.30	0.59	
C	Glyph	CG8	CG8.2	10.07	0.22	72.96	0.42	
C	Glyph	CG8	CG8.3	7.22	0.18	76.61	0.39	
C	Glyph	CG9	CG9.1	10.65	0.14	78.64	0.20	
C	Glyph	CG9	CG9.2	8.19	0.16	76.15	0.32	
C	Glyph	CG9	CG9.3	9.98	0.19	72.30	0.37	
C	Glyph	CG10	CG10.1	9.34	0.17	73.11	0.33	
C	Glyph	CG10	CG10.2	13.09	0.42	77.53	0.57	
C	Glyph	CG10	CG10.3	9.49	0.16	73.80	0.32	
C	Glyph	CG11	CG11.1	9.64	0.22	72.80	0.43	
C	Glyph	CG11	CG11.2	9.19	0.19	73.20	0.37	
C	Glyph	CG11	CG11.3	10.70	0.18	78.62	0.25	
C	Glyph	CG12	CG12.1	6.75	0.25	76.74	0.56	
C	Glyph	CG12	CG12.2	8.34	0.19	74.94	0.39	
C	Glyph	CG12	CG12.3	6.43	0.18	76.33	0.41	
C	Glyph	CG13	CG13.1	7.82	0.18	75.81	0.38	
C	Glyph	CG13	CG13.2	6.27	0.24	77.81	0.53	
C	Glyph	CG13	CG13.3	5.98	0.20	78.37	0.45	
C	Glyph	CG14	CG14.1	9.36	0.24	69.85	0.49	
C	Glyph	CG14	CG14.2	6.98	0.15	77.86	0.32	

Panel	Glyph or Reference?	Glyph/ Ref. No.*	Msmt. No.*	%Mn	+/-	%Fe	+/-	Unweathered (U) or Weathered (W) Reference?
C	Glyph	CG14	CG14.3	9.68	0.20	77.43	0.29	
C	Glyph	CG14	CG14.4	6.74	0.19	78.09	0.39	
C	Glyph	CG15	CG15.1	10.84	0.18	78.37	0.25	
C	Glyph	CG15	CG15.1	10.54	0.22	72.18	0.41	
C	Glyph	CG15	CG15.1	12.57	0.17	76.81	0.23	
C	Glyph	CG16	CG16.1	9.06	0.17	74.02	0.33	
C	Glyph	CG16	CG16.2	11.81	0.19	76.76	0.28	
C	Glyph	CG16	CG16.3	9.71	0.19	73.11	0.36	
C	Glyph	CG17	CG17.1	9.65	0.17	78.50	0.24	
C	Glyph	CG17	CG17.2	13.70	0.08	74.34	0.12	
C	Glyph	CG17	CG17.3	6.00	0.15	76.69	0.35	
C	Glyph	CG17	CG17.4	2.75	0.17	81.52	0.52	
C	Glyph	CG18	CG18.1	11.45	0.19	77.47	0.27	
C	Glyph	CG18	CG18.2	10.88	0.18	71.42	0.34	
C	Glyph	CG18	CG18.3	13.22	0.17	75.15	0.24	
C	Glyph	CG19	CG19.1	14.36	0.18	74.67	0.25	
C	Glyph	CG19	CG19.1	13.40	0.23	74.66	0.32	
C	Glyph	CG19	CG19.1	12.93	0.21	75.45	0.30	
C	Glyph	CG20	CG20.1	10.33	0.08	74.36	0.16	
C	Glyph	CG20	CG20.2	11.73	0.20	77.51	0.28	
C	Glyph	CG20	CG20.3	10.39	0.18	78.39	0.26	
C	Reference	CR1	CR1.1	2.32	0.14	81.25	0.47	U
C	Reference	CR2	CR2.1	2.31	0.14	81.57	0.48	U
C	Reference	CR3	CR3.1	2.38	0.15	80.89	0.50	U
C	Reference	CR4	CR4.1	18.32	0.14	71.70	0.18	W
C	Reference	CR5	CR5.1	15.26	0.15	72.34	0.21	W
C	Reference	CR6	CR6.1	20.24	0.15	69.86	0.20	W
C	Reference	CR7	CR7.1	14.12	0.08	70.57	0.12	W

Panel	Glyph or Reference?	Glyph/ Ref. No.*	Msmt. No.*	%Mn	+/-	%Fe	+/-	Unweathered (U) or Weathered (W) Reference?
C	Reference	CR8	CR8.1	17.99	0.13	72.69	0.18	W
C	Reference	CR9	CR9.1	18.84	0.08	59.65	0.13	W
C	Reference	CR10	CR10.1	20.02	0.16	70.77	0.21	W
C	Reference	CR11	CR11.1	17.69	0.11	72.55	0.15	W
C	Reference	CR12	CR12.1	20.20	0.13	69.85	0.17	W
C	Reference	CR13	CR13.1	19.43	0.15	70.26	0.21	W
C	Reference	CR14	CR14.1	20.10	0.25	70.01	0.53	W
C	Reference	CR15	CR15.1	18.78	0.20	71.11	0.37	W
C	Reference	CR16	CR16.1	16.58	0.11	73.00	0.15	W
C	Reference	CR17	CR17.1	14.41	0.13	73.97	0.18	W
C	Reference	CR18	CR18.1	23.94	0.11	66.30	0.15	W
C	Reference	CR19	CR19.1	27.57	0.16	61.66	0.23	W
C	Reference	CR20	CR20.1	19.29	0.13	70.24	0.17	W
D	Glyph	DG1	DG1.1	2.61	0.17	83.73	0.51	
D	Glyph	DG1	DG1.2	2.83	0.17	82.11	0.52	
D	Glyph	DG1	DG1.3	2.50	0.16	82.69	0.52	
D	Reference	DR1	DR1.1	25.43	0.18	63.53	0.25	W
D	Reference	DR2	DR2.1	28.07	0.07	47.09	0.11	W
D	Reference	DR3	DR3.1	26.30	0.17	62.86	0.24	W
E	Glyph	EG1	EG1.1	14.66	0.17	77.33	0.23	
E	Glyph	EG1	EG1.2	16.15	0.19	73.81	0.26	
E	Glyph	EG1	EG1.3	14.43	0.15	76.50	0.19	
E	Glyph	EG2	EG2.1	4.26	0.13	79.91	0.34	
E	Glyph	EG2	EG2.2	8.87	0.19	71.81	0.39	
E	Glyph	EG2	EG2.3	6.12	0.17	76.71	0.40	
E	Glyph	EG3	EG3.1	15.55	0.15	73.37	0.21	
E	Glyph	EG3	EG3.2	18.18	0.17	71.67	0.22	
E	Glyph	EG3	EG3.3	19.21	0.09	67.26	0.12	

Panel	Glyph or Reference?	Glyph/ Ref. No.*	Msmt. No.*	%Mn	+/-	%Fe	+/-	Unweathered (U) or Weathered (W) Reference?
E	Glyph	EG4	EG4.1	5.26	0.19	77.40	0.47	
E	Glyph	EG4	EG4.2	6.05	0.19	76.98	0.43	
E	Glyph	EG4	EG4.3	5.34	0.18	75.79	0.45	
E	Glyph	EG5	EG5.1	5.80	0.19	78.33	0.43	
E	Glyph	EG5	EG5.2	7.28	0.20	73.64	0.43	
E	Glyph	EG5	EG5.3	11.27	0.20	74.28	0.39	
E	Glyph	EG6	EG6.1	18.97	0.20	70.26	0.27	
E	Glyph	EG6	EG6.2	14.99	0.19	73.88	0.27	
E	Glyph	EG6	EG6.3	14.83	0.20	75.26	0.27	
E	Glyph	EG6	EG6.4	3.73	0.18	79.25	0.50	
E	Glyph	EG7	EG7.1	4.77	0.19	77.68	0.50	
E	Glyph	EG7	EG7.2	5.59	0.19	75.18	0.47	
E	Glyph	EG7	EG7.3	7.92	0.30	77.59	0.64	
E	Glyph	EG8	EG8.1	14.70	0.20	72.80	0.28	
E	Glyph	EG8	EG8.2	19.06	0.27	70.12	0.36	
E	Glyph	EG8	EG8.3	12.76	0.20	74.49	0.29	
E	Glyph	EG9	EG9.1	6.06	0.20	76.43	0.46	
E	Glyph	EG9	EG9.2	6.42	0.20	75.65	0.45	
E	Glyph	EG9	EG9.3	5.89	0.21	75.73	0.50	
E	Glyph	EG10	EG10.1	12.90	0.19	76.20	0.27	
E	Glyph	EG10	EG10.2	11.58	0.20	77.48	0.28	
E	Glyph	EG10	EG10.3	12.21	0.19	78.13	0.26	
E	Reference	ER1	ER1.1	25.09	0.13	63.45	0.19	W
E	Reference	ER2	ER2.1	29.92	0.14	56.15	0.22	W
E	Reference	ER3	ER3.1	29.18	0.19	58.33	0.29	W
F	Glyph	FG1	FG1.1	23.27	0.18	65.65	0.25	
F	Glyph	FG1	FG1.2	17.95	0.26	71.38	0.36	
F	Glyph	FG1	FG1.3	12.28	0.18	77.97	0.24	

Panel	Glyph or Reference?	Glyph/ Ref. No.*	Msmt. No.*	%Mn	+/-	%Fe	+/-	Unweathered (U) or Weathered (W) Reference?
F	Glyph	FG2	FG2.1	25.38	0.18	64.39	0.26	
F	Glyph	FG2	FG2.2	24.57	0.23	64.86	0.31	
F	Glyph	FG2	FG2.3	24.02	0.15	65.45	0.20	
F	Glyph	FG3	FG3.1	3.47	0.14	80.66	0.40	
F	Glyph	FG3	FG3.2	3.72	0.17	80.12	0.45	
F	Glyph	FG3	FG3.3	5.78	0.24	76.81	0.57	
F	Glyph	FG4	FG4.1	17.64	0.18	73.08	0.24	
F	Glyph	FG4	FG4.2	19.64	0.20	71.22	0.26	
F	Glyph	FG4	FG4.3	19.77	0.19	71.14	0.27	
F	Glyph	FG5	FG5.1	9.93	0.22	72.43	0.43	
F	Glyph	FG5	FG5.2	10.40	0.20	79.01	0.28	
F	Glyph	FG5	FG5.3	10.40	0.24	78.99	0.33	
F	Glyph	FG6	FG6.1	10.43	0.23	72.69	0.43	
F	Glyph	FG6	FG6.2	10.52	0.20	78.80	0.29	
F	Glyph	FG6	FG6.3	9.95	0.19	80.14	0.27	
F	Reference	FR1	FR1.1	19.77	0.19	71.14	0.27	W
F	Reference	FR2	FR2.1	18.27	0.21	71.56	0.30	W
F	Reference	FR3	FR3.1	23.12	0.17	67.22	0.23	W
F	Reference	FR4	FR4.1	2.39	0.12	82.02	0.40	U
F	Reference	FR5	FR4.2	2.12	0.14	83.14	0.48	U
F	Reference	FR6	FR4.3	2.36	0.17	82.14	0.46	U

*Identifying numbers for each measurement refer to panel, then glyph or reference indicator, then glyph or reference number, then measurement number. Thus measurement FG6.3 refers to Panel F, Glyph 6, third measurement.

Evaluation of Assumptions

As discussed above, four main assumptions needed to be met for PXRF to be successfully applied to the Hole-in-the-Ground petroglyphs. These assumptions were: 1) weathering of these rocks produces chemical change significant enough to be detectable by the

instrument, 2) baseline geology of the rocks is sufficiently uniform within a given panel, 3) within-panel weathering is sufficiently uniform, and 4) petroglyph pecking removed all prior patination.

Assumption 1 is clearly supported by the fact that, among reference measurements, the differences between average unweathered Mn and Fe scores and average fully weathered scores is substantial (Table 3-3), suggesting weathering produces a significant difference in rock Mn and Fe content by enriching Mn by a factor of 10 in the varnish atop weathered surfaces and diluting Fe by a smaller, though readily detectible amount.

Table 3-3. Average Relative Concentrations of Mn and Fe for PXRF Reference Measurements

Unweathered Rock		**Weathered Rock**		**Difference**	
Mn	**Fe**	**Mn**	**Fe**	**Mn**	**Fe**
2.37	81.39	20.26	68.59	17.89	12.80

Assumption 2 is supported for Mn data by the fact that one-way ANOVA testing of unweathered rock measurements grouped by source sample shows statistical compatibility at 95% confidence, meaning baseline Mn concentrations are statistically compatible among all unweathered rocks measured ($p = 0.093$). This assumption is not met for Fe data, for which ANOVA testing shows significant differences between distinct samples of unweathered rocks at 95% confidence ($p = 0.042$). Thus for these boulders, the Fe concentrations are likely too heterogeneous to provide a consistent baseline against which to compare relative weathering of glyphs.

Assumption 3 is not fully met by either Mn or Fe data. ANOVA testing of fully weathered surfaces shows significant differences in Mn ($p < 0.001$) and Fe ($p < 0.001$) when values are compared between panels, and values can even vary widely within some panels. This variance in weathering-related Fe is unsurprising given baseline variance in unweathered Fe values, although it means that Fe is clearly unsuitable for inferring relative weathering of these stones. Unfortunately, significant differences in Mn accumulation also suggest that Mn accumulation at Hole-in-the-Ground is not sufficiently uniform even when baseline geology is homogenous and slope and aspect are held constant. Again, significantly different boulder exposure ages are unlikely given the fact that boulders were most likely subaerially exposed during the Alvord Lake outburst flood. Instead, it is likely that some of the observed variation is a product of micro-topographic effects, as small black Mn-rich pits are clearly visible on the surfaces of many boulders at the site, leading to relatively high variability in Mn values even within panels. Some variation may also be due to differential rates of eolian erosion, as this would act to reduce apparent varnish accumulation and may not have impacted the irregular boulders/panels uniformly. Whatever the cause, however, it is clear that Mn varnish accumulation is not strictly a product of duration of exposure. On the other hand, the standard deviation in Mn scores across all measurements of fully weathered surfaces was 4.04, so on the whole error in fully weathered Mn values is about 20% of average measured Mn. Given this, Mn varnish accumulation for these stones can be assumed to be *predominantly* a function of time, with about 20% error resulting from other factors. Thus, provided this uncertainty is incorporated

into estimates of relative age (see below), valid results can be derived, even if they may be encumbered by additional uncertainties due to uneven Mn varnish accumulation over time.

Assumption 4 was partially addressed through sampling (discussed above) and analytical protocols (see below). Additionally, the data support the assumption that glyphs were well-pecked in antiquity. This is because the lowest Mn value of glyph DG1 (Mn = 2.50 +/- 0.16)—a glyph noted in the site report as having been defaced between 1976 and 1991—closely resembles the average Mn value of freshly broken reference samples (2.37 +/- 0.13), indicating that its rock's varnish was extensively removed during defacement. Glyph DG1 appears to be no more thoroughly pecked in terms of depth of varnish removal than most of the glyphs examined for this study, making it likely that these other glyphs also removed the preexisting varnish at the time of original glyph creation.

Scaled Weathering Scores

With the necessary assumptions largely met for Mn and the above caveats in mind, it was possible to derive tentative "scaled weathering scores" representing the degree to which each glyph's minimum measured Mn value differs from that of unweathered rock. Fe values were discarded for purposes of this analysis, since they violate multiple methodological assumptions. The scaled weathering scores are given for all measured glyphs in Table 3-4 and were derived using the following steps:

- The lowest Mn value of each glyph's group of measurements was selected. As discussed above, this value is arguably the most likely to represent the most well-pecked portion of the glyph, and therefore the best estimate of overall glyph age. In cases where glyphs

appeared to have been re-pecked (CG14, CG17, and EG6), a lowest Mn value was selected for both the "original" and the "re-pecked" portion of the glyph.

- An average Mn value with standard deviation (Mn = 2.37 +/-.13) was calculated using pooled data from all unweathered reference measurements. Because these samples' Mn values are statistically compatible, this average is a valid representation of Mn values for all unweathered reference samples.
- Each glyph's minimum Mn was then divided by the average Mn value for all unweathered rock. This conversion allows all glyphs to be standardized along a common scale, where the value of unweathered rock is given as 1.00 and glyph values are expressed as ratios to this value. Higher values along this scale represent more weathered rock. To the extent that Mn weathering is a direct representation of glyph age—again a tentative assumption in light of data suggesting weathering is not uniform even within panels—higher values also represent older glyphs. Thus glyph CG3, which has a score of 3.69, has roughly twice as much Mn accumulation as glyph EG2 score of 1.80, and by extension glyph CG3 is roughly twice as weathered and therefore likely to be at least twice[1] as old.
- Error terms in scaled weathering scores for glyphs were calculated using the following equation, where *Mn* reflects the scaled Mn score (multiplied by 0.2 to reflect the ~20%

[1] Assuming Mn accumulation per unit time follows a linear or sublinear (weathering slows over time) function. This assumption is supported by published research which suggests that weathering rates are typically sublinear (e.g., Colman 1981; Hunt 2015; Sak et al. 2004). The alternative—superlinearity—is unlikely, as this would imply that Mn accumulation progressively accelerates over time.

variance in Mn values for fully weathered surfaces), E_1 is the measurement error (as reported by the instrument) associated with the minimum Mn value for each glyph, E_2 is the standard deviation in all measurements of unweathered reference samples, and E_T is total error:

$$E_T = \sqrt{(0.2 * Mn)^2 + E_1^2 + E_2^2}$$

Importantly, error terms expressed here encompass only 1) the roughly 20% observed variance in fully weathered Mn values, 2) uncertainties reported by the instrument in counting of X-ray peaks in measurement of glyphs, and 3) uncertainties in characterizing the Mn signature of unweathered rocks. They *do not* account for uncertainties due to drift in instrumental performance (see Johnson 2014a), which would inflate instrumental error, and they should therefore be taken as slight underestimates of true errors. Fully accounting for true error terms is technically feasible, but would necessitate a major study far beyond the scope of the present inquiry due to the need for intensive re-measurement of sampled locations and calibration standards following strict and time-consuming protocols (see Johnson 2014a). In any case, the effect of possible additional instrumental error is unlikely to significantly influence results; observed errors are sufficiently high to render possible additional instrumental errors insignificant by comparison.

Table 3-4. Scaled Weathering Score for all Glyphs, Arranged in Ascending Order of Least-weathered to Most-weathered

Glyph	Scaled Weathering Score	+/-
Unweathered	1.00	0.18
DG1	1.06	0.32
CG17 (Re-pecked)*	1.16	0.34
FG3	1.47	0.37
EG6 (Re-pecked)*	1.58	0.41
AG6	1.79	0.40
EG2	1.80	0.42
EG4	2.22	0.52
EG7	2.36	0.54
CG2	2.38	0.53
EG5	2.45	0.56
EG9	2.49	0.57
CG13	2.53	0.57
CG4	2.70	0.59
CG12	2.72	0.60
CG14 (Re-pecked)*	2.85	0.63
CG6	2.86	0.62
CG5	3.03	0.64
CG8	3.05	0.66
AG4	3.39	0.72
CG7	3.44	0.73
CG9	3.46	0.73
CG3	3.69	0.78
BG1	3.78	0.80
CG16	3.83	0.80
CG11	3.88	0.80
AG5	3.88	0.82

Glyph	Scaled Weathering Score	+/-
CG10	3.95	0.83
CG14 (Original)*	3.96	0.82
BG2	3.96	0.84
CG17 (Original)*	4.08	0.85
FG5	4.20	0.89
FG6	4.21	0.88
CG1	4.25	0.88
CG20	4.37	0.90
CG15	4.46	0.94
AG1	4.51	0.94
CG18	4.60	0.95
BG3	4.79	0.99
EG10	4.89	1.02
FG1	5.19	1.07
EG8	5.39	1.11
CG19	5.47	1.13
EG1	6.10	1.24
EG6 (Original)*	6.27	1.28
EG3	6.57	1.34
AG3	7.03	1.43
FG4	7.46	1.51
AG7	7.57	1.53
AG2	7.76	1.57
FG2	10.15	2.04

* Glyphs which were re-pecked (CG14, CG17, and EG6) have both an “original” and a “re-pecked” score listed. “Unweathered” refers to the average Mn value of unweathered stone references.

Discussion of Results

A graphical display of scaled weathering scores with 1σ errors for all measured glyphs is given by Figure 3-7; glyphs are arrayed from lowest Mn score to the highest. In general, glyphs show a wide range of scaled weathering scores, with varnish covering the most-weathered glyph containing more than 10 times the estimated Mn as the least-weathered glyph. Given this, it is highly probable that glyphs at Hole-in-the-Ground were produced over a very long period of time. Glyph FG2, for example, is likely several millennia in age, as its observed Mn values were as high as most fully weathered reference measurements, while glyph DG1 is statistically indistinct from sampled unweathered rock, and is therefore essentially modern in its Mn content. Further, glyphs EG1, EG6 (Original), EG3, AG3, FG4, AG2, and FG2 have all accumulated between roughly 85 and 150 times as much Mn as glyph DG1 has accumulated since its defacement between 1976 and 1991 (24 to 37 years ago), implying that these glyphs are *at least*[2] somewhere between 2,000 and 5,500 years in age (24*85 = 2,040 and 37*150 = 5,550). While these rough age figures represent pure extrapolations of high-error estimates along an inaccurate mathematical model for Mn weathering, and are therefore *extremely* tentative, they are in accord with the known ~5,300-year occupational history of nearby Birch Creek Site (35ML181). Thus, although independent dates at Hole-in-the-Ground are needed before definitive claims as to the

[2] Using (for argument's sake) a linear accumulation of Mn over time. Again, Mn accumulation is likely to slow with time, increases in Mn concentrations are likely to yield incrementally diminishing observed PXRF intensities due to progressive detector saturation, so the assumption of linearity as a model, although technically inaccurate, provides a conservative (i.e., minimum) estimate of glyph age. Maximum glyph ages are governed by the exposure age of the boulders themselves, which in this case is likely to be the timing of the Alvord outburst floods (~14,000 years ago).

age of the petroglyphs can be made, current data suggest that production of these glyphs may have spanned most or all of the known local occupational history. The establishment of the antiquity of these glyphs is a significant result given the lack of independent chronometric data at Hole-in-the-Ground and the challenges in directly dating glyphs in general.

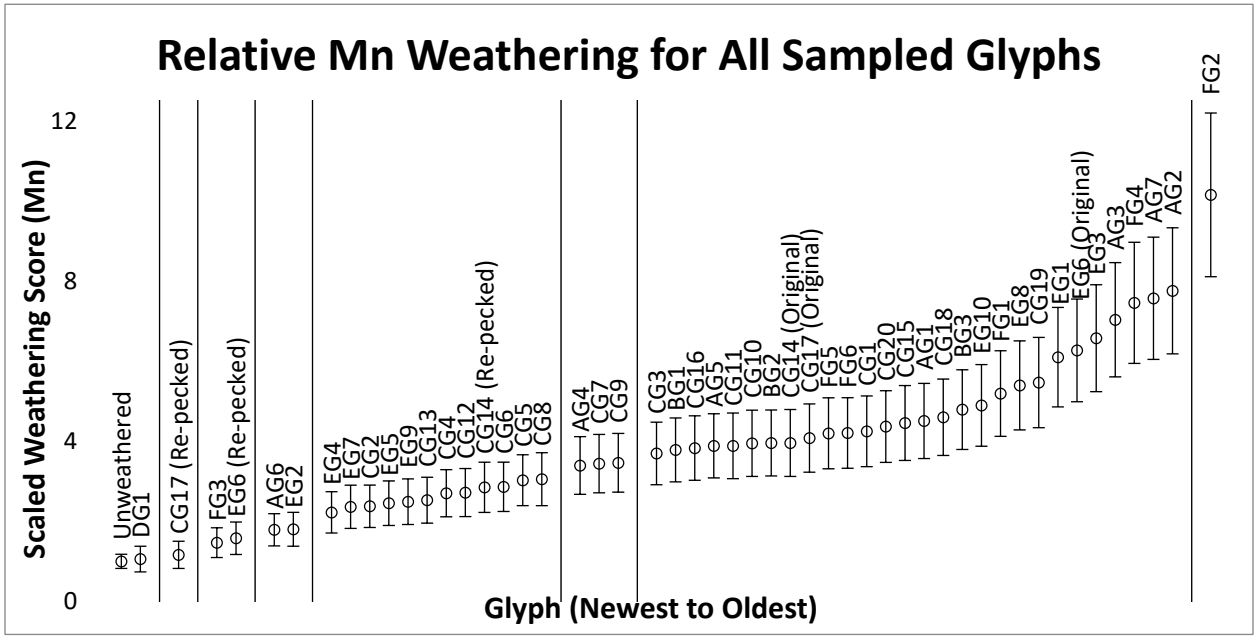

Figure 3-7. Display of Mn scaled weathering scores for all measured glyphs, with 1σ error terms. Tall vertical lines between glyphs indicate statistically significant differences in Mn scores at 90% confidence or higher.

Additionally, when arrayed by Mn score, t-testing[3,4] shows that some adjacent Mn scores along this array are statistically distinct at 90% confidence; this includes gaps between Mn scores for glyphs DG1 and CG17 (Re-pecked), glyphs CG9 and CG3, and glyphs AG2 and FG2. Further, Mn scores are likewise distinct at 99% confidence between glyphs CG17 (Re-pecked) and FG3, glyphs EG6 (Re-pecked) and AG6, glyphs EG2 and EG4, and glyphs CG8 and AG4. These statistical "seams" in the sample data suggest that there are likely at least eight distinct groups of sampled glyphs in terms of Mn varnish accumulation, and therefore inferred age. This is a strong result given the relatively high error terms which accompany estimated Mn scores (average 21.7% error; this is largely a product of the uncertainty in baseline Mn accumulation rates), although the partition of glyphs into eight groups—or perhaps "phases"—must remain tentative until a more comprehensive sample of Hole-in-the-Ground glyphs are measured, as Mn scores of as-yet unsampled glyphs may ultimately fill the statistical "seams" between groups. For measured glyphs which are statistically compatible in terms of Mn content, some within-group ordinal reversals may exist in the array provided by Figure 3-7, and the order in which these results are arrayed is therefore also a tentative rather than definitive ordering of Mn values.

Still, it is apparent from these data that individual Hole-in-the-Ground panels were re-used by the area's inhabitants over long periods of time. That is, PXRF results suggest that large panels such as Panel A, C, E, and F were not decorated once and then subsequently left

[3] Assuming infinite degrees of freedom, since thousands of instrument counts inhere in each measured Mn value.

[4] A more exhaustive analysis of distinct age groups in the form of "structural breaks" (e.g., Zeileis et al. 2001) was not merited given the small number of glyphs sampled and the imprecision incurred in estimating Mn scores.

unaltered, but were instead repeatedly altered over several centuries or millennia, perhaps during chronometrically distinct phases of use. In some sense, then, the XRF data show that panels represent a continuing decorative dialogue among ancient peoples. This dialogue has sometimes included revision or re-pecking of much older glyphs. Sometimes this re-pecking represents relatively modern vandalism, as is likely the case with glyph CG17 (Re-pecked) and possibly EG6 (Re-pecked), but sometimes this re-pecking was clearly performed in antiquity as with glyph CG14 (Re-pecked). Unfortunately, the number of sampled glyphs was too small to examine frequency changes in decorative motifs over time, but such a study appears to be within the realm of technical possibility moving forward.

On the whole, then, the application of PXRF to the Hole-in-the-Ground glyphs can be deemed a moderate success. The basic assumptions needed to productively apply the technique to the measurement of Mn varnish accumulations were addressed through strategic sampling and/or satisfied by empirical results, although the same cannot be said for Fe accumulations. Variance in Mn accumulation rates of un-pecked rock surfaces introduced uncertainty into results and were therefore somewhat damaging to the strength of conclusions, but in the end they did not entirely obscure age-related patterns in Mn weathering, and they were not a major hindrance to the construction of a relative chronology for glyphs. If an absolute chronology had been the goal, as with previous research (Lytle el al. 2009; McNeil and Cecil 2009), variation in varnish accumulation rates at Hole-in-the-Ground would have greatly impeded results, but this was not the case for this study. Instead, our efforts were able to produce a relative chronology of a substantial subset of Hole-in-the-Ground glyphs even under difficult and truncated field

conditions, while also allowing analysis of the possibilities for future archaeological fieldwork at Hole-in-the-Ground, including possible future applications of PXRF at the site.

FUTURE WORK

Given the preceding discussion, it is clear that the application of portable XRF to relative dating of petroglyphs holds promise. At Hole-in-the-Ground, achieving this promise will require comprehensive measurement of all glyphs—or at least those which appear to have been well pecked and have sufficient geometry to accommodate the instrument—to allow full examination of temporal patterns in panels and motifs. As part of this effort, a subsample of glyphs and reference locations should be measured following strict protocols outlined elsewhere (Johnson 2014a) as a means of more rigorously estimating actual instrumental errors, and comparing these errors with those derived from the methods used in this study.

Ideally, this work would be accompanied by broader fieldwork at Hole-in-the-Ground to produce 1) comprehensive documentation of site glyphs, 2) a broader understanding of the functional context of the site, and 3) independent dates for occupation of the site. Comprehensive and systematic documentation of all boulders, panels, and glyphs and would include mapping, photography, and sketches, and would facilitate detailed and reproducible analysis of time-transgressive patterns in glyphs. If possible, a high-resolution scanning device such as LiDAR should be used to create 3D digital documentation of glyphs (for examples, see Lerma et al. 2010; Stance et al. 2009; Vogt and Edsall 2010); this would help detect and document glyphs which are not readily detectable by the naked eye due to faint pecking or extensive weathering, including varnish accumulation. Examination of the site's use context and derivation of

independent dates for this use would likely require subsurface investigation, but such work is warranted given the array of interesting material remains at the site, which include not only glyphs but also chipped-stone tools, fire-modified rock, bedrock mortars (Figure 3-8), and others. Further, the apparent removal of glyphs—and therefore likely other artifacts—by modern site visitors travelling along the Owyhee only adds incentive to conduct this work before artifact removal and glyph defacement progress further.

Once this is done, the Hole-in-the-Ground site and petroglyphs should be compared with similar sites in the area, including the many other petroglyph sites nearby, to allow a broader regional analysis of the functional and temporal relationships which linked these sites in antiquity. Importantly, while relative dating of petroglyphs using XRF may be a useful tool within a given site, it will remain very difficult to integrate relative petroglyph ages between sites unless rock weathering can be shown to be identical between multiple sites. This is unlikely given the variation within Hole-in-the-Ground alone, and the possibility of development of an absolute method for dating petroglyphs with XRF seems even more remote. In the future it may be possible to date rock surfaces such as petroglyphs directly using luminescence dating (e.g., Freiesleben 2014; Sohbati et al. 2012), but the destructive nature of such techniques will likely make them a poor fit for petroglyphs. Still, XRF relative dating of petroglyphs within each site can help flesh out a picture of a given site's use history, and an investigation as to whether certain motifs are consistently relatively old or young across many sites is a worthwhile objective. If such efforts are anchored by independent absolute dates where available—

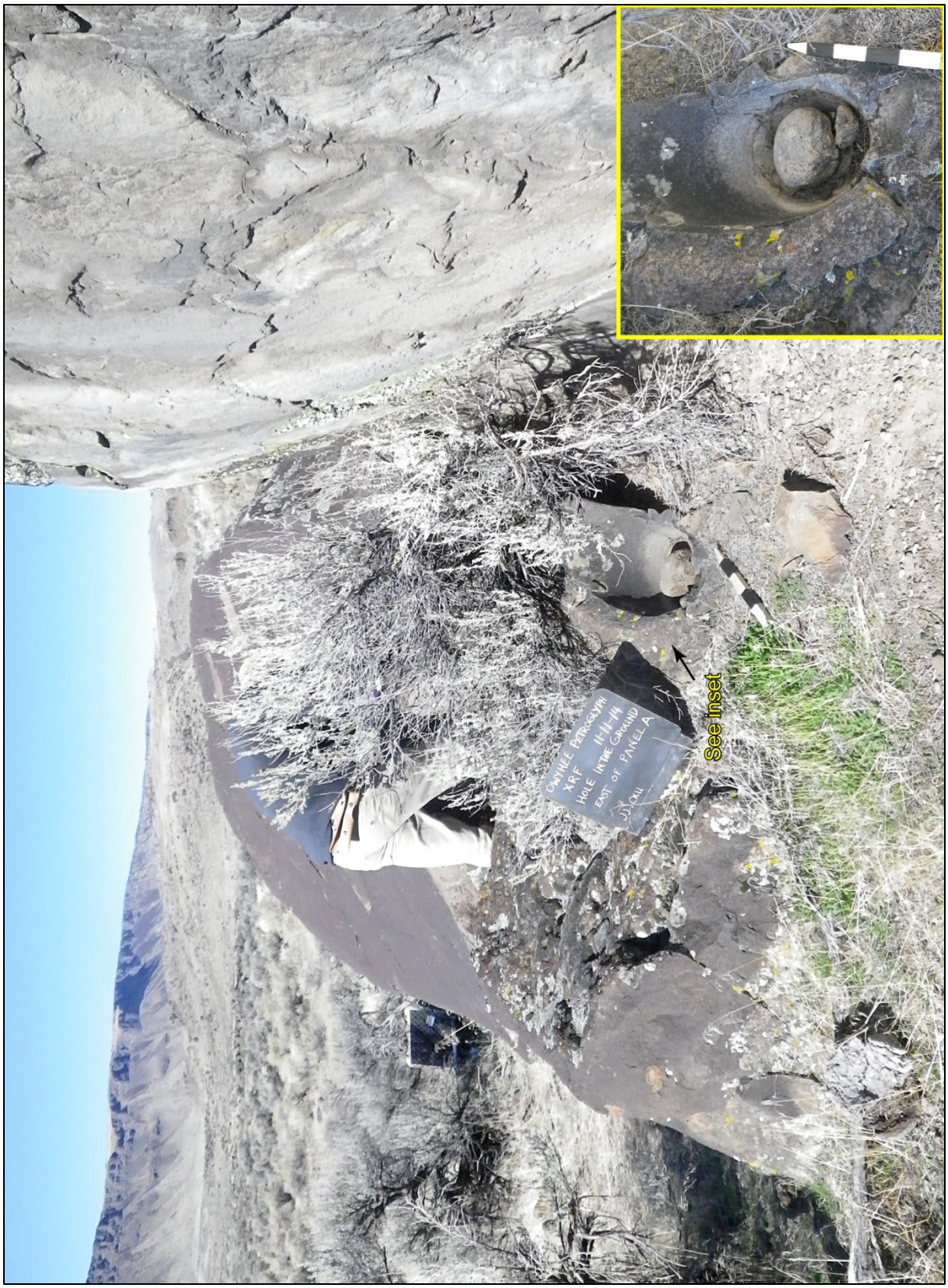

Figure 3-8. Bedrock mortar east of Panel A, view northwest.

for example, at the Bobcat Beach Camp and Petroglyphs site (35ML1168), which exhibits petroglyph overlain by sediment (Sudman 2005:32) which is likely to be dateable using luminescence dating (see Lian and Roberts 2006)—refinement of a detailed chronology of petroglyph ages may yet be possible.

ACKNOWLEDGMENTS

An earlier version of this research was submitted to the BLM Vale District Office of the as part of SWCA Cultural Resources Report No. 14-675 (Johnson 2014b). I am thankful to Cheryl Bradford of the Vale District Office for procuring funding for this work and for allowing the resulting report to be reproduced here. I am also thankful to SWCA for connecting me with this work, and for their subsequent support, and to the University of Washington Student Technology Fee Committee for providing funding to purchase the necessary XRF equipment. Cyrena Undem of SWCA was an indispensable partner in the field, and Dan Thomas of the BLM guided Cyrena and I into the wilderness of the Owyhee and evacuated us just in time to avoid almost certain death from exposure during Winter Storm Bozeman. Rhiannon Held and Johonna Shea provided valuable technical assistance and graphics. Thanks are also due to Jim Feathers for thoughtful comments on earlier drafts.

REFERENCES

Aikens, M.C, T.J. Connolly, and D.L. Jenkins
2011 *Oregon Archaeology*. Oregon State University Press, Corvallis.

Andrefsky, W., Jr.
2008 Projectile Point Provisioning Strategies and Human Land Use. In *Lithic Technology*, edited by William Andrefsky Jr., pp. 195–215. Cambridge University Press, New York.

2010 Human Land Use Strategies and Projectile Point Damage, Resharpening and Discard Patterns. *Human Evolution* 25(1–2):13–30.

Carter, D.T., L.L. Ely, J.E. O'Connor, and C.R. Fenton
2006 Late Pleistocene outburst flooding from pluvial Lake Alvord into the Owyhee River, Oregon. *Geomorphology* 75(3–4):346–367.

Colman, S.M.
1981 Rock-weathering rates as functions of time. *Quaternary Research* 15(3):250–264.

Craig, N, R.J. Speakman, R.S. Popelka-Filcoff, M.D. Glascock, J.D. Robertson, S.M. Shackley, and M.S. Aldenderfer
2007 Comparison of XRF and PXRF for analysis of archaeological obsidian from southern Perú. *Journal of Archaeological Science* 34(12):2012–2024.

Cremaschi, M.
1996 The Rock Varnish in the Messak Settafet (Fezzan, Libyan Sahara), Age, Archaeological Context, and Paleo-Environmental Implication. *Geoarchaeology* 11(5):393–421.

Evershed, R.P.
2008 Organic Residue Analysis in Archaeology: The Archaeological Biomarker Revolution. *Archaeometry* 50(6):895–924.

Feathers, J.K.
2003 Use of luminescence dating in archaeology. *Measurement Science and Technology* 14: 493–1509.

Fisher, P.R.
2010 Points in Time: Stone Tool Analysis as an Indication of Group Movement at the Birch Creek Site (35ML181), Southeastern Oregon. Unpublished Master's thesis. Washington State University, Pullman.

Freiesleben, T.
2014 Developing a Method for Luminescence Dating of Rock Surfaces. Unpublished Master's thesis. Roskilde University, Roskilde, Denmark.

Goodale, N., D.G. Bailey, G.T. Jones, C. Prescott, E. Scholz, N. Stagliano, and C. Lewis
2012 pXRF: a study of inter-instrument performance. *Journal of Archaeological Science* 38(3):875–883.

Gjesfjeld, E.
2014 Of pots and people: investigating hunter-gatherer pottery production and social networks in the Kuril Islands. Unpublished Ph.D. dissertation. University of Washington, Seattle.

Hungate, B., A. Danin, N.B. Pellerin, J. Stemmier, P. Kjellander, J.B. Adams, and J.T. Staley
1987 Characterization of manganese-oxidizing (MnII to MnIV) bacteria from Negev Desert rock varnish: implications in desert varnish formation. *Canadian Journal of Microbiology* 33(10):939–943.

Hunt, A.G.
2015 Predicting Rates of Weathering Rind Formation. *Vadose Zone Journal* 14(7). doi:10.2136/vzj2014.09.0123.

Johnson, J.A.
2014a Accurate Measurements of Low *Z* Elements in Sediments and Archaeological Ceramics Using Portable X-Ray Fluorescence (PXRF). *Journal of Archaeological Method and Theory* 21(3):563–588.

2014b *Exploratory Portable X-Ray Fluorescence Measurement of the Hole-in-the-Ground Petroglyphs, Malheur County, Oregon*. Report No. 14-675. Prepared for Bureau of Land Management, Vale District Office. SWCA Environmental Consultants, Seattle.

Lachance, G.R., and F. Claisse
1995 *Quantitative X-ray Fluorescence Analysis*. Wiley, Chichester, UK.

Lerma, J.L., S. Navarro, M. Cabrelles, and V. Villaverde
2009 Terrestrial laser scanning and close range photogrammetry for 3D archaeological documentation: the Upper Paleolithic Cave of Parpallo as a case study. *Journal of Archaeological Science* 37(3):499–507.

Lian, O.B., and R.G. Roberts
2006 Dating the Quaternary: progress in luminescence dating of sediments. *Quaternary Science Reviews* 25(19–20):2449–2468.

Liu, T., and W.S. Broecker
2000 How Fast Does Rock Varnish Grow? *Geology* 28(2):183–186.

Lytle F., M. Lytle, A. Rogers, A. Garfinkel, C. Maddock, W. Wright, and C. Cole
2008 *An Experimental Technique for Measuring Age of Petroglyph Production: Results on Coso Petroglyphs*. Paper presented at the 31st Great Basin Anthropological Conference, Portland, Oregon.

McNeil, J.A., and F.E. Cecil
n.d. *X-ray Fluorescence Measurements of Manganese in Petroglyphs and Graffiti in the Bluff, Utah Area*. Unpublished monograph. http://inside.mines.edu/~jamcneil/XRF_Report_Bluff_Ut.pdf, accessed December 7, 2014.

Nagy, B., L.A. Nagy, M.J. Rigali, W.D. Jones, D.H. Krinsley, and N.A. Sinclair
1991 Rock varnish in the Sonoran Desert: microbiologically mediated accumulation of manganiferous sediments. *Sedimentology* 38(6):1153–1171.

Nazaroff, A.J., K.M. Prufer, and B.L. Drake
2009 Assessing the applicability of portable X-ray fluorescence spectrometry for obsidian provenance research in the Maya lowlands. *Journal of Archaeological Science* 37(4):885–895.

Noll, C.D.
2009 Late Holocene Occupation of the Birch Creek Site (35ML181), Southeastern Oregon. Unpublished Master's thesis. Washington State University, Pullman.

Phillips, S.C. and R.J. Speakman
2009 Initial source evaluation of archaeological obsidian from the Kuril Islands of the Russian Far East using portable XRF. *Journal of Archaeological Science* 36(6):1256–1263.

Sak, P.B., D.M. Fisher, T.W. Gardner, and K. Murphy
2004 Rates of weathering rind formation on Costa Rican basalt. *Geochimica et Cosmochimica Acta* 68(7):1453–1472.

Shackley, S.M., ed.
2011 *X-ray Fluorescence Spectrometry (XRF) in Geoarchaeolgy*. Springer.

Sohbati, R., A.S. Murray, M.S. Chapot, M. Jain, and J. Pederson
2012 Optically stimulated luminescence (OSL) as a chronometer for surface exposure dating. *Journal of Geophysical Research* 117(B9). doi:10.1029/2012JB009383.

Stance, H., E. Louden, and E. Johnson
2009 Unraveling the Cowhead Mesa Petroglyphs with 3-d Laser Scanning and High-resolution Digital Photography. *Rock Art Research: The Journal of the Australian Rock Art Research Association (AURA)* 26(1):43–53.

Sudman, N.
1993 *Oregon Cultural Resource Site Record for the Hole in the Ground Petroglyphs, site numbers 35ML169 and 35ML169*. Prepared for the Bureau of Land Management, Vale District, Jordan R.A.

2005 *A Cultural Resources Inventory and Monitor of the Lower Section of the Owyhee River: Rome to Birch Creek, Oregon*. BLM Report J-05-13. Prepared for the Bureau of Land Management, Vale District, Jordan R.A.

Vogt, B.J., and R.M. Edsall
2010 Terrestrial laser scanning and exploratory spatial analysis for the mapping of weathering forms on rock art panels. *Geocarto International* 25(5):346–367.

Zeileis, A., F. Leisch, K. Hornik, and C. Kleiber
2001 *strucchange: An R package for testing for structural change in linear regression models*. SFB Adaptive Information Systems and Modelling in Economics and Management Science 55. WU Vienna University of Economics and Business, Vienna.

Chapter 4. Luminescence Dating of Sediments at the Bear Creek Site (45KI839), Redmond, Washington

ABSTRACT

Excavations at the Bear Creek Site (45KI839) in Redmond, Washington provided a rare opportunity to investigate human occupation during the Late Pleistocene-Holocene transition in the Puget Lowland. As part of a multidisciplinary effort to examine this site, an extensive multi-proxy effort was undertaken to refine site chronometry. This paper describes the application of luminescence dating to Bear Creek sediments, as well the Bayesian statistical integration of dates derived from luminescence, radiocarbon, and Mazama ash into a single model of site chronology.

Eleven samples for luminescence dating were collected from sediments at Bear Creek, including three samples of Stratum Vc, four samples of Stratum III, and four samples of Stratum II. Twelve dates were derived from these samples using a combination of IRSL measurement of single-grain feldspars and OSL measurement of fine-grained quartz to estimate equivalent dose values, and a combination of alpha counting, beta counting, and X-ray fluorescence for dose rate measurements.

Results show that, while fine-grained samples exhibited some evidence of partial bleaching, luminescence ages of Bear Creek samples generally match dates based on independent methods while providing novel information on depositional rates at the site. Further, Bayesian analysis using OxCal 4.2 shows luminescence dates are statistically and stratigraphically compatible with independent dates, including radiocarbon dates and published

ages for Mazama ash, strengthening interpretations of site antiquity and overall depositional chronology.

