Received: 20 September 2023 | Revised: 16 May 2024 | Accepted: 8 July 2024

DOI: 10.1002/gea.22018

RESEARCH ARTICLE

Geoarchaeology An International Journal

WILEY

Non-destructively characterizing sandstones, orthoquartzites, agates, and petrified wood for provenance research: Perspectives from the Southeastern Coastal Plain, United States

Simon P. Sherman III[1,2] | **Ryan M. Parish**[1] | **Youngsang Kwon**[1] | **Steven Meredith**[3] | **David Johnson**[4]

[1]Department of Earth Sciences, University of Memphis, Memphis, Tennessee, USA

[2]Department of Anthropology, Washington University in St. Louis, St. Louis, Missouri, USA

[3]Cedars Consulting, LLC, Epes, Alabama, USA

[4]Alabama Archaeological Society, Livingston, Alabama, USA

Correspondence
Simon P. Sherman III, Department of Anthropology, Washington University in St. Louis, St. Louis, MO 63130, USA.
Email: Simonsherman13@gmail.com

Scientific editing by Sarah Sherwood.

Abstract

Siliceous sandstone (including quartzites), petrified wood, and agates located in Alabama and Mississippi were utilized as a toolstone resource during every recognized cultural period in the Lower Mississippi Valley region of the Southeastern United States. Regrettably, these materials have not been the focus of many provenance-related investigations. Recent analyses of quartzite and sandstone from other regions in North America and from the Pyrenees were successful in discriminating sources using petrographic techniques. The current study examines the application of visible/near-infrared reflectance and Fourier transform infrared reflectance (FTIR) spectroscopy on sourcing siliceous materials besides chert, particularly sandstones, orthoquartzites (quartz sandstone), petrified woods, and agates. This source characterization investigation focuses on a case study involving materials gathered from eight distinct collection sites, encompassing nine different siliceous resources collected in Alabama and Mississippi. These materials were sourced from two distinct geological formations: the Hattiesburg and Tallahatta. Results demonstrate the ability of non-destructive reflectance spectroscopy and introduces a new outlier modeling method that detects, clusters, and separately models outliers with their own set of basis vectors. Principal component analyses, least absolute shrinkage and selection operator regression, linear discriminant function analysis (LDA), and random forest classification are used in this paper to better identify outlier elements as well as discriminate for stone materials accurately (between 67% and 100%). Although this is the first reflectance spectroscopy investigation used to characterize these materials for provenance applications, the preliminary results compare favorably with other provenance techniques whose aim is to quantify between-formation (inter) and within-formation (intra) outcrop variation. The quantified and differentiated sources, based on the hyperspectral signatures of the material, will provide a better understanding of prehistoric reliance on these lithic resources and produces a proxy to determine mobility, social interaction, and other past behavior.

KEYWORDS
Lower Mississippi Valley, machine learning, provenance, quartzite, reflectance spectroscopy, sandstone, toolstone

1 | INTRODUCTION

Sandstone and quartzite utilization and their distribution by prehistoric groups are comprehensive and extend from the Paleolithic through the Holocene, as evidenced by archaeological deposits discovered in Africa (Hay, 1976; Willoughby, 2007), Australia (Holdaway & Douglass, 2015), in Asia (China) (Wang et al., 2014), as well as Belgium (Blomme et al., 2012; Cnudde et al., 2013; Veldeman et al., 2012) and Spain (Prieto et al., 2019, 2021a, 2021b) in Europe.

In Europe, preliminary structural and chemical studies have been conducted on sedimentary quartzites (Cnudde et al., 2013), in addition to sourcing studies of archaeological quartzites at Mesolithic sites in Belgium (Blomme et al., 2012; Veldeman et al., 2012). Other explorations into sandstone and quartzite characterization have taken place on outcrops, as well as archaeological assemblages from the southeastern Pyrenees mountains (Roy Sunyer et al., 2017), including similar studies of the adjoining region of Picos de Europa of the Cantabrian Mountains of northern Spain (Prieto et al., 2019, 2021a, 2021b). These research projects have integrated petrography, digital image processing, and X-ray fluorescence (XRF) as analytical tools, to better model the procurement strategies, mobility, and territorial management of middle Paleolithic groups.

In North America, hunter-gatherers were quarrying, working, and exchanging these types of materials to make stone implements up to 9000 years ago, especially in portions of east-central Mississippi and southwestern Alabama, and until the 19th century, where early settlers used siliceous sandstone for millstones (McGahey, 1999; McGahey et al., 1992). Studies conducted recently in both the western and eastern regions of the United States have effectively identified and traced the origins of these material types (Dalpra, 2016; Dalpra & Pitblado, 2016; Dunning, 1964; Ensor, 1981; Peacock, 1995; Pitblado et al., 2013; Price et al., 2009; Raber, 2021; Starnes, 2015a, 2015b, 2016, 2019a, 2019b; Starnes & Leard, 2020).