BACKGROUND

Test excavations conducted at the Bear Creek Site (45KI839) in 2008 by Northwest Archaeological Associates, Inc., revealed a buried peat layer radiocarbon-dated to between 8000 and 10,000 calibrated years before present (cal BP) superimposed atop chipped-stone artifacts thought to represent human occupation coeval with the Late Pleistocene-Holocene (LPH) transition (Kopperl et al. 2010). In general, archaeological evidence of LPH human occupation in western Washington is very sparse, and is limited to a few isolated finds of fluted projectile points (Carlson 1990; LeTourneau 2010; Meltzer and Dunnell 1987; Osborne et al. 1956) and possible finds of human hunting/butchering of extinct fauna such as mastodon (Gustafson et al. 1979; Waters et al. 2011) or *Bison antiquus* (Kenady et al. 2007), although the former case remains problematic (Grayson and Meltzer 2015). Given this, the assemblage and intact stratigraphy at the Bear Creek Site hold significant potential for examining human occupation in western Washington during the LPH.

Because of this potential, SWCA Environmental Consultants (SWCA) returned to the Bear Creek Site in 2013 to undertake extensive data recovery excavations as part of mitigation efforts associated with rehabilitation and restoration of the creek and the surrounding wetland. This excavation recovered around 3,600 chipped stone artifacts from the Bear Creek Site, roughly 90% of which derive from the layer thought to be LPH in age (Taylor and Beck 2016).

The overarching research goals of this excavation were multiple, but included analysis of the Bear Creek lithic assemblage and a comparison of this assemblage with established Paleoarchaic (*sensu* Willig 1989; Willig and Aikens 1988) lithic technologies such as the Western Stemmed Tradition (WST) (Bryan 1980; Willig and Aikens 1988) and the later Old Cordilleran Tradition (OCT) (Butler 1961; Leonhardy and Rice 1970), a geoarchaeological analysis of site formation processes, and interpretation of the changing local environment throughout the Holocene. To achieve these research goals, the development of a more refined and comprehensive site chronology was also necessary.

The University of Washington Luminescence Dating Laboratory (UWLDL) was contracted by SWCA to address this need, and was tasked with three specific chronometric objectives. First, SWCA sought independent testing of radiocarbon ages for Stratum Vc, the artifact-bearing layer thought to date to before 10,000 cal BP, as a means of more fully assessing evidence for the antiquity of the earliest cultural deposits at the site. Second, SWCA sought dates for sediments in which samples for radiocarbon dating were scarce, absent, or arguably detrital, including Strata III and II (channel alluvium and diatomaceous earth, respectively) as a means of detailing and refining site depositional history through the integration of multiple dates derived from multiple independent methods. Third, SWCA sought statistical integration of dates from all proxies within a Bayesian framework, and (if possible) refinement of the precision of individual dates within this framework. To this end, I participated in 8 weeks of SWCA's 2013 Bear Creek excavations, collected 11 samples for luminescence dating, and performed all dating of these

samples as well as Bayesian analysis at UWLDL. Following a brief general introduction to luminescence dating at Bear Creek, sample collection and analysis are detailed below.

APPLYING LUMINESCENCE DATING AT BEAR CREEK

Luminescence dating provides a method for estimation of the depositional age of sediments containing common crystalline minerals such as quartz or feldspar, and therefore has obvious potential for addressing the need for independent dating of the variety of sediments at Bear Creek. This approach relies on the fact that 1) defects in these crystalline minerals progressively accumulate a charge in direct proportion to the intensity of ambient radiation, 2) accumulated charge is removed by exposure to light of sufficient intensity, leading to "bleaching" of the crystal lattice, and 3) rates of ambient radiation and the charge "trapped" in defects can be measured and calibrated through careful experimentation, allowing for estimation of duration of charge accumulation since the last "bleaching" event, and therefore time of most recent exposure to sufficient light. In principle this process is straightforward, but in practice sediment samples exhibit a wide array of luminescence properties, and as a result luminescence dating often involves a lengthy process of experimentation before a given sample can be successfully and reliably dated. Aitken (1985, 1998) provides detailed discussion of the fundamentals of luminescence dating, and Nelson and colleagues (2015) provide an excellent modern review of its practical application to a variety of archaeological contexts (see also Roberts et al. 2015).

Given the preceding, successful application of luminescence dating to the Bear Creek sediments requires 1) sampling of sediments likely to have been "bleached" in antiquity but not

since, 2) measurement of the trapped charge within each sample; this value is called the "natural signal," 3) calculation of the radiation needed to account for each observed natural signal; this value is called the "equivalent dose" (D_e), and is the numerator of the age equation, 4) calculation of the rate of radiation (in Gy/year) affecting each sample; this value is called the "dose rate" (D_r), and is the denominator of the age equation, 5) statistical assessment of the degree to which resultant data reflect a "well-bleached" sample of uniform luminescence age, and 6) statistical isolation of the well-bleached portion of samples for calculation of age estimates if "partial bleaching" (retention of charge accumulated prior to the exposure event of interest) is evident.

Sampling, preparation, luminescence measurement, and analysis were conducted with these objectives in mind.

SAMPLING

Extensive geoarchaeological analysis and interpretation of Bear Creek sediments was performed by SWCA and is exhaustively detailed elsewhere (Hodges et al. 2016), but a summary is useful here to provide a basis for understanding the priorities which guided luminescence sampling. Briefly, Bear Creek deposits were divided into six distinct strata (Figure 4-1); these were assigned numerals as Strata I through VI. Strata lithology and organic inclusions were meticulously examined and used to infer the ancient depositional environment represented by each stratum (Table 4-1).

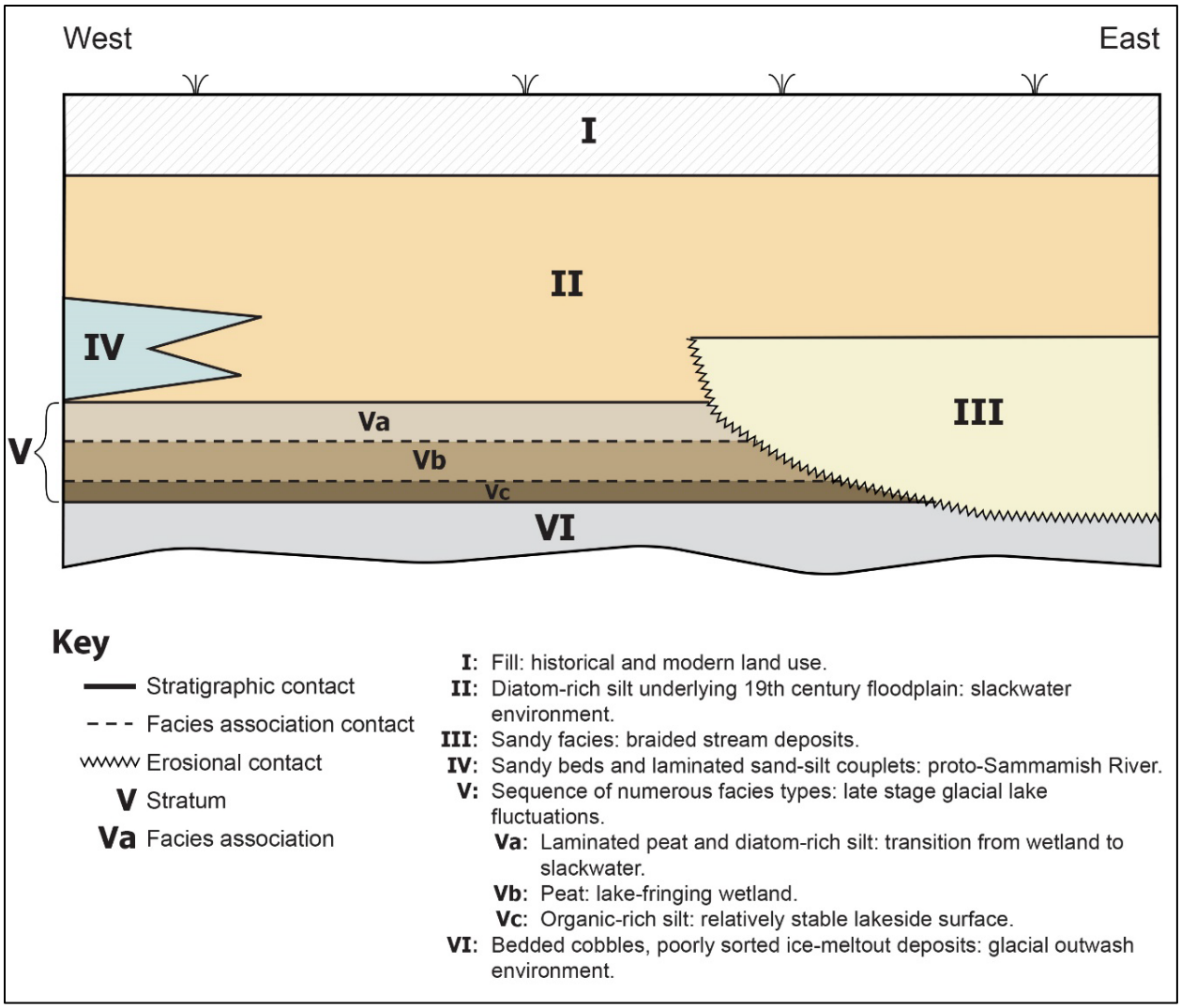

Figure 4-1. Summarized site-wide stratigraphy and relationships between major stratigraphic units. Image modified from Kopperl (2016) and used courtesy of SWCA Environmental Consultants.

Table 4-1. Summarized Site-wide Stratigraphy and Inferred Depositional Environment

Stratum	Facies Associations	Inferred Depositional Environment
I	N/A	Deposits associated with historical and modern land use.
II	N/A	A 6-foot-thick sequence of diatom-rich silts (diatomaceous earth) interpreted as later Holocene Sammamish River slackwater alluvium which underlies the 19th century floodplain.
III	a, b	Ancestral Bear Creek alluvium represented by amalgamated braided channel deposits including ripple-laminated, horizontally laminated, and trough-bedded sand with numerous scour-and-fill structures and mud drapes. The basal unconformable contact to Strata V and VI passes to a conformable upper bounding contact with the overlying Stratum II. Stratum IIIa appears to be the western leading edge of a much later channel, which is mostly east of the site.
IV	N/A	Sand beds and laminated sand-silt couplets along with discontinuous soil bodies preserved in topographic lows in the southwest portion of the site intercalated with Stratum II diatomaceous earth. This stratum is well expressed west of the site and is considered to have been deposited by the ancestral Sammamish River.
V	a	Well-laminated clay, silt, and organic matter representing the transition from the wetlands complex of Stratum Vb to the slackwater depositional system of Stratum II.
	b	A wetland complex consisting of a thick sequence of interbedded fibrous peat and sedimentary peat at the east end of the site laterally transitioning westward to thin fibrous and soil-affected peat at the west end of site; represents a relative rise in water level.
	c	A thin, artifact-bearing, upland nearshore depositional system characterized by lateral facies shifts from predominantly silty sediments at the east end of the site to very fine sandy sediments at the west end of the site. Represents a period of relatively lower water level. Conformably overlain by the Stratum Vb wetlands facies association.
VI	N/A	Well-bedded, cobble-gravel, glacial outwash at depth veneered in places with poorly sorted and fabric-supported, reworked, glacially derived diamict inferred to be ice floe melt-out.

Source: Reformatted version from Hodges et al. 2016; used with permission of SWCA Environmental Consultants.

Stratum VI, the earliest of the excavated sequence, is composed of bedded cobble and gravel glacial outwash. This stratum has been interpreted as part of the paleoshoreline of Lake Sammamish as ice melt swelled the lake at the end of the Pleistocene (Hodges et al. 2016).

Stratum V overlies Stratum VI and is a complex of three distinct facies—Vc, Vb, and Va (hereafter referred to as separate strata for brevity)—each of which consists of finer particles overall than Stratum VI. Stratum Vc consists predominantly of very dark brown silt, segueing to very fine sandy silt toward the western extent of the site. This thin layer represents an environmental transition from the outwash-derived Stratum VI to the wetland represented by Stratum Vb, and is thought to also reflect a recession in the water levels of ancient Lake Sammamish. Importantly, this stratum also contains the cultural materials thought to date to the LPH, as early occupants of the area took advantage of the surface exposed by the receding waterline of Lake Sammamish. Stratum Vb consists predominantly of reddish-brown interbedded fibrous peat and sedimentary peat, and is thought to represent a rise in local water level and the existence of a wetland at Bear Creek. Stratum Va is composed of various combinations of well-laminated clay, silt, and organic matter, and represents the transition from the wetland represented by Stratum Vb to the slackwater (lacking flow) depositional system represented by Stratum II. Stratum IV is described by Hodges and colleagues (2016:79) as "a large allostratigraphic unit composed of sediments deposited in several environments including marsh, channel, and deltaic sediments lying off-site between the modern Sammamish River channel and the western site boundary." As Stratum IV was outside the portion of the site excavated in 2013, this stratum is irrelevant to the remainder of this study. Stratum III is a predominantly sandy bedded deposit largely confined to the central portion of the site. It represents a fluvial channel—likely braided—which periodically flowed through the site. This stratum contained cultural materials which were likely in secondary context as a result of

entrainment during periods of high streamflow. Stratum II is a thick diatom-rich silt deposit representative of "accumulation of sediment during periods of slackwater flooding in either the lower reach of Bear Creek when the Sammamish River flooded, or when Sammamish floodwaters accessed the lower reach of Bear Creek" (Hodges et al. 2016:89). Importantly, Mazama ash is visible near the lower boundary of this stratum, and published ages (e.g., Egan et al. 2015) for the eruption of Mt. Mazama provide an independent age control on the transition between Stratum Va and Stratum II. Stratum I consists largely of modern fill; this stratum was removed prior to 2013 excavations, and is therefore not relevant to this study.

Sampling for luminescence dating targeted Strata Vc, III, II, and the Va/II interface. Stratum Vc, representing the earliest human occupation at the site, was targeted to provide independent dates for comparison with radiocarbon ages (provided by SWCA; see below) of the site's antiquity. The Strata Va/II interface was likewise targeted to provide an independent check on radiocarbon dates for the wetland/slackwater transition. Stratum III was targeted to facilitate a detailed examination of flood timing and stream avulsion, as any charcoal in this depositional context was likely to be detrital, and therefore of questionable value in dating deposition itself. Finally, Stratum II was targeted for two reasons: first, because it lacked material suitable for radiocarbon dating; second, to allow modeling of within-stratum accumulation rates, and therefore inference of changing flood sedimentation rates at Bear Creek.

Importantly, all Bear Creek luminescence samples were collected in collaboration with site geoarchaeologists to target strata and locations in which age mixing[1,2] was unlikely to be a factor. For example, Stratum III samples were collected from locations exhibiting fine bedding, indicating little post-depositional alteration. The sampled portion of the Stratum Va/II transition exhibited interlaminated lenses of clay, organic-rich clay, and diatomaceous earth, and thus mixing is also unlikely to be an issue in this sample. Some evidence for minor turbation due to root penetration and/or incipient pedogenesis was extant in portions of Strata Vc and II, especially at the eastern margins of the site (for Stratum V facies) and at the margins of Stratum III (for Stratum II) (Hodges et al. 2016). On the other hand, these strata also exhibited a high degree of bedding overall, and strata boundaries were strongly pronounced, suggesting any post-depositional turbation was localized rather than pervasive across any given stratum's horizontal extent. Given this, samples for Strata Vc and II were only collected where no macroscopically visible indices of such turbation were present. Hence, the potential for turbation-related age mixing sufficient to systematically pollute luminescence dates for Strata Vc and II was minimal, and for purposes of this study all samples were thus considered to be unmixed.

[1] Chemical analysis by the Washington State University Microbeam Laboratory indicates this tephra is in fact Mazama ash (Foit 2014). Micromorphological analysis of the tephra layer (Rinck 2016) identified angular silica particles as well as carbonized particles characteristic of unweathered tephra; at the western extent of the excavated site this tephra layer is massive (exhibiting no bedding) and stratigraphically bracketed by finely bedded diatomaceous silt, suggesting tephra deposition as airfall ash within a slackwater environment.

[2] A mixed-age sample contains grains of different luminescence ages as a result of post-depositional turbation.

Table 4-2 provides a summary of all samples collected from Bear Creek; Figure 4-2 shows all sample locations relative to site excavations. Eleven total samples were collected, including three samples each from Strata Vc and II, one sample from the Stratum Va/II interface, and four samples from Stratum III.

Table 4-2. Sample Names and Provenience Information

Field Sample	Provenience	Stratum	UW Sample
Bear Creek OSL1	EU109-28, South Wall	Vc (basal)	UW3102
Bear Creek OSL2	EU105-50, South Wall	III (basal)	UW3103
Bear Creek OSL3	EU95-08, South Wall	III (scour into V)	UW3104
Bear Creek OSL4	EU114-15, North Wall	III (basal)	UW3105
Bear Creek OSL5	EU80-84, North Wall	Vc (basal)	UW3106
Bear Creek OSL6	EU118-26, North Wall	III (basal)	UW3107
Bear Creek OSL7	TU-122, South Wall	Vc (basal)	UW3108
Bear Creek OSL8	Western Trench, North Wall	II (basal)	UW3109
Bear Creek OSL9	Western Trench, North Wall	II (intermediate)	UW3110
Bear Creek OSL10	Western Cut Face, Western Wall	II (intermediate)	UW3111
Bear Creek OSL11	Western Cut Face, Western Wall	II (upper)	UW3112

In sampling Stratum Vc, preference was given to sampling locations where nearby radiocarbon samples were available. One sample each was drawn from the eastern (UW3102), central (UW3106), and western (UW3108) portions of the excavated site area to help examine spatio-temporal patterning of deposition across the site as a whole. Additional samples of Strata VI and Vb were collected at each location to aid in dose rate estimation, as radiation from these strata contributed to the natural luminescence signal in Stratum Vc.

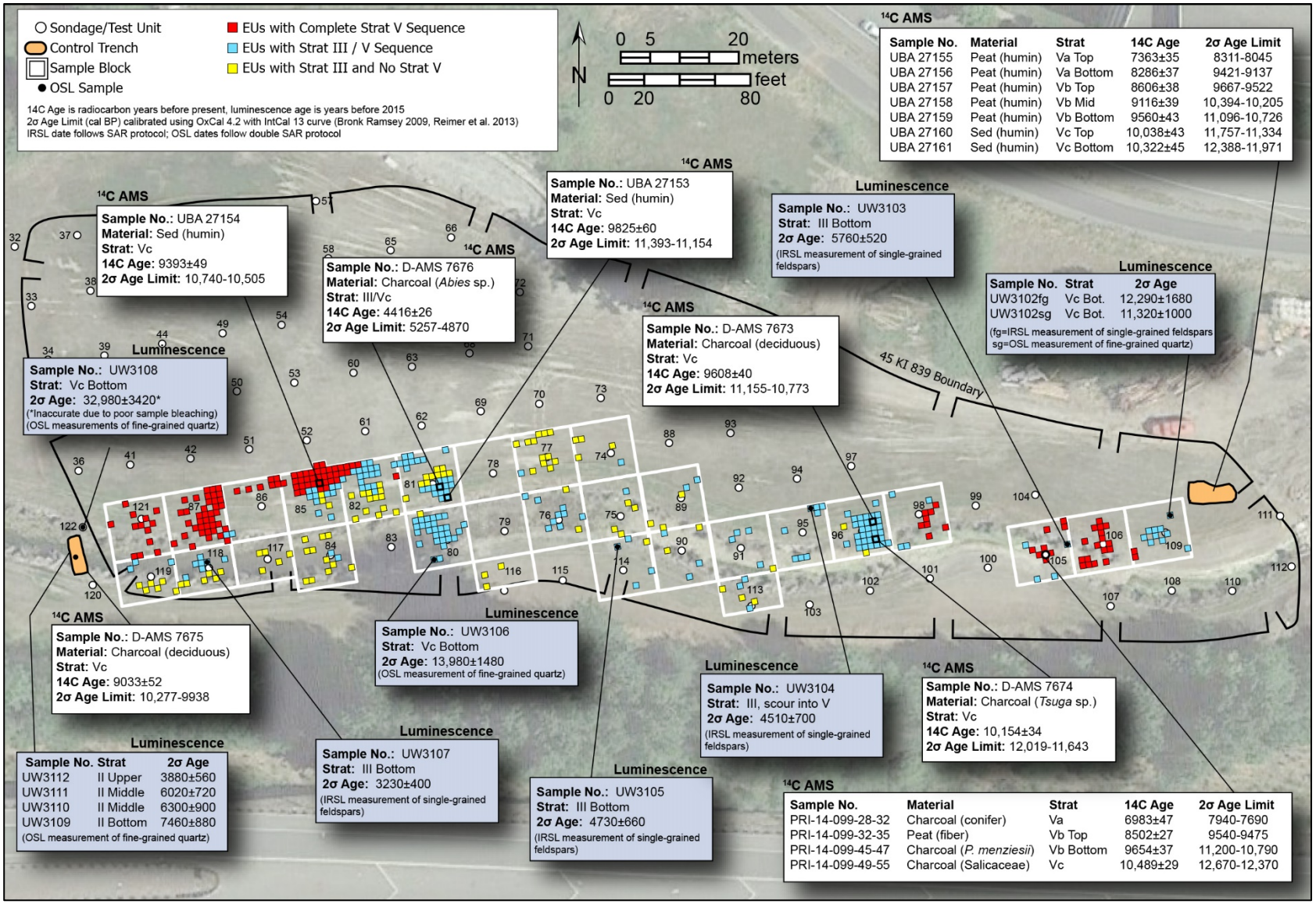

Figure 4-2. Overview of Bear Creek Site showing details for all luminescence and radiocarbon samples in relation to excavation units. Image from Kopperl 2016 and used courtesy of SWCA Environmental Consultants.

Stratum III sampling targeted small scour pockets containing finely bedded Stratum III deposits which plummeted into Stratum V; these small pockets would have been sheltered from subsequent erosion, and are thus the portions of Stratum III most likely to represent initial deposition of Stratum III sediments. As above, samples of Stratum III were selected from across the site to help evaluate time-transgressive movement of channel flows, with samples collected from the eastern (UW3103), east-central (UW3104, UW3105), and western (UW3107) portions of Stratum III. Sample UW3103 was collected from a portion of Stratum III which rests atop intact Mazama ash, providing an upper age limit of 7815 cal BP +/- 50 (Egan et al. 2015) for this sample.

Most of Stratum II was removed prior to data recovery, especially toward the eastern portion of the excavation area, and was therefore unavailable for sampling. In areas where Stratum II remained at the time of 2013 excavations, an indeterminate amount of its uppermost portion was also disturbed by hydroseeding[3] which took place prior to data recovery; this disturbed portion was not sampled for luminescence dating, as it was expected to contain minerals of mixed luminescence ages. Underlying this disturbed uppermost portion of Stratum II, a 110-cm-thick deposit of undisturbed Stratum II was accessible within and nearby the western test trench. Four samples spanning the depositional sequence of this undisturbed portion were collected to allow assessment of variation in the rate of accumulation of Stratum II over time. UW3109 represents the transition from Stratum Va to Stratum II and therefore initiation of

[3] A form of planting designed to aid in erosion control.

Stratum II deposition (the basal ~3 cm of the available 110-cm sequence), and UW3112 represents the uppermost datable deposition of Stratum II (the uppermost ~3 cm of the undisturbed deposit). UW3110 (~37–40 cm above Stratum V) and UW3111 (~74–77 cm above Stratum V) provide samples between these two endpoints. In the sampled location, Mazama ash was visible within Stratum II at 10 cm above the Stratum Va/II interface, providing an additional independent date between UW3109 and UW3110.

Samples of Strata Vc and III were collected from SWCA excavation units by creating photographic "dark room" conditions in the field by placing an opaque canopy over sample locations prior to sampling. Head lamps equipped with filtered red light were used exclusively during sampling; exposed portions of sampled profiles were removed prior to sample collection and saved for dose rate measurement. Samples were triple-bagged to preserve moisture content, then placed into light-proof containers before removal from the canopy and transport to the laboratory. Samples of Stratum II were collected as cohesive blocks of sediment, then labeled, wrapped, and sealed within watertight containers for transport to the laboratory. Placement of field dosimeters was prohibited by the development timeline for the site, which called for complete removal of all remaining site sediments within the area of 2013 data recovery immediately following the completion of excavation.

EQUIVALENT DOSE MEASUREMENT

Estimation of D_e, the numerator of the age equation, was performed on either coarse (180- to 212-µm) grains or fine (1- to 8-µm) grains, depending on sampled sediment composition. Samples of sediments in which silt was the predominant component (Strata Vc,

Va/II, and II) were initially prepared for fine-grain dating, although partial bleaching identified in Strata Vc (see below) made it necessary to also prepare a sample of Stratum Vc for dating of coarse-grains to facilitate single-grain resolution in D_e measurement as a means of statistically isolating well-bleached grains for age estimation. Samples of Stratum III were predominantly sandy alluvium, and were therefore prepared for dating of coarse grains at single-grain resolution.

All samples were treated with 20% hydrochloric acid and 30% hydrogen peroxide to remove carbonates and organics. For fine-grained samples, the desired size (1- to 8-µm) fraction was isolated using Stokes' Law to settle samples for specified intervals of time. For coarse-grained samples, nested sieves were used to isolate the desired size fraction. Prepared fine-grain aliquots contained both quartz and feldspar minerals, although an attempt was made to measure each mineral separately (see below). For coarse-grain samples, grains of potassium feldspar were separated from grains of quartz and other heavier minerals with the use of a lithium metatungstate solution with a specific gravity of 2.58 g/mL.

Measurement of coarse-grain samples used the SAR protocol (Murray and Wintle 2000) to measure single grains of feldspar (Auclair et al. 2003). Measurements were made using Risø TL/OSL DA-20 reader with an IR single-grain attachment. Stimulation used a 150 mW 830 nm IR laser set at 30% power and passed through an RG 780 filter. Luminescence emissions were collected by a photomultiplier through a blue-filter pack which allowed transmission in the 350- to 450-nm range. IRSL measurements were made at 50°C, and a preheat of 250°C for 1 minute

at 5°C/s proceeded each measurement. Exposure of single-grains was for 0.8 s, using the first 0.06 s for analysis and the last 0.15 s for background.

Feldspars were selected for single-grain measurement of coarse grains because past experience with quartz dating in the region has shown that local quartz typically has poor sensitivity for single-grain dating, with a very small proportion of grains exhibiting luminescence properties useful for dating (e.g., Feathers and Tunnicliffe 2011; Tunnicliffe et al. 2012). Local feldspars, by contrast, tend to produce a higher yield of grains amenable to single-grain dating, making feldspar measurement the preferred approach at Bear Creek.

Measured coarse-grained feldspars exhibited anomalous fading (Wintle 1973), a spontaneous depletion of natural signal common to feldspars which can lead to underestimation of sample age if uncorrected. Anomalous fading of Bear Creek feldspars was measured using procedures set forth by Auclair and colleagues (2003) on single grains. Age was corrected following Huntley and Lamothe (2001) using storage times of up to 3–5 days, a 50-sβ test dose, and 200-sβ irradiations. Table 4-3 displays details of D_e measurement for all samples.

Table 4-3. Display of Equivalent Dose Methods and Sample Size Details

UW Sample	Fraction Prepared	Fraction Dated	Measurement Type	n Measured	n Dated
UW3102	Fine-grain polymineral	Fine-grain quartz	double-SAR	52	52
	Single-grain feldspar	Single-grain feldspar	SAR (IRSL)	400	90
UW3103	Single-grain feldspar	Single-grain feldspar	SAR (IRSL)	600	37
UW3104	Single-grain feldspar	Single-grain feldspar	SAR (IRSL)	400	63
UW3105	Single-grain feldspar	Single-grain feldspar	SAR (IRSL)	600	48
UW3106	Fine-grain polymineral	Fine-grain quartz	double-SAR	48	45
UW3107	Single-grain feldspar	Single-grain feldspar	SAR (IRSL)	600	66
UW3108	Fine-grain polymineral	Fine-grain quartz	double-SAR	48	42

UW Sample	Fraction Prepared	Fraction Dated	Measurement Type	n Measured	n Dated
UW3109	Fine-grain polymineral	Fine-grain quartz	double-SAR	44	44
UW3110	Fine-grain polymineral	Fine-grain quartz	double-SAR	48	35
UW3111	Fine-grain polymineral	Fine-grain quartz	double-SAR	48	47
UW3112	Fine-grain polymineral	Fine-grain quartz	double-SAR	48	31

Measurement of fine-grain polymineral aliquots followed the double-SAR protocol (Banerjee et al. 2001; Roberts and Wintle 2001) using a Risø TL/OSL Reader Model TL-DA-15 and detection through 7.5 mm of Hoya U 340 (UV) filter, a 240°C preheat for 60 s after regeneration doses, a 200°C cut heat, a 30 sβ (~2.5 Gy) test dose, and regeneration doses of 0–300 sβ (0- ~25 Gy). Feldspar stimulation used a 45-mW/cm^2 880-nm IR diode for 100 s at 60°C, followed by quartz stimulation with a 35-mW/cm^2 470-nm blue diode for 100 s at 125°C. This protocol helped allow the quartz and feldspar fractions of each aliquot to be stimulated and measured separately, facilitating isolation of quartz for dating purposes. This is advantageous because quartz typically bleaches more rapidly than feldspar (Godfrey-Smith et al. 1988), making it more likely to provide an accurate luminescence age in case sediments were partially bleached at time of deposition. Because each fine-grained aliquot contains numerous quartz grains, the relatively poor luminescence properties of local quartz are not detrimental to measurement of D_e for aliquots of this type.

Because IR stimulation may not completely remove the trapped feldspar signal, fading tests were performed for fine-grained aliquots using the same instrument setup as quartz D_e measurement, with a 50-sβ (~4.17-Gy) test dose and a 200-sβ (~16.67-Gy) regeneration dose; no

evidence of fading was detected in the blue-stimulated signal, indicating that IR stimulation successfully negated the influence of feldspathic grains susceptible to fading. This, in turn, indicates that the protocols used for dating fine-grained aliquots successfully isolated the sample fraction dominated by quartz—and therefore the fraction most likely to be well bleached—for use in D_e estimation. Again, Table 4-3 displays details of D_e measurement for all samples.

As a check on the validity of protocols used for D_e measurement, dose recovery testing was performed on three aliquots of each fine-grained sample and one disc (100 aliquots) for each coarse-grain sample to evaluate whether these protocols successfully "predicted" a known administered dose of 200 sβ (~16.67 Gy). The administered dose was successfully recovered for both the quartz (mean recovered dose of 202.68 +/- 12.99 sβ) and the feldspar fraction (mean recovered dose of 204.23 +/- 7.60 sβ) of all tested samples, indicating D_e was accurately estimated with the protocols used for measurement.

DOSE RATE MEASUREMENT

A portion of each dated sample and each environmental reference sample was pulverized to the consistency of a fine flour to facilitate alpha counting, and each sample underwent a battery of measurements to produce estimates of sample dose rate, the denominator of the age equation.

Rates of terrestrial alpha radiation were estimated for each sample by using plexiglass containers with ZnS:Ag screens and a Daybreak alpha counter to count total alpha emissions as well as the fast and slow pairs over a period of 96 hours. The pairs are used to distinguish the proportion of the total counts coming from either the U or Th decay chains. All samples were

sealed for at least one month before alpha counting. Terrestrial beta radiation rates were measured on four aliquots of each sample using a Riso GM-25-5 Beta Multicounter System and a counting time of at least 24 hours per sample. Average sample beta count rate was converted to beta dose rate following Bøtter-Jensen and Mejdahl (1988). ^{40}K, a radioactive species of potassium, is an important contributor to the accumulation of natural signal (Aitken 1985), and sample potassium content was therefore directly measured using a Bruker Tracer III-V portable X-ray fluorescence (XRF) spectrometer and strict parameters and protocols developed for this application and extensively detailed elsewhere (Johnson 2014); ^{40}K was estimated from total K content using rates of natural abundance. Rates of terrestrial gamma radiation were derived from data produced by alpha counting and potassium measurement. Cosmic radiation was estimated using sample provenience and burial depth data provided by SWCA in combination with global values for cosmic radiation provided by Prescott and Hutton (1994). Sample water content, which can affect sample D_r, was measured directly for all samples by calculating water loss as a product of drying in a 50°C oven for at least 48 hours. Alpha efficiency in fine-grained samples was calculated through measurement of b-values (Aitken 1985) on three aliquots of each fine-grained sample using the same equipment and protocol used in D_e measurement. Results of b-value estimation (Table 4-4) exhibit values in range expected for quartz aliquots, indicating that quartz was likely successfully isolated by the protocols used. Radioactivity concentrations were translated into dose rates following Guérin and colleagues (2011). Because many samples derive from dose environments in which multiple strata were affecting sample D_r—for example, samples of Stratum Vc were typically within the effective range of radiation produced by

Stratum VI and Stratum Vb sediments—D_r estimates used for dating made use of fractional dose calculations set forth by Aitken (1985:Appendix H) to account for all in situ radiation relevant to sample luminescence. Final D_r estimates are set forth in Table 4-4 for all samples.

Table 4-4. Dose Rate Data for all Samples

UW Sample	B-value	(1σ)	α-Dose (Gy/ky)	(1σ)	β-Dose (Gy/ky)	(1σ)	γ-Dose (Gy/ky)	(1σ)	Cosmic Dose (Gy/ky)	(1σ)	Total Dose (Gy/ky)	(1σ)
UW3102fg	0.517	0.036	0.0010	0.0005	0.5770	0.0413	0.3557	0.0213	0.1401	0.0288	1.0738	0.0547
UW3102sg	N/A	N/A	0.0958	0.0579	1.2002	0.1340	0.3054	0.0200	0.1268	0.0261	1.7282	0.1497
UW3103	N/A	N/A	0.0958	0.0579	1.3093	0.1376	0.3643	0.0292	0.1268	0.0261	1.8962	0.1543
UW3104	N/A	N/A	0.0958	0.0579	1.2214	0.1325	0.3387	0.0261	0.1401	0.0288	1.7960	0.1497
UW3105	N/A	N/A	0.0958	0.0579	1.2231	0.1325	0.3529	0.0265	0.1401	0.0288	1.8119	0.1498
UW3106	0.730	0.102	0.0073	0.0032	0.5475	0.0480	0.3567	0.0324	0.1268	0.0261	1.0396	0.0636
UW3107	N/A	N/A	0.0958	0.0579	1.2164	0.1332	0.3863	0.0320	0.1401	0.0288	1.0383	0.0636
UW3108	0.790	0.091	0.0079	0.0034	0.7844	0.0556	0.5357	0.0317	0.1268	0.0261	1.4548	0.0692
UW3109	0.852	0.105	0.0085	0.0037	0.6672	0.0488	0.4977	0.0456	0.1268	0.0261	1.3003	0.0717
UW3110	0.863	0.063	0.0087	0.0036	0.5644	0.0489	0.5247	0.0399	0.1401	0.0288	1.2378	0.0695
UW3111	0.859	0.112	0.0086	0.0038	0.6436	0.0528	0.5359	0.0481	0.1401	0.0288	1.3283	0.0771
UW3112	0.827	0.058	0.0083	0.0035	0.5104	0.0479	0.3878	0.0393	0.1401	0.0288	1.0466	0.0685

Note: “UW3102fg” refers to the portion of the sample used for dating fine-grained aliquots, while “UW3102sg” refers to the portion of this sample used for SAR dating of single-grains of this sample.

ASSESSMENT OF SAMPLE BLEACHING

Statistical isolation of a well-bleached and unmixed portion of samples is essential for accurate age estimation. As discussed above, all sample collection targeted locations where stratigraphic evidence strongly suggested post-depositional turbation was negligible, and it is therefore highly unlikely that mixing was sufficient to pollute luminescence dates for any of the

collected samples[4]. Mixing was therefore considered to be a nonfactor for these samples. Partial bleaching, on the other hand, is a relatively common issue for sediments deposited in sub-aqueous environments with fluvial inputs (e.g., Jain et al. 2004; Olley et al. 1998, 2004; Stokes et al. 2001), as at Bear Creek. Given this, partial bleaching was a potential confounding factor in dating the Bear Creek samples.

Addressing partial bleaching involves diagnosis of sample bleaching characteristics followed by the application of corrective statistical formulae such as the "minimum age model" (MAM) (Galbraith 2005; Galbraith et al. 1999; Galbraith and Roberts 2012) in cases where distributions of sample aliquot D_e or age values exhibit evidence of partial bleaching in the form of overdispersion[5] (OD) in the "central age model," (CAM) high outliers, or in some cases pronounced differences between D_e distributions of distinct mineral fractions[6] (Murray et al.

[4] This assumption is empirically supported by the fact that application of the finite mixture model (FMM) (see Galbraith and Roberts 2012) to D_e distributions (for quartz fine-grained aliquots) and age distributions (for single-grain feldspar aliquots) exhibited single-aged distributions for samples UW3102 through UW3107, as well as UW3110 and UW3111, suggesting these samples do not contain aliquots of mixed ages. Further, UW3108, UW3109, and UW3112, which exhibited two single-aged populations in the FMM, yielded FMM (young component) ages component which were not statistically distinct at 95% confidence from ages predicted by the minimum age model (MAM; see below). Thus mixing is unlikely in these samples, and in the three cases where scant evidence for mixing is extant the MAM produces a valid result, rendering the FMM irrelevant to the Bear Creek samples.

[5] Overdispersion (OD) is the variance parameter in the central age model (Galbraith and Roberts 2012). Dose recovery testing of fine-grained aliquots yielded an average OD value of 6.2% for single-aged samples; fine-grained D_e distributions higher than double this value (12.4%) were considered partially bleached for the Bear Creek samples. Dose recovery testing yielded an average OD value of 13.1% for single-grain samples. To account for additional OD due to possible heterogeneity in paleodose, single-grain Bear Creek samples exhibiting 20% OD or higher were considered partially bleached. This is a conservative approach which avoids the assumption that dose recovery testing incorporates all possible sources of variance in samples (see Galbraith and Roberts 2012).

[6] Quartz typically bleaches more rapidly than feldspars (Godfrey-Smith et al. 1988), so samples exhibiting markedly higher D_e values for feldspars than those for quartz are more likely to have been poorly bleached overall.

2012). In general, Bear Creek samples exhibited varying degrees of partial bleaching depending on the stratum from which the samples were drawn. Each stratum is discussed separately below.

Stratum II and the Va/II Interface: UW3109, UW3110, UW3111, UW3112

Dated fine-grain samples from Stratum II and the Stratum Va/II interface generally show some evidence of partial bleaching. OD values (Table 4-5) are more than twice those produced by dose recovery testing, suggesting sample variance is not reflective of single-aged samples. Radial plots of samples UW3109, UW3110, UW3111, and UW3102 (Figures 4-3 to 4-6, respectively) depict distributions of aliquot D_e values which exhibit moderate scatter in the quartz D_e fraction, including a tendency toward high outliers; these attributes support the interpretation that insufficiently bleached aliquots are present. This is particularly true for UW3109 and UW3112. On the other hand, quartz D_e values are distinctly higher than feldspar D_e values for these samples[7], which likely indicates that many aliquots are likely to have been sufficiently bleached for dating of the quartz fraction to proceed, provided the MAM is applied.

Table 4-5. Luminescence Ages of all Samples

Stratum	UW Sample	Central Age (ky*)	1σ	%OD	Partial Bleaching?	Minimum Age (ky*)	1σ
V	UW3102fg	16.87	0.88	15.7	Yes	12.29	0.84
	UW3102sg	11.32	0.50	18.3	No	N/A	N/A
	UW3106	19.78	0.94	12.5	Yes	13.98	0.74
	UW3108	36.85	1.79	13.3	Yes	32.98	1.71

[7] D_e of feldspars diminishes over time following a power law functional relationship as a result of anomalous fading, whereas quartz D_e typically remains stable. Hence fully bleached samples should exhibit lower De values for the feldspar fraction than those for the quartz fraction.

Stratum	UW Sample	Central Age (ky*)	1σ	%OD	Partial Bleaching?	Minimum Age (ky*)	1σ
III	UW3103	5.76	0.26	12.9	No	N/A	N/A
	UW3104	4.51	0.35	19.7	No	N/A	N/A
	UW3105	4.73	0.33	17.1	No	N/A	N/A
	UW3107	3.23	0.20	16.6	No	N/A	N/A
II	UW3109	10.63	0.60	28.7	Yes	7.46	0.44
	UW3110	9.62	0.54	14.4	No	6.30	0.45
	UW3111	6.90	0.39	14.9	No	6.02	0.36
	UW3112	4.38	0.60	19.6	Yes	3.88	0.28

Note: Ages are expressed in thousands of years before the year 2015. Errors are expressed in thousands of years and represent 1σ values. %OD represents dated fraction overdispersion in the central age model. Shaded values represent best estimates of luminescence age, as determined by diagnosis of sample bleaching. “UW3102fg” refers to the portion of the sample used for dating fine-grained aliquots, while “UW3102sg refers to the portion of this sample used for SAR dating of single-grains of this sample.

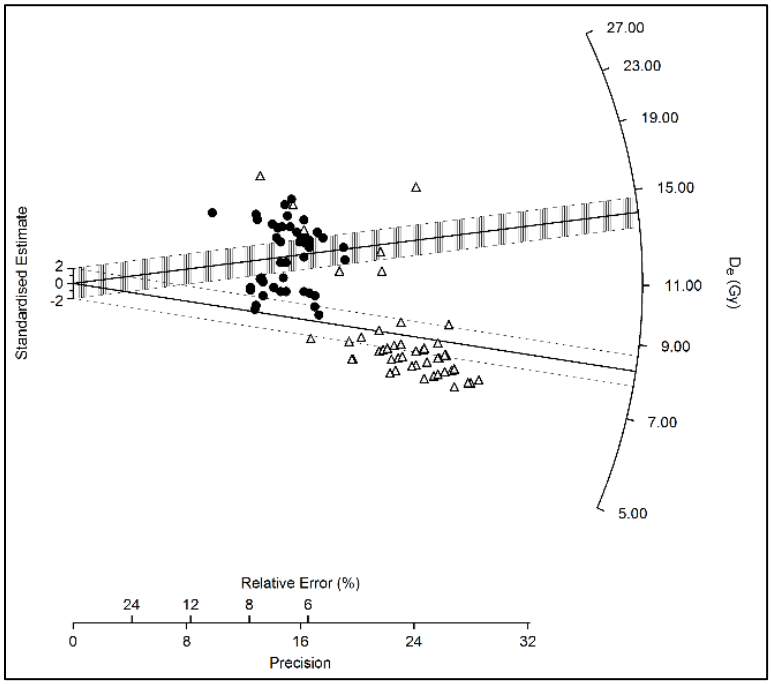

Figure 4-3. Radial plot for Sample UW3109’s double-SAR equivalent dose data. Black dots represent quartz value, and white triangles represent uncorrected feldspar values. The shaded and unshaded bands are centered on the mean quartz and feldspar D_e values, respectively, as derived from the “central age model” and represent 2σ errors about the mean.

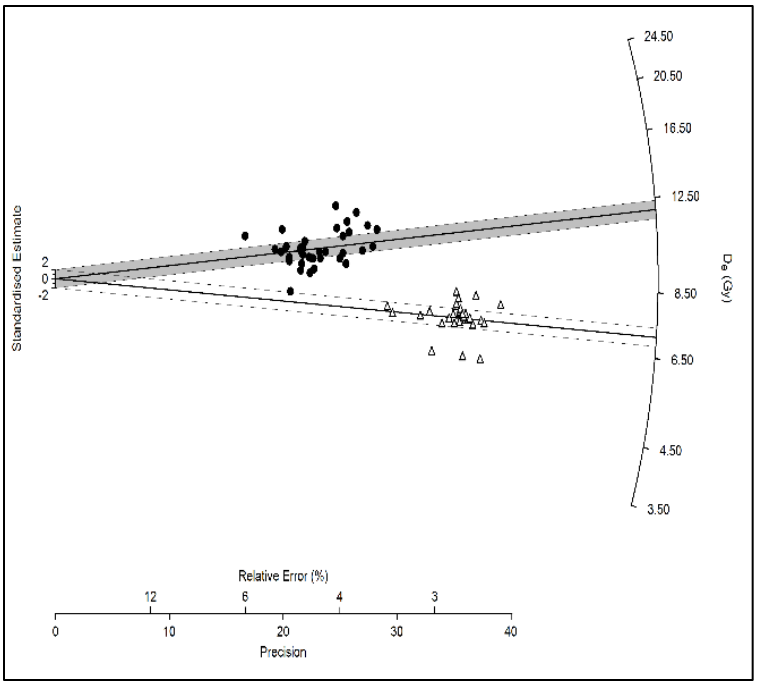

Figure 4-4. Radial plot for Sample UW3110's double-SAR equivalent dose data. See Figure 4-3 for description of symbols.