Sandstone (including highly cemented quartzite), petrified wood, and agate (banded chalcedony) have been found in the toolkits of many of the recognized cultural groups in the Southeastern United States and their use extends from the Paleoindian (13,000–10,000 B.P.) and Archaic (10,000–2500 B.P.) periods, to the end of the Woodland (2500–1000 B.P.) period. These material types naturally occur in surficial deposits over a wide portion of the Coastal Plain (Brooms, 1977; Dunning, 1964; Thomas et al., 1982; Walthall, 1980), moving westward into the Tombigbee River Valley (Bense, 1983; Ensor, 1981; Jenkins, 1982), then in an eastern direction toward the Chattahoochee River Valley (Lloyd et al., 1983; White, 1981) (Figure 1). Notable geoarchaeological investigations in eastern North America include examinations on the Tallahatta and Kosciusko sandstone and quartzite sources within Alabama and Mississippi (Dunning, 1964; McGahey, 1999; McGahey et al., 1992; Peacock, 1995; Starnes, 2015a, 2019a, 2019b), the Cockfield and Hattiesburg formations in Mississippi (Starnes, 2019b; Starnes & Leard, 2020), Sioux orthoquartzite (Starnes, 2015b, 2019a, 2019b), the South Bay quarries in Massachusetts (Lemire, 1975), and Lower Juniata and Susquehanna Valleys in Pennsylvania (Raber, 2021). In the American West, archaeological investigations were successful in chemically and petrographically discriminating distinct quartzite sources in the Upper Gunnison Basin, Colorado, using a host of analytical applications (Dalpra, 2016; Dalpra & Pitblado, 2016; Pitblado et al., 2013).

Lithics are arguably the most-recovered artifact type in archaeological assemblages, and cherts and flints are commonly the most encountered material. Used to produce stone tool implements, sandstone and quartzite are frequently overlooked, although they have the potential to produce a strong and long-lasting edge when conchoidally fractured. Both materials are ideal for creating flake tools that serve a variety of purposes for cutting and scraping (Raber, 2021). Further, Tallahatta sandstone is special among coastal plain rocks for its hardness and durability.

There has been limited research focused on the origin of quartzites, sandstones, agates, and petrified woods, with only a few analytical techniques applied to these materials. Fortunately, the few published efforts to discriminate among sources have been largely successful (e.g., Blomme et al., 2012; Church, 1994, 1996; Cnudde et al., 2013; Dalpra, 2016; Dalpra & Pitblado, 2016; Ebright, 1987; Pitblado et al., 2013; Prieto et al., 2019, 2021a, 2021b; Stross et al., 1988; Veldeman et al., 2012). These attempts to define and characterize siliceous sandstone, quartzite, agate, and petrified wood in provenance studies, including metamorphic and sedimentary siliceous rocks, increase the range of raw materials studied by archaeologists. Provenance studies on these materials are relevant to areas such as the Southeastern United States in the Coastal Plain, where they occur within different geologic formations across southwestern Alabama, as well as across southwestern and east-central Mississippi (Figure 1). These formations contain primary outcrops, secondary deposits, and prehistoric quarries. The present-day deposits are available for collection and subsequent examination in some areas. In this paper, results are reported of the first application of non-destructive reflectance spectroscopy on 284 samples of siliceous sandstone, orthoquartzite, petrified wood, and agate from eight collection loci in the Lower Mississippi Valley (LMV). Previous investigations across many continents suggest that a degree of separation is possible between various collection locales. Accurate discrimination between these raw material sources using non-destructive analytical techniques and multiple classification models, as well as machine learning algorithms, has global implications on archaeological sourcing studies of quartzite, siliceous sandstone, agate, and petrified wood.