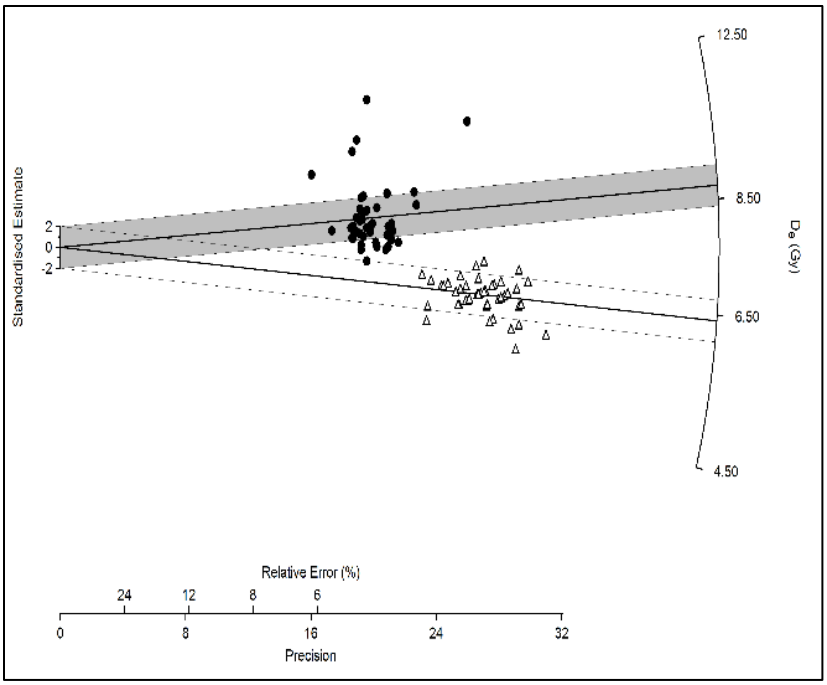

Figure 4-5. Radial plot for Sample UW3111's double-SAR equivalent dose data. See Figure 4-3 for description of symbols.

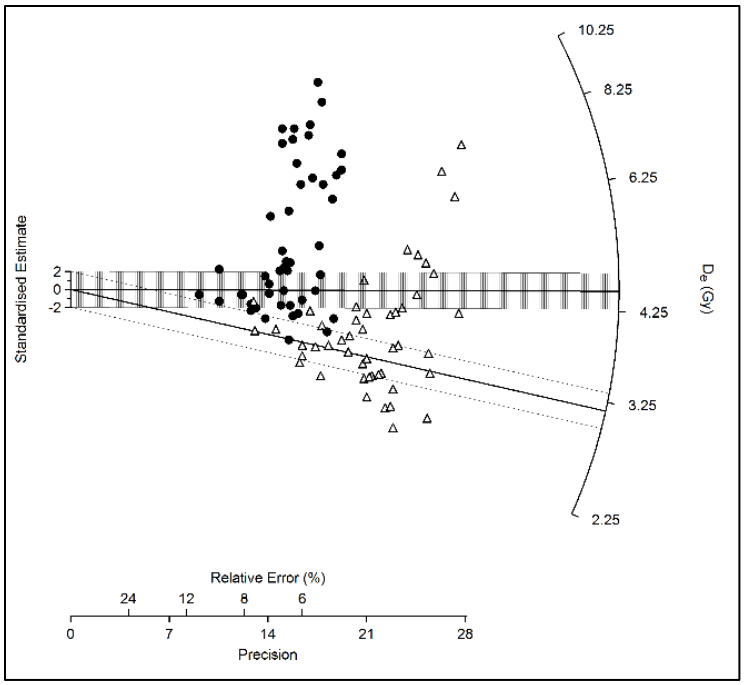

Figure 4-6. Radial plot for Sample UW3112's double-SAR equivalent dose data. See Figure 4-3 for description of symbols.

Stratum III: UW3103, UW3104, UW3105, UW3107

Dated coarse-grain feldspar samples from Stratum III show no evidence of partial bleaching. OD values are below 20%, suggesting samples are not likely to include significant numbers of poorly bleached grains. Radial plots of single-grain ages of samples UW3103, UW3104, UW3105, and UW3107 (Figures 4-7 to 4-10, respectively) exhibit low scatter, strong convergence within 2σ of the "central age" mean, and few outliers. When coupled with the fact that these sediments were finely bedded in situ, indicating they were deposited relatively gradually and gently, it is probable that all three Stratum III samples were well bleached in antiquity.

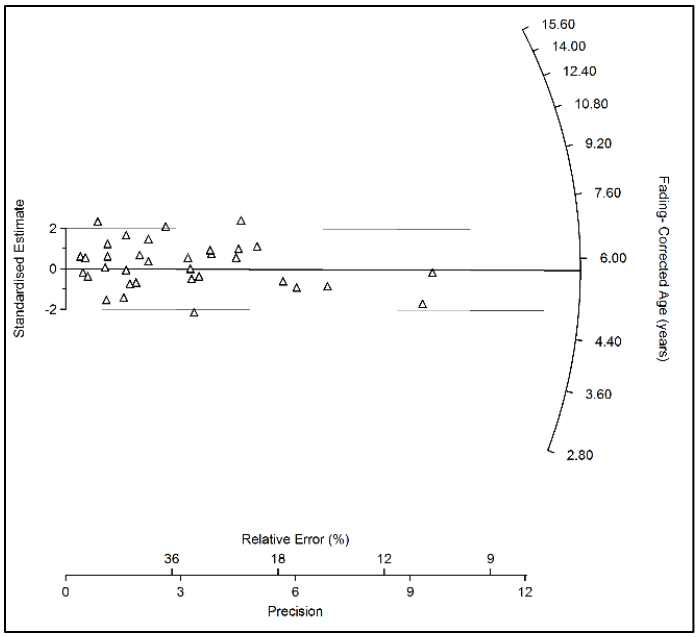

Figure 4-7. Radial plot for Sample UW3103's fading-corrected single-grain ages. Each white triangle represents a single feldspar aliquot. The unshaded band is centered on the mean age derived from the "central age model" and represents 2σ errors about the mean.

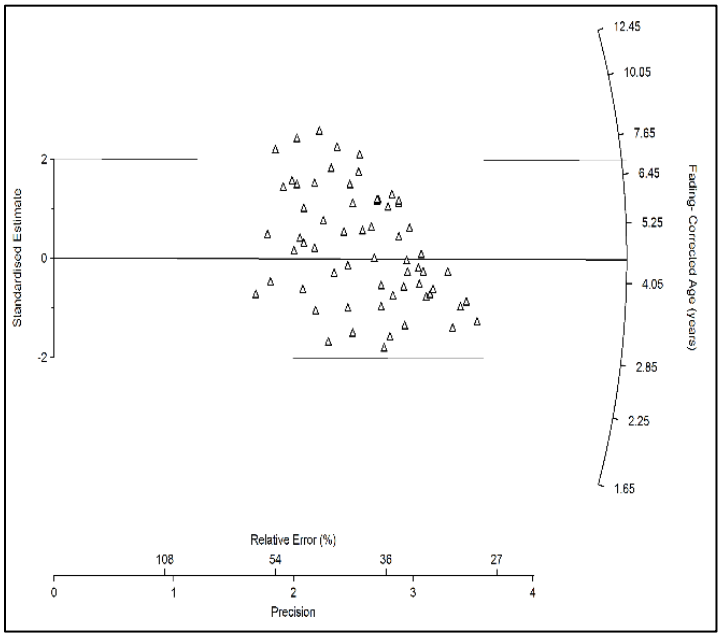

Figure 4-8. Radial plot for Sample UW3104's double-SAR equivalent dose data. See Figure 4-7 for description of symbols.

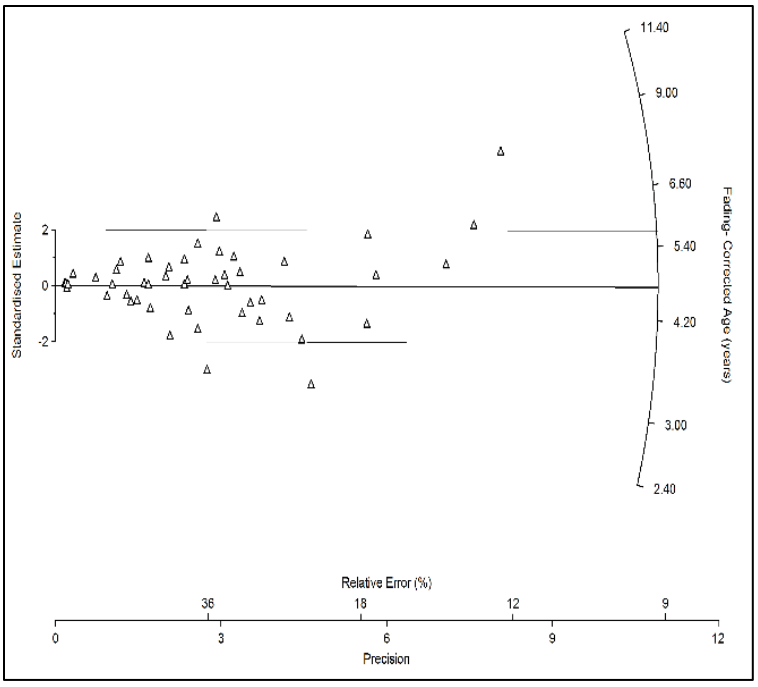

Figure 4-9. Radial plot for Sample UW3105's double-SAR equivalent dose data. See Figure 4-7 for description of symbols.

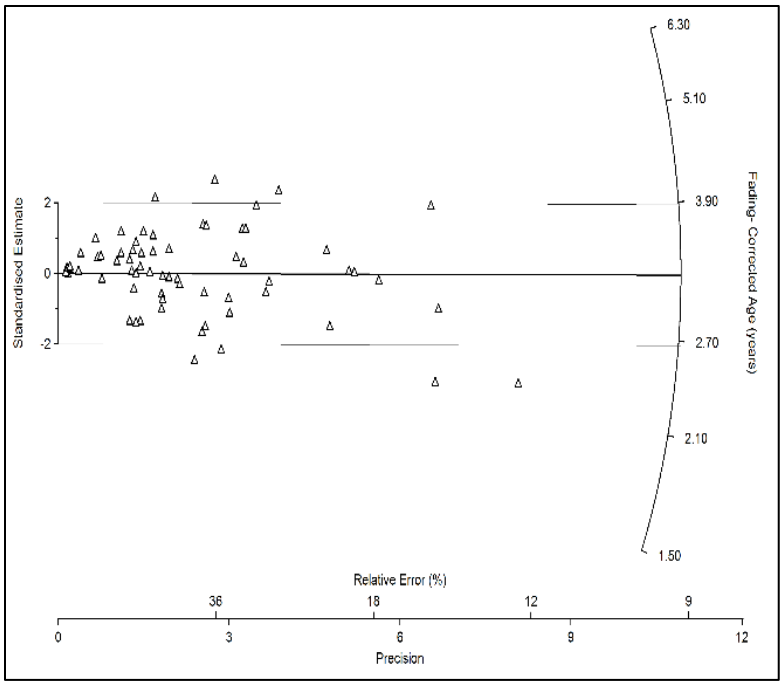

Figure 4-10. Radial plot for Sample UW3107's double-SAR equivalent dose data. See Figure 4-7 for description of symbols.

Stratum Vc: UW3102, UW3106, UW3108

Dated fine-grain samples from Stratum Vc exhibit partial bleaching. OD values are more than double those observed during dose recovery testing, and radial plots of samples UW3102, UW3106, and UW3108 (Figures 4-11 to 4-13, respectively) include distributions of quartz and feldspar D_e values with high scatter (UW3108) and elevated quartz outliers (UW3102). Further, feldspar D_e values are consistently much higher than quartz D_e values, suggesting ancient exposure was not sufficient to bleach the feldspar fraction of the samples. On the other hand, the quartz fraction may nonetheless have been sufficiently bleached for luminescence dating of UW3102 and UW3106 using the MAM. Given the immense scatter of UW3108 and the troubling heteroscedasticity of its D_e distributions, however, it is unlikely that this sample was sufficiently bleached for luminescence dating, and results of this date should therefore be interpreted with caution.

Given both 1) the relevance of Stratum Vc samples to the establishment of the antiquity of the Bear Creek Site, and 2) the specter of partial-bleaching in the dating of the fine-grained component of this stratum, single-grain dating of UW3102 was also performed as a means of further evaluating sample bleaching. Single-grain resolution is appropriate for this objective because it removes averaging between grains which may obscure heterogeneity in bleaching and/or impede isolation of well-bleached individual aliquots. As with Stratum III samples, feldspars were used for single-grain dating due to issues with quartz sensitivity in the region of interest; coarse-grained feldspars were extracted from the bulk sample as described above.

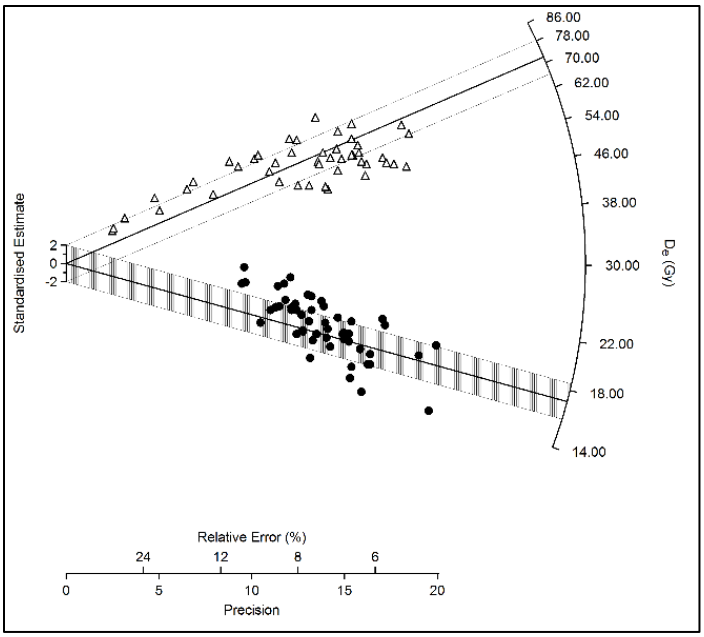

Figure 4-11. Radial plots of Sample UW3102's double-SAR equivalent dose data. Black dots represent quartz values, and white triangles represent uncorrected feldspar values. The shaded and unshaded bands are centered on the mean quartz and feldspar D_e values, respectively, as derived from the "central age model" and represent 2σ errors about the mean.

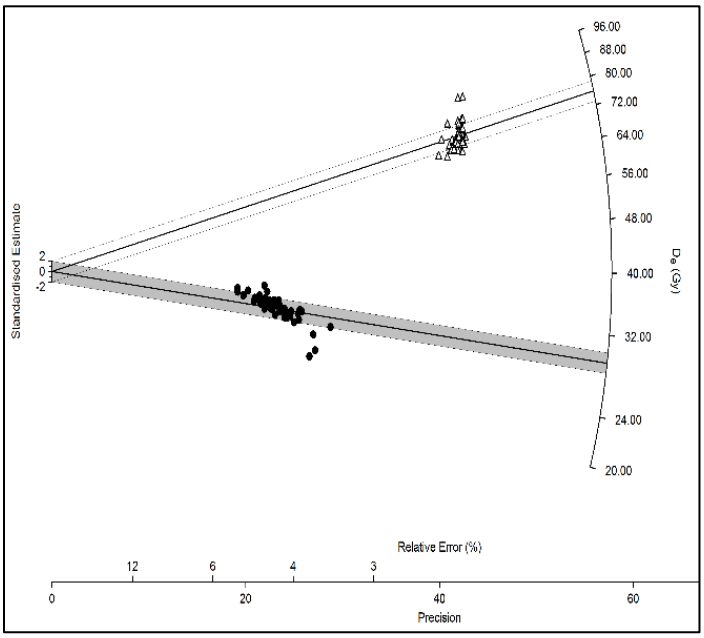

Figure 4-12. Radial plots of Sample UW3106's double-SAR equivalent dose data. See Figure 4-11 for description of symbols.

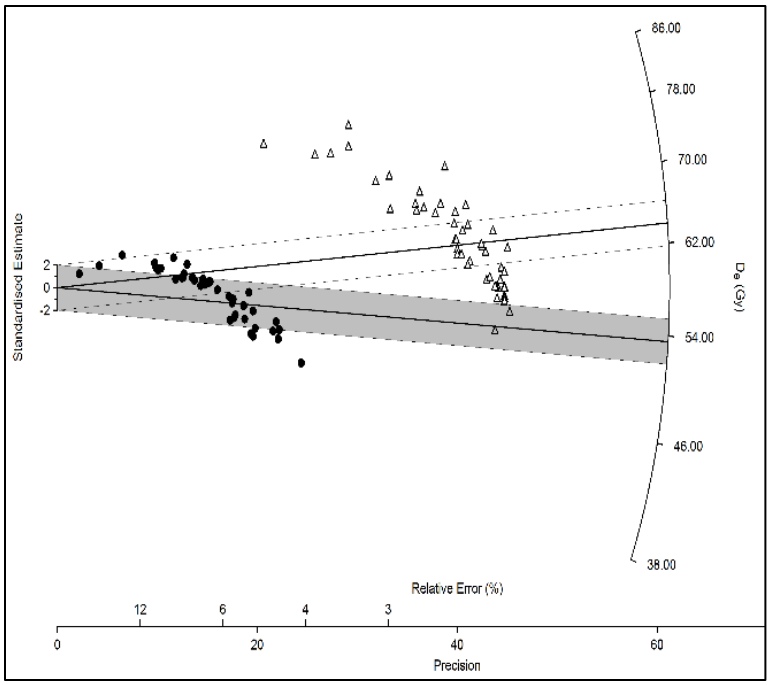

Figure 4-13. Radial plots of Sample UW3106's double-SAR equivalent dose data. See Figure 4-11 for description of symbols.

Figure 4-14 gives a radial plot of single-grain ages from measured aliquots of UW3102, and it shows that coarse grains of UW3102 were sufficiently bleached for dating purposes; few unbleached grains exhibiting elevated ages are present, OD is below the 20% threshold, and the finite mixture model (FMM) indicated the presence of only one single-aged population. A few low outliers exist, perhaps indicating the presence of a few later, intrusive grains, but these outliers had almost no statistical effect on estimated age given the large sample size and the relatively strong modal convergence. As a result, the single-grain portion of UW3102 supports the accuracy of ages produced by the application of the minimum age model to fine-grain data from UW3012 and UW3106. Additionally, the single-grain portion of UW3102 provides an

extra “semi-independent”[8] date for Stratum Vc for use in Bayesian modeling and in interpreting site chronology.

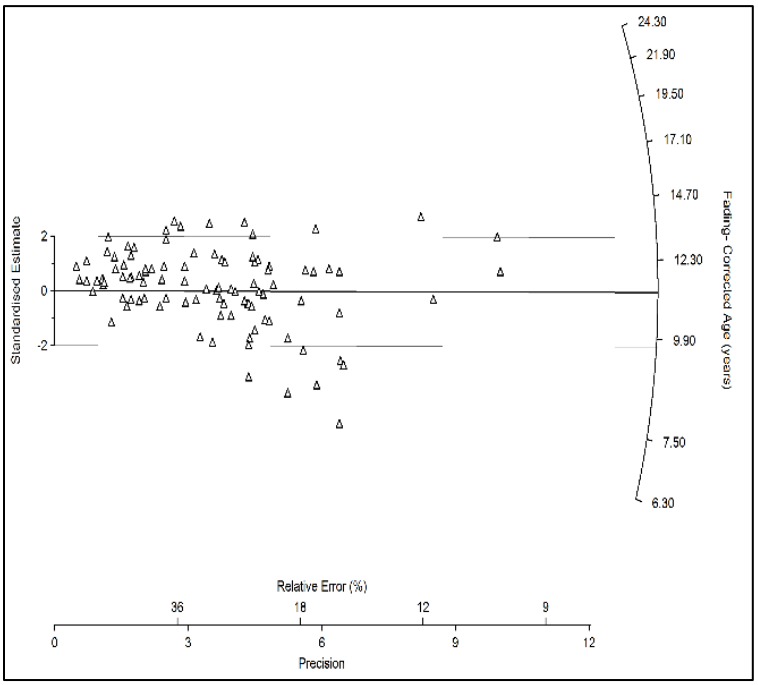

Figure 4-14. Radial plot of Sample UW3102 fading-corrected single-grain ages. Each white triangle represents a single feldspar aliquot. The unshaded band is centered on the mean age derived from the “central age model” and represents 2σ errors about the mean.

[8] The single-grain date for UW3102 used the same basic radiation measurements used to produce the fine-grain date for UW3102, so the two ages for UW3102 were produced with independent D_e data but the same baseline D_r data (adjusted for differences between coarse and fine grains in the contribution of the alpha radiation component).

LUMINESCENCE AGE ESTIMATES

Table 4-5 displayed luminescence ages calculated for all samples. Values in this table are expressed in thousands of years before 2015, and error terms are given at 1σ. Age estimates for sample UW3102 (single-grain) and the Stratum III samples (UW3103, UW3104, UW3105, and UW3107) were derived from the CAM, as these samples exhibited no evidence of partial bleaching. The other samples showed evidence of partial bleaching, so the MAM was used to estimate their luminescence ages; for these samples a hypothetical CAM date is provided to give a reference as to the degree to which the use of the MAM has affected estimated sample age.

These luminescence ages are generally stratigraphically compatible both internally and with available independent dates, including both radiocarbon dates and the presence of Mazama ash, suggesting the application of luminescence dating at Bear Creek was successful overall for all samples but UW3108, which exhibited poor bleaching and produced a minimum age of ~33,000 cal BP. With the exception of this sample, the MAM (when necessary) successfully isolated quartz aliquots which were sufficiently bleached to produce accurate depositional ages; if the CAM had been used to date poorly bleached samples, resultant dates would not be stratigraphically compatible with the larger body of dates at Bear Creek. As a whole, then, these luminescence dates should be considered accurate, with the exception of UW3108, which is too poorly bleached to yield an accurate age.

BAYESIAN DEPOSITIONAL MODELING

Bayesian analysis of dates for the Bear Creek depositional sequence, including all radiocarbon dates furnished by SWCA and the known age of Mazama ash, was undertaken using the "Sequence" program available online through OxCal 4.2 (Bronk Ramsey 2009). This type of analysis makes use of ordinal stratigraphic data as a statistical constraint (see Bronk Ramsey 2008) to facilitate 1) more rigorous empirical evaluation of the stratigraphic compatibility of a sequence of many dates (where an overall "A value" of less than 60 indicates stratigraphic incompatibility), 2) identification of individual dates which are incompatible with the overall sequence (having an individual "A value" of less than 60), and 3) (provided dates were found to be stratigraphically compatible by the model) refinement of the precision of individual dates within the sequence. "Sigma Boundaries" were used as upper and lower constraints on depositional sequences to allow results to be governed by actual measured error terms. As an initial step, dates for each stratum—and for Stratum V, each facies—were modeled separately to enable a narrow, detailed evaluation of statistical compatibility within strata. Dates found to be compatible were then incorporated into a "master" sequence useful for guiding interpretation of site chronology as a whole. Each step of this analysis is detailed below.

Modeling Stratum II

Table 4-6 displays samples (in stratigraphic order) used for OxCal modeling of Stratum II, including the Stratum Va/II transition. The uppermost radiocarbon date for Stratum Va was also included to help constrain the initiation of Stratum II deposition, and to provide an initial check on whether dates are compatible between strata. Model input code is provided by Code 1

(see OxCal 4.2 Code section). Radiocarbon dates were entered into the model as radiocarbon years BP using the "R_Dates" function to allow calibration within the model (using default OxCal 4.2 settings for calibration), while dates for Mazama ash and luminescence ages were entered into the model as years BC using the "C_Dates" function, as these two types of dates are already calibrated into calendar years, and model software requires use of the BC/AD system[9].

Table 4-6. OxCal Bayesian Model of Stratum II Including all Available Dates

Strat	Phase	Sample	Type	Location	mE	^{14}C Age	SD	Cal BC	(2σ)	Model Cal BC	(2σ)	A Score
II	Top II	UW3112	Fine-Grain Quartz	W. Cont. Trench	5	N/A	N/A	1331	720	1435	360	97.7
	69th %ile II	UW3111	Fine-Grain Quartz	W. Cont. Trench	5	N/A	N/A	4006	720	3904	317	104.3
	35th %ile II	UW3110	Fine-Grain Quartz	W. Cont. Trench	5	N/A	N/A	4284	899	4490	363	105.1
M	9th %ile II	Mazama Ash	Multiple; see Egan et al. 2015 for summary	W. Cont. Trench, E. Cont. Trench, EU 105-50	5, 175, 200	N/A	N/A	5687	50	5685	25	100
Va/II	Top Va/ Bottom II	UW3109	Fine-Grain Quartz	W. Cont. Trench	5	N/A	N/A	5447	870	5923	130	85.1
Va	Top Va	UBA 27155	Peat (Humin)	E. Cont. Trench	200	7363	35	6229	133	6230	66	100.2
												A_O = 96.9

Note: See Code 1 below for OxCal code used to produce this result, and Figure 4-15 for a display of arrayed dates resulting from the model. Sample depth information within this sequence was available, allowing modeling of accumulation rates for Stratum II (Figure 4-16). "Phase" indicates discrete depositional units entered into OxCal; phase designations were determined by sample substratum provenience data. Sample location and meters east of the westernmost extent of excavation blocks 120 and 122 are included for reference. "^{14}C" Age provides sample age in uncalibrated radiocarbon years with a standard deviation of SD. "Cal BC" provides calendar age with 2σ errors; BC/AD ages are used to maintain consistency with model input code, which uses the BC/AD system. "Model Cal BC" provides sample age in modeled calendar years, with 2σ errors. "A Score" provides the statistical compatibility of sample age with the overall modeled sequence; a score of >60 is considered compatible. "A_O" provides the overall A score for the modeled sequence; a score of >60 indicates a sequence which is stratigraphically compatible.

[9] The BC/AD system is used to report Bayesian results of this study for this reason, as it allows the reader to cross-examine input code and results without the need to convert between distinct timescales.

Results of OxCal modeling of Stratum II are indicated in the right-hand shaded columns of Table 4-6, and are illustrated by Figure 4-15. All Stratum II dates are stratigraphically compatible, the model produces a valid result ($A_O = 96.9$), and some refinement of the precision of the luminescence dates is evident, especially for UW3109. Additionally, because Stratum II samples were collected from a single exposure and relative depths within this exposure were recorded, a graphical display of modeled dates relative to depth (Figure 4-16) allows interpretation of changes in Stratum II accumulation rates over time. Because the model was successful and all dates are statistically compatible, all Stratum II dates were entered into the master sequence (below) without alteration.

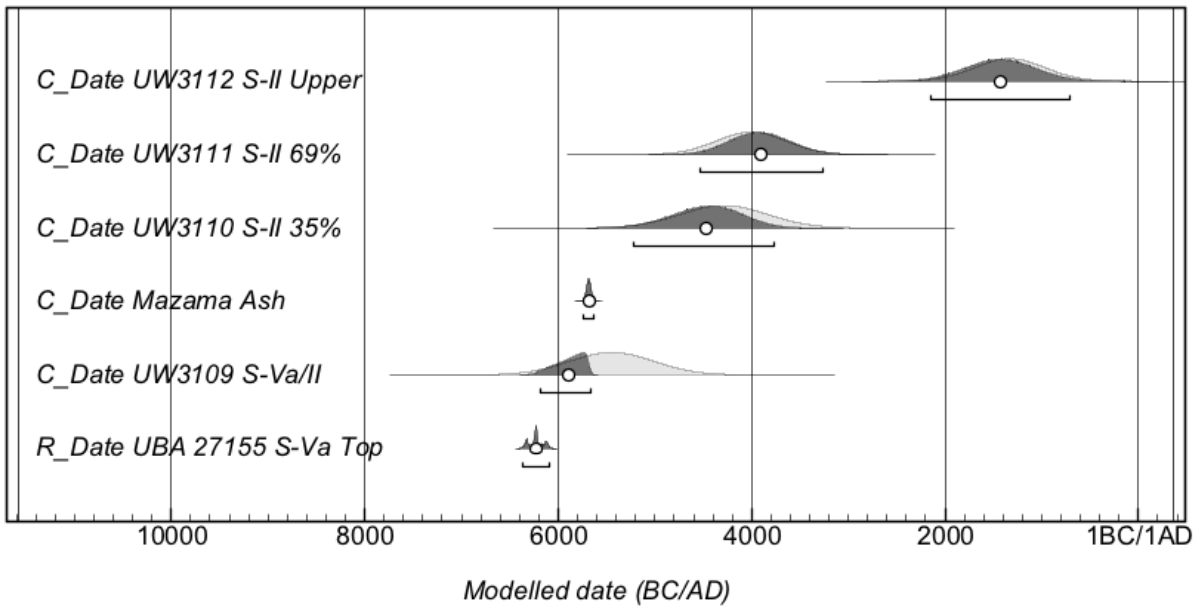

Figure 4-15. Results of OxCal Bayesian modeling of Stratum II samples. Light shaded areas represent age ranges at 2σ prior to modeling; dark shaded areas represent modeled 2σ age ranges. See Code 1 below for OxCal input code and Table 4-6 for a display of numerical results.

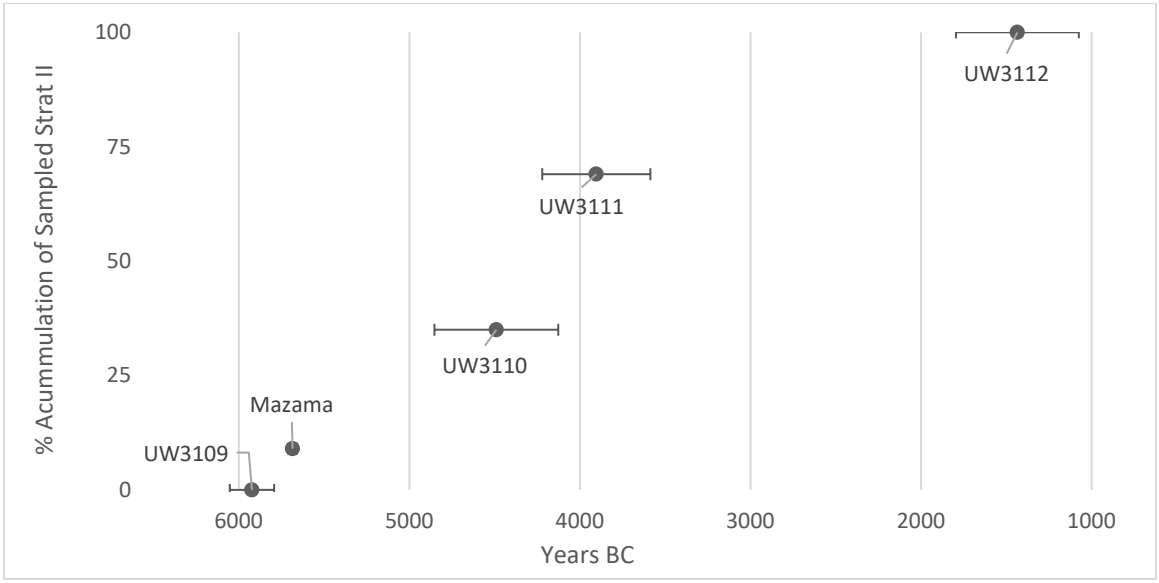

Figure 4-16. Depositional model of Stratum II displaying sample depth versus Bayesian modeled sample age. Error bars show 2σ age ranges. Accumulation of Stratum II appears to have been rapid between about 4500 BC (UW3110) and 3900 BC (UW3111).

Modeling Stratum III

Table 4-7 displays samples used for OxCal modeling of Stratum III. Here, the age of the Mazama ash, which occurs immediately below the sampled location of UW3103 in Excavatino Unit (EU) 105-50, was used as a lower constraint. Other model parameters (see Code 2) were identical to those described for Stratum II, with the exception that samples from Stratum III were entered in order of their measured age, as no superposition existed between sampled Stratum III locations. One radiocarbon sample collected from the contact between Strata III and V was included in this analysis.

Table 4-7. OxCal Bayesian Model of Stratum III Including all Available Dates

Strat	Phase	Sample	Type	Location	mE	^{14}C Age	SD	Cal BC	2σ	Model Cal BC	2σ	A Score
III	Bottom	UW3107	Single-Grain Feldspar	EU 118-26	36	N/A	N/A	1214	400	1245	198	99.8
	Bottom	UW3104	Single-Grain Feldspar	EU 95-08	138	N/A	N/A	2494	700	2340	267	107.2
	Bottom	UW3105	Single-Grain Feldspar	EU 114-15	104	N/A	N/A	2714	660	2693	194	119.6
	Bottom	D-AMS 7676	Charcoal (*Abies* sp.)	EU 81-67	77	4416	26	3115	194	3116	97	100.6
	Bottom	UW3103	Single-Grain Feldspar	EU 105-50	175	N/A	N/A	3747	520	3737	319	103.6
M	Mazama	Mazama Ash	Multiple; see Egan et al. 2015 for summary	W. Cont. Trench, E. Cont. Trench, EU 105-50	5, 175, 200	N/A	N/A	5685	50	5684	25	100
												A_O = 112.5

Note: See Code 2 below for OxCal code used to produce this result, and Figure 4-17 for a display of arrayed dates resulting from the model. See Table 4-6 note for explanation of column headings. "A_O" provides the overall A score for the modeled sequence; a score of >60 indicates a sequence which is stratigraphically compatible.

Results of OxCal modeling of Stratum III are indicated in the right-hand shaded columns of Table 4-7, and are illustrated by Figure 4-17. All Stratum III dates are stratigraphically compatible, the model produces a valid result (A_O = 112.5), and some refinement of the precision of the luminescence dates is evident. Stratum III dates were entered into the master sequence (below) without alteration.

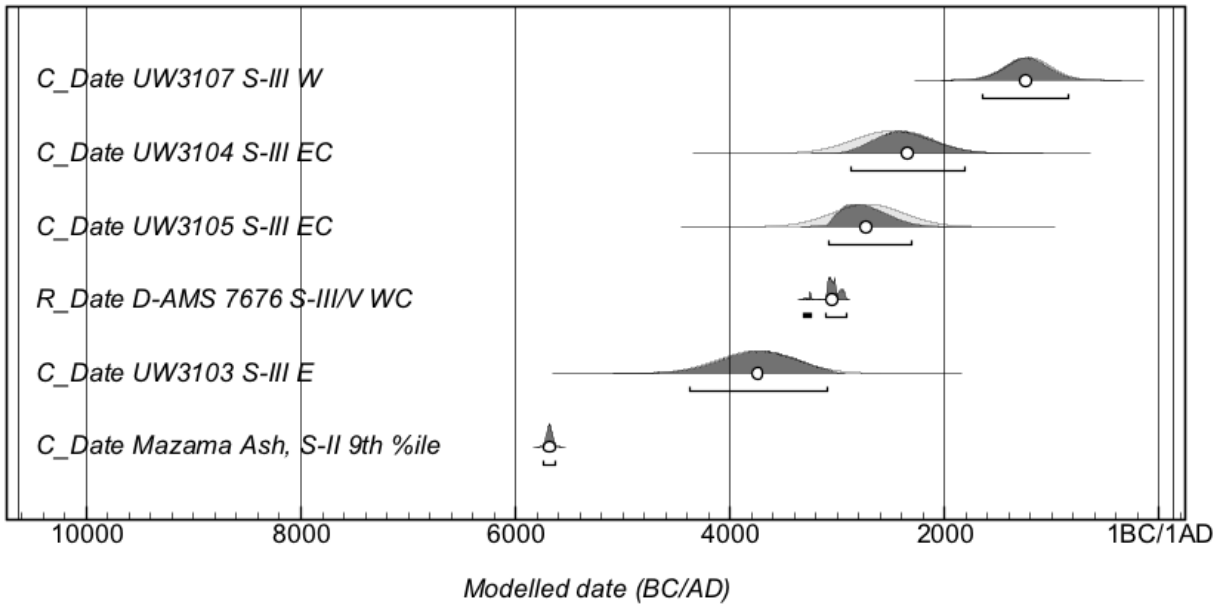

Figure 4-17. Results of OxCal Bayesian modeling of Stratum III samples input in chronological order. Light shaded areas represent age ranges at 2σ prior to modeling; dark shaded areas represent modeled 2σ age ranges. See Code 2 below for OxCal input code and Table 4-7 for a display of numerical results.

Modeling Stratum Va

Stratum Va was modeled using the parameters described above. In this case, Mazama ash was used as an upper constraint, and one of the radiocarbon dates from the top portion of Vb (sample PRI-14-099-32-35) was used as a lower constraint. Radiocarbon sample PRI-14-099-28-32, which was collected from an unspecified portion of Stratum Va, was assumed to be intermediate to samples UBA 27156 and UBA 27155, which were collected from the top and bottom of Va, respectively. Table 4-8 displays samples used for OxCal modeling of Va, and Code 3 displays the input code.

Table 4-8. OxCal Bayesian Model of Stratum Va Including all Available Dates

Strat	Phase	Sample	Type	Location	mE	^{14}C Age	SD	Cal BC	2σ	Model Cal BC	2σ	A Score
M	Mazama	Mazama Ash	Multiple; see Egan et al. 2015 for summary	W. Cont. Trench, E. Cont. Trench, EU 105-50	5, 175, 200	N/A	N/A	5685	50	5684	25	100.2
Va/II	Va/II Contact	UW3109	Fine-Grain Quartz	W. Cont. Trench	5	N/A	N/A	5447	879	5941	136	80.7
Va	Top	UBA 27155	Peat (Humin)	E. Cont. Trench	200	7363	35	6229	133	6232	66	99.9
	Unspec.	PRI-14-099-28-32	Charcoal (conifer)	SON 105	175	6983	47	7557	29	7462	39	4.7
	Bottom	UBA 27156	Peat (Humin)	E. Cont. Trench	200	8286	37	7330	142	7513	18	5.3
Vb	Top	PRI-14-099-32-35	Peat (Fiber)	SON 105	175	8502	27	7557	29	7558	14	100.3
												A_O = 8.0

Note: See Code 3 below for OxCal code used to produce this result. “Unspec.” indicates samples which were collected without provenience to substratum. See Table 4-6 note for explanation of column headings. “A_O” provides the overall A score for the modeled sequence; a score of >60 indicates a sequence which is stratigraphically compatible. Sample PRI-14-099-28-32 was not compatible with the sequence.

The initial attempt to model Stratum Va produced a model with poor validity (A_O = 8.0) due primarily to the incompatibility of PRI-14-099-28-32 (A = 4.7), which again was collected from an unspecified portion of Stratum Va. This age of this sample is therefore a statistical outlier compared to other Stratum Va dates, as it is significantly younger than the other two radiocarbon dates available for this stratum. This sample was subsequently excluded to enable an additional model of Stratum Va (Table 4-9, Code 4). This second attempt successfully produced a statistically valid model (A_O = 90.5) displayed in Figure 4-18. As a result, this incompatible sample was excluded from the master chronology detailed below; the other two Stratum Va samples were used for the master chronology, however.

Table 4-9. OxCal Bayesian Model of Stratum Va Including all Stratigraphically Compatible Dates

Strat	Phase	Sample	Type	Location	mE	^{14}C Age	SD	Cal BC	2σ	Model Cal BC	2σ	A Score
M	Mazama	Mazama Ash	Multiple; see Egan et al. 2015 for summary	Western Trench, Eastern Trench, 105-50	5, 175, 200	N/A	N/A	5685	50	5684	25	100
Va/II	Va/II Contact	UW3109	Fine-Grain Quartz	W. Cont. Trench	5	N/A	N/A	5447	879	5939	135	81.5
Va	Top	UBA 27155	Peat (Humin)	E. Cont. Trench	200	7363	35	6229	133	6232	66	100
	Bottom	UBA 27156	Peat (Humin)	E. Cont. Trench	200	8286	37	7330	142	7328	70	98.8
Vb	Top	PRI-14-099-32-35	Peat (Fiber)	SON 105	175	8502	27	7557	29	7557	15	99.5
												A_O = 90.5

Note: See Code 4 below for OxCal code used to produce this result, and Figure 4-18 for a display of arrayed dates resulting from the model. See Table 4-6 note for explanation of column headings. "A_O" provides the overall A score for the modeled sequence; a score of >60 indicates a sequence which is stratigraphically compatible.

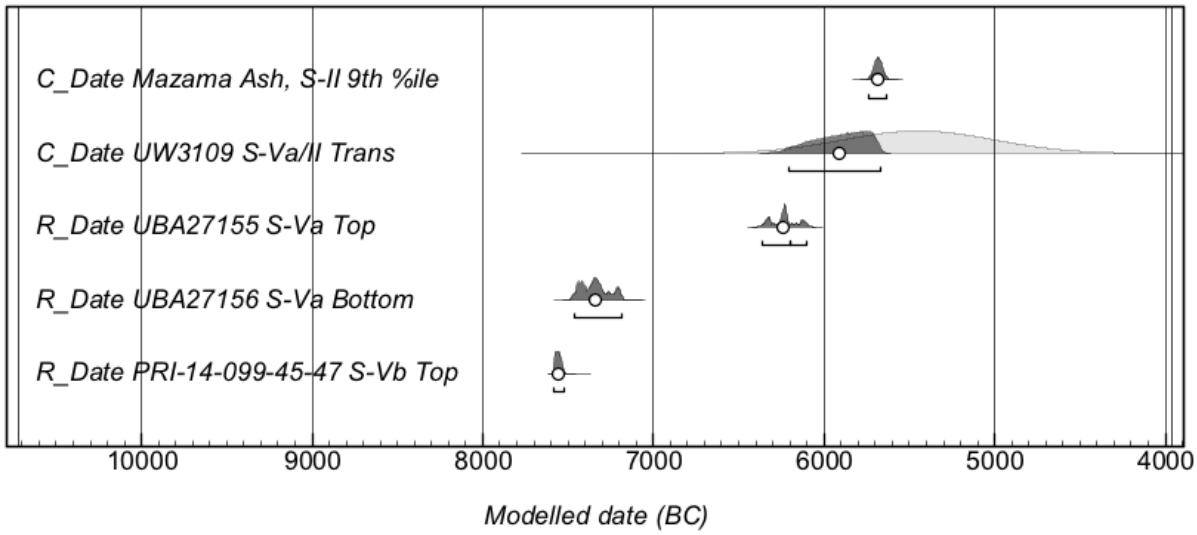

Figure 4-18. Results of OxCal Bayesian modeling of Stratum Va samples, excluding radiocarbon sample PRI-14-099-28-32, which was sampled from an unspecified height within Stratum Va and was statistically incompatible (using Code 3) with the rest of the stratigraphic sequence for Va. Light shaded areas represent age ranges at 2σ prior to modeling; dark shaded areas represent modeled 2σ age ranges. See Code 4 below for OxCal input code and Table 4-9 for a display of numerical results.

Modeling Stratum Vb

Table 4-10 displays samples used for OxCal modeling of Stratum Vb. Here, sample UBA 27156 (bottom of Va) was used as an upper constraint, and sample UBA 27160 (top of Vc) was used as a lower constraint. Input code (see Code 5) was similar to above, with one modification. Specifically, because multiple samples were collected from the same portions of Stratum Vb (e.g., two samples from the bottom of the stratum, and another two samples from the top), dates within a designated portion (e.g., "bottom") were entered as a single "phase" within the OxCal code to avoid imposing assumptions about the stratigraphic order of samples within stratum sub-portions (i.e., to avoid assumptions about which Vb "bottom" sample is earliest). In other words, multiple samples entered within a "phase" are considered by the model to be unordered within the phase, although stratigraphic order still applies between phases. Results of modeling (Table 4-10, Figure 4-19) show that all dates for Stratum Vb are statistically compatible ($A_O = 98.8$).