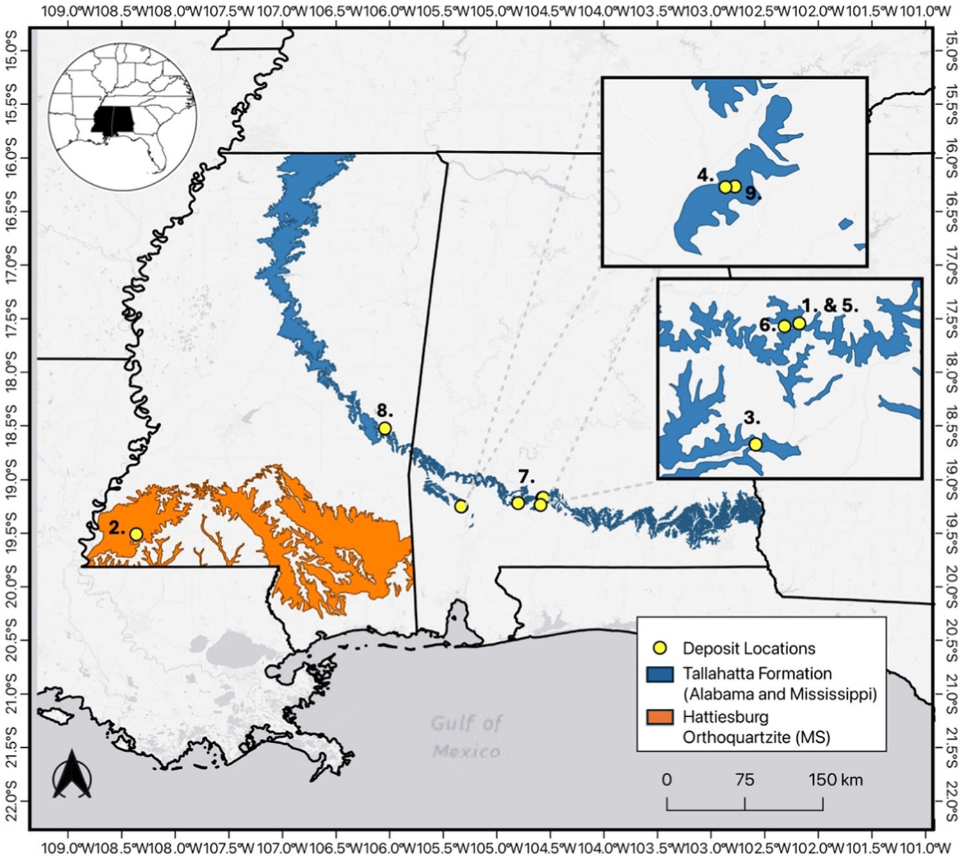

FIGURE 1 Sample locations of Coastal Plain raw material deposits sampled in the source project. Reference and artifact deposits are: 1. Petrified wood (AL); 2. Hattiesburg orthoquartzite (quartz sandstone) (MS); 3. Peterman sandstone (AL); 4. Jackson sandstone (AL); 5. Monroe sandstone 1 (AL); 6. Monroe sandstone 2 (AL); 7. Holly Mill Creek sandstone (AL); 8. Meridian sandstone (MS); and 9. Tallahatta agate (AL).

1.1 | Defining sandstone and quartzites

Formal definitions and characterization of siliceous sandstone and quartzite are problematic as the term is used loosely with ambiguous definitions. The use of the terms "sandstone" and "quartzite" is often referenced based on their geological origins, either metamorphic or sedimentary processes.

Quartzite and siliceous sandstones are alike in that both contain quartz grains; however, in quartzite, mineralization occurs due to the regrowth of silica that fill empty spaces between the grains or from the formation of cement such as opal ($SiO_2 \cdot nH_2O$), chalcedony, chert, and flint, and other silica varieties (Longwell & Flint, 1962, p. 47; Starnes, 2015a). Quartzite is a metamorphic rock derived from metamorphosed orthoquartzite, and often, the distinction between metamorphic quartzite and sedimentary orthoquartzite can appear difficult. A particular type of sandstone, quartzite is a highly lithified quartz. An informal way to recognize the characteristically high lithification of orthoquartzites is their mechanical behavior; the rock breaks through or across the grains rather than around the grains, as is the case in common sandstone (McGahey et al., 1992; Raber, 2021; Starnes, 2015a).

Sometimes, quartzite is a term used both for sedimentary and metamorphic rocks. Quartzites are organized into two categories: orthoquartzites and metaquartzites. The former are formed at temperatures below 200°C (392°F) and low depths, while the latter form at higher temperatures and higher depths (>3 km [1.86 miles]). The definitions used here for sandstones and quartzites are rocks mainly composed of sand and quartz grain modified by high pressure and heat, but it is important to note that orthoquartzite is simply a synonym for quartz sandstone, while metaquartzite is a synonym for quartzite (sensu stricto) (Crandell, 2021). It is crucial to highlight that sandstones do not necessarily contain quartz and they may have undergone lithification from sand without high temperatures and pressures. In this raw material resource investigation, a regional study was conducted to assess the potential for discriminating between different sources of sandstone from other siliceous sandstones (orthoquartzites) in addition to agate and petrified wood.

2 | STUDY AREA

2.1 | Physical setting: The Coastal Plain

The Coastal Plain of the Eastern and Southeastern United States is a low-relief physiographic region, formed during the early Mesozoic era (252 million to 66 million years ago), as a by-product of the breakup of the supercontinent Pangea. The Southeastern Coastal Plain is composed of young Cenozoic (last 66 million years) sediments that dip gently off the continent in a seaward direction into both the Atlantic Ocean and the Gulf of Mexico (National Park Service, 2022). The Coastal Plain forms a low fringe around the edge of the North American continent from New York to Texas and southward into Mexico covering over 3540 km (2200 miles). The Coastal Plain incorporates the eroded edges of the much older rocks of the Appalachian Mountains and extends southwestward into the wide lowland of the Mississippi River, known as the Mississippi Embayment. Most of the Coastal Plain material is composed of poorly consolidated clay, silt, and sand, but in the southeast, it also includes limestone (Florea & Vacher, 2009). There is no western equivalent of the coastal plain, because the younger western coast ranges drop steeply into the Pacific Ocean, making this physiographic region characteristically unique. Today, the coastal plain is drained by four major rivers: The Tombigbee and Black Warrior in the western portion; Alabama, flowing across the central portion; and the Chattahoochee in the east (Johnson et al., 2017). All 284 samples analyzed come from eight geologic deposits representing two different Cenozoic period formations—the Hattiesburg and Tallahatta formations (Table 1 and Figure 1).

2.2 | Tallahatta sandstones (Alabama and Mississippi, USA)

The Tallahatta formation is a Middle Eocene-aged (42 million years ago) geologic formation located in Alabama (AL), Florida, Georgia, and Mississippi (MS) and is part of the Claiborne Group. A cuesta, which is a hill or ridge with a gentle slope on one side, is formed by erosion-resistant Tallahatta lithologies and occurs across the region. The cuesta extends from the formation's type locality in the Tallahatta Hills of southern Alabama northwestward through Lauderdale County and into north-central Mississippi (McGahey et al., 1992, p. 38). The Tallahatta formation occurs in a thin band across the Gulf Coastal Plain in present-day south-central Alabama extending northwestward into Mississippi. Tallahatta sandstone is a grainy, sandy, micaceous sedimentary rock found outcropping in riverbeds and on ridge tops in a belt across the state's southern portion (Price et al., 2009) (Figure 1).

Tallahatta derives its name from the Creek (Muscogee [Mvskoke]) Indian term "Tallahatta," signifying "a town has been taken with many springs." Alternatively, it may originate from the Choctaw Indian words "táli," meaning "rock," and "háta," signifying "white" (Bright, 2004, p. 475; Seale, 1939, p. 178). This visual characterization is mostly accurate as the material has a gray-to-white sugary texture with scattered dark

TABLE 1 Coastal Plain raw material collection contexts.

Sampled locations with name	Description	*n*	Town, province, county, and state	Adjacent drainage and context	Geologic age	Geologic formation
Petrified wood	Deposit	34	Beatrice, Monroe County, Alabama	Private property	Between Paleozoic: Pennsylvanian and Mesozoic: Cretaceous periods	Found within deposits along the Tallahatta
Hattiesburg orthoquartzite (quartz sandstone)	Procurement site	30	Rosetta, Wilkinson County, Mississippi	Dry Creek	Cenozoic: Middle Miocene	Hattiesburg orthoquartzite
Peterman sandstone	Procurement site	30	Peterman, Monroe County, Alabama	Brushy Creek	Cenozoic: Middle Eocene	Tallahatta
Jackson sandstone	Procurement site	30	Jackson, Clarke County, Alabama	Private Property: Located on Hatchetigbee Dome	Cenozoic: Middle Eocene	Tallahatta
Monroe sandstone 1	Procurement site	30	Beatrice, Monroe County, Alabama	Private property	Cenozoic: Middle Eocene	Tallahatta
Monroe sandstone 2	Procurement site	30	Beatrice, Monroe County, Alabama	Private property	Cenozoic: Middle Eocene	Tallahatta
Holly Mill Creek sandstone	Secondary deposit site	36	Vrendenburgh, Monroe County, Alabama	Holly Mill Creek	Cenozoic: Middle Eocene	Tallahatta
Meridian sandstone	Procurement site	30	Meridian, Lauderdale County, Mississippi	Construction quarry	Cenozoic: Middle Eocene	Tallahatta
Tallahatta agate	Procurement site	34	Jackson, Clarke County, Alabama	Private Property: Located on Hatchetigbee Dome	Cenozoic: Middle Eocene	Found within deposits along the Tallahatta
	Total (*n*)	284				