Table 4-10. OxCal Bayesian Model of Stratum Vb Including all Available Dates

Strat	Phase	Sample	Type	Location	mE	^{14}C Age	SD	Cal BC	2σ	Model Cal BC	2σ	A Score
Va	Bottom	UBA 27156	Peat (Humin)	E. Cont. Trench	200	8286	37	7330	142	7333	72	101
Vb	Top	PRI-14-099-32-35	Peat (Fiber)	SON 105	175	8502	27	7557	29	7557	15	100
		UBA 27157	Peat (Humin)	E. Cont. Trench	200	8606	38	7646	73	7647	36	98
	Mid	UBA 27158	Peat (Humin)	E. Cont. Trench	200	9116	39	8351	95	8351	48	100
	Bottom	PRI-14-099-45-47	Charcoal (*P. menziesii*)	SON 105	175	9654	37	9045	201	9044	100	98
		UBA 27159	Peat (Humin)	E. Cont. Trench	200	9560	43	8962	185	8958	93	100
Vc	Top	UBA 27160	Sed (Humin)	E. Cont. Trench	200	10038	43	9597	212	9590	106	101
												A_O = 98.8

Note: See Code 5 below for OxCal code used to produce this result, and Figure 4-19 for a display of arrayed dates resulting from the model. See Table 4-6 note for explanation of column headings. "A_O" provides the overall A score for the modeled sequence; a score of >60 indicates a sequence which is stratigraphically compatible.

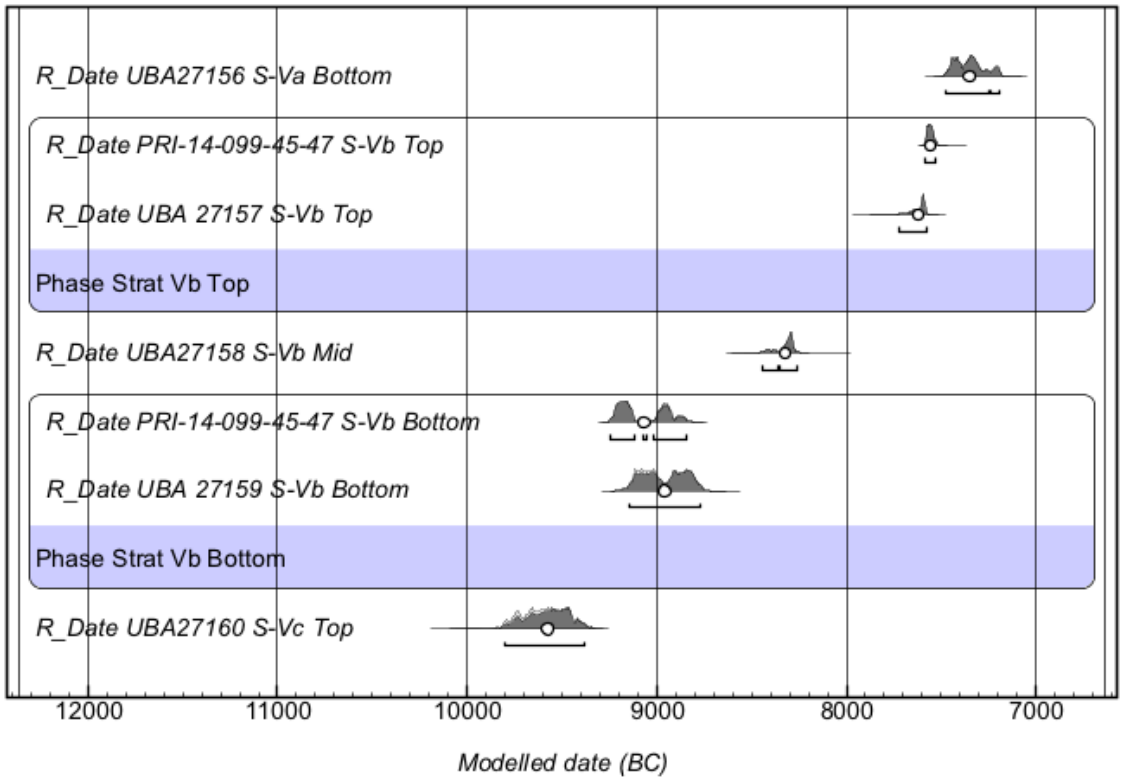

Figure 4-19. Results of OxCal Bayesian modeling of Stratum Vb samples. Light shaded areas represent age ranges at 2σ prior to modeling; dark shaded areas represent modeled 2σ age ranges. Multiple dates within a single portion of a sub-stratum (e.g., two or more dates within the bottom of Stratum Vc) were input as a single "phase" into OxCal, and are displayed within a box delineating the phase; the code/display order of dates within each phase is not statistically meaningful. See Code 5 below for OxCal input code and Table 4-10 for a display of numerical results.

Modeling Stratum Vc

Modeling of Stratum Vc used methods described above (Table 4-11 and Code 6). As with Stratum Va, Stratum Vc samples which were not designated as collected from the top, bottom, etc. of Stratum Vc were assumed to be "middle" samples stratigraphically. As with Stratum Vb, multiple samples within a designated sub-portion (e.g., "bottom") were entered as a single phase.

Table 4-11. OxCal Bayesian Model of Stratum Vc Including all Available Dates

Strat	Phase	Sample	Type	Location	mE	^{14}C Age	SD	Cal BC	2σ
	Top	UBA 27160	Sed (Humin)	E. Cont. Trench	200	10038	43	9596.5	211.5
	Unspec.	D-AMS 7675	Charcoal (Deciduous)	W. Cont. Trench	5	9033	52	8158.5	169.5
	Unspec.	UBA 27154	Sed (Humin)	EU 85-27	57	9393	49	8673.5	117.5
	Unspec.	UBA 27153	Sed (Humin)	EU 81-88	78	9825	60	9324.5	119.5
	Unspec.	D-AMS 7673	Charcoal (Deciduous)	EU 96-48	138	9608	40	9015	191
Vc	Unspec.	D-AMS 7674	Charcoal (*Tsuga* sp.)	EU 96-78	138	10154	34	9882	188
	Unspec.	PRI-14-099-49-55	Charcoal (Salicacaeae)	SON 105	175	10489	29	10455	154
	Bottom	UW3106	Fine-Grain Quartz	EU 80-84	74	N/A	N/A	11964	740
	Bottom	UW 3102	Single-Grain Feldspar	EU 109-28	198	N/A	N/A	9304	1000
	Bottom	UW 3102	Fine-Grain Quartz	EU 109-28	198	N/A	N/A	10278	1676
	Bottom	UBA 27161	Sed (Humin)	E. Cont. Trench	200	10322	45	10230.5	208.5
									$A_O = 0$

Note: See Code 6 below for OxCal code used to produce this result. “Unspec.” indicates samples which were collected without provenience to substratum. See Table 4-6 note for explanation of column headings. “A_O” provides the overall A score for the modeled sequence; a score of >60 indicates a sequence which is stratigraphically compatible. Because $A_O = 0$, the model could not be completed, and modeled dates are not provided for this sequence.

Results indicate stratigraphic incompatibility between samples ($A_O = 0$) and an utter failure of the model to produce any modeled dates. This incompatibility is likely due in large part to the general pattern of Stratum Vc producing slightly younger dates towards the west of the site and older dates towards the east, although the large number of dates from “unspecified” portions of Stratum Vc hinder a full assessment of this assertion. Still, given the data and the depositional context it is reasonable to interpret these dates as indicative of a time-transgressive east-west trend in initial deposition of Stratum Vc. In any case, it is worth noting that the antiquity of Stratum Vc is not impugned by the result of this model, as 7 of the 11 dates for this stratum are compatible with an age of 11,450 cal BP (9500 BC) or older.

Because of the above difficulty, however, the full suite of dates from Stratum Vc needed to be pared down to a subset of stratigraphically compatible dates for inclusion into the master chronology so that the wider depositional history of Bear Creek could be encompassed and described. Dates from the eastern portion of the site were selected for this purpose, as these dates 1) are generally the earliest portion of the site, and thus best encompass its full history, and 2) are most closely comparable to the sequence of eight dates (including the Mazama ash) from the Eastern Trench, which were collected with precise stratigraphic control and therefore offer a high degree of certainty in model construction. Samples collected from "unspecified" locations within Stratum Vc were also excluded to avoid assumptions regarding how they fit into the East Trench sequence. The resulting basic model of Stratum Vc (Table 4-12, Figure 4-20, Code 7) therefore included only three dates for the bottom phase (UBA 27161, plus the single-grain date and the fine-grain date derived from UW3102) and one date (UBA 27160) for the top of Vc. Results show a valid model ($A_O = 96.8$), as well as refinement of the precision of the luminescence ages.

Table 4-12. OxCal Bayesian Model of Eastern Stratum Vc Dates with Known Substratigraphic Context

Strat	Phase	Sample	Type	Location	mE	^{14}C Age	SD	Cal BC	2σ	Model Cal BC	2σ	A Score
Vc	Top	UBA 27160	Sed (Humin)	E. Cont. Trench	200	10038	43	9597	212	9595	106	100.2
	Bottom	UW 3102	Single-Grain Feldspar	EU 109-28	198	N/A	N/A	9304	1000	9920	228	74.5
	Bottom	UW 3102	Fine-Grain Quartz	EU 109-28	198	N/A	N/A	10278	1676	10295	406	125.2
	Bottom	UBA 27161	Sed (Humin)	E. Cont. Trench	200	10322	45	10231	209	10222	108	100.3
												A_O = 96.8

Note: See Code 7 below for OxCal code used to produce this result, and Figure 4-20 for a display of arrayed dates resulting from the model. See Table 4-6 note for explanation of column headings. "A_O" provides the overall A score for the modeled sequence; a score of >60 indicates a sequence which is stratigraphically compatible.

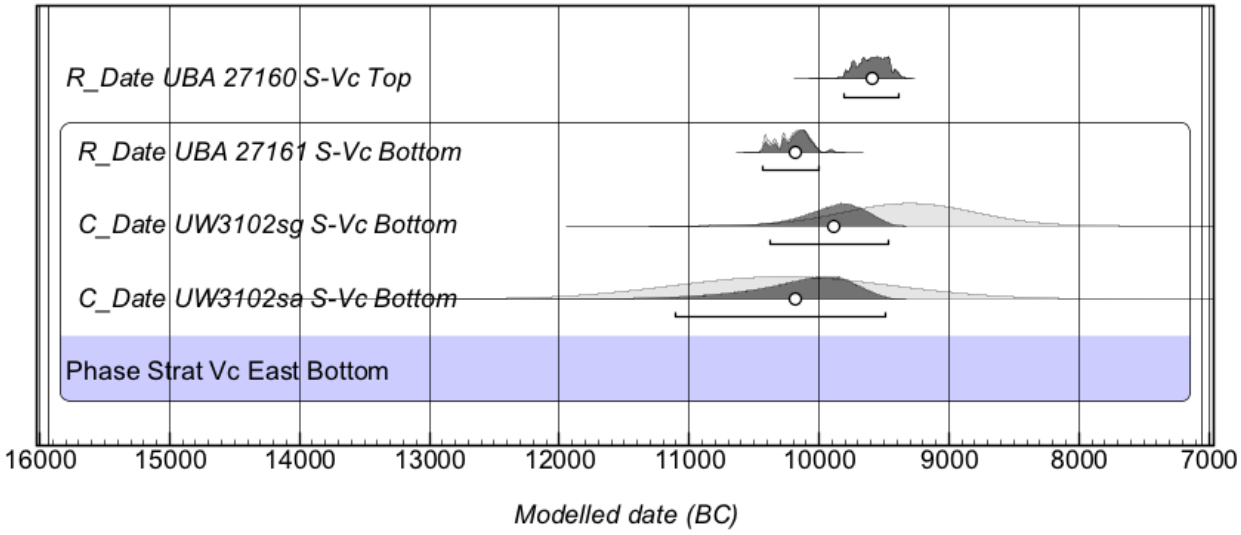

Figure 4-20. Results of OxCal Bayesian modeling of selected Stratum Vc samples from the eastern portion of the site. Modeling of all dates for the entire Stratum Vc sequence (Table 4-11, Code 6) yielded an A_O value of 0; dates from the western portion of the site are stratigraphically incompatible with those from the east, and dates without a recorded substratum (e.g., top, middle, bottom) are also incompatible with dates for which such information is available. Samples in Figure 4-16 were therefore limited to those samples with both 1) a recorded substratum, and 2) samples from the eastern quarter of the site extent. Light shaded areas represent age ranges at 2σ prior to modeling; dark shaded areas represent modeled 2σ age ranges. Multiple dates within a single portion of a sub-stratum (e.g., two or more dates within the bottom of Stratum Vc) were input as a single "phase" into OxCal, and are displayed within a box delineating the phase; the code/display order of dates within each phase is not statistically meaningful. See Code 7 below for OxCal input code and Table 4-12 for a display of numerical results.

Modeling the Master Bear Creek Chronology

All dates which were included in the successful models detailed above were integrated into a single, master sequence which employed the same methods described above. To create this model, a decision as to how to represent the relative order of Strata II and III was necessary. Although no stratigraphic information is available to determine precisely where Stratum III fits within the dated Stratum II sequence, which exhibits overlapping ages with Stratum III, statistical comparison between dates for UW3111 (69^{th} percentile of Strat II accumulation) and UW3103 (the earliest Stratum III date) indicates with 99% confidence (t-test result 2.652, 82 df) that UW3103 is younger. Given this, Stratum III dates were entered into the model as stratigraphically younger than UW3111. Because the terminal date for sampled Stratum II sediment (UW3112) overlaps with the Stratum III sequence, however, this date was combined with Stratum III dates into a single phase to reflect the simultaneity of the two depositional processes which produced these strata. In this way, the model for the master chronology incorporated available data without imposing undue assumptions about the relative ordering of the terminal dates for each stratum.

Input code for the modeled master chronology is given by Code 8, and resultant data are displayed in Table 4-13 and illustrated by Figure 4-21. The model was successful ($A_O = 99.5$), and some gains in precision over raw dates are again apparent, especially for the luminescence dates, which exhibit an average reduction of 2σ error by 522 years (SD = 289.5); average 2σ error of radiocarbon dates was reduced by 74 years (SD = 32.4). Notably, precision gains for Stratum III sediments were lower in the master chronology than in the stand-alone Stratum III

model described above due to the inclusion of these dates as a phase in the master chronology; this is a direct product of the relaxed model assumptions about stratigraphic ordering within a phase, limiting the power of the model to constrain dates within the phase itself. In light of this fact as well as the preceding discussion, it is clear that each of the Bayesian models produced by this analysis has inherent value, as each has slightly different assumptions, inputs, limitations, and strengths.

Table 4-13. OxCal Bayesian Model of Bear Creek Master Chronology

Strat	Phase	Sample	Type	Location	mE	^{14}C Age	SD	Cal BC	2σ	Model Cal BC	2σ	A Score
II/III	Top II	UW3112	Fine-Grain Quartz	W. Cont. Trench	5	N/A	N/A	1329	720	1401	358	99.4
	Bottom III	UW3107	Single-Grain Feldspar	EU 118-26	36	N/A	N/A	1214	400	1236	200	99.8
		UW3104	Single-Grain Feldspar	EU 81-67	138	N/A	N/A	2494	700	2545	347	99.8
		UW3105	Single-Grain Feldspar	EU 114-15	104	N/A	N/A	2714	660	2752	327	99.9
		D-AMS 7676	Charcoal (*Abies* sp.)	EU 95-08	77	4416	26	3115	194	3115	97	99.7
		UW3103	Single-Grain Feldspar	EU 105-50	175	N/A	N/A	3747	520	3674	230	103.9
	69th %ile II	UW3111	Fine-Grain Quartz	W. Cont. Trench	5	N/A	N/A	4004	720	4075	255	114.4
	35th %ile II	UW3110	Fine-Grain Quartz	W. Cont. Trench	5	N/A	N/A	4284	900	4578	339	102.3
M	9th %ile II	Mazama Ash	Multiple; see Egan et al. 2015 for summary	W. Cont. Trench, E. Cont. Trench, EU 105-50	5, 175, 200	N/A	N/A	5685	50	5683	25	100
V/II	Top Va/ Bottom II	UW3109	Fine-Grain Quartz	W. Cont. Trench	5	N/A	N/A	5445	870	5929	132	84.4
Va	Top Va	UBA 27155	Peat (Humin)	E. Cont. Trench	200	7363	35	6229	133	6231	66	100
	Bottom Va	UBA 27156	Peat (Humin)	E. Cont. Trench	200	8286	37	7330	142	7330	71	99.7

Strat	Phase	Sample	Type	Location	mE	^{14}C Age	SD	Cal BC	2σ	Model Cal BC	2σ	A Score
Vb	Top Vb	PRI-14-099-32-35	Peat (Fiber)	SON 105	175	8502	27	7557	29	7557	15	99.4
		UBA 27157	Peat (Humin)	E. Cont. Trench	200	8606	38	7646	73	7646	36	99.1
	Mid Vb	UBA 27158	Peat (Humin)	E. Cont. Trench	200	9116	39	8351	95	8350	47	100.3
	Bottom Vb	PRI-14-099-45-47	Charcoal (*P. menziesii*)	SON 105	175	9654	37	9045	201	9045	101	98.9
		UBA 27159	Peat (Humin)	E. Cont. Trench	200	9560	43	8962	185	8960	93	99.8
Vc	Top Vc	UBA 27160	Sed (Humin)	E. Cont. Trench	200	10038	43	9597	212	9566	102	100.8
	Bottom Vc	UW 3102	Single-Grain Feldspar	EU 109-28	198	N/A	N/A	9304	1000	9904	242	81.1
		UW 3102	Fine-Grain Quartz	EU 109-28	198	N/A	N/A	10278	1676	10505	522	116.7
		UBA 27161	Sed (Humin)	E. Cont. Trench	200	10322	45	10231	209	10229	105	100.2
												A_O = 99.5

Note: See Code 8 below for OxCal code used to produce this result, and Figure 4-21 for a display of arrayed dates resulting from the model. See Table 4-6 note for explanation of column headings. "A_O" provides the overall A score for the modeled sequence; a score of >60 indicates a sequence which is stratigraphically compatible.

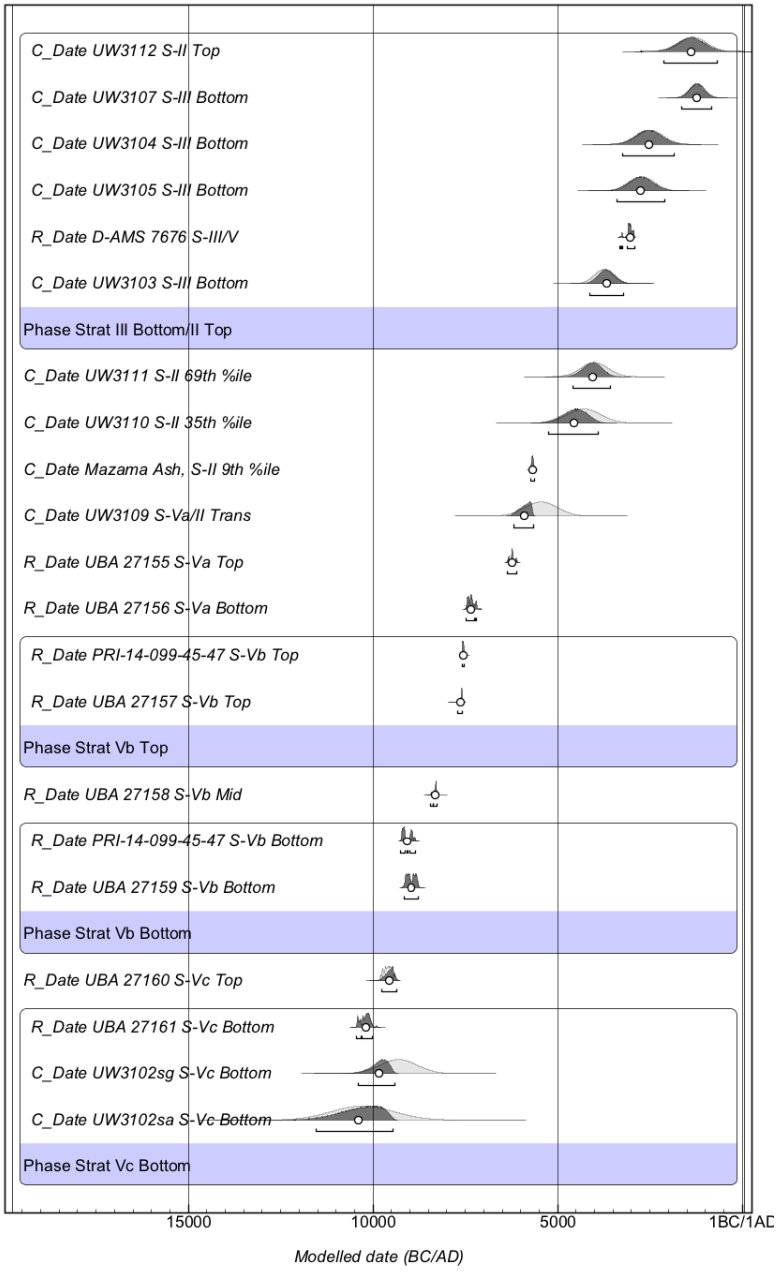

Figure 4-21. Results of OxCal Bayesian modeling of all selected samples. All samples included in successful models (e.g., Figures 4-12 to 4-16) were included. Light shaded areas represent age ranges at 2σ prior to modeling; dark shaded areas represent modeled 2σ age ranges. Multiple dates within a single portion of a sub-stratum (e.g., two or more dates within the bottom of Stratum Vc) were input as a single "phase" into OxCal, and are displayed within a box delineating the phase; the code/display order of dates within each phase is not statistically meaningful. See Code 8 below for OxCal input code and Table 4-13 for a display of numerical results.

DISCUSSION AND CONCLUSIONS

Several salient points emerge from the preceding analysis. First, the luminescence dating of Bear Creek sediments was successful as whole. Sample ages consistently exhibited general stratigraphic compatibility both between individual luminescence dates and between luminescence dates and independent dates, including radiocarbon dates produced by three different laboratories as well as known ages for Mazama ash. Only one luminescence sample, UW3108, failed to produce a reasonable age estimate, and double-SAR measurement of this sample yielded data which exhibited easily identifiable indices of partial bleaching, enabling rejection of the date on methodological grounds. Second, the luminescence dates add vital information to the interpretation of Bear Creek chronology, as these data were able to 1) provide ages for sediments which would have been extremely difficult or impossible to date extensively with other methods, including Strata II and III, which contained little material for radiocarbon dating besides occasional and scant detrital charcoal, and 2) provide independent support for radiocarbon estimates of the antiquity of Stratum Vc, and therefore the earliest cultural component of the Bear Creek Site. Third, the Bayesian modeling of the aggregate dataset was effective in allowing statistically rigorous integration of a suite of dates derived from multiple methods; this increased confidence in the compatibility of all Bear Creek dates, allowed for the identification of outliers, increased the precision of individual dates, and provided a baseline for discussing interesting aspects of time-transgressive processes of Bear Creek Site such as the probability that Stratum Vc deposition was earliest in the eastern portion of the site.

Perhaps more importantly, however, the ultimate result of the analysis undertaken here is the construction of a powerful tool for interpreting the process of formation of the Bear Creek Site as a whole. Given the Bayesian modeled dates for Stratum Vc, it is clear that the transformation from the paleoshoreline of LPH Lake Sammamish to a wetland environment began around 11,000 BC (12,950 cal BP) in the eastern portion of the excavated site, with this transformation affecting the western portion centuries later. During this time, ancient inhabitants of the Bear Creek Site moved onto the newly exposed ground and took extensive advantage of the wetland, producing most of the chipped-stone artifacts recovered during excavations. Evidence from analysis of residues on this assemblage indicates these people processed bear, artiodactyl, bison, mountain or bighorn sheep, and some variety of salmonid fish (Kopperl et al. 2016), among other resources. By 9000 BC (10,950 cal BP), deposition of fibrous peat was underway, indicating rising water levels which once again drowned the wetland and pushed out human occupation.

Deposition of the fibrous peat continued until sometime between 7650 BC and 7500 BC (9600 and 9450 cal BP), at least in the eastern portion of the site where samples of Vb were collected, indicating water levels remained high—and the land surface at Bear Creek submerged and uninhabitable—for approximately 2,500 years. By 7300 BC (9250 cal BP) deposition of this fibrous peat had abated, and from this time until around 6000 BC (7950 cal BP) the local environment gradually transitioned from a wetland directly connected with ancient Lake Sammamish to a slackwater pond separated from the lake and fed only by occasional flooding.

A layer of Mazama ash was deposited at around 5680 BC (7630 cal BP), although deposition of Stratum II continued until at least 1400 BC (at which point intact Stratum II sediment was unavailable for sampling), with a pronounced increase in rate of deposition between roughly 4500 BC and 3900 BC (6450 and 5850 cal BP) (see Fig 14), likely reflecting a period of high flood activity. Excavations revealed no evidence for human occupation in this isolated slackwater environment.

Shortly after the increase in Stratum II deposition, at around 3700 BC (5650 cal BP), the fluvial system which produced Stratum III began to meander through the site, bisecting the slackwater with occasional heavy flows which reached Lake Sammamish further to the south. This fluvial system was probably initiated in the eastern quarter of the site, but by 3100 BC (5050 cal BP) it had moved to the west-central quarter of the site. It then moved back to the east-central quarter of the site by around 2700 BC (4650 cal BP), where it seems to have lingered until about 1200 BC (3190 cal BP), when it appears in the western quarter of the site. This fluvial system, interpreted by site geoarchaeologists as a possible "proto–Bear Creek," eroded and reworked existing deposits, including Stratum Vc artifacts, depositing these artifacts in secondary contexts within Stratum III.

Taken together, then, the chronometric data from Bear Creek inform interpretations of nearly 9,000 years of site formation—from around 10,000 BC to around 1200 BC (12,000 to 3200 cal BP)—in the ancient local landscape, which was a resource-rich wetland around 10,000 BC (12,000 cal BP), but was inhospitable to sustained human occupation after about 9000 BC (11,000 cal BP).

ACKNOWLEDGEMENTS

I am thankful to Bob Kopperl and Chris Miss for inviting me to participate in this project, and to Jim Feathers for supporting my efforts in the laboratory. Thanks are also due to the City of Redmond, SWCA Environmental Consultants, the Muckleshoot Indian Tribe, the Snoqualmie Tribe, the Stillaguamish Tribe of Indians, and the Tulalip Tribes for their aid in this project and for allowing me to disseminate the work detailed here as part of my dissertation. Finally, thanks are due to Johonna Shea and Rhiannon Held for their technical assistance with graphics and copy editing, and to the Bear Creek field crew for a great—if exhausting—summer.

REFERENCES

Aitken, M.J.
1985 *Thermoluminescence Dating*. Academic Press, London.

1998 *An Introduction to Optical Dating*. Oxford University Press, Oxford.

Auclair, M., M. Lamothe, and S. Huot
2003 Measurement of anomalous fading for feldspar IRSL using SAR. *Radiation Measurements* 37:487–492.

Banerjee, D., A.S. Murray, L. Bøtter-Jensen, and A. Lang
2001 Equivalent dose estimation using a single aliquot of polymineral fine grains. *Radiation Measurements* 33:73–93.

Bøtter-Jensen, L., and V. Mejdahl
1988 Assessment of beta dose-rate using a GM multi-counter system. *Nuclear Tracks and Radiation Measurements* 14:187–191.

Bronk Ramsey, C.
2008 Deposition models for chronological records. *Quaternary Science Reviews* 27:42–60.

2009 Bayesian analysis of radiocarbon dates. *Radiocarbon* 51:337–360.

Bryan, A.L.
1980 The Stemmed Point Tradition: An Early Technological Tradition in Western North America. In *Anthropological Papers in Honor of Earl H. Swanson, Jr.*, edited by L. B. Harten, C. N. Warren, and D. R. Tuohy, pp. 77–107. Idaho Museum of Natural History, Pocatello, Idaho.

Butler, B.R.
1961 *The Old Cordilleran Culture in the Pacific Northwest*. Occasional Papers No. 5. Idaho State College Museum, Pocatello, Idaho.

Carlson, R.L.
1990 History of Research in Archaeology. In *Handbook of North American Indians, Volume 7: Northwest Coast*, edited by W. Suttles, pp. 107–115. Smithsonian Institution, Washington D.C.

Egan, J., R. Staff, and J. Blackford
2015 A high-precision age estimate of the Holocene Plinian eruption of Mount Mazama, Oregon, USA. *The Holocene* 25:1054–1067.

Feathers, J.K., and J. Tunnicliff
2011 Effect of single-grain versus multi-grain aliquots in determining age for K feldspars from Southern British Columbia. *Ancient TL* 29:53–58.

Foit, F.F.
2014 Washington State University Microbeam Laboratory report on results of analysis of Bear Creek tephra samples. Unpublished report transmitted to SWCA Environmental Consultants on November 4, 2014. Available in *Results of Data Recovery at the Bear Creek Site (45KI839), King County, Washington, Volume 2: Appendices*, edited by R. Kopperl, 2016. Prepared for City of Redmond, Washington and David Evans and Associates, Inc. SWCA Environmental Consultants, Seattle, Washington.

Galbraith, R.F.
2005 *Statistics for Fission Track Analysis*. Chapman and Hall, London.

Galbraith, R.F., and R.G. Roberts
2012 Statistical aspects of equivalent dose and error calculation and display in OSL dating: an overview and some recommendations. *Quaternary Geochronology* 11:1–27.

Galbraith, R.F., R.G. Roberts, G.M. Laslett, H. Yoshida, and J.M. Olley
1999 Optical Dating of Single and Multiple Grains of Quartz from Jinmium Rock Shelter, Northern Australia: Part I, Experimental Design and Statistical Models. *Archaeometry* 41:339–364.

Godfrey-Smith, D.I., D.J. Huntley, and W.H. Chen
1988 Optical Dating Studies of Quartz and Feldspar Sediment Extracts. *Quaternary Science Reviews* 7:373–380.

Guérin, G., N. Mercier, and G. Adamiec
2011 Dose-rate conversion factors: update. *Ancient TL* 29:5–8.

Gustafson, C.E., R. Daugherty, and D.W. Gilbow
1979 The Manis Mastodon Site: Early Man on the Olympic Peninsula. *Canadian Journal of Archaeology* 3:157–164.

Grayson, D.K., and D.J. Meltzer
2015 Revisiting Paleoindian exploitation of extinct North American mammals. *Journal Archaeological Science* 56:177–193.

Hodges, C., B. Rinck, J. Johnson, S. Johannessen, R. Kopperl, K. Troost, and E. Strata
2016 Chapter 7 – Site Formation. In *Results of Data Recovery at the Bear Creek Site (45KI839), King County, Washington, Volume 1: Report*, edited by R. Kopperl, pp. 65–148. Prepared for City of Redmond, Washington and David Evans and Associates, Inc. SWCA Environmental Consultants, Seattle, Washington.

Huntley, D.J., and M. Lamothe
2001 Ubiquity of anomalous fading in K feldspars, and measurement and correction for it in optical dating. *Canadian Journal of Earth Sciences* 38:1093–1106.

Jain, M., A.S. Murray, and L. Botter-Jensen
2004 Optically stimulated luminescence dating: how significant is incomplete light exposure in fluvial environments? *Quaternaire* 15:143–157.

Johnson, J.
2014 Accurate measurements of low *Z* elements in sediments and archaeological ceramics using portable X-ray fluorescence (PRXF). *Journal of Archaeological Method and Theory* 21:563–588.

Kenady, S., M.C. Wilson, and R.F. Schalk
2007 Indications of Butchering on a Late-Pleistocene *Bison antiquus* from the Maritime Pacific Northwest. *Current Research in the Pleistocene* 24:167–170.

Kopperl, R., C. Hodges, A. Taylor, C. Beck, K.M. Ames, and C.J. Miss
2016 Chapter 10 – Conclusions. In *Results of Data Recovery at the Bear Creek Site (45KI839), King County, Washington, Volume 1: Report*, edited by R. Kopperl, pp. 261–376. Prepared for City of Redmond, Washington and David Evans and Associates, Inc. SWCA Environmental Consultants, Seattle, Washington.

Kopperl, R.E., C.J. Miss, and C.M. Hodges
2010 Results of Testing at Bear Creek, Site 45KI839, Redmond, King County, Washington. Report WA09-13. Prepared for David Evans and Associates, Inc. Northwest Archaeological Associates, Seattle, Washington.

Leonhardy, F.C., and D.G. Rice
1970 A proposed Culture Typology for the Lower Sanke River Region, Southeastern Washington. *Northwest Anthropological Research Notes* 4(1):1–29.

LeTourneau, P.D.
2010 A Clovis Point from the Pacific Northwest Coast. *Current Research in the Pleistocene* 27:115–117.

Meltzer, D.J., and R.C. Dunnell
1987 Fluted Points from the Pacific Northwest. *Current Research in the Pleistocene* 4:64–67.

Murray, A.S., K.J. Thomsen, N. Masuda, J.P. Buylaert, and M. Jain
2012 Identifying well-bleached quartz using the different bleaching rates of quartz and feldspar luminescence signals. *Radiation Measurements* 47(9):688–695.

Murray, A.S., and A.G. Wintle
2000 Luminescence dating of quartz using an improved single-aliquot regenerative-dose protocol. *Radiation Measurements* 32:57–73.

Nelson, M.S., H.J. Gray, J.A. Johnson, T.M. Rittenour, J.K. Feathers, and S.A. Mahan
2015 User guide for luminescence sampling in archaeological and geological contexts. *Advances in Archaeological Practice* 3:166–177.

Olley, J., G. Caitcheon, and A. Murray
1998 The distribution of apparent dose as determined by Optically Stimulated Luminescence in small aliquots of fluvial quartz: Implications for dating young sediments. *Quaternary Science Reviews* 17(11):1033–1040.

Olley, J.M., T. Pietsch, and R.G. Roberts
2004 Optical dating of Holocene sediments from a variety of geomorphic settings using single grains of quartz. *Geomorphology* 60(3–4):337–358.

Osborne, D., W.W. Caldwell, and R.H. Crabtree
1956 The Problem of Northwest Coast-Interior Relationships as Seen from Seattle. *American Antiquity* 22(2):117–128.

Prescott, J.R., and J.T. Hutton
1994 Cosmic ray contributions to dose rates for luminescence and ESR dating: large depths and long time durations. *Radiation Measurements* 23:497–500.

Rinck, B.A.
2016 Micromorphological Analysis of Thin Sections from Bear Creek (45KI839), Redmond, King County, Washington. In *Results of Data Recovery at the Bear Creek Site (45KI839), King County, Washington, Volume 2: Appendices*, edited by R. Kopperl. Prepared for City of Redmond, Washington and David Evans and Associates, Inc. SWCA Environmental Consultants, Seattle, Washington.

Roberts, R.G., Z. Jacobs, B. Li, N.R. Jankowski, A.C. Cunningham, and A.B. Rosenfeld
2015 Optial dating in archaeology: thirty years in retrospect and grand challenges for the future. *Journal of Archaeological Science* 56:41–60.

Stokes, S., H.E. Bray, and M.D. Blum
2001 Optical resetting in large drainage basins: tests of zeroing assumptions using single-aliquot procedures. *Quaternary Science Reviews* 20(5–9):879–885.

Taylor, A., and C. Beck
2016 Chapter 8 – Lithic Technology. In *Results of Data Recovery at the Bear Creek Site (45KI839), King County, Washington, Volume 1: Report*, edited by R. Kopperl, pp. 149-226. Prepared for City of Redmond, Washington, and David Evans and Associates, Inc. SWCA Environmental Consultants, Seattle, Washington.

Tunnicliffe, J., M. Church, J.J. Clague, and J.K. Feathers
2012 Postglacial sediment budget of Chilliwack Valley, British Columbia. *Earth Surface Processes and Landforms* 37:1243–1262.

Waters, M.R., T.W. Stafford Jr., H.G. McDonald, C. Gustafson, C.M. Rasmussen, E. Cappellini, J.V. Olsen, D. Szklarczyk, L. Juhl Jensen, M.T.P. Gilbert, and E. Willerslev
2011 Pre-Clovis Mastodon Hunting 13,800 Years Ago at the Manis Site, Washington. *Science* 334:351–353.

Willig, J.A.
1989 Paleo-Archaic Broad-Spectrum Adaptations at the Pleistocene-Holocene Boundary. Unpublished Ph.D. Dissertation, Department of Anthropology, University of Oregon, Eugene, Oregon.

Willig, J.A., and C.M. Aikens
1988 The Clovis-Archaic Interface in far western North America. In *Early Human Occupation in Far Western North America: The Clovis-Archaic Interface*, edited by J. A. Willig, C. M. Aikens and J. L. Fagan, pp. 1–40. Anthropological Papers No. 21. Nevada State Museum, Carson City, Nevada.

Wintle, A.G.
1973 Anomalous fading of TL in mineral samples. *Nature* 245:143–144.

OXCAL 4.2 CODE

Code 1: Used to model Stratum II (see Table 4-6, Figure 4-15, and Figure 4-16)

```
Sequence("Bear Creek Strat II")
{
Sigma_Boundary();
R_Date("UBA 27155 S-Va Top",7363,35);
C_Date("UW3109 S-Va/II",-5445,435);
C_Date("Mazama Ash",-5685,25);
C_Date("UW3110 S-II 35%",-4284,450);
C_Date("UW3111 S-II 69%",-4004,360);
C_Date("UW3112 S-II Upper",-1329,360);
Sigma_Boundary();
};
```

Code 2: Used to model Stratum III (see Table 4-7 and Figure 4-17)

```
Sequence("Bear Creek Strat III")
{
Sigma_Boundary();
C_Date("Mazama Ash, S-II 9th %ile", -5683, 25);
C_Date("UW3103 S-III E", -3747, 360);
R_Date("D-AMS 7676 S-III/V WC", 4416, 26);
C_Date("UW3105 S-III EC", -2714, 330);
C_Date("UW3104 S-III EC", -2494, 350);
C_Date("UW3107 S-III W", -1214, 200);
Sigma_Boundary();
};
```

Code 3: Used to model Stratum Va (Model unsuccessful; see Table 4-8)

```
Sequence("Bear Creek Va")
{
Sigma_Boundary();
R_Date("PRI-14-099-45-47 S-Vb Top", 8502, 27);
R_Date("UBA27156 S-Va Bottom", 8286, 37);
R_Date("PRI-14-099-32-35 S-Va Unspc.", 8502, 27);
R_Date("UBA27155 S-Va Top", 7363, 35);
C_Date("UW3109 S-Va/II Trans", -5445, 440);
C_Date("Mazama Ash, S-II 9th %ile", -5683, 25);
Sigma_Boundary();
};
```

Code 4: Used to model Stratum Va; excludes incompatible date from unspecified substratum location (see Table 4-9 and Figure 4-18)

```
Sequence("Bear Creek Va")
{
Sigma_Boundary();
R_Date("PRI-14-099-45-47 S-Vb Top", 8502, 27);
R_Date("UBA27156 S-Va Bottom", 8286, 37);
R_Date("UBA27155 S-Va Top", 7363, 35);
C_Date("UW3109 S-Va/II Trans", -5445, 440);
C_Date("Mazama Ash, S-II 9th %ile", -5683, 25);
Sigma_Boundary();
};
```

Code 5: Used to model Stratum Vb (see Table 4-10 and Figure 4-19)

```
Sequence("Bear Creek Vb")
{
Sigma_Boundary();
R_Date("UBA27160 S-Vc Top", 10038, 43);
Phase("Strat Vb Bottom")
{
R_Date("UBA 27159 S-Vb Bottom", 9560, 43);
R_Date("PRI-14-099-45-47 S-Vb Bottom", 9654, 37);
};
R_Date("UBA27158 S-Vb Mid", 9116, 39);
Phase("Strat Vb Top")
{
R_Date("UBA 27157 S-Vb Top", 8606, 38);
R_Date("PRI-14-099-45-47 S-Vb Top", 8502, 27);
};
R_Date("UBA27156 S-Va Bottom", 8286, 37);
Sigma_Boundary();
};
```

Code 6: Used to model Stratum Vc (Model unsuccessful; see Table 4-11)

```
Sequence("Bear Creek Vc")
{
Sigma_Boundary();
Phase("Strat Vc Bottom")
{
C_Date("UW3102sa S-Vc Bottom", -10278, 840);
C_Date("UW3102sg S-Vc Bottom", -9304, 500);
C_Date("UW3106 S-Vc Bottom", -11964, 320);
R_Date("UBA27161 S-Vc Bottom", 10322, 45);
};
Phase("Strat Vc Unspc.")
{
R_Date("PRI-14-099-49-55 S-Vc Unspc.", 10489, 29);
R_Date("D-AMS 7674 S-Vc Unspc.", 10154, 34);
R_Date("D-AMS 7673 S-Vc Unspc.", 9608, 40);
R_Date("UBA 27153 S-Vc Unspc.", 9825, 60);
R_Date("UBA 27154 S-Vc Unspc.", 9393, 49);
R_Date("D-AMS 7675 S-Vc Unspc.", 9033, 52);
};
R_Date("UBA27160 S-Vc Top", 10038, 43);
Sigma_Boundary();
};
```

Code 7: Used to model the eastern portion of Stratum Vc; excluded date from unspecified substratum location (see Table 4-12 and Figure 4-20)

```
Sequence("Bear Creek Vc")
{
Sigma_Boundary();
Phase("Strat Vc East Bottom")
{
C_Date("UW3102sa S-Vc Bottom", -10278, 840);
C_Date("UW3102sg S-Vc Bottom", -9304, 500);
R_Date("UBA 27161 S-Vc Bottom", 10322, 45);
};
R_Date("UBA 27160 S-Vc Top", 10038, 43);
Sigma_Boundary();
};
```

Code 8: Used to model the master chronology of Bear Creek Site (see Table 4-13 and Figure 4-21)

```
Sequence("Bear Creek Master")
{
Sigma_Boundary();
Phase("Strat Vc Bottom")
{
C_Date("UW3102sa S-Vc Bottom", -10278, 840);
C_Date("UW3102sg S-Vc Bottom", -9304, 500);
R_Date("UBA 27161 S-Vc Bottom", 10322, 45);
};
R_Date("UBA 27160 S-Vc Top", 10038, 43);
Phase("Strat Vb Bottom")
{
R_Date("UBA 27159 S-Vb Bottom", 9560, 43);
R_Date("PRI-14-099-45-47 S-Vb Bottom", 9654, 37);
};
R_Date("UBA 27158 S-Vb Mid", 9116, 39);
Phase("Strat Vb Top")
{
R_Date("UBA 27157 S-Vb Top", 8606, 38);
R_Date("PRI-14-099-45-47 S-Vb Top", 8502, 27);
};
R_Date("UBA 27156 S-Va Bottom", 8286, 37);
R_Date("UBA 27155 S-Va Top", 7363, 35);
C_Date("UW3109 S-Va/II Trans", -5445, 440);
C_Date("Mazama Ash, S-II 9th %ile", -5683, 25);
C_Date("UW3110 S-II 35th %ile", -4284, 450);
C_Date("UW3111 S-II 69th %ile", -4004, 360);
Phase("Strat III Bottom/II Top")
{
C_Date("UW3103 S-III Bottom", -3747, 260);
R_Date("D-AMS 7676 S-III/V", 4416, 26);
C_Date("UW3105 S-III Bottom", -2714, 330);
C_Date("UW3104 S-III Bottom", -2494, 350);
C_Date("UW3107 S-III Bottom", -1214, 200);
C_Date("UW3112 S-II Top", -1329, 360);
};
Sigma_Boundary();
};
```

Chapter 5. A Luminescence-Based > 2,000-year Record of El Niño–driven Terrain Change in the Chao and Santa Valleys, Perú

ABSTRACT

Archaeologists have long argued that pronounced short-term environmental changes caused by El Niños impacted ancient coastal Peruvian societies, but local, long-term, well-dated proxies for ancient El Niño activity have remained elusive, hindering such analysis. Because El Niño activity in the area produces identifiable earthen deposits, luminescence dating of local sediments offers the opportunity to remedy this situation. As a first step toward harvesting this opportunity, local effects of El Niño–driven deposition are discussed in detail, field collection and laboratory luminescence analysis of 28 sediments derived from Chao Valley laminar flood deposits and Santa Valley beach ridges are described, and a luminescence-based chronology for major late Holocene flood and combined flood/earthquake events is set forth. This chronology is then evaluated and refined using stratigraphic information, Bayesian analysis in OxCal, and additional sediment composition data derived from X-ray fluorescence, granulometry, and magnetic susceptibility. Finally, a detailed local El Niño–driven depositional history is set forth, and possible implications for interpreting the effects of this hypothesized history on ancient coastal Peruvians is briefly discussed.