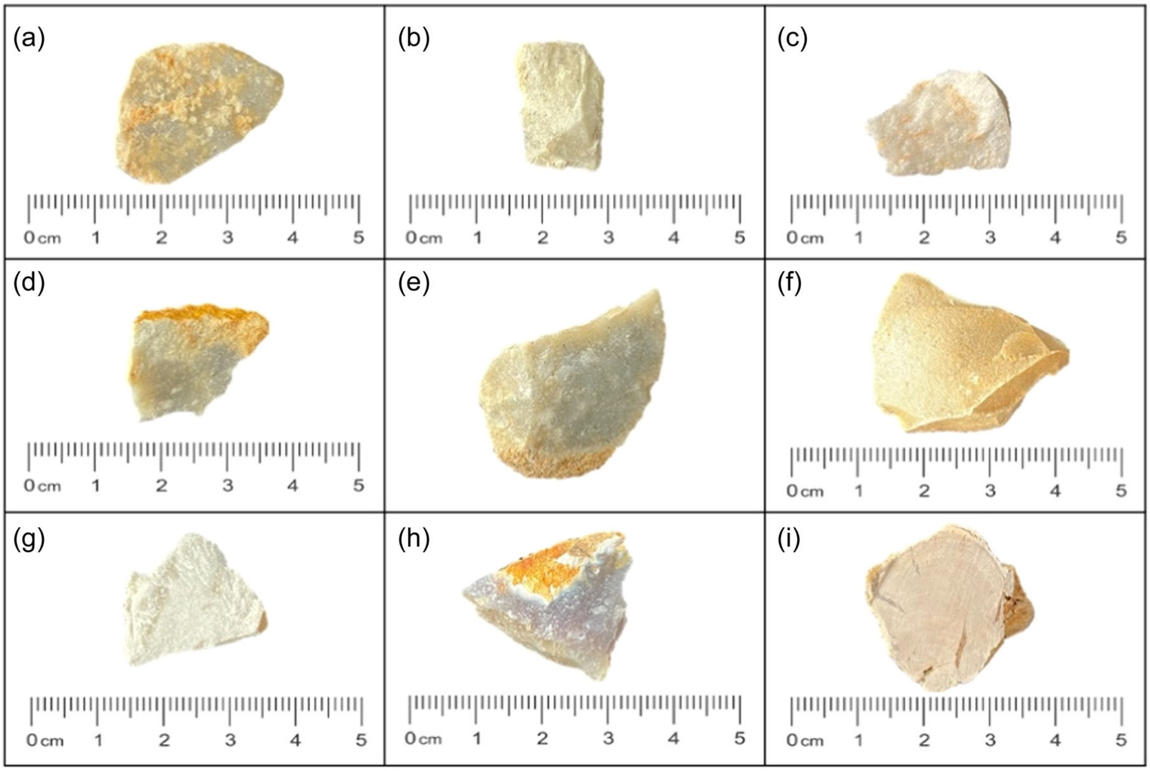

FIGURE 2 Picture of the nine different materials analyzed for this study. The material sources outlined: (a) Holly Mill Creek sandstone (AL); (b) Monroe sandstone 1 (AL); (c) Monroe sandstone 2 (AL); (d) Peterman sandstone (AL); (e) Meridian sandstone (MS); (f) Hattiesburg orthoquartzite (MS); (g) Jackson sandstone (AL); (h) Tallahatta agate (AL); and (i) petrified wood.

grains of glauconite (McGahey et al., 1992) (Figure 2a–e,g). Lenses composed of various minerals, such as chert (which is, in fact, a fine-grained rock comprised of low-crystallinity quartz), have been observed within the Tallahatta formation. These lenses show a range of sizes, spanning from a fraction of a centimeter to several centimeters in thickness (Lloyd et al., 1983). Other formations in Alabama besides the Tallahatta bear sandstone that can be used for many purposes. These sandstones have been shaped into grindstones and metates, anvil stones, abraders, stone bowls, axe heads, saws, gorgets, arrow/spear/knife points, and other utilitarian items (Johnson et al., 2017). Additionally, tractable material suitable for knapping has been identified only in Washington, Clarke, Choctaw, Monroe, and Conecuh counties of Alabama (Miller et al., 2004; Price et al., 2009) (Table 1 and Figure 1).