INTRODUCTION

El Niño/Southern Oscillation (ENSO) events currently represent the largest source of inter-annual climate variation in the world and cause billions of dollars of damage every time a "strong" or "very strong" event occurs (Glantz 1996). On the arid coast of Perú, modern El Niño events often bring powerful floods (Lavado Casimiro et al. 2013; Morera et al. 2017), disease (e.g., Helguero Santin et al. 2018; Ramirez et al. 2013), and economic hardship (Lavado

Casimiro et al. 2013), and a significant body of coastal Perúvian research demonstrates that El Niño events have impacted people in the area for millennia (e.g., Bawden 1996; Billman 1996; Brooks et al. 2005; Craig and Shimada 1986; Dillehay and Kolata 2004; Huckleberry and Billman 2003; Keefer et al. 2003; Moseley 2001; Moseley et al. 1992; Nialls et al. 1979a, 1979b Sandweiss 1986, 2003; Sandweiss et al. 2009; Shimada 1994; Uceda and Tufinio 2010; Wells 1990), especially when coupled with local seismic activity (e.g., Keefer et al. 2003; Moseley et al. 1992; Sandweiss 1986; Sandweiss et al. 2009) or other environmental challenges such as droughts (e.g., Shimada et al. 1991).

Variability and complexity in Holocene ENSO dynamics and teleconnections (e.g., Carre et al. 2014), when coupled with marked geographic differences in the effects of individual ENSO events within coastal Perú (Morera et al. 2017) and differences in effects between distinct El Niños even when location is held constant (e.g., Lavado Casimiro et al. 2013), necessitate the use of local paleo-Niño proxies if the local cultural effects of past ENSO events are to be examined in detail. Ideally, such local proxies would also provide an ENSO record which is continuous, long term, well dated, and derived from a depositional context which is highly relevant to those local cultural practices most likely to have been significantly affected by ENSO events. Because of the arid local environment, available data on paleo-Niño activity in north coastal Perú are limited in one or more of these dimensions, either by geographic distance, time depth, depositional continuity, and/or chronometry, leaving modern researchers without a strong baseline understanding of local north coastal ENSO-driven environmental change over the last several thousand years.

A geoarchaeological approach which makes use of luminescence dating of sediment deposition is highly suited to this research context and its incumbent challenges. Evidence for

ENSO-driven sedimentation is extant in the area, ENSO-driven depositional events can be directly calendrically dated with luminescence dating, and the systematic study of these deposits can also provide a direct index of long-term variation in ENSO effects on local landforms, and by extension cultural resources such as domestic architecture and subsistence infrastructure.

The research presented here describes an initial effort to undertake such research in the Santa and Chao Valleys on the northern coast of Perú, with the objective of the creation of an integrated luminescence-based chronology of natural ENSO-driven local depositional change for much of the late Holocene. This chronology is focused on temporal variation in Chao Valley Niño flood alluvium and in Santa Valley beach ridge formation and coastal progradation as a result of interactions between seismic events and El Niño–driven erosion and deposition. After this chronology is set forth and evaluated, a detailed depositional history of the area is set forth, followed by a brief discussion of potential implications for the use of this depositional history to interpret ancient cultural changes in the area.

PROJECT BACKGROUND

Natural Setting

The northern coast of Perú consists of a series of arid valleys bounded by the Andes to the east, the Pacific Ocean to the west, and numerous *cerros* (foothills) running in between. It is a region of extremes; the second-highest mountain range in the world sits adjacent to one of the most nutrient-rich stretches of ocean in the world, with a narrow strip of desolate, arid terrain in between. The rising Andes are due to ongoing subduction of the Nazca plate under the South American plate, and powerful earthquakes are endemic to the area. The richness of the ocean is due to the combination of the Perú-Chile trench and the Humboldt Current, which causes a strong upwelling of cold, nutrient-rich waters along the length of the Peruvian coast. The aridity

of the coastal plain, which was formed by aggradational deposition of Andean glacial outwash, is caused by a combination of the Andes' rain shadow and the cold Pacific water, which act together to shield the area from rainfall, while the intense tropical sun causes the local air to expand and withdraw moisture from the coastal terrain.

Under normal conditions, precipitation in the north Perúvian coastal desert is typically around 20 mm/year or less, with roughly 75% of annual precipitation occurring between January and March (Rau et al. 2017; Tapley and Whalen 1990), typically as the product of *garúa* (fog/mist) rather than rainfall. While garúa discharge is insufficient to create runoff along the desert, upland communities of xerophytic shrubs, known locally as *lomas*, subsist on the garúa moisture. At higher altitudes to the east, this moisture condenses into rainfall which feeds a series of west-flowing rivers which dissect the desert lowland, although periodic upland droughts have likely affected discharge of these rivers for millennia (Shimada et al. 1991). A riparian zone follows these rivers, concentrating naturally occurring lowland terrestrial biomass within relatively narrow corridors separated by expanses of parched terrain. River flow recharges local aquifers, which in turn feed small springs and a series of small, localized oases at valley margins (Clement and Moseley 1991).

Normal conditions are upended roughly every 7 to 10 years by the onset of an El Niño event, a part of the ENSO weather phenomenon which brings drastic changes to the coastal Peruvian environment (Morera et al. 2017). During an El Niño event, northward-moving coastal currents and upwelling abate along the Peruvian coast, and the cold, nutrient-rich water is displaced by warmer equatorial waters pushing further south, resulting in a huge net loss of marine biomass in the region (Barber and Chavez 1983; Glynn 1988), rising local sea levels, and temporary extirpation of some shellfish species (Arntz et al. 1987). Warmer local sea surfaces

also allow moisture-rich air to reach the arid coast, resulting in heavy rains—sometimes up to 170 times the annual mean within a few weeks (Morera et al. 2017)—which cause concomitant erosion and deposition of terrestrial sediments, and a brief "desert bloom" of terrestrial flora (Jaksic 2001). As experienced on the Peruvian coast, the effects of individual El Niño events vary widely in terms of strength, extent, periodicity, and speed and duration of onset, among other dimensions (Lavado Casimiro et al. 2013; Morera et al. 2017).

Published proxies have identified millennial- or century-scale shifts in ENSO activity (e.g., Carre et al. 2014; Sandweiss 2003; Zhang et al. 2014), although comparing and reconciling distinct proxies remains a daunting exercise due to differences in proxy type, location, model assumptions, and dating resolution (Sadekov et al. 2012). Recent work in the Galapagos Islands suggests ENSO activity in the eastern Pacific was erratic in the early Holocene (9200–5600 radiocarbon years before present [BP]) and relatively weak in the mid-Holocene (5600–3500 BP), while strengthening to its modern pattern from 3500 BP onward, with a temporary weakening from about 1500 to 1000 BP (Zhang et al. 2014). These "phases" of distinct ENSO activity are roughly congruent with those derived by Sandweiss' (2003) examination of anthropogenic shell deposits along the Peruvian coast, and therefore likely provide a highly reliable local index of broad periods of ENSO activity for the area. In light of this fact, it is probable that the modern climatic regime and natural setting—both marine and terrestrial—of coastal Perú has been in place for around 3,600 years, although Late Holocene variations in the strength and periodicity of El Niño impacts on the Peruvian coast remain relatively poorly characterized due to the absence of well-dated, long-term local proxies.

Depositional Setting

Given the aridity of the north coast of Perú, under normal (non–El Niño) conditions sediment transport and deposition over most of the terrain is driven primarily by a combination of marine and eolian processes, although fluvial processes also play a role. Rivers discharge sediment into the Pacific Ocean, and strong upwelling and longshore currents move this sediment northward, redepositing it along the coast. Prevailing winds then transport this coastal sediment northeastward, and immense dunefields migrate across large swaths of the pampa, with longitudinal linear dunes on the lee sides of large cerros giving way to barchan dunes on the open pampa. On the upland cerros, regolith is rapidly produced by extremely intense salt weathering of exposed bedrock—a result of intense diurnal temperature fluctuations evaporating the salty garúa (e.g., Goudie et al. 2002)—to produce grus, which then further breaks down into finer particles. This grus-derived regolith undergoes intense chemical weathering (Huckleberry and Billman 2003) in the form of oxidation, clay coatings, and the formation of evaporites such as caliche, although the resultant regolith is typically poorly consolidated and easily transported by eolian processes, especially on windward slopes of cerros. Eventually, eolian sediments reach either leeward slopes, the western flanks of the Andes (where prevailing winds can push them inland no further), or the next river to the north, where they are carried again to the Pacific Ocean. Given this, pampa dunefields are more prominent in the southern half of each coastal valley.

Under El Niño conditions, this depositional regime is upended, and colluvial and fluvial processes become predominant. Torrential rains rapidly erode the accumulated, unconsolidated regolith atop cerros, and accumulated discharge flows through *quebradas* (ephemeral streams) towards the nearest river, where annual fluvial suspended sediment yield (SSY) in this area sometimes increases sixty-fold during very strong El Niño events (Morera et al. 2017). The

resulting flash flooding causes a number of pronounced erosional and depositional effects on river channels, riparian zones, and the nearby pampa and coast. For example, a single El Niño flood event can cause pronounced channel incision (Huckleberry and Billman; Manners et al. 2007; Tote et al. 2011), avulsion (Huckleberry and Billman 2003), lateral erosion (Manners et al. 2007), and/or aggradation of alluvial fans which dam or redirect rivers temporarily or permanently (Moseley et al. 1992; Sandweiss 1986). El Niño floodwaters also overtop riverbanks, thus aggradationally depositing entrained flood sediments across vast areas of the desert pampa (Keefer et al. 2003; Manners et al. 2007; Wells 1990). At river mouths, the combination of increased flood discharge (Morera 2017; Tote et al. 2011) and higher El Niño–related sea-surface levels rapidly builds large ephemeral alluvial fans, drastically reshaping the local terrestrial, littoral, and subaqueous environments (Moseley et al. 1992; Rogers et al. 2004).

After El Niño conditions abate, riverine sediment flux plummets, and it can take years to return to normal levels after the scouring caused by El Niño floods (Tote et al. 2011). Decades may pass before sufficient deposition narrows river channels to pre-flood dimensions (Manners et al. 2007). Rates of salt weathering, and therefore regolith generation, also take time to rebound due to inheritance effects (Goudie et al. 2002). Longshore erosion of coastal ENSO alluvium gradually feeds coastal progradation (Moseley et al. 1992; Rogers et al. 2004; Rollins et al. 1986), and these coastal sediments then (eventually) act as source material for the dunefields which inundate pampas and riparian zones south of each coastal river (Moseley et al. 1992; Sandweiss 1986). Given this, the geomorphic effects of El Niño events on cerros, quebradas, pampas, rivers, riparian zones, and littoral areas continue to unfold over multiple timescales, making El Niños the primary engine of natural geomorphic change in the area.

The powerful earthquakes of this region (magnitude [M] > 7.5), which have historically occurred roughly twice per century (Dorbath et al. 1990), also feed erosion and deposition, including El Niño–driven processes. For example, tectonic uplift due to earthquakes can change local hydrology by redirecting watercourses (e.g., Ortloff et al. 1983) or accelerating channel incision. Earthquakes also destabilize the regolith and cause structural failures such as landslides, surface cracking, and riverbank collapse, thereby amplifying local erosion and deposition during subsequent El Niño flooding (Keefer and Moseley 2004). Such amplified erosion and deposition has been described as "mass wasting" of the local terrain (Keefer et al. 2003), and historically observed mass wasting events have been directly linked to pronounced spikes in local fluvial erosion and deposition, as well as punctuated increases in coastal progradation in the form of beach ridge-building events (Keefer et al. 2003; Moseley et al. 1992; Rogers et al. 2004; Sandweiss 1986). Tsunamis caused by subduction-zone earthquakes, although much more common further south, also occur on the north coast of Perú, causing marked geomorphic change of affected nearshore areas (Bourgeois et al. 1999).

Cultural Setting

Given the dynamism of the local depositional environment, a number of researchers have argued that ancient earthquakes and/or El Niño floods directly impacted key elements of the infrastructure upon which the subsistence and politico-religious economies of ancient north coastal Peruvians relied. A brief review of the locally relevant arguments is offered here, and partitioned into three themes: 1) posited environmental impacts on coastal subsistence resources, 2) posited environmental impacts on terrestrial subsistence infrastructure, and 3) posited environmental impacts on the domestic sphere, including domestic infrastructure.

The Peruvian north coastal littoral and near-shore marine environment is extremely rich in biomass (e.g., Food and Agriculture Organization of the United Nations 2018), and these resources have been a key component of local diet since first settlement in the area occurred at least 15,000 years ago (Dillehay et al. 2017; Moseley 1975; Tykot et al. 2006). In terms of biomass, the key marine species in this area is the Peruvian anchoveta (*Engraulis ringens*); El Niño conditions in this area result in a precipitous decline in the availability of this species as well as others adapted to the normally cold offshore waters, which in turn effects temporary population crashes in a number of culturally important predator species such as seabirds and seals (Glynn 1988; Jaksic 2001). Availability of culturally important littoral species such as crustaceans and mollusks (e.g., Moseley 1975; Moseley and Feldman 1988; Pozorski 1979; Shady et al. 2001) also dramatically declines during and immediately following El Niño events (Glynn 1988), and studies of ancient molluscan assemblages from local archaeological sites have argued that ancient changes in ENSO activity have in some cases greatly affected distributions of these species over the long term, and therefore their local availability to ancient Peruvians (e.g., Sandweiss 2003). The onset of El Niño conditions on the ancient Peruvian coast would therefore have greatly diminished harvesting of coastal resources over the short term, as well as the short-term cultural redistribution of these resources into the coastal interior, where they were also an important dietary component (e.g., Shady et al. 2001). Over the longer term, the depositional consequences of combined earthquakes and El Niños likely caused abandonment of shellfish harvesting locations in some areas such as the Ostra Base Camp site in the Santa Valley, as littoral resources became locally unavailable (e.g., Sandweiss 1986).

A wide array of terrestrial resources was also used since initial settlement (Dillehay et al. 2017), and by the late Preceramic period an increasing reliance on agriculture, with maize as a

significant dietary component, had caused increased population density in the coastal interior of this area (Haas et al. 2013; Shady et al. 2001). Between roughly 1800 BC and Spanish conquest, terrestrial economic production became more reliant on intensive irrigation agriculture of maize, and large networks of irrigation canals were eventually essential to agricultural production throughout the region (e.g., Billman 2002; Eling 1988; Hayashida 2006; Netherly 1984), reaching their maximum extent by around AD 1400 (Shimada 2000). These networks were most likely built, administered, and maintained through communal labor organized and mobilized by means of a hierarchical kinship system (Netherly 1984), but they were also highly susceptible to damage from El Niño flooding, which typically caused erosion of canals and/or deposition within canals, requiring organized cleanouts, reconstruction, or construction of new intakes after channel incision. Archaeological studies have documented instances of El Niño flooding impacting north coastal irrigation networks; examples include canals in the Lambayeque Valley (Huckleberry et al. 2012) and the Moche Valley (Nials et al. 1979a, 1979b). Other studies have also suggested that uplift events occasionally affect north coastal canal hydrology; Otloff and colleagues (1982, 1983), for example, have argued that the Moche-Chicama Intervalley Canal was damaged by uplift during construction, rendering it useless. Over the short term, damages to canal infrastructure would at minimum necessitate repair and/or replacement of affected canal segments and intakes, especially those near the main river channel and flood-activated quebradas, where flood erosion and deposition would be most pronounced. In cases of severe or ill-timed floods, entire crop yields may have also been lost, exposing local communities to a higher risk of famine and reducing the availability of essential goods such as cotton, which was needed for the manufacture of fishing implements such as nets. When coupled with the increased risk of disease due to contaminated and/or stagnant pools of water (e.g., Gerszten et al. 2012;

Ramirez et al. 2013) which can occur as a result of El Niño flooding, the effects of such food shortages may have occasionally been severe.

Monumental mounds have been extant on the coast of Perú for at least 6,700 years (Dillehay et al. 2017), and urban occupations date as far back as 2600 BC (Shady et al. 2001); many of the hundreds of known ancient coastal settlements incorporated monuments, which acted as local foci for social, political, and religious activities ever since (see Moseley 2001). Most coastal architecture—domestic, monumental, or otherwise—is constructed in whole or in part from unfired earthen materials such as adobe or rammed earth, and many were embellished with sculptural elements and/or elaborate painted murals. Coastal architecture as a whole, then, was highly susceptible to structural damage and defacement during heavy rains, requiring some form of repair before the full resumption of intended use. Published documentation of El Niño damage to ancient north coastal architecture abounds. For example, Dillehay and Kolata (2004) document El Niño damage to various sites in the Jequetepeque and adjacent valleys, Shimada (1994) and Bawden (1996) document El Niño damage to architecture in the Lambayeque Valley and the Moche Valley, respectively, which may represent the same AD 500–600 El Niño event, and Sandweiss and colleagues (2009) document ancient El Niño damage to architecture in the Supe Valley. Documented instances of earthquake damage to buildings are more rare, but nonetheless significant; Sandweiss and colleagues (2009) detail earthquake damage at both Aspero and Caral, for example. When such damage occurred, whether due to El Niños, earthquakes, or both, mobilization of the labor necessary to repair local domestic infrastructure may have been taxing for local politico-religious elites, who probably bore a spiritual responsibility for appeasement of divine forces, as well as the health and prosperity of the community as a whole (e.g., Bawden 1996; Moseley 2001). In cases where El Niño damages

were sufficiently severe, the local population was sufficiently desperate, or the political economies of local elites sufficiently exhausted, social tensions were likely heightened between communities, kin groups, and/or elites.

Ancient Peruvians employed many technologies which may have helped to cope with the impacts of these events, including structures designed to hold floodwaters at bay (Brooks et al. 2005), a fractal-like organization of the local labor force for flexible deployment in times of need (Netherly 1984; Quilter and Koons 2012), adaptation of irrigation networks to the changing environment (Dillehay and Kolata 2004), storage facilities for curation of food in times of need (e.g., Anders 1981; Billman 1996), and extensive trade networks to help diversify resource procurement (e.g., Sayre et al. 2013; Pillsbury 2017; Shimada and Merkel 1991), but it remains an open question whether these strategies were employed and/or effective as mitigative approaches on a case-by-case basis. Certainly, there were cases in which ancient Peruvians successfully coped with the impact of El Niño and earthquake erosion and deposition, but there were also probably cases in which they did not, and a cultural catastrophe ensued. Discerning the occurrences of each scenario demands a detailed understanding of both the natural changes caused by these events, as well as the cultural strategies employed to cope with them. In examining either type of change, detailed long-term records are essential to interpreting the degree to which a given episode of change represents an unusual environmental or cultural event.

Given this need, the remainder of this study attempts to build a record to facilitate deeper exploration of human-environmental dynamics in this area. In this environment, extant sediments resulting from El Niño and earthquake activity provide a clear index of major natural changes, as well as a point of connection between the natural environment and the local resource infrastructure which co-evolved with this environment. Given this, this study examines the

depositional changes in the north coast of Perú by attempting to construct a detailed chronology of earthquake/ENSO deposition along the Santa Valley coastline and El Niño flooding in the southwestern Chao Valley.

FIELD METHODS

Sampling Strategy

The Chao and Santa Valleys were excellent locations for this research due to the presence of geomorphic indices of El Niño and earthquake activity in these areas, as well as their proximity to areas of intensive ancient cultural activity. Although archaeological study of ancient cultural change in this region has traditionally been more focused on the Moche and Chicama Valleys, continuous, long-term geomorphic indices of past ENSO activity are much more limited in these valleys due to extensive terrain alterations resulting from ancient and modern urban and agricultural development (e.g., Proyecto Chavimochic 2017). By contrast, cultural ground disturbance was far less extensive in 2007 (the period of fieldwork) in the Chao[1] and Santa Valleys, making geomorphic indices of past ENSO activity readily available in these areas, which are governed by the same climatic and rainfall regimes and terrestrial depositional environments as the Viru, Moche, and Chicama Valleys within 150 km to the north (Lavado Casimiro et al. 2012; Rau et al. 2017). Further, in many ways the Rio Santa is a significant contributor to long-term depositional change throughout this area. Discharge from the Santa River is the highest in the area, and second-highest on the Peruvian coast as a whole (Lavado Casimiro et al. 2012, 2013). Sediment yield is extremely high in the Rio Santa during El Niño events (Morera et al. 2017), supplying much of the coastal sediment for the Chao, Viru, Moche, and Chicama Valleys to the north, and the subsequent eolian activity eventually derived from

[1] This studied portion of the Chao Valley has since been disturbed by the creation of agricultural furrows.

this coastal deposition. The Santa Valley is also home to extensive and well-studied beach ridges (e.g., Moseley et al. 1992; Sandweiss 1986; Sandweiss et al. 1983; Wells 1996), which provide the most well-documented geomorphic evidence of the interplay between seismic events and El Niño deposition for the region (e.g., Moseley et al. 1992). The Chao Valley, which lies between the Santa and Viru Valleys, contains extant laminar accumulations of El Niño flood sediments which are marginal to zones of cultural disturbance and active fluvial erosion. Together, then, sediment deposits in the Chao and Santa Valleys provide an excellent starting point for examining ENSO-driven flooding and coastal change in the area.

Santa Valley Sample Collection

Within the Santa Valley, the extant beach ridges north of the river mouth were chosen as a focus for this study to enable measurement of changing rates of progradation and ridge formation, thereby providing direct evidence of earthquakes followed by extreme instances of ENSO-driven erosion and deposition (Sandweiss et al. 1983; Sandweiss et al. 1998; Sandweiss 1986; Moseley et al. 1992; see also Rogers et al. 2004)—a rare combination of events which has profound impacts on coastal morphology. Published radiocarbon dates for the earliest ridges (Sandweiss et al. 1983) also allowed independent evaluation of luminescence data, facilitating more robust interpretation of project data.

Santa beach ridge morphology has been extensively discussed elsewhere (Moseley et al. 1992; Sandweiss 1986; Sandweiss et al. 1998; Wells 1996), but in general the Santa ridges are a series of linear topographic features which trend northwest away from the mouth of the Rio Santa, and along the Playa Tres Chozas (Figure 5-1). Inland (older) ridges are larger amplitude, higher volume, and more stratigraphically complex, typically containing multiple distinct depositional events within a given ridge structure, although individual deposits—where visible in

profile—are predominantly poorly sorted, cobble dominant, and clast-supported, with silt/sand matrix between cobbles. Brackish wetlands with tall, dense vegetation occupy topographic troughs which represent relict shorefaces and separate individual ridges or sets of ridges. Seaward (younger) ridges are generally volumetrically smaller, with no visible internal stratigraphy, and often occur in sets of roughly three to eight low individual ridges atop a local topographic high. The most seaward of the Santa ridges often contain no sand/silt matrix between clasts. The change in ridge morphology between the earliest ridges and all later ridges is likely due in part to relative changes in local depositional geometry as a result of tectonic uplift (Sandweiss et al. 1983).

Full estimation of the number of depositional events represented by the Santa beach ridge system was impossible within the scope of this study for three reasons. First, ascertaining the depositional relationships among the seaward ridges is a very complex enterprise; these ridges tend to converge and/or diverge with adjacent within-cluster ridges at some point along their linear extents. The total number of extant ridges therefore varies by distance from the Rio Santa, and detailed geomorphic mapping of the entire ridge complex would be required to fully ascertain depositional relationships among these seaward ridges. This research lacked the technical or financial resources to undertake this type of mapping with sufficient precision and comprehensiveness. Second, the inland ridges are so massive and cobbly that sufficient hand excavation was impossible, and machine excavation was also impossible given available resources, permissions, and concerns for preservation. Comprehensive exposure and study of extant deposits was therefore not possible for these ridges. Third, some large inland ridges are currently used for economic activity such as agriculture and transportation, and it is likely that

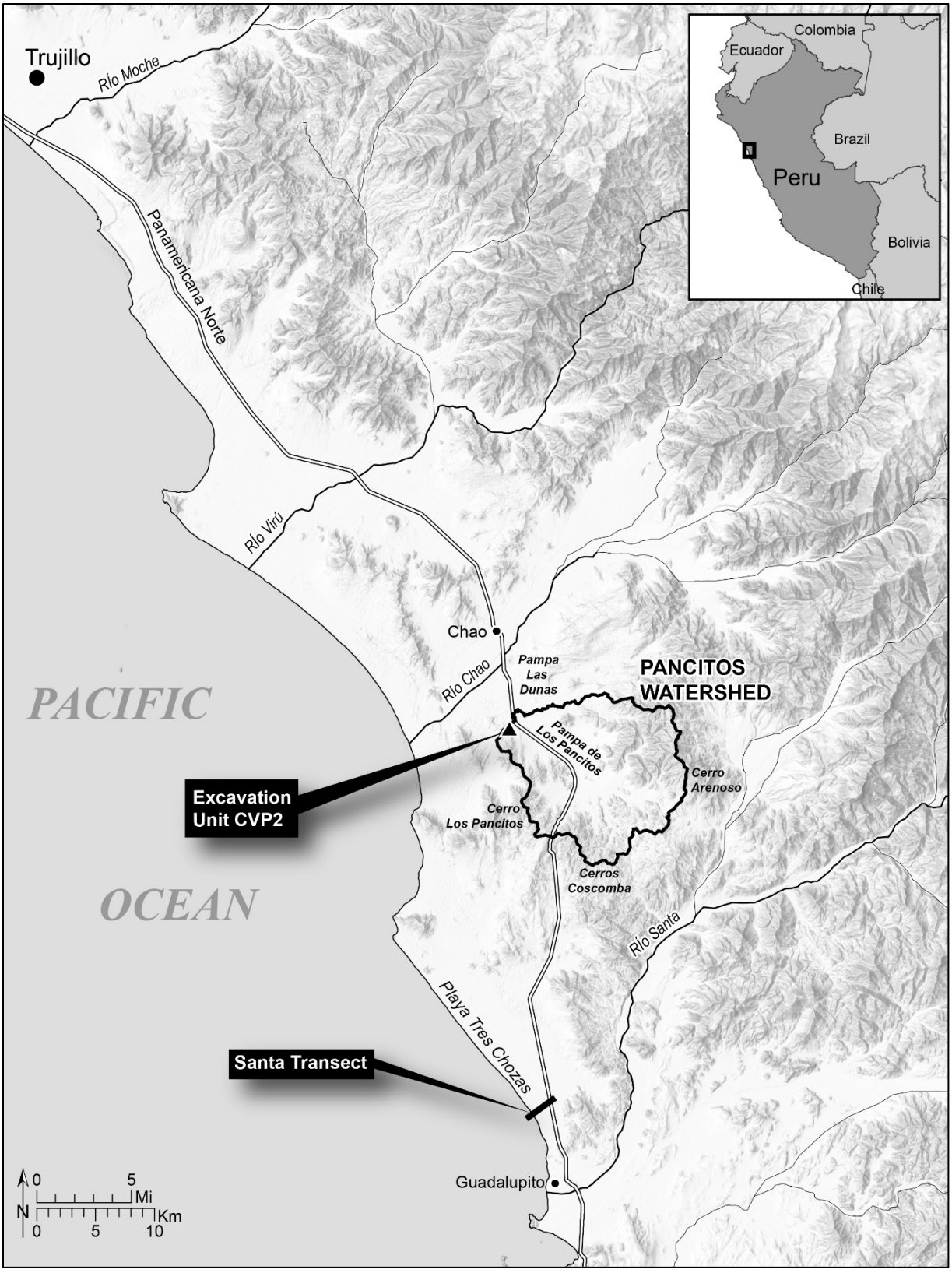

Figure 5-1. The study area in detail, including sampled locations.

modern surface alterations have affected the preservation of stratigraphic integrity atop these ridges.

Because comprehensive systematic examination of depositional events in the area was impossible, surface ridges were surveyed along a transect[2] (see Figure 5-1) to provide a simplified model of discrete phases of local progradation, and therefore aggregate changes in ridge-building activity over time. This simplified model is likely to be a good first-order representation of progradation along the Playa Tres Chozas as a whole, because the volumetric majority of progradational deposition is sub-aqueous rather than sub-aerial and atop the topographic highs of the actual ridges themselves. Further, it must be remembered that extant ridges are not in any case a complete record of individual powerful earthquake/ENSO events due to the subsequent reworking of these deposits (e.g., Moseley et al. 1992). Thus, the primary story of changing aggregate El Niño–driven deposition over time is told by the westward movement of the Playa Tres Chozas rather than the sub-aerial accumulation of the ridges themselves, which provide valuable delineation of the most pronounced episodes of deposition that remain extant in the present.

Fourteen sets of beach ridges were identified along the studied transect[3] (Table 5-1); Ridge 1 in this study corresponds with Wells' (1996) Morphostratigraphic Group II and with the location of the ^{14}C date for early ridge formation reported by Sandweiss and colleagues

[2] The transect began at Panamerican Highway marker km 457 and bore 232° from true north. The easternmost terminus was chosen for its proximity to the sampling location for ^{14}C dates detailed by Sandweiss and colleagues (1983). The westernmost terminus was at the 2007 shoreline. This transect location and bearing roughly approximated the approach used in Wells' (1996) study.

[3] Sandweiss (1986) identifies 8 ridges, and Wells (1996:7) identifies roughly double that number, but does not provide a total count. Differences in ridge counts are likely due to differences in transect locations. Sandweiss (1986:20) used a more northerly transect than this study; the location of Wells' (1996) transect is unknown. Generally, however, more ridges are extant closer to the mouth of the Rio Santa due to the complex interplay between depositional agents in this location and movement of the river mouth over time. Agricultural development after the 1980's prevented use of Sandweiss' transect for this study.

(1983:286). Ridges 2 through 8 gradually become topographically lower and represent Wells' Morphostratigraphic Group III. Ridges 9 through 14 are sets of compound smaller ridges and represent Wells' Morphostratigraphic Group IV. Ridges 11 through 14 trend slightly more northward than earlier ridges and are somewhat concave relative to the modern shoreline as a result of a slightly altered coastal depositional geometry in the studied location after the establishment of Ridge 9.

Table 5-1. Location of Identified Santa Valley Beach Ridges and Samples

Ridge	E	N	m Inland	% Dist. Inland	Samples	MS Group
14	757430	9014036	0	0.0	OS-72851	IV
13	757535	9014108	127	5.0		IV
12	757679	9014217	308	12.2	UW1815, OS-72850	IV
11	757735	9014258	377	14.9		IV
10	757825	9014324	489	19.3		IV
9	757991	9014439	690	27.3		IV
8	758169	9014578	917	36.3	UW1845, OS-72846	III
7	758252	9014615	1005	39.8		III
6	758409	9014754	1214	48.0		III
5	758769	9014984	1640	64.9	UW1844	III
4	758893	9015060	1786	70.7	UW1843, OS-72843	III
3	759044	9015162	1968	77.9	UW1839	III
2	759226	9015294	2193	86.8	UW1840	III
1	759492	9015498	2528	100.0	UW1814, SI-4957 (Sandweiss et al. 1983)	II

Note: Eastings and northings refer to meters in WGS 84 UTM Zone 17L. Meters inland and percent distance inland refer to distances along the observed transect. UW sample numbers are luminescence samples; all others are ^{14}C samples. "MS Group" refers to "Morphostratigraphic Groups" set forth by Wells (1996).

Given the size and composition of the ridges and the difficulties of hand excavation, samples were collected opportunistically from open gravel and/or borrow pits along or near the survey transect when such pits were present and exhibited deposits with a sufficient sand

component to facilitate single-grain luminescence dating[4]. Where open pits were absent but ridges were small enough to allow excavation of sandy sediments, samples were collected from hand-excavated pits placed at the ridges' seaward shoulders (former beach faces) to target the timing of terminal deposition of sampled ridges. A total of seven beach ridge luminescence samples were collected for this study. Each sample represents a distinct beach ridge, and Ridges 1, 2, 3, 4, 5, 8, and 12 were sampled for luminescence dating; luminescence sampling contexts were not available for other ridges due to the aforementioned practical limitations on sampling. A full list of documented Santa ridges and collected Santa samples is provided by Table 5-1.

Radiocarbon samples of unidentified plant debris were also collected from Ridges 4, 8, 12, and 14 to provide independent checks on luminescence data. Samples collected from Ridges 4, 8, and 12 were collected from the same exposures used for luminescence sampling, and the sample from Ridge 14 was collected from within the ridge's cobble matrix, which contained insufficient sand inclusions to facilitate luminescence sampling; this radiocarbon sample was expected to yield a modern date based on prior studies (see Moseley et al. 1992), constraining the overall sequence of dates while providing a check on other ^{14}C samples recovered from earlier ridges of unknown ages. Published ^{14}C dates for the earliest portion of the Santa ridge sequence also provide additional checks on the data produced by this study; the luminescence sample from Ridge 1 was collected from the same profile and stratum ^{14}C dated by Sandweiss and colleagues (1983; Sample SI-4957).

[4] Samples in this context were expected to be poorly bleached (see below), eliminating the possibility of using fine-grained polymineral aliquots for use in luminescence dating.

Chao Valley Sample Collection

The portion of the Chao drainage examined in this study sits at about 8°36'46"S latitude, 78°40'12"W longitude within the western part of the coastal plain known as the Pampa los Pancitos (see Figure 5-1). This area is a visible flood surface located at the eastern foot of Cerro los Pancitos, about 3.3 km southeast of the riparian zone. The area lies at about 30 m above the nearest (southern) bank of the Rio Huamansaña, and a large longitudinal northeastward-moving dune to the north acted as a topographic barrier between the flood surface and the river itself (Figure 5-2), impounding floodwaters and making the immediate area an ephemeral endorheic basin fed by slopewash from a relatively small (157.4 km^2) watershed until 1920s construction of the Panamerican Highway bisected the dune, allowing subsequent El Niño floodwash to more easily reach the Huamansaña.

Figure 5-2. The Pampa las Dunas (left) meeting the Pampa los Pancitos (center and right). View facing northeast, overlooking the location of CVP5 (not visible). Linear dune pictured at left extends northeastward from the northern spur of Cerro los Pancitos to asparagus fields across the Panamerican Highway; Pampa los Pancitos alluvium (light-colored surface at center and right) was deposited by El Niño floodwaters impounded by this dune.

Between formation and bisection of the dune, however, this portion of the Pampa los Pancitos provides a relatively continuous aggradational sequence of El Niño flooding of its immediate catchment, which confers several practical benefits to excavation and analysis of the area. First, excavation is easier here than in flood sediments fed by larger watersheds, because flood sediment layers are simply less voluminous in this context. More interestingly, since this catchment is also spatially removed from the nearest river, it is highly likely to have been relatively insulated from the erosive effects of the most powerful El Niño floodwaters, especially at its topographic margins, where no ephemeral channels are extant. Of course, some erosion—deflation due to eolian activity, for example—is still occurring even in this "sheltered" portion of the Chao Valley, and these sediments cannot be considered entirely free of post-depositional alteration. Nor are they a perfectly complete record of all sediments deposited by all floods, as there should be a clear bias in favor of the preservation of the thickest sediments representative of the most powerful events. On the other hand, such gradual processes are unlikely to significantly bias relative accumulation rates of sediments in this context, and local El Niño strength as manifested by local rainfall and flash flooding is almost certainly the primary depositional/erosive agent governing this system over the long term.

A total of five excavation units were placed in the Chao study area (see Figure 5-1) to examine sediment stratigraphy and provide a basis for selection of optimal sampling conditions. Unit CVP2 was selected for intensive sampling due to its abundant extant flood sediments, its marginal location and spatial distance from extant ephemeral channels, and the fact that maximum flood sediment depth was about 2 m in this location, allowing for hand excavation and sampling of a continuous sequence of all extant local flood sediments. Given this, CVP2 will be the focus of discussion for the remainder of the discussion of the Chao Valley in this study.

CVP2 was excavated in natural levels by gradually peeling away blocks of cohesive flood sediment to reveal the underlying strata. A pedestal of each stratum was preserved in situ to allow interpretation of the full sequence of flood deposits in the finished excavation unit. Occasionally, strata exhibited strong evidence of prior surface exposure in the form of surface cracking infilled by subsequent flood sediment (Figure 5-3) and/or a hardened, buckled cap due to surface desiccation and subaerial weathering. A total of 21 flood strata were encountered, implying that at least 21 distinct flood events are represented by the CVP2 sequence (Figure 5-4). A total of 21 luminescence samples was collected and analyzed from this sequence; one for each observed pre-modern flood surface, along with a basal sample of non-flood sediment to help constrain the initiation of the flood sequence. Each sample was collected at a recorded depth to allow estimation of gross accumulation rates in addition to the timing of each distinct flood deposit.

DATING METHODS

Radiocarbon samples were sent to the NOSAMS laboratory at Woods Hole for AMS analysis. Calibration used the SHCal13 curve (Hogg et al. 2013). All luminescence preparation and analysis was performed at the University of Washington Luminescence Dating Laboratory. The luminescence dating process as a whole involves several steps; these are discussed in detail below.

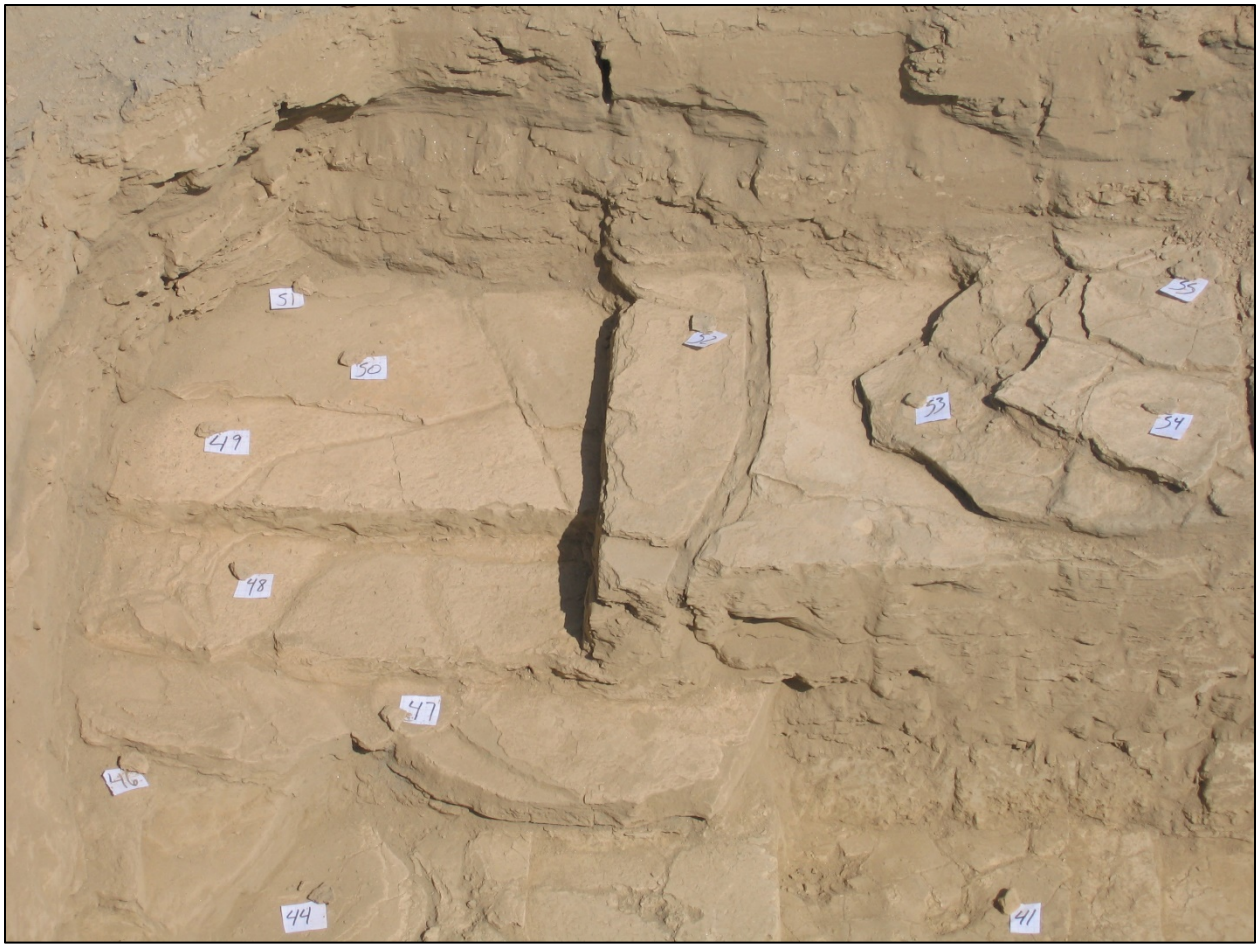

Figure 5-3. View of laminar Chao Valley flood sediments and surface cracking revealed during excavation of CVP2. Number labels denote individual peds identified during excavation; flood surfaces were inferred from visible surface cracks, and most contained multiple peds. Number label 52 sits adjacent to a visible surface crack infilled by subsequent flooding; this surface was designated flood surface #F17 and sampled as UW2217.

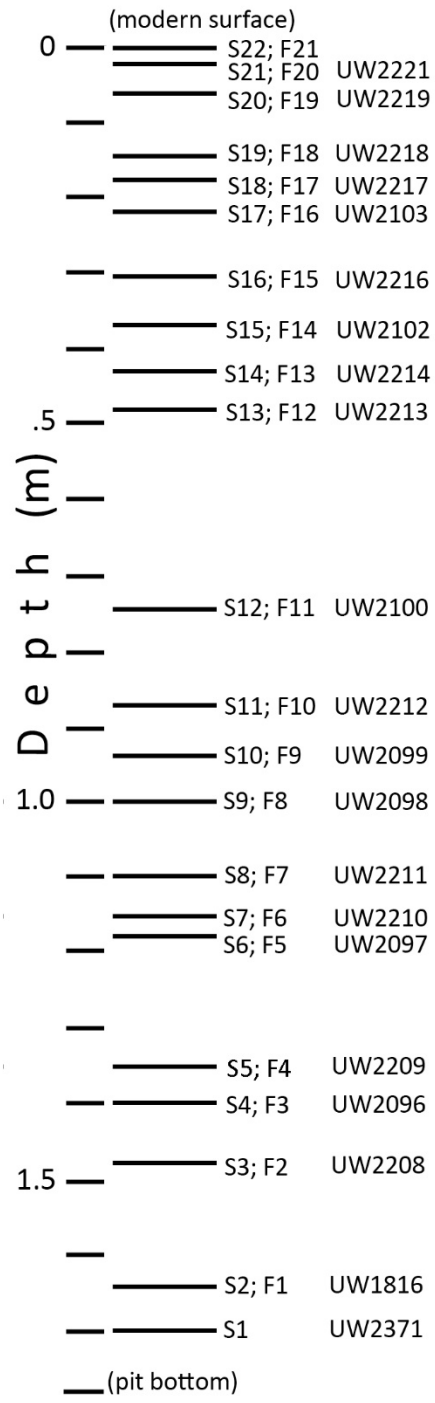

Figure 5-4. Observed former surfaces versus CVP2 excavated depth, indicating the locations from which luminescence samples were collected.

Equivalent Dose Measurement

The term "equivalent dose" (D_e) is the numerator of the luminescence age equation Age = D_e/D_r. This term refers to the amount of radiation, measured in Grays (Gy), necessary to account for the intensity of the luminescence signal observed in a given sample. Estimation of D_e is itself a complex process which can proceed along different pathways depending on sample composition, inferred depositional agency, instrumentation, or other variables. Alternate pathways confer distinct costs and benefits in terms of analysis time, data resolution, and in some cases the accuracy of age estimates.

In the case of sampled Chao flood sediments, three complementary ways of estimating D_e were attempted, each with inherent strengths and weaknesses. These methods differ primarily in the portion of bulk samples used for analysis—the first strategy used fine-grain (1–8 μm) polymineral aliquots, the second strategy used single-grain (90–212 μm) quartz aliquots, and the third used single-grain (90–212 μm) potassium (K) feldspar aliquots. Initial testing showed the second strategy to be fruitless, as tested quartz aliquots yielded insufficient luminescence signals in this mineral fraction[5]. This result fits with past experience showing general insensitivity of Peruvian quartz to luminescence (Feathers et al. 2008; Huckleberry et al. 2012), and given this no further measurement of single-grain quartz was pursued for these samples. All subsequent discussion will therefore focus on the use of fine-grained polymineral aliquots and single-grain K feldspar aliquots.

Fine-grain dating has been successfully used in some cases to date laminar fine-grained waterborne sediments (e.g., Londoño et al. 2012; Waters et al. 2011), and given this was applied

[5] Single-grain quartz dating was tested on samples UW2100 and UW2103 (100 aliquots each); no aliquots yielded luminescence signals which met requirements set forth by standard protocols set forth by Murray and Wintle (2000).

here. Importantly, this method also enjoys at least two practical advantages over single-grain dating for analysis of Chao flood sediments. First, the composition of these sediments is predominantly fine-grained, so this fraction is the most abundant for analysis. Second, fine-grained aliquots, as composites of multiple minerals and multiple individual grains, are relatively likely to contain materials with measurable luminescence signals. In other words, luminescence measurement of fine grains yields D_e data at a high rate of return (Chao aliquots produced useful D_e measurements at a rate of 97%), especially in contrast to the low rate of return often observed during single-grain dating of Peruvian sediments (under 20% for Chao sample single-grain feldspars). This, in turn, confers obvious benefits in terms of efficient measurement of large numbers of samples, and given this it is worthwhile to assess the potential of fine-grain dating for future use in dating coastal Peruvian floods.

On the other hand, in some cases fine-grained dating can mask meaningful variability within samples and lead to inaccurate results. If, for example, sediments were not homogenously exposed to sunlight at the time of deposition, fine-grained aliquots will yield elevated D_e estimates due to the conflation of grains "bleached" at the time of deposition and grains retaining older, "unbleached" luminescence signals. Resultant fine-grain ages will in turn be too old. Importantly, sediments resulting from energetic flows are generally likely to exhibit such "partial bleaching," including many flood sediments. Ultimately, however, it must be empirically determined whether these Chao sediments, which represent relatively low-energy flood deposits given their small catchment and peripheral location, face this issue.

Single-grain dating allows this empirical determination by generating D_e data for each analyzed grain of sand, which in turn facilitates the identification and statistical isolation of grains which were fully exposed to sunlight at the time of deposition. Thus although more

intensive as a process, single-grain dating can help diagnose partial bleaching while producing age estimates robust to this problem. This method was therefore used when possible in conjunction with fine-grained dating to help inform the extent to which sampled sediments were exposed at the time of deposition, and to allow direct comparison of the utility of each approach for fine-grained flood sediments in coastal Perú. For this reason, single-grain dating of K feldspars, which often contain useful luminescence properties even when local quartz does not, was also employed for dating of the Chao Valley samples. The Santa Valley beach ridge samples were dated exclusively with this technique due to their sandy composition and the apparent lack of sensitivity of the local quartz.

Protocols for preparation, measurement, and analysis are detailed below for each method. All stages of D_e preparation were conducted under filtered red light in a photographic dark room. Exterior portions of samples likely to have been exposed to sunlight during collection were removed from use in D_e measurement and set aside for use in dose rate (D_r) estimation.

Polymineral Fine-grained Aliquots

Exposed portions of samples were removed for dose rate estimation (see below). Unexposed portions of samples were treated with washes of 20% HCl and 30% H_2O_2 in sequence to remove carbonates and organic matter. Importantly, no samples reacted chemically to these treatments, meaning no organic matter was available for radiocarbon dating of bulk sediments to serve as an independent check on luminescence dates. The 1- to 8-µm fraction was then isolated by settling of treated samples in an acetone suspension and decanting at intervals specified by Stokes' Law (Gibbs et al. 1991). This fraction was then settled onto discs for measurement. Each disc is one aliquot for use in D_e measurement, and D_e data for each aliquot represents an average value for all grains on the disc.

Measurement followed the double "single-aliquot regeneration" (double SAR) protocol set forth by Bannerjee and colleagues (2001; see also Roberts 2007). Measurements made use of a Risø model TL-DA-15 TL/OSL reader, a 240°C 1-minute preheat, a 30-sβ test dose, OSL stimulation at 125°C via 100 s of exposure to an 80% strength blue diode, IRSL stimulation at 60°C via 100 s of exposure to a 90% strength infrared diode, a Hoya U-340 UV filter, and a regeneration dose schedule of 160, 80, 240, 320, 400, 0, and 160 sβ. Fine-grained aliquots were measured for 20 samples total, including all collected Chao Valley samples except UW2371, as the coarser-grained and less well-sorted composition of this sample undermines the assumption that it was well bleached during deposition. Between 22 and 56 fine-grained aliquots were measured for each Chao sample. B-values were measured on the quartz fraction of a subset of aliquots (2 to 10 per sample) using the same instrument settings after 100 minutes of exposure to an Am-241 α source; mean B-values of 1.05 +/- 0.06 were calculated following Lai and colleagues (2008) and used for the estimation of quartz fraction of the fine-grain Chao Valley sample D_e.

Again, Santa Valley beach ridge samples were coarse-grained and therefore unsuited to this approach; only single-grain feldspar measurement was undertaken for Santa samples.

Single-grain Feldspar Aliquots

Coarse grains for single-grain dating were prepared by sieving the unexposed fraction of samples to isolate grains between 90 and 212 μm in diameter. Typically, a narrower range of grain sizes such as 180 to 212 μm is preferred, as this helps reduce variance in dose rate estimation (Aitken 1985). In this case, however, the fine-grained composition of bulk samples forced the use of a broader grain-size range in order to obtain sample volumes sufficient for analysis. Once isolated, the 90- to 212-μm fraction was exposed to washes of HCl and H_2O_2 as

above. Again, neither carbonate nor organic material was present. K feldspar grains were isolated using an LMT heavy liquid with a specific gravity of 2.58. Feldspar grains were then placed on discs machined to accommodate 100 aliquots each for measurement. Here, each aliquot therefore represents either a single grain (for grains around 212 μm in size) or a composite of two or three grains (as multiple 90 μm grains will sometimes fall into disc aliquot dimples). Functionally, however, the resolution of measurements is effectively single-grain, as the low overall rate of measured aliquots responsive to IRSL dating (under 20%) made consistent co-occurrence of responsive grains within an aliquot unlikely. In other words, even if some aliquots consist of two or three grains, the vast majority of data still reflect single-grain resolution, allowing for the assessment of sample bleaching during analysis.

D_e values of feldspar grains were measured with infrared-stimulated luminescence (IRSL) using the SAR protocol (Auclair et al. 2003; Murray and Wintle 2000, 2003) on a Risø DA-20 TL/OSL reader, a 250°C 1-minute preheat, a 50-sβ test dose, stimulation at 50°C via 100-s exposure to an 830-nm IR laser passed through an RG-780 filter, a 7.5-mm blue filter applied to emissions, and variable regeneration dose schedules, ranging from 0 sβ up to 200 sβ as a maximum administered dose. A total of about 4,400 single grains from 17 Chao Valley samples (100–400 grains per sample) and 3,900 grains from 7 Santa Valley samples (200–700 grains per sample) were measured for single-grain D_e data. Many Chao Valley samples yielded insufficient material for single-grain dating. In addition, many of the 17 Chao Valley samples measured yielded very few aliquots with useful luminescence signals, and given this statistical sample size is a possible issue for some sample ages discussed below.

Dose Rate Measurement

Dose rate (D_r), the denominator of the age equation, is an estimate of the amount of radiation (in Gy) to which each sample was exposed annually. The total budget of relevant radiation includes both cosmic and terrestrial sources. Estimates of annual cosmic radiation doses were derived from recorded sample global location and burial depth following Prescott and Hutton (1994). Terrestrial radiation, which is composed of alpha (α), beta (β), and gamma (γ) types, was measured for exposed portions of bulk samples in a variety of ways. Samples were prepared for this battery of measurements by first drying them at 50°C for 48 hours. Moisture content before and after drying was measured to allow quantification of attenuation of radiation as a function of moisture. Dried samples were then pulverized into a fine flour with a tungsten-carbide rock mill; this homogenized samples and facilitated α counting.

Alpha radioactivity was measured directly using either Daybreak or Littlemore α counters to record total, fast paired and slow paired emissions over 96-hour window following Aitken (1985). Estimated concentrations of uranium (U) and thorium (Th)— two elements whose radioactive isotopes contribute significantly to D_r estimates, especially for fine-grained samples—were derived using the pairs technique from α counts. Estimates of α, ß and γ activity for the U and Th decay chains were then calculated from these estimates.

Beta radioactivity was counted directly for four aliquots of each sample using a Risø GM-25-5 Beta Multicounter System and a 24-hour counting window. In addition, because potassium-40 (^{40}K) is a major source of beta radiation, particularly for K feldspars, K concentrations were measured directly using X-ray fluorescence, and ^{40}K concentrations were derived using natural abundance information. A Bruker Tracer III-V with a 1-mil titanium filter and partial vacuum, a 20-μA stimulation at 15 keV, and 1-minute assays was used for this.

Protocols matched those described in Johnson (2014), and each sample was measured 30 times and repeated local calibration and interpolation were performed using statistical resampling.

The ß dose rate derived from beta counting and from alpha counting/flame photometry were compared to verify harmonious convergence and to check for secular disequilibrium at the top of the ^{238}U decay chain, as this situation can affect dosimetry. No significant differences in ß dose rate was observed, and final D_r estimates were derived using formulae set forth by Aitken (1985).

Fading Correction

Often feldspars exhibit "anomalous fading" (Huntley and Lamothe 2001; Wintle 1973). When present, this phenomenon spontaneously diminishes the natural luminescence signal over time, resulting in systematic underestimates of D_e, and therefore sample ages. Consequently, it was necessary to identify and correct for anomalous fading in the single-grain dating of Chao and Santa samples. Fading for all single-grain aliquots was measured following protocols set forth by Auclair et al. (2003) using storage times of up to 10 days. Fading correction followed Huntley and Lamothe (2001). By contrast, fine-grained aliquots incorporate both quartz and feldspar but are exposed to analytical conditions in the double SAR protocol which attempt to eliminate the effects of feldspathic signals upon quartz measurements, and thus fine-grained measurements are not necessarily influenced by anomalous fading. Given this, fading tests were only performed for the quartz fraction of a limited subset of fine-grained aliquots (n = 44); no fading was detected for the quartz fraction, indicating quartz was successfully isolated in fine-grained D_e measurement.

Dose Recovery Testing

Dose recovery testing following Murray and Wintle (2003) was performed to 1) demonstrate protocols used in D_e measurement derive accurate results for these samples, and 2) characterize the latent degree of inter-aliquot variation expected from fully bleached single-grain samples, often reported as the sample overdispersion[6] (or %OD) value for a sample. Such characterization is of critical importance when applying luminescence dating to depositional contexts in which full bleaching in antiquity cannot necessarily be assumed (Thomsen et al. 2007), as is the case for contexts sampled here. After removal of the natural luminescence signal using the same stimuli detailed above for De measurement, aliquots were exposed to 50-sβ irradiation, then D_e measurement following original protocols tested whether these protocols successfully "predicted" the administered dose.

Thirty fine-grained aliquots were measured in this way, including at least one aliquot for each of the 20 samples used for fine-grain dating. Dose recovery testing of Chao Valley fine-grain aliquots successfully recovered experimental doses (mean 48.90 sβ +/- 1.73), indicating D_e measurement protocols accurately estimated luminescence inherent in these samples, and revealed a baseline OD of 9.1% for fine-grain aliquots. Single-grain material was less available for dose recovery testing of Chao Valley flood sediment samples, as sand grains from all but six of these samples were exhausted during D_e measurement; dose recovery testing was therefore limited to one disc (100 aliquots each) of these six samples and two discs of UW2371. Dose recovery testing again successfully returned the known administered dose (mean of 49.92 sβ +/- 1.55) for single-grain Chao Valley Samples, with a mean OD value of 15.1% +/- 3.2. Dose recovery testing was performed on 100 aliquots each for all Santa Valley beach ridge single-

[6] A measure of scatter in sample aliquot D_e statistical distribution.

grain samples, and was likewise successful (mean of 49.34 sβ +/- 1.54), while exhibiting a mean OD value of 18.6% +/- 3.9. A summary of dose recovery data is displayed in Table 5-2.

Table 5-2. Dose Recovery Data

Sample	Location	Type	n Aliquots		Recovered Dose		%OD
			Input	Valid	sβ	1σ	
Varous	Chao	FG	30	30	48.90	1.73	9.1
UW2099	Chao	SG	100	87	48.98	1.11	16.6
UW2208	Chao	SG	100	44	52.22	1.58	11.6
UW2211	Chao	SG	100	28	50.56	1.90	14.6
UW2216	Chao	SG	100	39	48.90	2.59	17.7
UW2217	Chao	SG	100	49	50.03	1.09	10.8
UW2218	Chao	SG	100	40	47.64	1.75	19.7
UW2327	Chao	SG	200	124	51.13	0.89	14.8
UW1845	Santa	SG	100	31	52.10	1.66	13.1
UW1844	Santa	SG	100	35	48.39	1.23	16.7
UW1843	Santa	SG	100	33	47.80	1.50	21.6
UW1840	Santa	SG	100	48	50.53	0.92	17.9
UW1839	Santa	SG	100	56	49.47	1.37	23.0
UW1815	Santa	SG	100	25	49.11	1.32	15.2
UW1814	Santa	SG	100	42	47.98	2.12	22.7

Note: All samples were administered an experimental dose of 50-s exposure to beta radiation. “Valid” aliquots are those which met requirements of the SAR protocol used, yielding useful luminescence data. Observed overdispersion values are at right.

Age Estimation

Multiple statistical models are commonly used for the estimation of luminescence sample ages (e.g., Galbraith 2005; Galbraith and Roberts 2012; Galbraith et al. 1999); deployment of the appropriate technique for a given sample of sediment depends on a suite of theoretical and empirical criteria related to depositional context, sample composition, and observed overdispersion. The “central age” model assumes differences between aliquots to be the product of statistical variance between aliquots of identical depositional ages, and is therefore most

appropriate to depositional contexts in which thorough bleaching[7] of sediments can be inferred from the depositional context, in cases in which observed sample OD is commensurate with OD values documented through dose recovery testing, and in cases in which polymineral fine grain aliquots are used[8]. When depositional context indicates rapid, high-energy deposition, or when observed sample OD values are high relative to OD values observed in dose recovery testing, techniques such as the "minimum age" model or the "finite mixture" model are more appropriate.

The finite mixture model is appropriate when age distributions exhibit statistically significant multi-modality as a result of the conflation of multiple populations of single-aged grains[9]; this method also allows the statistical isolation of the subset population of well-bleached grains for use in age estimation. The minimum age model is best applied when the finite mixture model is not applicable for statistical reasons, including instances in which aliquot D_e and/or age[10] distributions do not exhibit sufficient multimodality for the finite mixture model to identify and isolate distinct statistical populations of single-aged grains. The minimum age model is better-suited to address instances in which the distribution of these values is relatively continuous but skewed due to partial bleaching[11], exhibiting a dominant mode representative of most recent exposure to sunlight and a tail towards older ages, representing partially bleached aliquots.

[7] Resetting of the luminescence signal at the time of deposition, such that the experimentally observed luminescence signal is the product only of accumulation since the time of deposition.

[8] The use of fine grain aliquots analytically conflates the luminescence signals of multiple grains during measurement, embedding the assumption that all grains were equally bleached at the time of deposition.

[9] Grains which were fully exposed at the time of deposition, producing a normal statistical distribution of aliquot D_e and/or age values.

[10] For quartz aliquots, the D_e distribution is used for this analysis. For feldspar aliquots, fading-corrected age distributions are used instead.

[11] Grains which were exposed to an extent sufficient to remove some (but not all) prior luminescence signal, with some residual signal remaining, inflating apparent luminescence age of deposition.

Each model includes diagnostic tools to explore the degree to which the model fits aliquot distribution data given expected %OD values of well-bleached samples. The task of the luminescence analyst is to use depositional information as well as empirical D_e and/or age data to infer, document, and diagnose the degree of bleaching of each sample and apply the appropriate model on a case-wise basis.

In the case of the Santa Valley beach ridge samples, the depositional context clearly suggested a high likelihood of incomplete bleaching as a result of sub-aqueous entrainment and high depositional energy. Overdispersion data (see Table 5-2), however, illustrate that some Santa samples (UW1814, UW1839, and UW1843) exhibit the hallmarks of well-bleached distributions, with age distribution OD values (22.0% OD on average) roughly commensurate with OD values observed during dose recovery testing. This likely reflects a situation in which grains were well bleached before or during erosion and entrainment and prior to high-energy deposition in beach ridges. The central age model was therefore used to estimate ages of these samples. The remaining Santa samples (UW1815, UW1840, UW1843, UW1844, UW1845) exhibited high OD values (54.7% on average). Such high OD values are likely indicative of the presence of multiple age populations within these samples, and given this the finite mixture model was applied using 20% assumed OD values[12] to mimic values for single-aged populations observed during dose recovery testing[13]. Application of this model identified distinct age populations within each sample for which it was used; the youngest age component of each

[12] The finite mixture model was also applied using assumed values of 25% and 30% OD to test model robusticity. The model successfully identified multiple statistically distinct populations of grain ages for all samples at 30% OD, indicating robust application of the model.

[13] The assumed 20% single-grain OD value is slightly higher than the 18.1% value observed during dose recovery testing to ensure that model application does not falsely impose specious structure upon aliquot distributions.

sample best represents most recent exposure—that which took place during ridge formation—and was therefore used for sample age estimation.

In the case of the Chao Valley flood samples, D_e measurement was performed on two types of aliquots, including the OSL signal of polymineral fine-grained aliquots and single grains of K feldpsars. For fine-grained aliquots, the central age model was applied as above in instances for which observed sample OD was commensurate with OD values observed during dose recovery testing. The fine-grained component of samples UW1816, UW2098, UW2099, UW2102, UW2208, UW2211, UW2214, and UW2217 was dated in this way. Because luminescence measurement of fine-grained aliquots does not yield the single-grain precision necessary to apply the finite mixture model, the minimum age model was applied when OD values of fine-grained samples were high; the fine-grained component of all remaining Chao Valley flood samples was dated using this model. Dating of the single-grain K feldspar component of the Chao Valley samples was performed as detailed for the Santa Valley samples above. Only one Chao single-grain sample, UW2371 (the pre-flood surface), exhibited a low OD value, allowing application of the central age model. Three other samples—UW1816, UW2217, and UW2218—exhibited high OD values but did not exhibit multiple single-aged populations of grains in the finite mixture model; the minimum age model was therefore used to date these samples. The remaining Chao Valley single-grain samples were dated using the youngest component of the finite mixture model as described above.

LUMINESCENCE RESULTS

Table 5-3 gives details on age model used and resultant estimated ages for all Chao Valley samples, and Table 5-4 provides the same details for Santa Valley samples.

Table 5-3. Luminescence Results for all Chao Valley Samples, Including Both Single-grain Feldspar and Fine-grained Polymineral Dates

Sample	Provenience			Polymineral Fine-Grain							Single-Grain K Feldspar						
				Aliquots	Central Age			Estimated Sample Age			Aliquots	Central Age			Estimated Sample Age		
	Surface#	Flood#	%Depth		Date	Error	%OD	Date	Error	Type		Date	Error	%OD	Date	Error	Type
UW2221	21	20	1.2	22	-160	180	22.8	810	520	Minimum	25	880	350	33.1	1830	70	Finite
UW2219	20	19	4.1	22	-2990	550	20.1	-1750	560	Minimum	85	280	190	79.0	1520	70	Finite
UW2218	19	18	8.8	22	-8330	1780	37.5	-3620	570	Minimum	59	1290	50	31.9	1490	60	Minimum
UW2217	18	17	10.6	22	-1390	310	11.2	-1390	310	Central	63	1080	60	27.5	1380	90	Minimum
UW2103	17	16	12.9	22	-190	240	38.1	1360	80	Minimum	26	-1070	400	43.5	270	600	Finite
UW2216	16	15	18.2	22	-820	210	25.6	-830	190	Minimum	40	540	150	50.9	1140	380	Finite
UW2102	15	14	21.8	22	-1060	310	8.6	-1060	310	Central	63	280	170	51.6	1070	170	Finite
UW2214	14	13	25.3	22	-2280	410	12.5	-2280	410	Central	0		N/A			N/A	
UW2213	13	12	28.8	22	-4100	470	32.3	1060	680	Minimum	42	-70	230	51.9	1120	210	Finite
UW2100	12	11	43.5	22	-1080	360	43.1	-290	170	Minimum	14	-540	660	67.6	790	370	Finite
UW2212	11	10	51.2	22	-9200	2100	39.8	-3780	1610	Minimum	0		N/A			N/A	
UW2099	10	9	55.3	22	-210	150	6.9	-210	150	Minimum	102	-300	130	36.8	470	190	Finite
UW2098	9	8	58.8	22	-2730	410	13.3	-2730	410	Central	56	-1980	280	35.9	-200	310	Finite
UW2211	8	7	64.7	22	-4700	530	6.0	-4700	530	Minimum	54	-3040	500	60.0	620	210	Finite
UW2210	7	6	68.2	22	-5260	900	21.4	-2700	750	Minimum	87	-2690	260	34.0	130	530	Finite
UW2097	6	5	69.4	22	-5140	810	40.2	-1400	310	Minimum	0		N/A			N/A	
UW2209	5	4	79.4	22	-12200	1660	20.8	-7860	1250	Minimum	22	-3430	940	35.7	470	380	Finite
UW2096	4	3	82.4	22	-1110	260	22.5	60	190	Minimum	0		N/A			N/A	
UW2208	3	2	87.1	22	-4590	610	10.3	-4590	610	Central	63	-1880	360	53.8	410	180	Finite
UW1816	2	1	97.1	55	-1460	230	11.1	-1460	230	Central	28	-1750	270	41.3	270	140	Minimum
UW2371	1	N/A	100.0	0		N/A			N/A		61	-1800	160	20.7	-1800	160	Central

Note: Estimates derived from the central age model are shown for all samples for reference. “Estimated Sample Age” reflects the best age estimate for each sample fraction after the application of appropriate statistical models as discussed in the text. Negative numbers reflect years BC. Error terms are 1σ. The modern surface (S22, F21) was not sampled and is therefore excluded.

Table 5-4. Luminescence and Radiocarbon Results for all Santa Valley Samples

Sample	Provenience		Radiocarbon Samples				Luminescence Samples (Single-Grain K Feldspar Aliquots)						
	Ridge No.	% Distance Inland	^{14}C Years BP	Error	Cal Date	Error	Aliquots	Central Age			Estimated Sample Age		
								Date	Error	%OD	Date	Error	Type
OS-72851	14	0.0	0	N/A	1980	30	N/A		N/A			N/A	
UW1815	12	14.4	N/A		N/A		76	1510	50	77.8	1890	30	Finite
OS-72850	12	14.4	240	25	1720	70	N/A		N/A			N/A	
UW1845	8	37.4	N/A		N/A		68	1440	50	62.1	1720	40	Finite
OS-72846	8	37.4	230	25	1720	70	N/A		N/A			N/A	
UW1844	5	65.0	N/A		N/A		90	600	80	69.0	600	80	Finite
UW1843	4	71.6	N/A		N/A		93	10	70	20.9	10	70	Central
OS-72843	4	71.6	2290	35	-330	70	N/A		N/A			N/A	
UW1839	3	77.8	N/A		N/A		124	-450	150	21.4	-450	150	Central
UW1840	2	87.5	N/A		N/A		135	-230	110	40.2	-300	170	Finite
UW1814	1	100.0	N/A		N/A		143	-1920	160	23.6	-1920	160	Central
SI-4957	1	100.0	4235	115	-2820	190	N/A		N/A			N/A	

Note: Estimates derived from the central age model are shown for all luminescence samples for reference. “Estimated Sample Age” reflects the best age estimate for each luminescence sample after the application of appropriate statistical models as discussed in the text. Negative numbers reflect years BC. Error terms are 1σ.

Chao Valley polymineral fine-grain aliquots yielded an array of dates which are stratigraphically incompatible (Figure 5-5), and subject to enormous uncertainty on several occasions. This result implies fine-grain samples of Chao flood sediments is unreliable overall, although fine-grain dates from individual samples may be viable if additional data from other methods are available to corroborate them.

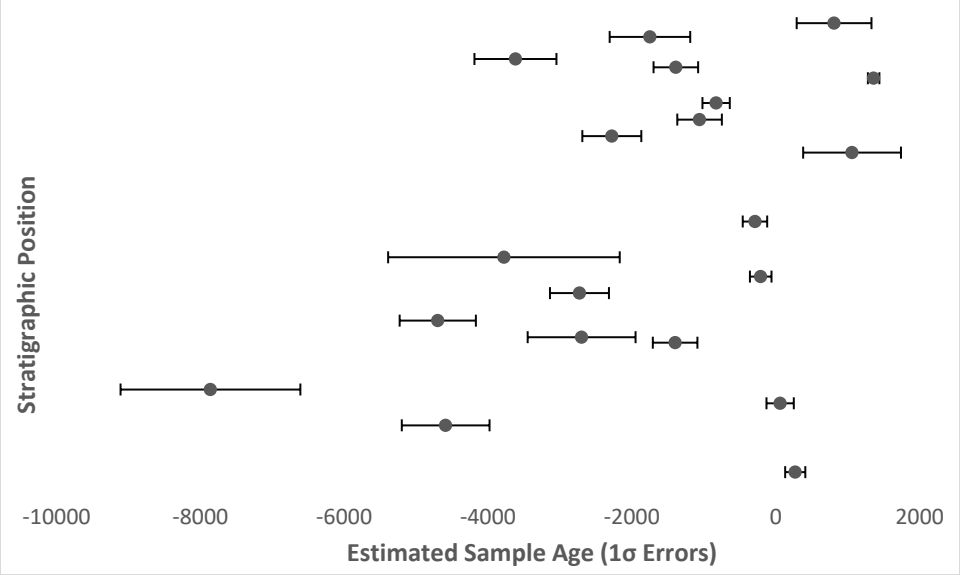

Figure 5-5. Fine-grained polymineral luminescence dates for all Chao Valley flood samples versus stratigraphic position, with stratigraphically lower samples at bottom. Dates derived using this method are clearly stratigraphically incompatible.

Chao Valley single-grain feldspar dates are much more stratigraphically consistent (Figure 5-6). High uncertainties nonetheless plague several of these samples, likely due to the low number of aliquots available for measurement as a result of the overall fine-grained composition of sampled flood deposits. Still, these dates provide a promising basis upon which to interpret Chao Valley flood activity using Bayesian modeling of the overall depositional sequence (see below). The superior performance of single-grain dating in the case of the Chao Valley samples highlights the value of single-grain resolution for dating alluvial deposits, even when these deposits are well sorted and stem from relatively low-energy entrainment and deposition in small catchments, and stands in stark contrast to other documented instances in

which the use of fine-grain aliquots was sufficient for the dating of similar deposits (e.g., Waters et al. 2011).

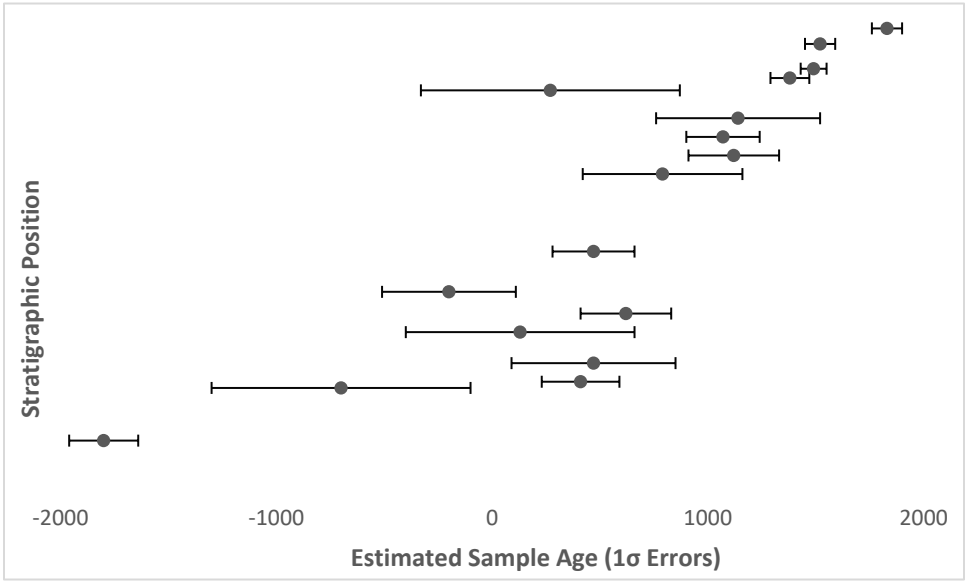

Figure 5-6. Single-grain feldspar luminescence dates for all Chao Valley flood samples versus stratigraphic position. Dates derived using this method show better stratigraphic compatibility than dates derived from fine-grained polymineral aliquots. Negative values represent years BC.

Santa Valley luminescence dates are compatible across the observed transect (Figure 5-7), exhibiting no reversals within the progradational sequence of Santa Ridges. Further, these dates are also compatible with available ^{14}C dates. For cases in which paired luminescence and ^{14}C dates are available for a given ridge, in three instances—Ridges 1, 4, and 12—the ^{14}C dates are clearly older than the luminescence dates, but this gap is not surprising due to the detrital nature of the sampled organic material in this context. In the other case of paired dates—Ridge 8—the ^{14}C dates and the luminescence dates are essentially identical. Luminescence dates for Santa ridges exhibit relatively high precision compared to those for Chao floods, likely as a

result of the higher prevalence of single-grain aliquots available in the sandy Santa deposits. The high overall quality of the Santa dating sequence makes it highly likely that this sequence provides a reliable basis for interpreting Santa Valley beach ridge progradation using Bayesian modeling (see below).

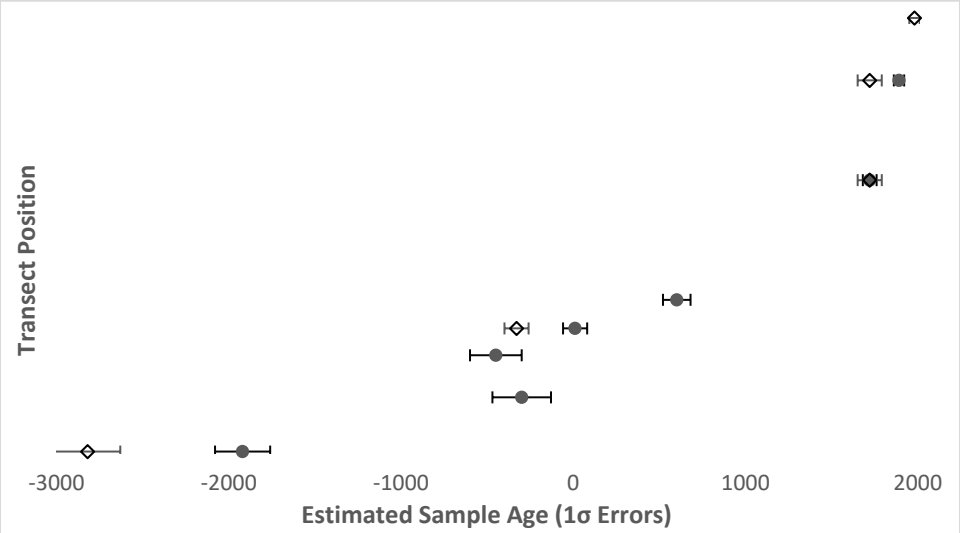

Figure 5-7. Santa Valley dates arrayed by distance along transect. Inland samples at bottom. Black dots represent single-grain feldspar luminescence dates, and hollow diamonds represent radiocarbon dates. Negative values represent years BC.

Bayesian Analysis

Once sample ages were estimated by luminescence dating, Bayesian modeling in OxCal 4.3.2 (Bronk Ramsey 2009) was applied to make use of stratigraphic information[14] to integrate ages into depositional models. This objective of this process was to first systematically identify

[14] In this case, ordinal depositional information. For the Santa Samples, the lateral position of ridges provides ordinal information which can reliably act as a surrogate for stratigraphic superposition.

outlier dates using explicit empirical criteria, then make use of compatible dates to derive refined luminescence-based depositional models for Chao and Santa sediments.

To model the Santa Valley beach ridge sediments, the "Sequence" command was used (see Code 1). This command was selected for its conservatism in its assumptions and parameters; it imposes only the assumption that sediment dates should array in stratigraphic order, which is likely for both depositional contexts under examination. "Sigma Boundary" commands were used to bracket the sequence (Bronk Ramsey 2008) to avoid the imposition of artificial statistical constraints on the ages of the depositional system. This same command was also used to represent undated ridges within the sequence to make use of the stratigraphic information these ridges provide. In cases where both ^{14}C and luminescence dates were available for single ridge deposit, the ^{14}C date was entered as a distinct stratum underlying the luminescence date to reflect the expected age relationship between the detrital ^{14}C sample and the luminescence sample. Because the ^{14}C date for Ridge 14, the youngest observed ridge, was "modern," this date was entered with the "C Date" command, with a mean age of 1980 +/- 30[15] to make use of the sample age information as a model constraint without forcing undue precision on the model. Percent distance along the surveyed transect was entered as sample "depth" in the model, with Ridge 1 representing 100% distance.

OxCal model results for the Santa Valley beach ridge sequence are displayed in Table 5-5, rounded to the nearest decade (see also Figure 5-8). On the whole, model results show

[15] 1980 was estimated as a rough mean between 1950, the year of the ^{14}C "present," and 2007, the year of observation. Error terms were estimated as roughly half of the timespan between 1950 and 2007, rounded up.

excellent stratigraphic agreement between all input dates; the A_O value was 70.6, and no individual sample A value was below 66.3[16]. This result suggests the combined ^{14}C and luminescence dates provide a robust depositional model for the Santa Valley beach ridge sequence, with no identified statistical outliers.

Table 5-5. OxCal 4.3.2 Results for Dated Santa Samples

Sample	Sample Type	Provenience		OxCal Model Results		
		Ridge No.	% Distance Inland	Cal BC/AD	+/- 2σ	A
OS-72851	^{14}C	14	0.0	1970	50	102.2
UW1815	SG IRSL	12	12.2	1890	70	76.1
OS-72850	^{14}C	12	12.2	1870	90	66.3
UW1845	SG IRSL	8	36.3	1720	80	81.9
OS-72846	^{14}C	8	36.3	1720	80	105.9
UW1844	SG IRSL	5	64.9	600	160	100.0
UW1843	SG IRSL	4	70.7	10	140	100.0
OS-72843	^{14}C	4	70.7	-310	100	80.8
UW1839	SG IRSL	3	77.9	-430	150	120.3
UW1840	SG IRSL	2	86.8	-520	210	68.2
UW1814	SG IRSL	1	100.0	-1900	320	100.0
SI-4957	^{14}C	1	100.0	-2870	400	100.9

Note: Negative ages represent years BC. Error terms are 2σ. The model A_O = 70.6.

[16] A_O provides an index of the overall stratigraphic compatibility of the modeled depositional sequence; an A_O value below 60 indicates a model which exhibits poor stratigraphic compatibility overall. A sample's A value provides a similar index of a that sample's individual stratigraphic compatibility; an A value of below 60 indicates a statistical outlier in the overall sequence.

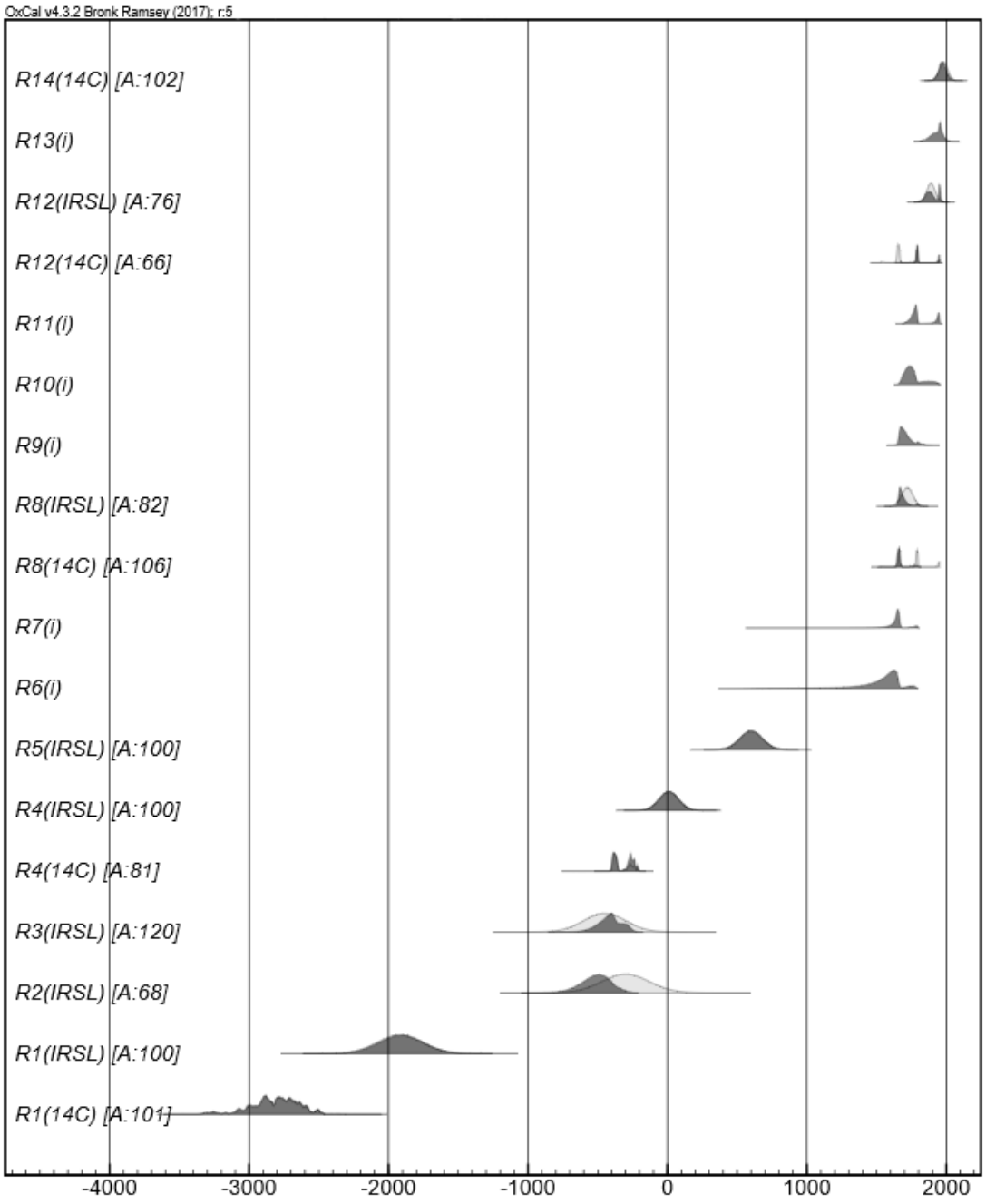

Figure 5-8. OxCal 4.3.2 model output showing Santa Valley dates—including single-grain feldspar luminescence and radiocarbon samples—arrayed versus percent sampled depth. Observed ridge numbers are at left. Ridge numbers followed by an "(i)", e.g., R6(i), were not sampled for dating; their age distribution is therefore a product of their stratigraphic position within the model. Sample A values are displayed next to their surface numbers.

Depositional modeling of Chao Valley flood sediments followed the same basic approach described above, with two notable exceptions. First, no radiocarbon dates exist for this sequence. Second, percent depth in the observed CVP2 profile was used as the model depth input parameter.

As noted above, fine-grain dates for the Chao Valley sediments are clearly stratigraphically problematic (A_O = null); this is another indication that the problem of poor bleaching likely plagues many of these dates. Available single-grain dates for Chao Valley samples[17] exhibit better agreement in OxCal modeling (A_O = 59.3), but the presence of two outlier dates—UW2103 (A = 34.9) and UW2098 (A = 8.5)—limit the viability of the overall model (Code 2, Table 5-6, and Figure 5-9).

Given this, the Chao Valley depositional model was re-run in a modified format to exclude these two outlier dates. Because the fine-grain date for UW2103 (AD 1360 +/- 80) is one of the few that exhibits reasonable stratigraphic agreement with adjacent single-grain dates, the fine-grain date for this sample was used in the next model iteration instead of the incompatible single-grain date (see Code 3). The fine-grain date for UW2098 (4590 BC +/- 610) was clearly stratigraphically incompatible with single-grain data overall, and thus its stratum was entered an undated boundary as above.

[17] Not all Chao Valley samples yielded sufficient material for single-grain dating. Strata for which no single-grain data were available were entered into OxCal as described for undated Santa Valley ridge deposits.

Table 5-6. OxCal 4.3.2 Results for all Chao Valley Single-grain Luminescence Samples

Sample	Sample Type	Provenience		OxCal Model Results		
		S#, F#	% Depth	Cal BC/AD	+/- 2σ	A
UW2221	SG IRSL	21, 20	1.2	1780	140	84.4
UW2219	SG IRSL	20, 19	4.1	1560	110	103.1
UW2218	SG IRSL	19, 18	8.8	1480	100	110.5
UW2217	SG IRSL	18, 17	10.6	1390	130	114.4
UW2103	SG IRSL	17, 16	12.9	1270	190	34.9
UW2216	SG IRSL	16, 15	18.2	1190	210	134.5
UW2102	SG IRSL	15, 14	21.8	1110	220	114.5
UW2213	SG IRSL	13, 12	28.8	940	250	19.2
UW2100	SG IRSL	12, 11	43.5	790	260	133
UW2099	SG IRSL	10, 9	55.3	600	220	103.6
UW2098	SG IRSL	9, 8	58.8	570	210	8.5
UW2211	SG IRSL	8, 7	64.7	541	200	120.3
UW2210	SG IRSL	7, 6	68.2	520	210	107.4
UW2209	SG IRSL	5, 4	79.4	450	220	135.7
UW2208	SG IRSL	3, 2	87.1	310	220	107.4
UW1816	SG IRSL	2, 1	97.1	170	220	95.9
UW2371	SG IRSL	1, N/A	100.0	-1780	320	99.1

Note: Negative ages represent years BC. Error terms are 2σ. The model A_O = 59.3. Shaded samples have individual A values below 60, signaling stratigraphic incompatibility of these samples' dates.

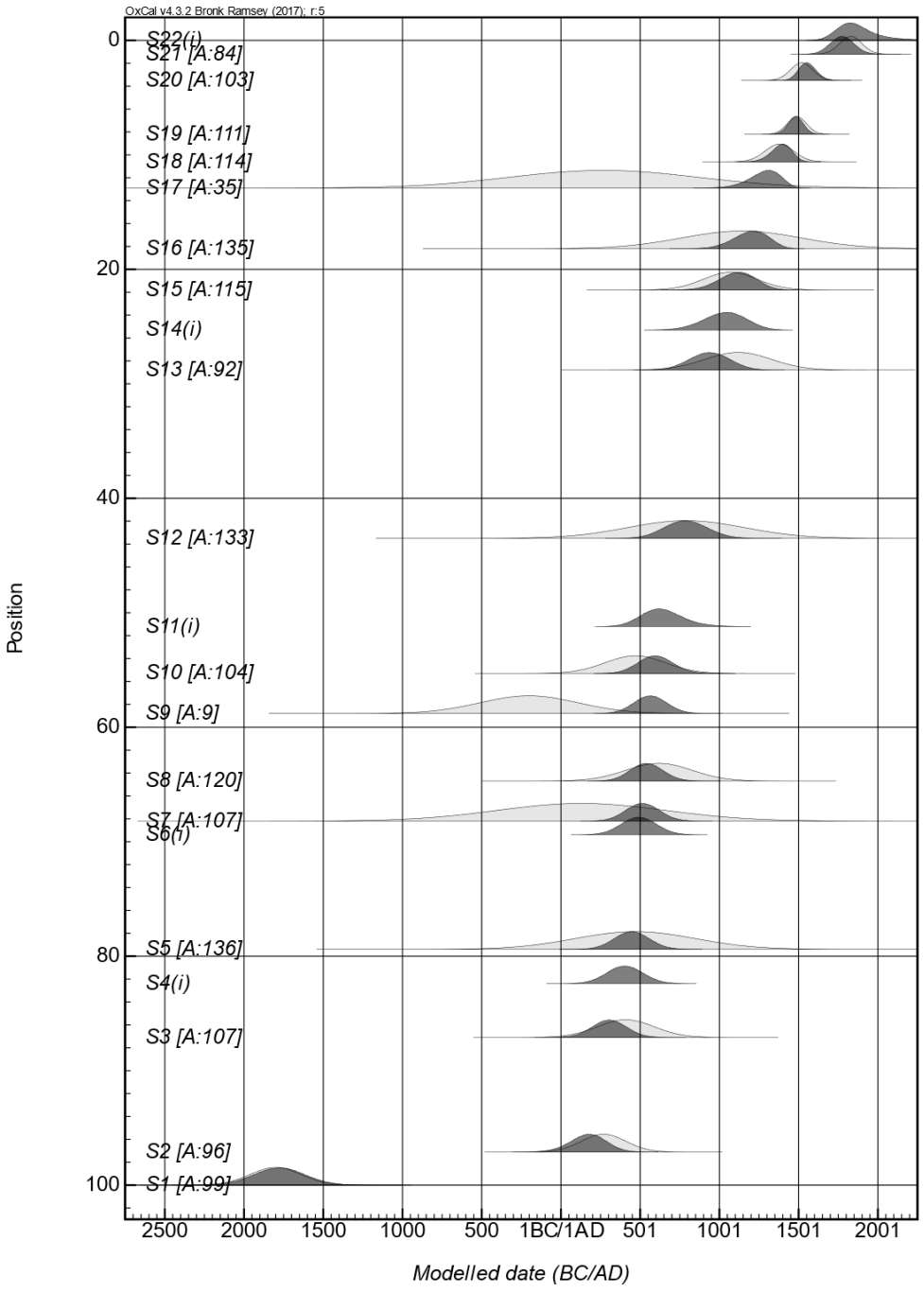

Figure 5-9. OxCal 4.3.2 model output showing all Chao Valley single-grain feldspar luminescence dates arrayed versus percent sampled depth. CVP2 observed surface numbers are at left, and include the last pre-flood deposit. Surface numbers followed by an “(i)”, e.g., S4(i), did not yield sufficient coarse grains for single-grain dating; their age distribution is therefore a product of their stratigraphic position within the model. Sample A values are displayed next to their surface numbers.