2.3 | Hattiesburg orthoquartzite (Mississippi, USA)

Orthoquartzite (quartz sandstone) in the southeast can be differentiated from Tallahatta sandstone in Alabama and Mississippi by how well-cemented the grain boundaries are. The Hattiesburg quartz sandstone is a Middle Miocene-aged (16 million to 11.6 million years ago) known from only a few outcrops along the uplands overlooking the Homochitto River Valley near the Amite/Franklin County line in Mississippi (Starnes, 2016, p. 1; Starnes & Leard, 2020) (Table 1 and Figure 1). This material can be characterized as hard, gray-colored, opal-cemented, siltstone to fine-grained sandstone, with artifacts weathering differently or completely to a light gray to whitish color (Starnes & Leard, 2020) (Figure 2f). Despite the extensive activity depicted at some quarry sites in Mississippi, little is known concerning the distribution and cultural use of Hattiesburg quartz sandstone. Hammerstones made from metaquartzites from outside the region have been found at prehistoric Hattiesburg orthoquartzite quarry and reduction sites. It is believed that these hammerstones were used to quarry and process the raw material (Starnes, 2015b, 2016; Starnes & Leard, 2020).

2.4 | Tallahatta agates and Chalcedony's (Alabama, USA)

Coastal Plain agate, a translucent chalcedony variant, shows a distinctive cloudy, banded pattern featuring shades of blue, purple, gray, black, and pink (Figure 2h). Typically found in thin layers and cavities within materials like siltstone and Coastal Plain chert, it often coexists with Tallahatta sandstones, with agate forming the outer layer and occasionally containing marine fossil remnants. These agates are primarily located in the southern Coastal Plain of Alabama (Dunning, 1964; Ensor, 1981; Wise, 2016, p. 11).

Agate in the LMV goes by various names like Tallahatta chert, Alabama agate, Tallahatta agate, or chalcedony in archaeological references. It forms similarly to chert, via silicification of fossil-rich limestone in the Tallahatta formation. Tallahatta agate is an opaline-rich chalcedony, discovered in the lower Eocene (56–47.8 Ma) Tallahatta Formation in Mississippi and Alabama (Brown, 2008). Besides Tallahatta, other scholars (Dunning, 1964, p. 57; Meredith & Grunewald, 2005) mention a co-existing “chalcedony” as a “secondary agatized phase of the Tallahatta.” This variant material develops in vertical cracks within the geological formation through secondary silica transport. It is high quality and often found with Tallahatta sandstone in fractures, burrow casts, and fossil molds. Tallahatta agate ranges from translucent yellow to orange in color, with color variations due to mineralogical impurities (Brown, 2008, p. 4; Dunning, 1964; Priddy, 1961; Wise, 2016). This stone material resembles Florida's Agatized Coral but can be distinguished by its laminar bedding that affects knapping.

Dunning (1964) proposes that agate may have been specifically used for robust tools due to its durability, while Tallahatta was commonly used for a wide range of tool types, drawing from various sources in the surrounding region. Outcrops are usually thin, vertical slabs with alternating layers of limestone rind and agate material that occur both in the uplands and creek banks, with colors and textures varying greatly between formations. These colors range in appearance between reds, yellows, grays, browns, and blues, while textures are predominantly waxy. It is important to specify here that the Alabama agate investigated here is not made exclusively of agate or chalcedony, which are only the main components of the investigated chert. Chert is a term that designates the rock (or nodule of a limited size) that is made of micro and cryptocrystalline varieties of quartz (like chalcedony). The predominant color observed is a translucent dark brown to amber “root beer bottle” variety like that of Knife River Flint from North Dakota. It is also sometimes platy agate or banded with coarse and chalky varieties, as well as intractable cherts. This material specification is thought to be only located in the southwestern outcrop area of the Tallahatta formation of the Hatchetigbee Hills of Clarke, Choctaw, and Washington counties of Alabama (Johnson et al., 2017, p. 41) (Table 1 and Figure 1).

2.5 | Petrified wood (Alabama, USA)

Petrified wood is a fossil formed by the infiltration or replacement of minerals into cavities between and within cells of natural wood, usually by silica (silicon dioxide, SiO_2) or calcite (calcium carbonate, $CaCO_3$), chalcedony, or agate (cryptocrystalline quartz), but sometimes by apatite (Ca_5 (PO_4)), opal, and pyrite (FeS_2). The petrifaction process occurs beneath the earth's surface, where wood has become buried under sediment (Figure 2i).

This permineralization can be so pervasive that the internal replacement of organic tissue by mineral deposits can be so precise that the internal structures as well as the external shape of the wood can be perfectly preserved. For the Alabama petrified wood samples analyzed, ancient streams traveling through a swampy delta environment deposited the wood, sediment, and debris. The wood was then quickly buried in anaerobic conditions. This type of environment occurs during or just after a major depositional event such as a big flood. Knappable petrified wood found in southeastern North America has been fossilized by the replacement of wood structures with silica minerals, making it a form of chert (Johnson et al., 2017). All petrified wood found in this investigation is from the Coastal Plain, within the Tuscaloosa group, but much of the modifiable petrified wood can also be found in Catahoula (Louisiana and MS) and Upland Complex gravel sources (AL, MS) as well as from the Tallahatta formation (Figure 1 and Figure 2i).