Results of the modified OxCal Chao Valley depositional model (Table 5-7, Figure 5-10) show excellent stratigraphic agreement in sample ages (A_O = 133.4, with no sample A value below 68.6). This result supports the interpretation that single-grain IRSL dates generally provide viable estimates for the ages of the Chao Valley sediments once outliers are formally identified and discarded. Fine-grain luminescence ages are less useful in general, but individual fine-grain ages nonetheless remain useful on a case-by-case basis if deployed as auxiliary information to single-grain data.

For the Chao Valley sequence, additional sample composition information offered the possibility of further refinement of stratigraphic information used in OxCal modeling. Specifically, XRF data for Chao samples indicate Chao Valley Surfaces 9, 17, and 20 are enriched in calcium (Ca) relative to all other flood sediments (Table 5-8). In this context, the most plausible explanation for this enrichment is a relatively high post-depositional buildup of carbonate salts, most likely due to surface processes[18] such as sub-aerial salt weathering. Sample grain-size data[19] (Figure 5-11) indicate these same surfaces also exhibit high modal grain sizes relative to other flood deposits; the most likely explanation is that these deposits were exposed to a relatively high degree of post-depositional eolian deflation[20]. Taken together, the XRF data and

[18] The water table is many meters below the surface, so capillary action is unlikely in this context.

[19] Derived from a Micrometrics Saturn DigiSizer 5200, de-ionized water with sodium hexametaphosphate deflocculant and micro-bubbles removed, target measurement obscurity of ~15, an assumed refractive index of 1.55, and an assumed density of 1.

[20] Given the small size of the catchment, the catchment's spatial remove from higher-energy fluvial inputs, and the unremarkable mass of Surfaces 9, 17, and 20, it is unlikely that their relatively high modal grain size is due to differential inputs such as distinct sediment sources or higher-energy depositional agents or events.

Table 5-7. OxCal Results for Stratigraphically Compatible Chao Valley Luminescence Samples

Sample	Sample Type	Provenience		OxCal Model Results					
				Droughts Excluded			Droughts Included		
		S#, F#	% Depth	Cal BC/AD	+/- 2σ	A	Cal BC/AD	+/- 2σ	A
UW2221	SG IRSL	21, 20	1.2	1770	150	80.4	1820	100	109.5
UW2219	SG IRSL	20, 19	4.1	1560	110	103.0	1560	110	103.1
UW2218	SG IRSL	19, 18	8.8	1490	90	112.6	1480	100	109.6
UW2217	SG IRSL	18, 17	10.6	1410	110	117.0	1390	100	122.0
UW2103	FG IRSL	17, 16	12.9	1340	110	114.0	1260	60	73.1
UW2216	SG IRSL	16, 15	18.2	1240	170	132.2	1180	130	137.8
UW2102	SG IRSL	15, 14	21.8	1160	200	108.0	1110	170	120.9
UW2213	SG IRSL	13, 12	28.8	1010	230	112.2	960	210	101.9
UW2100	SG IRSL	12, 11	43.5	910	250	127.3	850	220	133.4
UW2099	SG IRSL	10, 9	55.3	710	230	68.6	660	110	93.2
UW2211	SG IRSL	8, 7	64.7	610	220	124.6	530	100	127.4
UW2210	SG IRSL	7, 6	68.2	570	230	99.0	500	120	109.2
UW2209	SG IRSL	5, 4	79.4	490	240	134.8	430	160	138.0
UW2208	SG IRSL	3, 2	87.1	330	230	111.8	310	190	110.5
UW1816	SG IRSL	2, 1	97.1	180	230	98.7	170	220	97.7
UW2371	SG IRSL	1, N/A	100.0	-1780	320	99.2	-1780	320	99.2

Note: Results are shown for varied model assumptions, including both the exclusion and the inclusion of inferred droughts as stratigraphic information. Negative ages represent years BC. Error terms are 2σ. A_O = 133.4 if droughts are excluded, and A_O = 143.2 if droughts are included. Dashed lines indicate the stratigraphic location of inferred droughts.

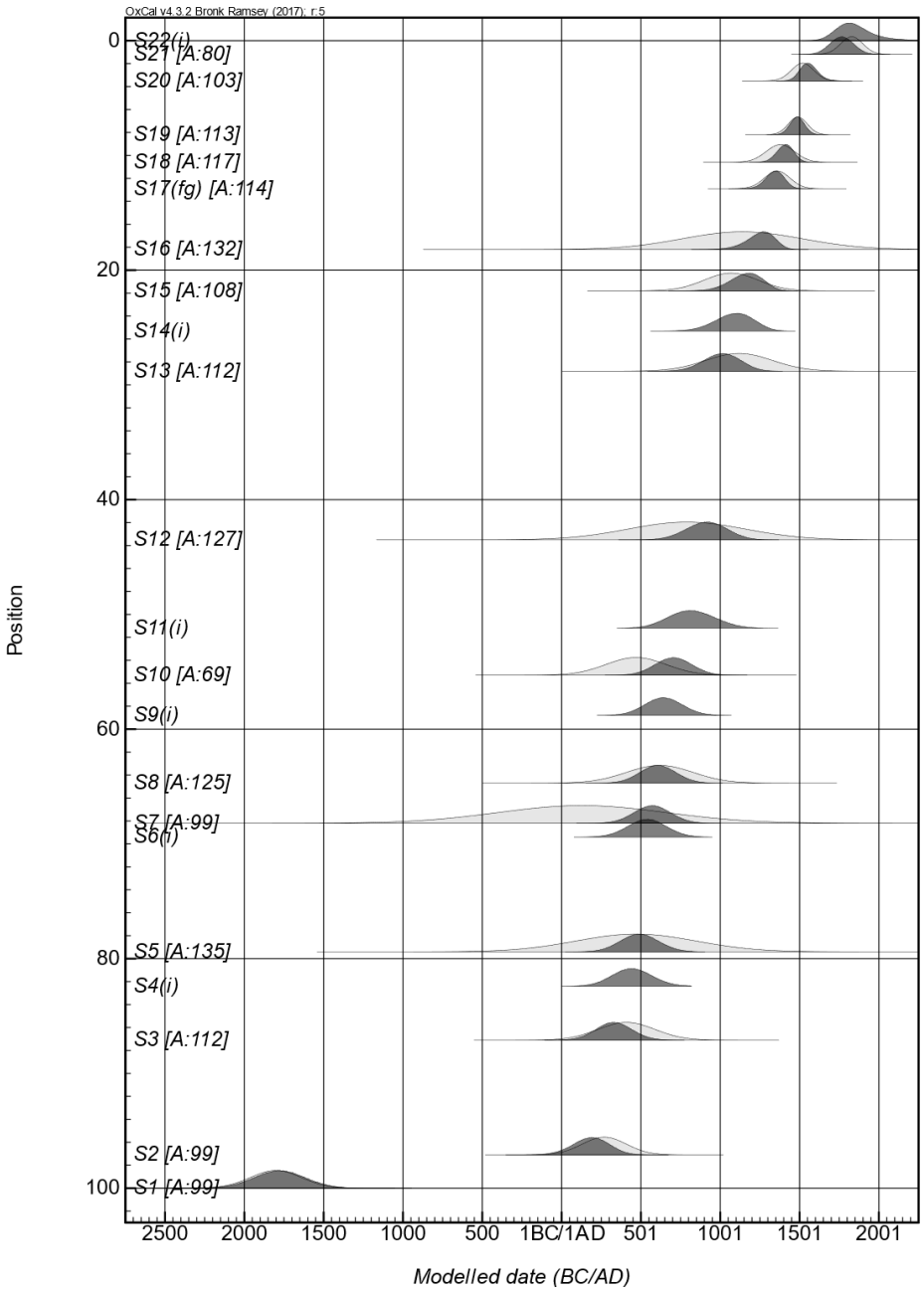

Figure 5-10. OxCal 4.3.2 model output showing Chao Valley single-grain feldspar luminescence dates arrayed versus percent sampled depth, with outliers removed and the fine-grained polymineral date for S17, F16 (UW2103) used instead of the single-grain feldspar date; this substitution is denoted by the "(fg)" label next to "S17". CVP2 observed surface numbers are at left, and include the last pre-flood deposit. Surface numbers followed by an "(i)", e.g., S4(i), did not yield sufficient coarse grains for single-grain dating; their age distribution is therefore a product of their stratigraphic position within the model. Sample A values are displayed next to their surface numbers.

Table 5-8. Results of Chao Valley Sample Composition Measurements, Including XRF, Magnetic Susceptibility, and Grain Size Analysis

Sample No.	Provenience		XRF			Magnetic Susceptibility				Modal Grain Size	
						Low Frequency		Coeff.			
	Surface No.	Flood No.	% Ca	+/-	Rank	Value	Rank	Value	Rank	Value	Rank
Undated	22	21	2.51	0.12	6	117.6	7	2.13	17	15.0	6
UW2221	21	20	2.12	0.12	18	145.4	3	1.667	19	14.1	7
UW2219	20	19	2.82	0.11	2	123.3	5	4.054	8	33.5	2
UW2218	19	18	2.41	0.10	9	132.2	4	3	15	11.2	11
UW2217	18	17	2.20	0.12	14	246.2	1	1.19	21	11.2	11
UW2103	17	16	2.97	0.16	1	112.8	9	4	11	29.9	4
UW2216	16	15	2.53	0.13	5	83.0	14	8.163	1	10.0	18
UW2102	15	14	2.63	0.11	4	73.2	21	6	3	11.9	8
UW2214	14	13	2.38	0.21	10	111.4	10	5.682	5	11.2	11
UW2213	13	12	2.29	0.13	12	113.7	8	4	9	11.9	8
UW2100	12	11	2.05	0.12	21	74.6	20	3.922	12	9.4	19
UW2212	11	10	2.18	0.16	17	81.5	15	3.704	13	10.6	17
UW2099	10	9	2.18	0.14	15	233.2	2	1.25	20	8.9	21
UW2098	9	8	2.71	0.11	3	77.0	19	3.704	14	50.1	1
UW2211	8	7	2.42	0.13	7	98.2	11	2.597	16	26.6	5
UW2210	7	6	2.35	0.12	11	93.9	13	5.128	7	33.5	2
UW2097	6	5	2.18	0.12	16	80.7	16	6.667	2	11.2	11
UW2209	5	4	2.08	0.12	19	77.3	18	4	10	9.4	19
UW2096	4	3	2.41	0.11	8	79.1	17	5.556	6	11.2	11
UW2208	3	2	2.08	0.12	20	120.9	6	2.128	18	11.9	10
UW1816	2	1	2.28	0.14	13	97.0	12	5.797	4	11.2	11
UW2371	1	N/A	1.80	0.16	N/A	-2.6	N/A	25.0	N/A	N/A	N/A

Note: Error terms are 1σ. For magnetic susceptibility data, measurements reflect exposure to 100 Hz current, and values are expressed in $10^{-8}m^3/kg$. “Coeff.” reflects the “coefficient of frequency dependency” and gives an index as to how much the susceptibility changes when the current is switched from 100 Hz to 1000 Hz.

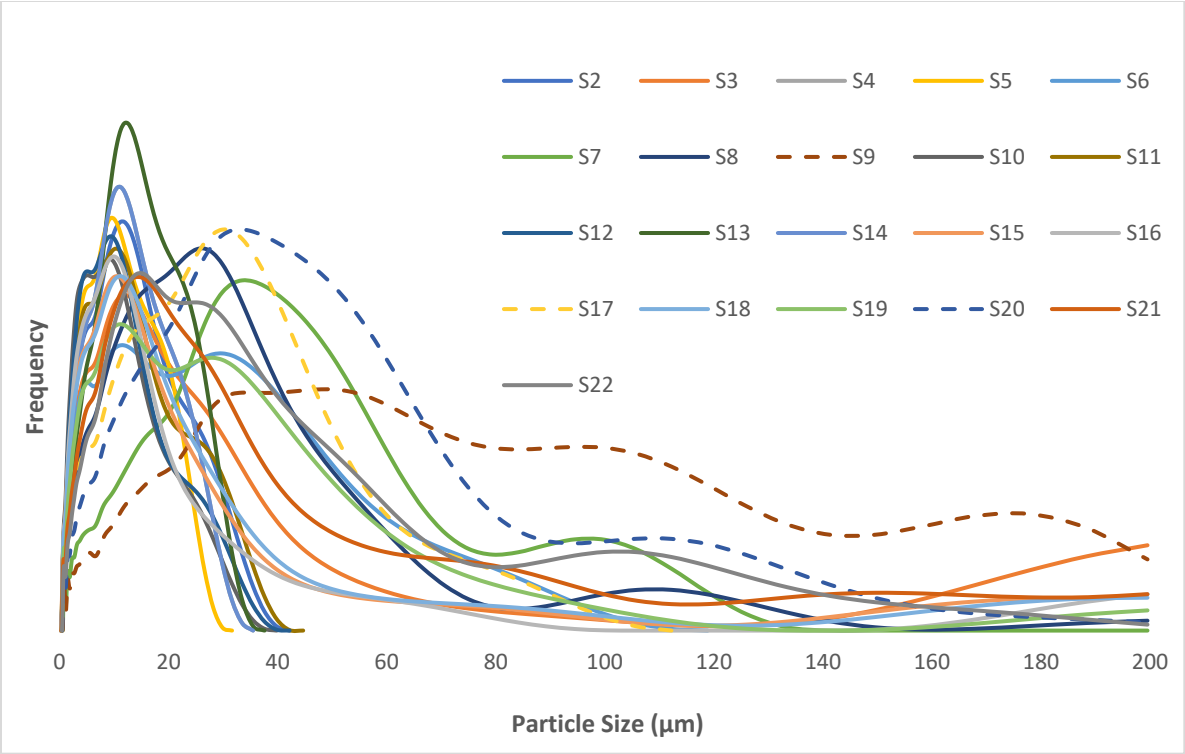

Figure 5-11. Grain size distributions of all sampled Chao Valley flood surfaces. Surfaces 9, 17, and 20 exhibit high modal grain size relative to other surfaces, possibly indicating eolian deflation as a result of exposure during a period of increased aridity.

the grain size data therefore indicate surfaces 9, 17, and 20 were probably exposed to a high degree of post-depositional alteration prior to the deposition of subsequent strata. Because 1) the deposition of these three surfaces was roughly contemporary with the three most significant highland dry periods identified in the high-resolution ~1,500-year Quelccaya ice core data (AD 570—AD 610, AD 1250–AD 1310, and AD 1720—AD 1860; Thompson et al. 1985), 2) ancient highland dry periods are broadly correlated across the upland Andes (Apaestegui et al. 2014) and 3) these dry periods typically have the effect of reducing flow in coastal rivers and springs, thereby greatly increasing the aridity of coastal irrigation agriculture zones (Apaestegui et al.

2014; Shimada et al. 1991), it is likely that the unique composition of Surfaces 9, 17, and 20 relative to other Chao Valley flood sediments is a result of surface exposure during periods of higher aridity as a result of upland drought. Magnetic susceptibility data[21] (see Table 5-8) support the interpretation that periods of increased aridity affected these surfaces; drought-subsequent surfaces 10, 18, and 21 exhibit elevated susceptibility and reduced coefficients of dependency, indicating weaker oxidation of iron-rich minerals, likely due to the reduced availability of ambient water to facilitate sub-aerial weathering prior to their deposition as alluvium. Further, Surface 9 exhibits post-depositional cracks of up to 5 cm (Figure 5-12), larger than the ~2-cm-wide cracks visible in other surfaces, indicating pronounced aridity after deposition of this surface.

Given this suite of convergent information indicating Surfaces 9, 17, and 20 are likely to represent drought-exposed surfaces, age ranges for Quelccaya dry periods (Thompson et al. 1985) were added to the Chao Valley OxCal depositional model using the "Top Hat" command (see Code 4) to represent flat probability density functions, depth values intermediate to adjacent strata[22], and estimated age midpoints and half-widths of AD 590 +/- 40, AD 1280 +/- 50, and AD 1720 +/- 90[23]. The resultant depositional model is shown on the right-hand side of Table 5-7 (see

[21] Derived following protocols outlined by Gale and Hoare (1991).

[22] In the "Sequence" model, the specific depth has no bearing on numerical results, provided the input command is entered in the correct order (i.e the assigned depth value is intermediate to adjacent strata depths).

[23] Shimada and colleagues (1991) relay an uncertainty of about 20 years in the Quelccaya ice core data, so an additional +/- 20 years was added to half-widths to avoid spurious precision in model inputs. This error exceeds age offsets found in a reanalysis of the Quelccaya data by Seimon (2003), and therefore encompasses errors in the ages of inferred droughts.

also Figure 5-13) alongside results that do not make use of inferred periods of drought as model constraints. The model shows excellent stratigraphic agreement (A_O = 143.2, all sample A values are 73.1 or higher), and improved precision in individual dates for floods adjacent to periods of inferred droughts as a result of the use of the relatively high-precision Quelccaya ages as a model input.

Figure 5-12. A ~5-cm-wide crack at the top of Chao Valley Surface 9 (Flood 8) which was infilled by subsequent flood sediment. The wide cracking and reddened color of this surface may indicate increased aridity during/after its deposition.

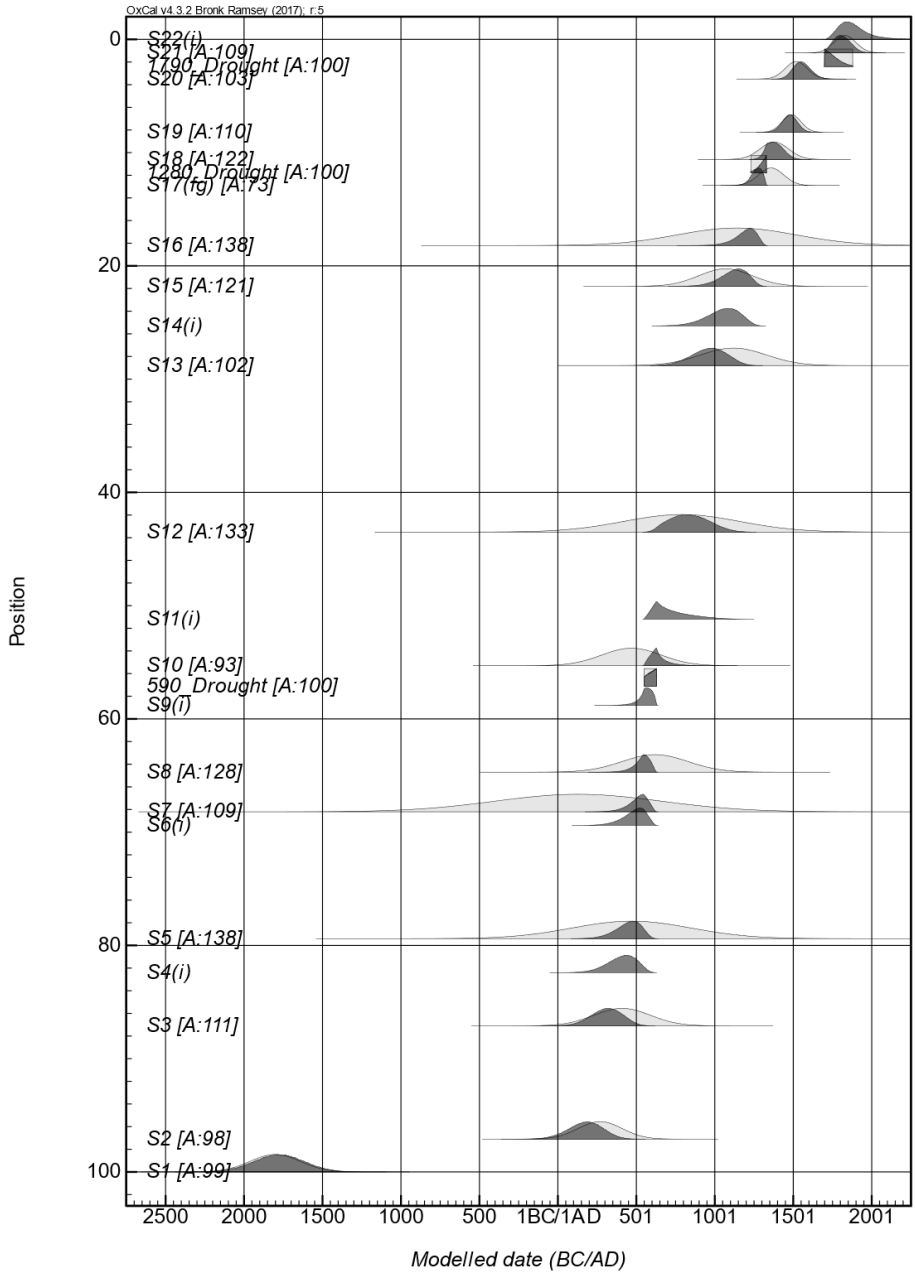

Figure 5-13. OxCal 4.3.2 model output showing Chao Valley single-grain feldspar luminescence dates arrayed versus percent sampled depth, with outliers removed, the fine-grained polymineral date for S17, F16 (UW2103) used instead of the single-grain feldspar date, and including drought information from Thompson and colleagues (1985). CVP2 observed surface numbers are at left, and include the last pre-flood deposit. Surface numbers followed by an “(i)”, e.g., S4(i), did not yield sufficient coarse grains for single-grain dating; their age distribution is therefore a product of their stratigraphic position within the model. Inferred periods of drought are also indicated at left. Sample A values are displayed next to their surface numbers.

DISCUSSION

The high number and density of dates produced by this study and the strong stratigraphic agreement of these dates in OxCal depositional modeling provide the opportunity to set forth a new and relatively robust interpretation of local ENSO-driven depositional change as evident in the roughly 2,000-year Chao Valley flood sequence and the roughly 4,000-year Santa Valley beach ridge sequence.

Santa Valley Coastal Depositional History

The Santa Valley beach ridge complex began forming after the occupation of the Ostra Base Camp site ended around roughly 4300 BC (Sandweiss 2003). Ridge 1 formed by around 1900 BC, and was the product of multiple earthquake/ENSO depositional episodes. During this interval, the paleoembayment encompassed by the pre-5000 BC shoreline was infilled completely, leaving the ancient Ostra sea cliff roughly 3 km inland. The ~2000 BC Salinas de Chao uplift event observed by Sandweiss and colleagues (1983) may have occurred during formation of this ridge. Observed earthquake damage to the important Late Preceramic archaeological sites of Aspero and Caral roughly 250 km south of the Santa Ridges (Sandweiss et al. 2009) is also roughly coeval with formation of this ridge. Further, evidence suggests that the Medio Mundo formation, which infilled mid-Holocene bays with sand and gravel along the coast of the Norte Chico area of Perú, is a likely depositional result of the ancient quake(s) which damaged Aspero and Caral (Sandweiss et al. 2009). Given these major coeval changes across the central and north-central coast of Perú at the end of the Late Preceramic, Ridge 1 was most likely

formed by powerful seismic activity which affected the entire region, greatly impacting local coastal peoples' domestic infrastructure and available shoreline resources.

Between Ridges 1 and 2, roughly 13.2% of the observed transect (which did not examine accumulation between the Ostra cliff and Ridge 1) progradation occurred, making the construction of Ridge 2 (modeled date of 520 BC +/- 210[24]) the second-largest Santa depositional event documented by this study[25]. Sandweiss (1986:23) observed an uplift event subsequent to Ridge 2, and Ridge 3 (modeled date 430 BC +/- 120; 8.9% of observed progradation) followed shortly thereafter, indicating major tectonic activity and coastal change during this period. This post-Ridge 2 uplift event may also be the Salinas de Chao uplift event observed by Sandweiss and colleagues (1983), but this remains uncertain due to limitations in dating of the Salinas de Chao event. If these events are the same, however, the age of the Salinas de Chao event is roughly 1,500 years younger than the age postulated by Sandweiss (1986). Either way, the 6th and 5th centuries BC were apparently a time of pronounced deposition in this system, accounting for as much as 22.1% of post–Ridge 1 progradation. Given the inherent bias towards extremely powerful events in this depositional record, the fact that Ridges 2 and 3 were so closely spaced and collectively powerful suggests that these depositional events were extremely powerful.

Ridge 4 (modeled date AD 10 +/- 140; 7.2% of observed progradation) and Ridge 5 (AD 600 +/- 160; 5.8% of observed progradation) exhibit more temporal spacing, and were less

[24] OxCal modeled dates with 2σ error terms are used throughout the Discussion section.
[25] An average of 7.7% +/- 3.9 progradation occurred between ridges.

powerful than either of the prior two events. The Ridge 5 date is significant here because it coincides with the timing of pronounced cultural changes in the Southern Moche heartland (see below). While this correlation is of course insufficient to fully explain local cultural changes during this period, it suggests that seismic activity followed by a major ENSO event—and the short-term impacts on local resources which would have resulted—are worth further exploration as part of a viable explanation for southern Moche cultural change.

Ridge 6 (no date; 16.9% progradation) represents the largest ridge-building event in this context, but remains undated, as does Ridge 7 (no date; 8.2% progradation). Given this, it is not possible to associate these ridges with specific culturally destructive events such as the seismic damage to the Moche-Chicama intervalley canal posited at around AD 1200 by Ortloff and colleagues (1982, 1983), which would likely be represented in this portion of the Santa Ridge sequence. On the other hand, sometime between the formation of Ridges 5 (AD ~600) and 8 (AD 1720 +/- 40; 3.5% progradation), it is clear from the Santa ridges that at least two major earthquake/ENSO events occurred, the earlier of which appears to have been the most powerful in the last four millennia. For the latter event, historical records of earthquakes in this area may be of aid. That is, given the horizontal position of Ridge 7 and the AD 1720 date for Ridge 8, it is likely that Ridge 7 represents ENSO erosion following either the AD 1619 quake event[26], which had an estimated strength of M_w of 7.7 to 8.0 (Dorbath et al. 1990) and caused damage as far

[26] The earliest local earthquake documented in written records.

south as Santa (Ferro 1978:22), or the AD 1678 quake event (estimated M_w of 8.4 [Dorbath et al. 1990]) whose epicenter was slightly south of the study area, or perhaps a combination thereof.

Ridge 8 (AD 1720 +/- 40) itself is well dated by both ^{14}C and luminescence, and is most likely associated with either the historically documented 1725 quake (estimated M_w of 7.5 [Dorbath et al. 1990]) in north-central coastal Perú followed by the 1728 "Very Strong" El Niño (Ortlieb 2000; see also Garcia-Herrera et al. 2008) or the 1678 quake (Dorbath et al. 1990) followed by one of the "Strong" El Niño events documented between 1687 and 1720 (Ortlieb 2000). Of these two possibilities, the former is more likely due to the local impacts of the 1725 event; Ferro's (1978:27) documentation of the 1725 event includes reference to a major flood in the upper Santa caused by displaced lake water resulting from glacial matter dislodged during the quake, and these floodwaters would have amplified erosion and discharge of quake-loosened sediment along the lower reaches of the river. In any case, the depositional effects of this quake/ENSO combination on the Santa Valley coastline were below average (3.5% progradation) for this system.

Ridges 9 through 11 (9%, 8%, and 4.4% of observed progradation, respectively) remain undated, again making direct connection to documented historical events difficult. Ridge 12 (modeled date of AD 1870 +/- 90), which was again dated with both ^{14}C and luminescence, provides an upper age limit on these deposits. Ridge 9 is most likely the product of the extremely powerful 1746 Callao earthquake (estimated M_w of 8.8 to 9.0, with an epicenter ~90 km north of Lima [Beck and Nishenko 1990; Jiminez et al. 2013]). This earthquake was the strongest central coast quake event recorded in the historical record, and it was followed by either the "Strong"

1747 El Niño (Ortlieb 2000) or an event in the 1760s (Garcia-Herrera et al. 2008), depending on historical source used. If the association of the 1746 quake with Ridge 9 is correct, the fact that a roughly $M_w = 9.0$ quake produced 9% of observed progradation gives a sense of proportion to the immense local erosive and depositional power of earlier events which produced Ridges 2 (13.2% progradation) and 6 (16.9% progradation); formation of these ridges must have been characterized by either increased local quake damage, stronger local El Niño flooding, or both. The alignment of this ridge and all subsequent ridges is distinct from their predecessors, possibly as a result of a changed location of the Santa River mouth, as argued by Sandweiss (1986).

None of the historically recorded quakes affecting the Santa area between 1746 and 1940 appear to have been particularly strong, but a number of them could be responsible for building the complex alignments of Ridges 10 and 11. For example, the 1759 earthquake directly damaged Villa de Santa (Ferro 1978:32), and the 1902, 1904, 1905, 1917, and 1937 events also caused documented quake tremors in Santa (Ferro 1978). The 1759 event was also followed by the 1761 "Strong" El Niño event, which flooded the Santa River and destroyed much of the town of Santa (Ortlieb 2000:229), and it is therefore plausible that Ridge 10 (8% of observed progradation) was formed after the 1761 El Niño. Unfortunately, no other quake/ENSO combination lends itself as strongly to the formation of Ridge 11 (4.4% of observed progradation). The 1877–1878 ENSO "Very Strong" El Niño event may be a contender, as it flooded the Santa River and destroyed a railway there (Ortlieb 2000:242; Garcia-Herrera et al. 2008), but it is not preceded by a local quake. No other contemporary "Strong" or "Very Strong" El Niño events during this interval are directly associated with documented evidence of heavy

rainfall, flooding, or erosion in the Santa Valley itself (Ortlieb 2000), so it seems likely that the multiple topographic highs constituting Ridge 11 were formed through the combination of a number of smaller quakes and lower-strength El Niño events in this period.

Ridge 12 (modeled date of AD 1870 +/- 90; 2.7% of observed progradation) is in roughly the same topographic position as Moseley and colleagues' (1992) 1944 coastline as documented by aerial imagery. This ridge is likely the result of the 1937 Trujillo earthquake (Ferro 1978:54–55) followed by the "Strong" 1941–1942 El Niño. Ridges 13 (no date; 7.1% of observed progradation) and 14 (modern ^{14}C date; 5% of observed progradation) reflect the 1970 and 1975/modern shorelines, respectively. Ridge 13 was therefore probably the result of the M_w 7.25 1946 highland earthquake followed by the "Moderate" 1951 El Niño event, and Ridge 14 the result of the M_w 7.7 1970 earthquake followed by the "Strong" 1972–1973 El Niño event, as described by Moseley and colleagues (1992).

In summary, the dated Santa beach ridge sequence generally accords well with 1) documentary evidence from aerial images and historic records of earthquakes and El Niño events in Perú and 2) published studies of ancient landscape evolution in the Santa Valley and adjacent areas. Data generated by this study also contribute new information to our understanding of this landscape by providing new dates for ridge-building events, age constraints for others, and a sense of proportion to the depositional strength of each ridge-building episode. In light of this information, it is clear that the Santa coastline has experienced the most progradation (~48% of the observed transect) after historical records of local earthquakes began in AD 1619 (Ridges 7–14), but the most powerful episodes of coastal change occurred prior to 2000 BC (Ridge 1),

between around 600 and 400 BC (Ridges 2 and 3), and sometime between roughly AD 600 and AD 1600 (Ridge 6). Two other ancient episodes at roughly AD 10 and AD 600 (Ridges 4 and 5) exhibit depositional magnitudes roughly commensurate with the strongest quake/ENSO combinations documented in historic records, and therefore likely represent very powerful prehistoric events.

Chao Valley Flood Sediment Depositional History

In the sampled portion of the Chao Valley, there was minimal accumulation of El Niño flood sediment between around 1800 BC and the early centuries AD. Because ENSO events were certainly occurring during this interval, it is likely that during this period local deposition was sufficiently minimal that subaerial erosion was able to match its pace. This erosional/depositional balance suggests that the linear dune which subsequently impounded flood sediments in this location was not in place until a short time prior to AD 170. Given the timing of the initiation of the Chao flood sequence and the fact that Santa sediment discharge is the primary source of Chao dune sediment, it is likely that the formation of this dune was an eventual product of the massive amounts of sediments made available by the extremely high rates of Santa sediment discharge which occurred during the roughly 600 to 400 BC progradational episodes represented by Santa Ridges 2 and 3 and/or the AD 10 Santa ridge-building event(s) represented by Ridge 4. Between dune formation and AD 170, only 2.9% of observed flood sediment accumulation occurred.

Between around AD 170 and AD 310, the impounding dune was evidently in place, but flood sediment accumulation in this location was low, as no visible flood surfaces are extant from this period. Between roughly AD 310 and AD 550, flood activity increased dramatically (Figure 5-14), with a high rate of accumulation of both flood sediment surfaces[27] and flood sediment volume. During this period, two of the three highest-volume flood sediment deposits[28] accumulated, with one at AD 310 +/- 190 (10% of observed accumulation[29]) and one at roughly AD 500 +/- 120 (11.2% of observed accumulation); these two events therefore represent extreme El Niño deposition in this location. When combined with the other five flood surfaces from this period, and the fact that this depositional sequence is likely representative of only the most erosive El Niño floods over the last 2,000 years, it is clear that the 4th through the 6th centuries would have presented a relatively difficult time to cope with the effects of El Niño erosion in the surrounding area. Further, the assertion of increased flood activity during this period is supported by evidence from the central coast of Perú, perhaps indicating a pattern of pervasive flooding along the coast at this time; Winsborough and colleagues (2012) analyzed cores of lagoon sediments at Pachacamac to document two periods of strong floods between roughly AD 260 and AD 651. This inferred pattern is further supported by coeval flood sediments at archaeological sites in both the Lambayeque and Moche Valleys (Bawden 1996; Shimada 1994).

[27] A proxy for minimum number of flood events.
[28] Accumulations of sediment between observed former surfaces.
[29] Average flood sediment accumulation is 4.8% +/- 3.3 per sampled surface.

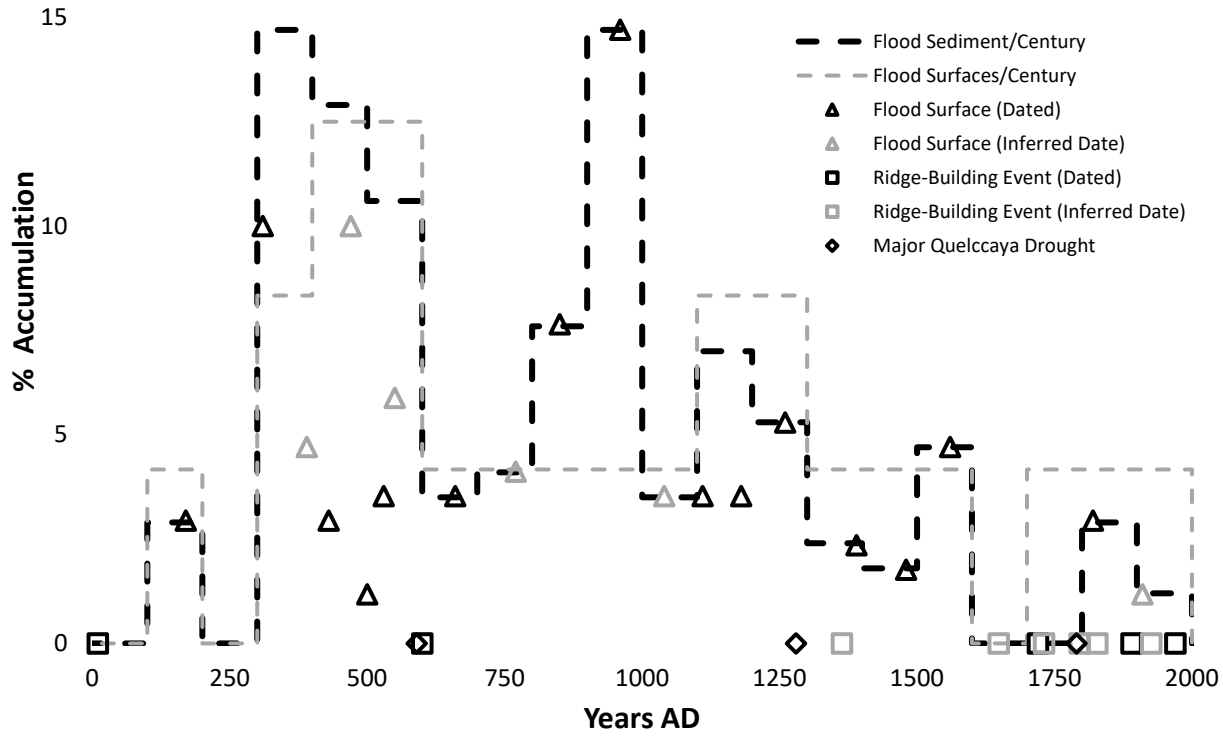

Figure 5-14. Modeled Chao Valley flood activity for the last 2,000 years. Ridge-building episodes and inferred droughts are included for reference. OxCal modeled mean dates were used; inferred dates are interpolated from stratigraphic position in the OxCal model and do not reflect measured age data. Sediment/Century rates reflect average accumulations within 100-year bins starting at year AD 1.

After the roughly AD 550 flood event in the Chao system, severe drought conditions likely persisted until shortly after AD 600 (Shimada et al. 1991). Substantial deposition of eolian deposits in irrigation canals are also argued to have occurred in the region during this time (Moseley 2001), a further indicator of increased landscape aridity and/or increased source sediment availability as the result of the apparent earthquake/El Niño event which occurred around this time, as evident in the Santa Valley beach ridge record.

Between roughly AD 600 and AD 850, only two flood surfaces are extant in the Chao sequence, and neither is representative of particularly high rates of sediment accumulation for

this system, as a total of only 7.6% of observed flood sediments were deposited during this interval. One event occurred at AD 660 +/- 110 (3.5% of observed accumulation) and the other at roughly AD 770 +/- 210[30] (4.1% of observed accumulation).

At about AD 850 +/- 220, a relatively powerful event occurred, accumulating roughly 7.6% of observed flood sediment, or the equivalent of the prior two flood surfaces combined. This was followed by another flood surface at AD 960 +/- 210 which accumulated 14.6% of observed Chao flood sediments, representing the most rapid period of accumulation in this sequence. Such rapid deposition is likely related to a uniquely destructive event, and may represent the legendary Naymlap Flood event, an extremely strong El Niño flood which has been documented in both northern Perú, where it was dated to around AD 1050 (Craig and Shimada 1986; Shimada 1990; Shimada and Wagner 2007), and at Pachacamac, where researchers interpret its age as AD 995–1008 (Winsborough et al. 2012).

After the AD 960 flood event, flood sediment accumulation rates were more moderate in this location for roughly the next 600 years, up to and including a flood event at around AD 1560 +/- 110. At least 7 flood surfaces are extant from this interval, but only the events at AD 1260 +/- 60 (5.3% of observed flood sediment) and at AD 1560 (4.7% of observed flood sediment) resulted in average or higher sediment accumulations for this system. A period of slightly elevated activity is evident in the three flood surfaces between AD 1110 +/- 170 and AD 1260

[30] OxCal modeled age of undated surface S11, F10; sample UW2212 was collected from this surface but failed to produce a stratigraphically compatible fine-grain date, and the sample yielded insufficient coarse grains for single-grain dating.

+/- 60, but activity during this interval is still modest relative to earlier periods. This AD 1260 surface is followed by an apparent drought interpreted to coincide with the roughly 1250–1310 highland dry period inferred by Thompson and colleagues (1985; see also Dillehay and Kolata 2004). Two additional flood surfaces representing relatively weak accumulation events, with one at AD 1390 +/- 100 (2.4% of observed flood sediment) and the other at AD 1480 +/- 100 (1.8% of observed flood sediment).

The timing of onset of this phase of lower activity is consistent with Galapagos Island proxies for ENSO activity in the eastern Pacific (Zhang et al. 2014), but unlike the Galapagos data the Chao Valley flood sediment data do not show increased activity in the 16th century, as no flood sediment accumulated in this system in the period between AD 1560 and roughly AD 1820 +/- 100 despite documented El Niño activity in the area during this period (Ortlieb 2000; Garcia-Herrera et al. 2008). This period coincides with the last major highland drought inferred from the Qualccaya data (Thompson et al. 1985), which extended between roughly AD 1650 and 1790, and during this period it is evident from the Chao sequence that local erosion equaled the pace of deposition during this time. This is somewhat unfortunate, because it means that fewer extant flood deposits can be compared with the historical record to attempt to connect observed flood accumulations with recorded El Niño events.

At around AD 1820 +/- 100, another 2.9% of observed sediment accumulated. Given the inherent bias in this record towards the strongest, most erosive El Niño events, this deposit is most likely associated with one or more of the five “Very Strong” El Niños documented by Ortlieb (2000) within this age range. These events occurred at roughly 1728, 1791, 1828, 1877–

1878, and 1891, and caused widespread destruction along the north coast of Perú. The final 1.2% of observed sediment accumulated sometime between this time and the present, most likely before the impounding dune was breached by road construction in the 1920s. This deposit is therefore most likely to represent one of the "Very Strong" El Niño events of the late 19th century or the "Very Strong" 1925 event (Takahashi and Martinez 2017). The fact that the six documented very strong 18th through early 20th century El Niños produced only 3.1% of the total accumulated sediment in this system is telling, since it implies that local eolian erosion is sufficient to remove much of the accumulated flood sediment in this area when historically observed El Niño conditions apply. By extension, past periods of relatively pronounced Chao Valley El Niño flood accumulation were probably due to El Niño floods which markedly exceeded the strength and/or frequency of historically observed floods in this area.

Cultural Implications

The detailed depositional histories outlined above provide a backdrop of El Niño–driven landscape change against which patterns in local cultural activity can be compared. In cases where pronounced depositional changes are coeval with inferred changes in cultural activity, the data presented here provide a good starting point for the exploration of environmental change as a possible factor in cultural change. Of course, correlation is not sufficient to demonstrate causation, but it is a necessary component in such a demonstration. In cases where pronounced depositional changes are coeval with periods of inferred relative cultural stasis, these data provide a good starting point for the exploration of cultural resilience to specific events.

A good example of each case is provided by the southern Mochica (or Moche) heartland centered on the Moche and Chicama Valleys. Here, a cultural florescence of distinctive elite goods and public architecture is thought to have occurred between roughly AD 200 and AD 600, after which local settlement patterns changed rapidly and dramatically. Shortly thereafter, significant cultural changes were afoot in this area. For example, Galindo rapidly arose as a significant population center alongside the Huacas de Moche (Lockard 2009), within the formerly preeminent Huacas de Moche complex Huaca de la Luna likely began to be eclipsed in importance by Huaca del Sol (Uceda 2010; Uceda and Tufinio 2003), and Moche V ceramics became pervasive (Koons and Alex 2014). By around AD 850, the southern Mochica culture was absent entirely, and cultural complexity of commensurate scale was subsequently absent for centuries. These changes have led many Mochica specialists to posit instances in which El Niño flooding, drought, and/or dune inundation may have been catalyzed the observed cultural changes (e.g., Bawden 1996; Moseley et al. 2008; Dillehay and Kolata 2004; Shimada 1994). While plausible, these arguments lack strong empirical backing in terms of baseline environmental data, as they were built on piecemeal inferences of natural changes observed from impacts of individual environmental events at specific locations and infrastructure rather than upon observations of long-term data sufficient to contextualize the severity of the observed environmental events. In absence of such long-term data, it has remained difficult to interpret the degree to which these isolated observations represent truly unusual natural activity, and in turn the possible cultural ramifications.

In light of data presented here, however, it is clear that the initiation of observed changes in the southern expression of Mochica culture were probably coeval with a period of pronounced environmental change of the sort that can be expected to have impacted significant cultural resources and activities such as irrigation canals, architecture, and marine harvesting. Specifically, between roughly AD 550 and 660 there were at least two major El Niño events in this area, accounting for roughly 9.4% of the total observed flood accumulation, including the fifth-strongest (in terms of accumulation) event over the last 2,000 years. These El Niño events were also coupled with a powerful local earthquake at around AD 600, which would have presented mitigation challenges in its own right, but which also would have spawned a period of increased eolian inundation as newly deposited coastal sediments were entrained into agricultural infrastructure on the drought-afflicted landscape. Altogether, then, the data presented by this study imply a suite of unusual and convergent changes to which may have been collective stimuli for the initiation of local cultural change, including strong flooding, a major earthquake, pronounced coastal deposition, drought, and likely eolian inundation. Further, another very strong event at roughly AD 850—the fourth-strongest of the last 2,000 years in terms of deposition—coincides with the end of the Mochica presence in this region, indicating a possible environmental contribution to this demise.

On the other hand, the Mochica archaeological tradition as a whole was clearly also resilient to very strong El Niño flooding during its early centuries of expression and expansion. For example, between roughly AD 310 and AD 530 flood sediment accumulation rates were among the highest observed over the last 2,000 years—and much higher than during the ensuing

period described above—with two of the three strongest observed individual events included in this period. In some sense, then, the people who produced Mochica artifacts can be seen as experts at mitigating the effects of extremely powerful El Niño floods, since they evidently successfully traversed more of these extreme events than any other distinctive cultural tradition in the area. Some may even argue that these extreme events actually encouraged Mochica cultural elaboration as an adaptive mechanism in the face of environmental uncertainty (e.g., Glatz and Plourde 2011; Kornbacher 1999; Nolan and Howard 2010). In any case, the apparent Mochica adroitness in coping with powerful El Niño floods, in turn, highlights the fact that El Niño flooding per se is highly unlikely to be the primary factor in initiating the changes to the southern Mochica tradition, since the Mochica had probably already endured far worse flooding than they experienced during the 7th through the 9th centuries AD.

Of course, because scholarly understanding of the full suite of cultural patterns is ever-evolving—especially in the case of the Mochica tradition, for which scholars are currently re-evaluating fundamental aspects of long-held assumptions such as regional chronology (e.g., Koons and Alex 2014) and social organization (e.g., Quilter and Castillo 2010; Quilter and Koons 2012)—the preceding temporal connections between cultural and natural changes may also need revision soon. When this time comes, however, the long-term data provided by this study will remain similarly useful for future comparisons, as well as comparisons with additional environmental proxies within north coastal Perú. For the time being, the data and hypothesized linkages detailed in this study help provide a solid basis for framing future investigations into the geographic, infrastructural, and resource dimensions of coastal Peruvian environmental impacts.

CONCLUSIONS

Luminescence dating coastal of Peruvian El Niño sediments required the use of single-grain dating of feldspars—a very labor-intensive approach—to enable the application of statistical models such as the finite mixture model as a means for coping with poor bleaching of sampled sediments as well as the low luminescence sensitivity of Peruvian quartz, but the results appear to be worth the effort. Derived luminescence dates exhibit good overall stratigraphic agreement, and in the case of the Santa Valley sequence strong agreement with radiocarbon data. Bayesian analysis in OxCal allowed the statistical identification of two outlier dates, and therefore empirical criteria for rejection of these dates. Bayesian modeling also allowed additional compositional data derived from field observations, XRF, magnetic susceptibility, and granulometry to augment stratigraphic data, thereby allowing for the refinement of the Chao Valley flood sequence, as well as increased precision in the dating of individual events within this sequence.

The resultant OxCal depositional models provide robust long-term proxies for local El Niño–driven depositional changes over the last ~2,000 years of flooding and the last ~4,000 years of combined major earthquake/El Niño. These models accord well with non-local proxies of century- or millennium-scale changes in El Niño activity, as well as individual instances of El Niño depositional change observed from archaeological contexts and historical records and published in prior studies. Given this, the data set forth by this study provide a key for refining current interpretations of the impacts of ancient El Niño–driven landscape change on ancient north coastal Peruvians. Implications of this chronology for interpreting the case of the southern

Mochica archaeological culture were briefly explored here, but because the data presented here are likely applicable to much of the north coastal region, additional possibilities for such comparisons abound.

Future investigations have the potential to refine both the depositional models set forth here and the interpretation of the effects of the changing depositional environment on ancient peoples in the area. Such investigations should be aimed at several major objectives.

First, the studied locations in the Chao Valley and Santa Valley should be revisited, with an eye toward more refinement of the current chronology through the collection of larger samples of Chao flood sediment (to allow luminescence dating of all extant floods and to provide a sufficient volume of single grains to increase sample size available for measurement) and comprehensive documentation and sampling of the Santa Valley beach ridge complex. This will enhance the chronology set forth by this study by increasing its comprehensiveness and refining the precision of individual dates within Bayesian depositional modeling.

Second, additional locations with long-term accumulations of laminar El Niño flood sediments should be sought and similarly studied as a means of comparison with data set forth here. Such efforts have the potential to augment data derived from the current models by filling in temporal gaps in depositional, especially flood activity before roughly AD 170, which the present study was unable to document. Such efforts will also allow inter-valley and within-valley comparison of the depositional activity of ancient El Niños, thereby adding resolution to the geographic component individual depositional events and allowing the effects of ancient El Niños to be understood more fully in terms of their impacts on specific sites and infrastructure.

Third, evidence for individual El Niño events which impacted culturally important architecture, irrigation canals, and littoral areas should be sought and directly dated with luminescence and/or radiocarbon techniques to flesh out the specific connection between these individual events and the developing long-term picture of local El Niño activity as a whole. This will foster stronger causal arguments by contextualizing these individual events within long-term proxies for El Niño depositional, allowing the relative natural strength of these individual El Niño events to be more fully understood. This will also help illuminate the ways in which individual events—strong or otherwise—directly affected cultural activity, helping researchers move from hypothetical natural-cultural connections to a more concrete understanding.

Fourth, a concerted effort should be made to refine available cultural chronologies using chronometric data. Ceramic typologies have long been used to estimate ancient assemblage and occupational ages, but recent analyses show that this practice is problematic for much of the local area due to significant differences in the timing of use of distinctive decorative styles in different locations. For example, Koons and Alex (2014) provide an excellent example of the issues presented by the uncritical use of ceramic typologies for understanding Mochica history. Luminescence dating may also be helpful in this endeavor, as it holds the potential to directly date the construction and modification of architecture as well as diagnostic ceramics, thereby yielding direct chronometric data for refining our understanding of site occupation histories and existing ceramic chronologies. Error terms in chronometric data will remain a limitation on interpretive resolution, but the use of Bayesian statistics, advancing computer analytical tools, and stratigraphic information likely hold the potential to narrow these error terms if sampled

contexts are chronometrically dated cleverly and through the use of multiple samples rather than through occasional opportunistic finds. In other words, if chronometry is given priority as a primary study objective rather than used as a small component in service of other site-level interests and objectives, our temporal resolution will improve. In the arid north coast of Perú, where so much of the ancient material culture consists of earthen material, luminescence dating will be key to this undertaking; all that remains is the mobilization of will and resources necessary to perform the analysis.

Finally, effort must be made to continually refine methods for systematically comparing proxies for environmental change with one another and with proxies for cultural change. This remains an immense empirical challenge given differences in proxy scale, resolution, and type. Setting forth compelling explanatory frameworks for environmentally related cultural change also poses a significant theoretical challenge, and meeting this challenge will require the illumination of myriad causal connections through the skillful deployment of astute archaeological theory in conjunction with excellent environmental and cultural data. As Contreras (2017) aptly notes, “correlation is not enough,” and explanation of past human-environmental dynamics should be the ultimate objective of the preliminary efforts set forth by this study.

ACKNOWLEDGEMENTS

This research was funded by NSF Doctoral Dissertation Improvement Grant #0731529. Dan Sandweiss and Michael Moseley gave essential advice at the outset of this project. Jeremy Innis and Rocio de la Fuente were essential for fieldwork, and Francisco Burga facilitated site

access. Jim Feathers graciously provided laboratory access and mentorship throughout. The intelligence, patience, kindness, and encouragement of Don Grayson, Gary Huckleberry, and Frances Hayashida were also critical to the completion of this study. Rhiannon Held and Johonna Shea provided valuable assistance with technical editing and graphics.

REFERENCES

Aitken, M.J.
1985 *Thermoluminescence Dating*. Academic Press, London.

Anders, Martha B.
1981 Investigation of State Storage Facilities in Pampa Grande, Perú. *Journal of Field Archaeology* 4:391–404.

Apaestegui, J., F.W. Cruz, Abdelfettah Sifeddine, Jhan Carlo Espinoza Villar, Jean-Loup Guyot, Myriam Khodri, N. Strikis, R.V. Santos, H. Cheng, L. Edwards et al.
2014 Hydroclimate variability of the South American Monsoon System during the last 1600 yr inferred from speleothem isotope records of the north-eastern Andes foothills in Perú. *Climate of the Past Discussions* 10:533–561.

Auclair, M., M. Lamothe, and S. Huot
2003 Measurement of anomalous fading for feldspar IRSL using SAR. *Radiation Measurements* 37:487–492.

Banerjee, D., Murray, A.S., Bøtter-Jensen, L., and Lang, A.
2001 Equivalent dose estimation using a single aliquot of polymineral fine grains. *Radiation Measurements* 33:73–93.

Barber, Richard T., and Francisco P. Chavez
1983 Biological Consequences of El Niño. *Science* 222(4629):1203–1210.

Bawden, Garth
1996 *The Moche*. Blackwell Publishers, Oxford.

Beck, Susan L., and Stuart P. Nishenko
1990 Variations in the mode of great earthquake rupture along the Central Perú Subduction Zone. *Geophysical Research Letters* 17(11):1969–1972.

Beresford-Jones, David, Alexander Pullen, George Chauca, Lauren Cadwallader, Maria Garcia, Isabel Salvatierra, Olivier Whaley, Victor Vasquez, Susana Arce, Kevin Lane, and Charles French
2017 Refining the *Maritime Foundations of Andean Civilization*: How Plant Fiber Technology Drove Social Complexity During the Preceramic Period. *Journal of Archaeological Method and Theory*. doi: 10.1007/s10816-017-9341-3.

Billman, Brian R.
1996 The Evolution of Prehistoric Political Organizations in the Moche Valley, Perú. Unpublished Ph.D. dissertation, University of California, Santa Barbara.

2002 Irrigation and the Origins of the Southern Moche State on the North Coast of Perú. *Latin American Antiquity* 13(4):371–400.

Bourgeois, Joanne, Catherine Petroff, Harry Yeh, Vasily Titov, Costas E. Synolakis, Boyd Benson, Julio Kuroiwa, James Lander, and Edmundo Norabuena
1999 Geologic Setting, Field Survey and Modeling of the Chimbote, Northern Perú, Tsunami of 21 February 1996. *Pure and Applied Geophysics* 154(3–4):513–540.

Arntz, W. E., T. Brey, J. Tarazona, and A. Robles
1987 Changes in the structure of a shallow sandy-beach community in Perú during and El Niño event. *South African Journal of Marine Science* 5(1):645–658.

Bronk Ramsey, Christopher
2008 Deposition models for chronological records. *Quaternary Science Reviews* 27(1–2):42–60.

2009 Bayesian analysis of radiocarbon dates. *Radiocarbon* 51(1):337–360.

Brooks, William E., Jason C. Willett, Jonathan D. Kent, Victor Vasquez, and Teresa Rosales
2005 The Muralla Pircada – an ancient Andean debris flow retention dam, Santa Rita B archaeological site, Chao Valley, Northern Perú. *Landslides* 2:117–123.

Carre, Matthieu, Julian P. Sachs, Sara Purca, Andrew J. Schauer, Pascale Braconnot, Rommel Angeles Falcon, Michele Julien, and Daniele Lavallee.
2014 Holocene history of ENSO variance and asymmetry in the eastern tropical Pacific. *Science* 29(345):1045–1048.

Contreras, Daniel A.
2017 Correlation Is Not Enough: Building Better Arguments in the Archaeology of Human-Environment Interactions. In *The Archaeology of Human-Environment Interactions: Strategies for Investigating Anthropogenic Landscapes, Dynamic Environments, and Climate Change in the Human Past*, edited by Daniel A. Contreras. Routledge, New York.

Clement, Christopher Ohm, and Michael E. Moseley
1991 The Spring-Fed Irrigation System of Carrizal, Perú: A Case Study of the Hypothesis of Agrarian Collapse. *Journal of Field Archaeology* 18(4):425–443.

Craig, Alan K., and Izumi Shimada
1986 El Niño deposits at Batán Grande, Northern Perú. *Geoarchaeology* 1:29–38.

Dillehay, Tom D., Herbert H. Eling Jr., and Jack Rossen
2005 Preceramic irrigation canals in the Peruvian Andes. *Proceedings of the National Academy of Sciences* 102(47):17241–17244.

Dillehay, Tom D., Steve Goodbred, Mario Pino, Victor F. Vasquez Sanchez, Teresa Rosales Tham, James Adovasio, Michael B. Collins, Patricia J. Netherly, Christine A. Hastorf, Katherine L. Chiou, Dolores Piperno, Isabel Rey, and Nancy Velchoff
2017 Simple technologies and diverse food strategies of the Late Pleistocene and Early Holocene at Huaca Prieta, Coastal Perú. *Science Advances* 3(5):e1602778. doi: 10.1126/sciadv.1602778.

Dillehay, Tom D., and Alan L. Kolata
2004 Long-term human response to uncertain environmental conditions in the Andes. *Proceedings of the National Academy of Sciences* 101(12):4325–4330.

Dorbath, L., A. Cisternas, and C. Dorbath
1990 Assessment of the Size of Large and Great Historical Earthquakes in Perú. *Bulletin of the Seismological Society of America* 80(3):551–576.

Eling, Herbert Henry
1988 The role of irrigation networks in emerging societal complexity during late prehispanic times: Jequetepeque Valley, North Coast, Perú. Unpublished Ph.D. thesis, University of Texas, Austin.

Feathers, James K., Jack Johnson, and Silvia Rodriguez Kembel
2008 Luminescence Dating of Monumental Stone Architecture at Chavin de Huantar, Perú. *Journal of Archaeological Method and Theory* 15(3):266–296.

Food and Agriculture Organization of the United Nations
2018 Fish, crustaceans, molluscs, etc. capture production by countries or areas. http://www.fao.org/fishery/docs/STAT/summary/a2.pdf, accessed July 16, 2018.

Galbraith, R.F.
2005 *Statistics for Fission Track Analysis*. London: Chapman and Hall.

Galbraith, R.F., and R.G. Roberts
2012 Statistical aspects of equivalent dose and error calculation and display in OSL dating: an overview and some recommendations. *Quaternary Geochronology* 11:1–27.

Galbraith, R.F., Roberts, R.G., Laslett, G.M., Yoshida, H., and J.M. Olley
1999 Optical Dating of Single and Multiple Grains of Quartz from Jinmium Rock Shelter, Northern Australia: Part I, Experimental Design and Statistical Models. *Archaeometry* 41:339–364.

Gale, Stephen J., and Peter G. Hoare
1991 *Quaternary Sediments: Petrographic Methods for the Study of Unlithified Rocks*. New York, Halsted Press.

Garcia-Herrera, R., H.F. Diaz, R.R. Garcia, M.R. Prieto, D. Barriopedro, R. Moyano, and E. Hernandez
2008 A Chronology of El Niño Events from Primary Documentary Sources in Northern Perú. *Journal of Climate* 21:1948–1962.

Gerszten, E., M.J. Allison, and B. Maguire
2012 Paleopathology in South American Mummies: A Review and New Findings. *Pathobiology* 79(5):247–256.

Gibbs, Ronald J., Martin D. Matthews, and David A. Link
1971 The relationship between sphere size and settling velocity. *Journal of Sedimentary Research* 41(1):7–18.

Glantz, Michael H.
1996 *Currents of Change*. Cambridge University Press, Cambridge.

Glatz, Claudia, and Aimee M. Plourde
2011 Landscape Monuments and Political Competition in Late Bronze Age Anatolia: An Investigation of Costly Signaling Theory. *Bulletin of the American Schools of Oriental Research* 361:33–66.

Glynn, Peter W.
1988 El Niño-Southern Oscillation 1982-1983: Nearshore Population Community, and Ecosystem Responses. *Annual Review of Ecology and Systematics* 19:309–345.

Goudie, Andrew S., Elaine Wright, and Heather A. Viles
2002 The roles of salt (sodium nitrate) and fog in weathering: a laboratory simulation of conditions in the northern Atacama Desert, Chile. *Catena* 48(4):255–266.

Grobman, Alexander, Duccio Bonavia, Tom D. Dillehay, Dolores R. Piperno, Jose Iriarte, and Irene Holst
2012 Preceramic maize from Paredones and Huaca Prieta, Perú. *Proceedings of the National Academy of Sciences* 109(5):1755–1759.

Haas, Jonathan, Winifred Creamer, Luis Huaman Mesia, David Goldstein, Karl Reinhard, and Cindy Vergel Rodriguez
2013 Evidence for maize (*Zea mays*) in the Late Archaic (3000-1800 B.C.) in the Norte Chico region of Perú. *Proceedings of the National Academy of Sciences* 110(13):4945–4949.

Hayashida, Frances M.
2006 The Pampa de Chaparri: Water, Land, and Politics on the North Coast of Perú. *Latin American Antiquity* 17(3):243–263.

Helguero Santin, L.M., L.K. Chininin-Yamo, L.A. Zeta, K. Perez Chuquihuanca, N. Melendres-Huaman, R.A.S. Niño-Garcia, B. Torres-Marigorda, and V.R. Ocana
2018 Thermal variability and incidence of febrile syndrome in a community in northern Perú. *International Journal of Infectious Diseases* 735:270.

Hogg, Alan G., Hua Quan, Paul G. Blackwell, Mu Niu, Caitlin E. Buck, Thomas P. Guilderson, Timothy J. Heaton, Jonathan G. Palmer, Paula J. Reimer, Row W. Reimer, Christian S.M. Turney, and Susan R.H. Zimmerman
2013 SHCAL13 Southern Hemisphere Calibration, 0-50,000 Years Cal BP. *Radiocarbon* 55(4):1889–1903.

Huckleberry, Gary, and Brian R. Billman
2003 Geoarchaeological Insights Gained from Surficial Mapping, Middle Moche Valley, Perú. *Geoarchaeology* 18(5):505–521.

Huckleberry, Gary, Frances Hayashida, and Jack Johnson
2012 New Insights into the Evolution of an Intervalley Prehistoric Irrigation Canal System, North Coastal Perú. *Geoarchaeology* 27:492–520.

Huntley, D.J., and M. Lamothe
2001 Ubiquity of anomalous fading in K feldspars, and measurement and correction for it in optical dating. *Canadian Journal of Earth Sciences* 38:1093–1106.

Huckleberry, Gary, and Brian R. Billman
2003 Geoarchaeological Insights Gained from Surficial Geologic Mapping, Middle Moche Valley, Perú. *Geoarchaeology: An International Journal* 18(5):505–521.

Jaksic, Fabian M.
2001 Ecological effects of El Niño in terrestrial ecosystems of western South America. *Ecography* 24:241–250.

Jimenez, Cesar, Nabilt Moggiano, Erick Mas, Bruno Adriano, Shunichi Koshimura, Yushiro Fujii, and Hideaki Yanagisawa
2013 Seismic Source of 1746 Callao Earthquake from Tsunami Numerical Modeling. *Journal of Disaster Research* 8(2):266–273.

Johnson, Jack
2014 Accurate measurements of low *Z* elements in sediments and archaeological ceramics using portable X-ray fluorescence (PRXF). *Journal of Archaeological Method and Theory* 21:563–588.

Keefer, David K. and Michael E. Moseley
2004 Southern Perú desert shattered by the great 2001 earthquake: Implications for paleoseismic and paleo-El Niño-Southern Oscillation records. *Proceedings of the National Academy of Sciences* 101(30):10878–10883.

Keefer, David K., Michael E. Moseley, and Susan D. deFrance
2003 A 38,000-year record of floods and debris flows in the Ilo region of southern Perú and its relation to El Niño events and great earthquakes. *Palaeogeography, Palaeoclimatology, Palaeoecology* 194:41–77.

Kornbacher, Kimberly D.
1999 Cultural Elaboration in Prehistoric Coastal Perú: An Example of Evolution in a Temporally Variable Environment. *Journal of Anthropological Archaeology* 18(3):282–318.

Koons, Michele L. and Bridget A. Alex
2014 Revised Moche Chronology Based on Bayesian Models of Reliable Radiocarbon Dates. *Radiocarbon* 56(3):1039–1055.

Lai, Z.P., L. Zöller, M. Fuchs, and H. Brückner
2008 Alpha efficiency determination for OSL of quartz from Chinese loess. *Radiation Measurements* 43:767–770.

Lavado Casimiro, W.S., O. Felipe, E. Silvestre, and L. Bourrel
2013 ENSO impact on hydrology in Perú. *Advances in Geosciences* 33:33–39.

Lavado Casimiro, Waldo Sven, Josyane Ronchail, David Labat, Jhan Carlo Espinoza, and Jean Loup Guyot
2012 Basin-scale analysis of rainfall and runoff in Perú (1969-2004): Pacific, Titicaca, and Amazonas drainages. *Hydrological Sciences Journal* 57(4):625–642.

Lockard, Gregory D.
2009 The Occupational History of Galindo, Moche Valley, Perú. *Latin American Antiquity* 20(2):279–302.

Londoño, Ana Cristina, Steven L. Forman, Timothy Eichler, and James Pierson
2012 Episodic eolian deposition in the past ca. 50,000 years in the Alto Ilo dune field, southern Perú. *Palaeogeography, Palaeoclimatology, Palaeoecology* 346–347:12–24.

Manners, R.B., F.J. Magilligan, and P.S. Goldstein
2007 Floodplain Development, El Niño, and Cultural Consequences in a Hyperarid Andean Environment. *Annals of the Association of American Geographers* 97(2):229–249.

Morera, Sergio B., Thomas Condom, Alain Crave, Phillippe Steer, and Jean L. Guyot.
2017 The impact of extreme El Niño events on modern sediment transport along the western Peruvian Andes (1968-2012). Scientific Reports 7. doi:10.1038/s41598-017-12220-x.

Moseley, Michael E.
1975 *The maritime foundations of Andean civilization*. Cummings Publication Company, Menlo Park.

2001 *The Incas and Their Ancestors: The Archaeology of Perú (Revised Edition)*. Thames and Hudson, New York.

Moseley, Michael E., and R.A. Feldman
1988 Fishing, farming and the foundations of Andean civilization. In *The archaeology of prehistoric coastlines*, edited by G. Bailey and J. Parkington, pp. 125–134. Cambridge University Press, Cambridge.

Moseley, Michael E., David Wagner, and James B. Richardson, III
1992 Space Shuttle Imagery of Recent Catastrophic Change Along the Arid Andean Coast. In *Paleoshorelines and prehistory: an investigation of method*, edited by Lucille Lewis Johnson, pp. 215–236. CRC Press, Boca Raton.

Murray, A.S., and A.G. Wintle
2000 Luminescence dating of quartz using an improved single-aliquot regenerative-dose protocol. *Radiation Measurements* 32:57–73.

2003 The single aliquot regenerative dose protocol: potential for improvements in reliability. *Radiation Measurements* 37(4–5):377–381.

Netherly, Patricia J.
1984 The Management of Late Andean Irrigation Systems on the North Coast of Perú. *American Antiquity* 49(2):227–254.

Nials, F.L., E.R. Deeds, M.E. Moseley, S.G. Pozorski, T. Pozorski, and R.A. Feldman
1979a El Niño: The catastrophic flooding of coastal Perú, I. *Field Museum of Natural History Bulletin* 50(7):4–14.

1979b El Niño: The catastrophic flooding of coastal Perú, I. *Field Museum of Natural History Bulletin* 50(8):4–10.

Nolan, Kevin C., and Steven P. Howard
2010 Using Evolutionary Archaeology and Evolutionary Ecology to Explain Cultural Elaboration: The Case of Middle Ohio Valley Woodland Period Ceremonial Subsistence. *North American Archaeologist* 31(2):119–154.

Ortlieb, Luc
2000 The Documented Historical Record of El Niño Events in Perú: An Update of the Quinn Record (Sixteenth through Nineteenth Centuries). In *El Niño and the Southern Oscillation: Multiscale Variability and Global and Regional Impacts*, edited by Henry F. Diaz and Vera Markgraf, pp. 207–295. Cambridge University Press, Cambridge.

Ortloff, Charles R., Michael E. Moseley, and Robert A. Feldman
1982 Hydraulic Engineering Aspects of the Chimu Chicama-Moche Intervalley Canal. *American Antiquity* 47(3):572–595.

1983 The Chicama-Moche Intervalley Canal: Social Explanations and Physical Paradigms. *American Antiquity* 48(2):375–389.

Pillsbury, Joanne
1996 The Thorny Oyster and the Origins of Empire: Implications of Recently Uncovered Spondylus Imagery from Chan Chan, Perú. *Latin American Antiquity* 7(4):313–340.

Pozorski, Shelia
1979 Prehistoric diet and subsistence of the Moche Valley, Perú. *World Archaeology* 11:163–184.

Prescott, J.R., and J.T. Hutton
1994 Cosmic ray contributions to dose rates for luminescence and ESR dating: large depths and long time durations. *Radiation Measurements* 23:497–500.

Proyecto Chavimochic
2017 Proyecto Chavimochic Website. http://www.chavimochic.gob.pe/, accessed April 28, 2018.

Quilter, Jeffrey, and Luis Jaime Castillo, editors
2010 *New Perspectives on Moche Political Organization*. Dumbarton Oaks, Washington, D.C.

Quilter, Jeffrey, and Michele L. Koons
2012 The Fall of the Moche: A Critique of Claims for South America's First State. *Latin American Antiquity* 23(2):127–143.

Ramirez, Ivan J., Sue C. Grady, and Michael H. Glantz
2013 Reexamining El Niño and Cholera in Perú: A Climate Affairs Approach. *Weather, Climate, and Society* 5:148–161.

Rau, Pedro, Luc Bourrel, David Labat, Pablo Melo, Boris Dewitte, Frederic Frappart, Waldo Lavado, and Oscar Felipe
2017 Regionalization of rainfall over the Peruvian Pacific slope and coast. *International Journal of Climatology* 37:143–158.

Roberts, Helen M.
2007 Assessing the effectiveness of the double-SAR protocol in isolating a luminescence signal dominated by quartz. *Radiation Measurements* 42:1627–1636.

Rogers, Stacy Shafer, Daniel H. Sandweiss, Kirk A. Maasch, Daniel F. Belknap, and Peggy Agouris
2004 Coastal Change and Beach Ridges along the Northwest Coast of Perú: Image and GIS analysis of the Chira, Piura, and Colan Beach-Ridge Plains. *Journal of Coastal Research* 20(4):1102–1125.

Sandweiss, Daniel H.
1986 The Beach Ridges at Santa, Perú: El Niño, Uplift, and Prehistory. *Geoarchaeology: An International Journal* 1(1):17–28.

2003 Terminal Pleistocene through Mid-Holocene archaeological sites as paleoclimatic archives for the Peruvian coast. *Palaeogeography, Palaeoclimatology, Palaeoecology* 194:23–40.

Sandweiss, D.H., K.A. Maasch, D.F. Belknap, J.B. Richardson III, and H.B. Rollins
1998 Discussion of: Lisa E. Wells, 1996. The Santa Beach Ridge Complex: Sea-Level and Progradational History of an Open Gravel Coast in Central Perú. *Journal of Coastal Research* 12(1), 1-17. *Journal of Coastal Research* 14(1):367–373.

Sandweiss, Daniel H., Harold B. Rollins, and James B. Richardson III
1983 Landscape Alteration and Prehistoric Human Occupation on the North Coast of Perú. *Annals of the Carnegie Museum* 52(12):277–297.

Sandweiss, Daniel H., Ruth Shady Solis, Michael E. Moseley, David K. Keefer, and Charles R. Ortloff
2009 Environmental change and economic development in coastal Perú between 5,800 and 3,600 years ago. *Proceedings of the National Academy of Sciences* 106(5):1359–1363.

Sayre, Matthew P., Melanie J. Miller, and Silvana A. Rosenfeld
2016 Isotopic evidence for the trade and production of exotic marine mammal bone artifacts at Chavin de Huantar, Perú. *Archaeological and Anthropological Sciences* 8(2):403–417.

Seimon, Anton
2003 Improving climatic signal representation in tropical ice cores: A case study from the Quelccaya Ice Cap, Perú. *Geophysical Research Letters* 30(14). doi:10.1029/2003GL017191.

Shady Solis, Ruth, Jonathan Haas, and Winifred Creamer
2001 Dating Caral, a Preceramic Site in the Supe Valley on the Central Coast of Perú. *Science* 292:723–726.

Shimada, Izumi
1990 Cultural continuities and discontinuities on the northern north coast, Middle-Late Horizons. In *The Northern Dynasties: Kingship and Statecraft in Chimor*, edited by Michael E. Moseley and Alana Cordy-Collins, pp. 297–392. Dumbarton Oaks, Washington, D.C.

1994 *Pampa Grande and the Mochica Culture*. University of Texas Press, Austin.

2000 The late-prehispanic coastal states. In *The Inca World: The development of pre-Columbian Perú A.D. 1000-1543*, edited by L. Laurencich Minelli, pp. 49–110. University of Oklahoma Press, Norman.

Shimada, Izumi, and John F. Merkel
1991 Copper-Alloy Metallurgy in Ancient Perú. *Scientific American* 265:80–86.

Shimada, Izumi, Crystal Barker Schaaf, Lonnie G. Thompson, Ellen Mosley-Thompson
1991 Cultural Impacts of Severe Droughts in the Prehistoric Andes: Application of a 1,500-Year Ice Core Precipitation Record. *World Archaeology* 22(3):247–270.

Shimada, Izumi, and Ursel Wagner
2007 A Holistic Approach to Pre-Hispanic Craft Production. *Archaeological Anthropology: Perspectives on Method and Theory*, edited by James Skibo, Michael W. Graves, and Miriam T. Stark, pp. 163–197. University of Arizona Press, Tucson.

Takahashi, Ken, and Alejandra G. Martinez
2017 The very strong coastal El Niño in 1925 in the far-eastern Pacific. *Climate Dynamics*, 1–27. doi:10.1007/s00382-017-3702-1.

Tapley, Thomas D., and Peter R. Waylen
1990 Spatial variability of annual precipitation and ENSO events in western Perú. *Journal of Hydrological Sciences* 35(4):429–446.

Thomsen, K.J., A.S. Murray, L. Bøtter-Jensen, and J. Kinahan
2007 Determination of burial dose in incompletely bleached fluvial samples using single grains of quartz. *Radiation Measurements* 42:370–379.

Thompson, L.G., E. Mosley-Thompson, and J.F. Bolzan
1985 A 1500-Year Record of Tropical Precipitation in Ice Cores from the Quelccaya Ice Cap, Perú. *Science* 229(4717):971–973.

Tote, Carolien, Gerard Govers, Stijn Van Kerckhover, Isabel Filiberto, Gert Verstraeten, and Herman Eerens
2011 Effect of ENSO events on sediment production in a large coastal basin in northern Perú. *Earth Surface Processes and Landforms* 36:1776–1788.

Tykot, Robert H., Nikolaas J. Van Der Merwe, and Richard L. Burger
2006 The Importance of Maize in Initial Period and Early Horizon Perú. In *Histories of Maize: Multidisciplinary Approaches to the Prehistory, Linguistics, Biogeography, Domestication, and Evolution of Maize*, edited by John E. Staller, Robert H. Tykot, and Bruce F. Benz, pp. 187–197. Academic Press, Burlington, Massachusetts.

Uceda, Santiago
2010 Theocracy and secularism: relationships between the temple and urban nucleus and political change at the Huacas de Moche. In *New Perspectives on Moche Political Organization*, edited by Jeffrey Quilter and Luis Jaime Castillo, pp. 132–158. Dumbarton Oaks, Washington, D.C.

Uceda, Santiago, and Moises Tufinio
2003 El Complejo Aquitectonico Religioso Moche de Huaca de la Luna: Una Aproximacion a Su Dinamica Ocupacional. In *Moche: Hacia el Final del Milenio*, pp. 179–228. Universidad Nacional de Trujillo and Pontificia Universidad Catolica del Perú, Trujillo and Lima.

Waters, Michael R., Steven L. Forman, Thomas A. Jennings, Lee C. Nordt, Steven G. Driese, Joshua M. Feinberg, Joshua L. Keene, Jessi Halligan, Anna Lindquist, James Pierson, Charles T. Hallmark, Michael B. Collins, and James E. Wiederhold
2011 The Buttermilk Creek Complex and the Origins of Clovis at the Debra L. Friedkin Site, Texas. *Science* 331(6024):1599–1603.

Wells, Lisa E.
1990 Holocene history of the El Niño phenomenon as recorded in flood sediments of northern coastal Perú. *Geology* 18:1134–1137.

1996 The Santa Beach Ridge Complex: Sea-Level and Progradational History of an Open Gravel Coast in Central Perú. *Journal of Coastal Research* 12(1):1–17.

Winsborough, Barbara M., Izumi Shimada, Lee A. Newsom, John G. Jones, and Rafael A. Segura
2012 Paleoenvironmental catastrophes on the Peruvian coast revealed in lagoon sediment cores from Pachacamac. *Journal of Archaeological Science* 39:602–614.

Wintle, A.G.
1973 Anomalous fading of thermo-luminescence in mineral samples. *Nature* 245:143–144.

Zhang, Zhaohui, Guillaume Leduc, and Julian P. Sachs
2014 El Niño evolution during the Holocene revealed by a biomarker rain gauge in the Galapagos Islands. *Earth and Planetary Science Letters* 404:420–434.

OXCAL 4.3.2 CODE

Code 1: Input code for modeling ages of Santa Valley beach ridges, including both single-grain luminescence dates and radiocarbon dates.

```
Plot()
{
Sequence("Santa Ridges")
{
Sigma_Boundary();
R_Date("R1(SI-4957)",4235,115)
{
z=100.0;
};
C_Date("R1(UW1814)",-1920,160)
{
z=100.0;
};
C_Date("R2(UW1840)",-300,170)
{
z=86.8;
};
C_Date("R3(UW1839)",-450,150)
{
z=77.9;
};
R_Date("R4(OS-72843)",2290,35)
{
z=70.7;
};
C_Date("R4(UW1843)",10,70)
{
z=70.7;
};
C_Date("R5(UW1844)",600,80)
{
z=64.9;
};
Boundary("R6(i)")
{
z=48.0;
};
```

```
Boundary("R7(i)")
{
z=39.8;
};
R_Date("R8(OS-72846)",230,25)
{
z=36.3;
};
C_Date("R8(UW1845)",1720,40)
{
z=36.3;
};
Boundary("R9(i)")
{
z=27.3;
};
Boundary("R10(i)")
{
z=19.3;
};
Boundary("R11(i)")
{
z=14.9;
};
R_Date("R12(OS-72850)",240,25)
{
z=12.2;
};
C_Date("R12(UW1815)",1890,30)
{
z=12.2;
};
Boundary("R13(i)")
{
z=5.0;
};
C_Date("R14(OS-72851)",1980,30)
{
z=0;
};
Boundary();
```

```
};
};
```

Code 2: Input code for modeling stratigraphic compatibility of all Chao Valley single-grain luminescence dates.

```
Plot()
{
Sequence()
{
Sigma_Boundary();
C_Date("S1",-1800,160)
{
z=100;
};
C_Date("S2",270,140)
{
z=97.1;
};
C_Date("S3",410,180)
{
z=87.1;
};
Boundary("S4(i)")
{
z=82.4;
};
C_Date("S5",470,380)
{
z=79.4;
};
Boundary("S6(i)")
{
z=69.4;
};
C_Date("S7",130,530)
{
z=68.2;
};
C_Date("S8",620,210)
{
```

```
z=64.7;
};
C_Date("S9",-200,310)
{
z=58.8;
};
C_Date("S10",470,190)
{
z=55.3;
};
Boundary("S11(i)")
{
z=51.2;
};
C_Date("S12",790,370)
{
z=43.5;
};
C_Date("S13",1120,210)
{
z=28.8;
};
Boundary("S14(i)")
{
z=25.3;
};
C_Date("S15",1070,170)
{
z=21.8;
};
C_Date("S16",1140,380)
{
z=18.2;
};
C_Date("S17",270,600)
{
z=12.9;
};
C_Date("S18",1380,90)
{
z=10.6;
```

```
};
C_Date("S19",1490,60)
{
z=8.2;
};
C_Date("S20",1520,70)
{
z=3.5;
};
C_Date("S21",1830,70)
{
z=1.2;
};
Boundary("S22(i)")
{
z=0;
};
Boundary();
};
};
```

Code 3: Input code for depositional modeling of Chao Valley single-grain luminescence dates with outliers (single-grain dates from UW2098 and UW2103) removed and the fine-grain date for UW2103 used instead.

```
Plot()
{
Sequence()
{
Sigma_Boundary();
C_Date("S1",-1800,160)
{
z=100;
};
C_Date("S2",270,140)
{
z=97.1;
};
C_Date("S3",410,180)
{
z=87.1;
```

```
};
Boundary("S4(i)")
{
z=82.4;
};
C_Date("S5",470,380)
{
z=79.4;
};
Boundary("S6(i)")
{
z=69.4;
};
C_Date("S7",130,530)
{
z=68.2;
};
C_Date("S8",620,210)
{
z=64.7;
};
Boundary("S9(i)")
{
z=58.8;
};
C_Date("S10",470,190)
{
z=55.3;
};
Boundary("S11(i)")
{
z=51.2;
};
C_Date("S12",790,370)
{
z=43.5;
};
C_Date("S13",1120,210)
{
z=28.8;
};
```

```
Boundary("S14(i)")
{
z=25.3;
};
C_Date("S15",1070,170)
{
z=21.8;
};
C_Date("S16",1140,380)
{
z=18.2;
};
C_Date("S17(fg)",1360,80)
{
z=12.9;
};
C_Date("S18",1380,90)
{
z=10.6;
};
C_Date("S19",1490,60)
{
z=8.2;
};
C_Date("S20",1520,70)
{
z=3.5;
};
C_Date("S21",1830,70)
{
z=1.2;
};
Boundary("S22(i)")
{
z=0;
};
Boundary();
};
};
```

Code 4: Input code for depositional modeling of the Chao Valley flood sequence with additional stratigraphic constraints derived from inferred periods of drought.

```
Plot()
{
Sequence()
{
Sigma_Boundary();
C_Date("S1",-1800,160)
{
z=100;
};
C_Date("S2",270,140)
{
z=97.1;
};
C_Date("S3",410,180)
{
z=87.1;
};
Boundary("S4(i)")
{
z=82.4;
};
C_Date("S5",470,380)
{
z=79.4;
};
Boundary("S6(i)")
{
z=69.4;
};
C_Date("S7",130,530)
{
z=68.2;
};
C_Date("S8",620,210)
{
z=64.7;
};
Boundary("S9(i)")
```

```
{
z=58.8;
};
Top_Hat("590_Drought",590,40)
{
z=57.1;
};
C_Date("S10",470,190)
{
z=55.3;
};
Boundary("S11(i)")
{
z=51.2;
};
C_Date("S12",790,370)
{
z=43.5;
};
C_Date("S13",1120,210)
{
z=28.8;
};
Boundary("S14(i)")
{
z=25.3;
};
C_Date("S15",1070,170)
{
z=21.8;
};
C_Date("S16",1140,380)
{
z=18.2;
};
C_Date("S17(fg)",1360,80)
{
z=12.9;
};
Top_Hat("1280_Drought",1280,50)
{
```

```
z=11.8;
};
C_Date("S18",1380,90)
{
z=10.6;
};
C_Date("S19",1490,60)
{
z=8.2;
};
C_Date("S20",1520,70)
{
z=3.5;
};
Top_Hat("1790_Drought",1790,90)
{
z=2.4;
};
C_Date("S21",1830,70)
{
z=1.2;
};
Boundary("S22(i)")
{
z=0;
};
Boundary();
};
};
```

Chapter 6. Afterword

The preceding pages have set forth an array of new geoarchaeological insights. Among them are 1) a characterization and diagnosis of limitations in the use of portable X-ray fluorescence (PXRF) for the measurement of accurate elemental concentrations, as well as novel strategies and tools useful for addressing these limitations, 2) novel strategies for the use of PXRF to provide relative ages of petroglyphs using manganese weathering rinds as a proxy for age, 3) rigorous documentation of the age and depositional history of the regionally significant Bear Creek archaeological site, and 4) rigorous documentation of late Holocene El Niño–driven depositional activity in the north coast of Perú. These insights have yielded significant new data which are useful their own right, and they also attempt to advance current geoarchaeological practice by creatively refining the ways geochronometric data are produced and evaluated through the application of currently available instrumentation and software tools. Due to the archaeometric approach of this work, it sacrifices investigative breadth on the altar of rigor to some extent, but it nonetheless succeeds in establishing research contributions which are multifaceted, substantial. More importantly, these contributions are practically useful to a variety of future applications of PXRF and geochronometry, as well as future archaeological investigations into the study areas and periods scrutinized here. Ultimately, the extent to which future research employs and expands upon the insights of this work will be the definitive test of its value and impact.

To date, this impact has been mixed. For example, the first paper in this volume has been frequently cited since its online publication in 2012, and has therefore clearly gained some

traction in the ongoing exploration of PXRF applications in archaeology and the geosciences. Cited uses of this research have not made use of the protocol and software developed by this study, however, and instead typically refer to this research to simply illustrate that PXRF application can be complex in practice. Thus while I am pleased that this work has helped convince others that PXRF use is complex, I am troubled that routine use of PXRF for archaeological applications appears to have nonetheless been relatively unaltered. On that front, the impact of this work has fallen short. As a result, many archaeologists are still likely to be churning out inaccurate PXRF data, because the central issue of quality control in our PXRF data is still not commonly being addressed with systematic efforts at remediation. This situation will likely continue until detailed protocols for each intended application are rigorously documented, strictly followed and evaluated during data collection, and detailed and cited in resulting publications.

The intellectual impacts of the second paper also remain quite limited, although realistically this is partly the result of the limited scope and distribution of this research to date. Still, this work has had practical impacts. For example, its documentation of distinct ages of petroglyphs—some very ancient—has helped support an "eligible" designation for the National Register of Historic Places under Criterion D, as it has shown that important information about the past can be gleaned from the petroglyphs themselves. Once this work is made more accessible to researchers by publication in this volume, I am hopeful that it will help stimulate similar applications to petroglyphs elsewhere. For my part, I have been discussing the possibility of additional case studies which apply PXRF to the study of petroglyphs, with possible

collaborators, and I am hopeful that major new investigations of petroglyphs along the Columbia River will take shape in the years to come. If so, this study will target larger research objectives, including the possibility of using PXRF to help seriate distinctive petroglyph motifs across a wider geographic area than the single site studied here.

The third paper, by contrast, was successful in firmly establishing the antiquity of the Bear Creek Site, but to date it has done little to influence geochronometric practice in Washington cultural resources management (CRM) contexts. In part, this limited impact on practice is due to limitations on opportunities; it is rare that a CRM project enjoys the resources and timeline necessary to enable detailed study of site formation processes and depositional chronology. In part, this is also due to the unique incentive structure which governs CRM investigations; a premium is placed of efficiency at the expense of thoroughness and innovation, so even readily available and no-cost techniques like Bayesian modeling in OxCal are adopted slowly, if at all. Finally, it is in part also due to insularity of the publications comprising the Washington CRM "gray literature"; professional CRM archaeologists rarely have time to widely read site reports as a means of engaging in regional synthesis or methodological advancement, but instead typically ingest these reports as a means of narrowly contextualizing the known cultural history of the next location and project on the docket. Efficient focus on paying objectives is simply too powerful a motive to leave much time for inessential tasks like wider gray literature review. Yet publication here may help garner some attention for this work, although the true test will arise the next time a major CRM excavation of a pre-contact site is undertaken in Washington State. In that circumstance, the convergence of budgets, timelines,

incentives, and analytical needs will again favor the intensive application of luminescence dating and Bayesian analytical techniques. The present study will hopefully provide a blueprint for such work when the opportunity arises.

The impact of the final paper is yet to be determined, as this paper is original to this volume, and its data and conclusions have not yet been widely circulated for review. Given the fact that it produces new lines of evidence useful for the examination of long-standing archaeological questions, however, I anticipate that this study will be received with interest by archaeologists working in Perú. I suspect that there will be some disappointment with error terms in dates—isn't there always?—but the simple fact is that slightly blurry data are miles better than no data at all, so I think these researchers will see the value in this work, as well as the analytical potential it can lend to their own interpretations and future work. Hopefully, this research will also convince others to use luminescence dating in their archaeological investigations in Perú. Despite the fact that it has been roughly 15 years since Jim Feathers and I dated our first Peruvian samples, the use of luminescence techniques in this region is still highly uncommon. It is time to address this situation given the geoarchaeological potential for this technique in the region.