3 | MATERIALS AND METHODS

3.1 | Sampling

All samples collected among the eight sampling locales were sampled, but nine different types of siliceous samples were collected from known precontact procurement sites, except for samples collected at Holly Mill Creek in Alabama, which was a secondary deposit along a large creek in south central Alabama. The collection sizes chosen for each sample location can vary depending on the presence and amount of variation within each deposit (Church, 1994; Luedtke & Myers, 1984). According to Drennan (1996, p. 109), a sample size of at least 30 units per location is sufficient for large populations. The 30 or more samples collected from each deposit in this investigation are potentially representative of the intra-outcrop variation present across the spatial extent of each deposit. Eight sampling locations over a linear distance of 376 km (233.6 miles) were sampled in the current study. Samples from these eight deposits cover an approximate area of 19,500 km^2 or 7500 square miles and collectively represent the analysis group of siliceous materials examined in the study (Figure 1). The geologic reference samples collected were split open by hard hammer stone percussion, revealing a less weathered conchoidal interior surface. In the case of petrified woods, a diamond wet saw was used to slice pieces of each sample, as knapping was difficult. In total, 284 samples were analyzed from eight deposits collectively representing different raw material groups based on both lithology and geological formation in the Southeastern United States (Table 1 and Figure 2). Most of the samples were derived from Cenozoic Middle Eocene to Miocene-aged formations, while some were either Paleozoic Cambrian or Pennsylvanian epoch-aged rock types, with raw material sample names reflecting their parent geological formation. The raw material types in this paper that were investigated were extensively utilized by prehistoric individuals and groups.

3.2 | VNIR and FTIR spectroscopy

The presence of diverse toolstone deposits and outcrops presents significant challenges for archaeologists due to their inherent variability. Archaeological scientists face challenges in characterizing the chemical and mineralogical variability of stone resources. Inconsistent patterns and source identification difficulties result from compositional variations, while the presence of multiple mineral phases, limited availability of well-preserved samples, and deposits complicate accurate analysis. Furthermore, the potential alteration of stone resources over time adds further complexity to understanding their variability. Other problems include the process of connecting the stone artifact to the likely geological point of origin (Luedtke & Meyers, 1984). These problems are multiplied when researchers are working in a region with diverse and abundant sources, such as what exists in the current study. Attempts to characterize the toolstone resources and source artifacts from Coastal Plain archaeological sites in the region have, until recently, relied solely on visual analysis.

Archaeologists have applied a variety of methods in chert source studies, including instrumental neutron activation analysis (Boulanger et al., 2015; Luedtke, 1978), X-ray diffraction (Olivares et al., 2009; Shackley, 1998), XRF (Mehta et al., 2017), inductively coupled plasma-mass spectrometry (ICP-MS) (Sánchez de-La Torre et al., 2017), and thin-section petrographic analysis (Dalpra, 2016). These petrographic and geochemical methods have been used to identify potentially diagnostic mineralogical and chemical compositions within cherts and other siliceous materials such as sandstones and orthoquartzites that are characteristic of their geological source and allow artifacts to be matched to a defined geographic source region. The formation of siliceous sandstone is a diagenetic process that occurs over time, increasing the probability of variation in outcrops. The broad range in variation makes the accurate source identification of the sandstone difficult (Parish, 2011). Siliceous stone resources like cherts and flints can show as much chemical variation within a single source as between different sources (Luedtke & Meyers, 1984). Detailed sampling of raw material source locales, and the continued refinement of methods, is needed to address these problems. The application of visible/near-infrared (VNIR) and Fourier transform infrared (FTIR) reflectance spectroscopy shows promise in archaeological research, especially in chert and flint sourcing studies (Barlow et al., 2020; Parish, 2011, 2013, 2016; Parish & Durham, 2015; Parish & Finn, 2016; Parish & Werra, 2018; Sherman et al., 2023), in addition to residue analyses (Hu et al., 2013; Manthey, 2006). With the utilization of reflectance spectroscopy, researchers can non-destructively characterize chert by individual deposit and by formation through statistically identifying diagnostic spectral features caused by molecular bonding, atomic configuration, and crystalline structure (Parish, 2016).

Reflectance spectroscopy shows great potential for archaeological research applications, especially in raw material sourcing studies. Different materials produce unique reflectance patterns, within the visible, near, and middle infrared areas of the electromagnetic spectrum, which collectively represent a material's atomic structure and molecular dipole bonding, which in turn is related to its mineralogical structure, crystalline matrix, and chemical composition (Parish, 2011, p. 3, 2013, 2016). By obtaining potentially diagnostic spectral features from chert, reflectance spectroscopy can characterize both in situ parent formation deposits and secondary residuum, as small impurities and matrices create potentially diagnostic features related to diagenetic processes that can be specific to a local deposit (Parish & Finn, 2016, p. 47). Non-destructive reflectance spectroscopy has been successful in accurately characterizing and predicting cherts and flints but has never been utilized to assess the variability in other silicious stone resource types, such as sandstones, siliceous sandstones (orthoquartzites), petrified wood, and agate.

3.3 | Spectral collection and processing

The PSR + 3500 spectrometer manufactured by Spectral Evolution Inc. was used to collect data in the visible and near-infrared regions and an Agilent 4300 FTIR spectrometer was used to collect spectral data in the middle-infrared region. These instruments collect electromagnetic reflectance data with the FTIR device picking up where the VNIR device stops (2500 nm). The VNIR spectrometer measures reflectance data from 350 to 2500 nm, while the FTIR instrument measures the middle-infrared region from 2500 to 16,000 nm. Following Parish (2013, 2016), a quartz-halogen bulb was used for illuminating the artifacts for VNIR analysis. It typically takes 1 minute to record a composite electromagnetic reflectance spectrum of the specimen. A white reference reading was taken after every 10 samples from a barium sulfate ($BaSO_4$) plate to recalibrate the instrument, thereby minimizing drift. The FTIR instrument produces a middle-infrared beam of radiation that stimulates a small, 1 cm, area of the artifact held to the detector orifice. A silver (Ag) standard was used as a calibration standard between each sample reading.

All raw spectral data were processed using conventional techniques to eliminate and reduce atmospheric interference, instrument noise, sample surface roughness, probe angular effects, and standardized measurements for comparison (Parish, 2013, 2016; Parish & Durham, 2015; Parish & Finn, 2016). After processing, the reflectance spectra were converted into absorbance values, normalized, and first derivative transformed to provide a more robust means for quantitative analysis as well as to highlight subtle hyperspectral slope changes. More detailed discussions of the spectral processing stages may be found in Parish (2013, p. 176–178, 2016). The final step in the hyperspectral processing stage is to combine both the VNIR and FTIR data sets by mending the latter onto the former in a spreadsheet. Each spectral reading of sandstones (both sandstone and orthoquartzite), petrified wood, or agate sample consists of 3048 reflectance values. Each reflectance value is potentially diagnostic for toolstone type and deposit location.

3.4 | Experimental multivariate data analysis (exploratory data analysis): Machine learning

To analyze large and complex spectral data sets such as those produced here, the use of numerical and statistical tools and methods such as dimensionality reduction and local spectroscopic modeling based on dissimilarity concepts should be implemented in the analysis

(Ramirez-Lopez et al., 2022). The first data treatment in this investigation grouped all 284 samples for the eight sampling locales before separation into respective geologic parent formation and focused on identifying outliers out of the 3048 individual spectral readings. Outliers are defined in terms of the robust Mahalanobis distance using the fast minimum covariance determinant algorithm as a robust estimator of the multivariate mean and covariance from which it is computed. Once outliers had been identified and removed from the principal component analyses (PCAs) model using the *resemble*, *ChemoSpec*, *ChemoSpecUtils*, *tidyr*, *amap*, *robustbase*, and *data.table* packages in R, the samples from the nine sampled rock types, representing eight deposits, could be grouped into four classes. It is worth mentioning that PCA is not a classification method; rather, it functions solely as a dimensionality reduction tool. The primary utility in using PCA lies in exploring data dispersion. By generating plots based on the main components, one can arbitrarily group the data for analysis.

The spectral data collected for the 284 geologic samples were first organized by material type by assigning each group a classification number of one through four, representing the four different material classes (petrified wood, orthoquartzite (quartz sandstone), sandstone, and agate) from nine different collection locales. We note that the Monroe sandstone 1 and 2 sampling sites, while both representative of the Tallahatta formation, were sufficiently distant from each other to preclude inclusion in the same sampling area. Later, only the Tallahatta sandstone and quartz sandstone (orthoquartzite) samples were organized by deposit to test the intraformational variation. The consolidation of the seven groups of sandstone found within the Tallahatta and Hattiesburg formations simplifies the number of groups identified in the model, as the petrified wood and agate samples undergo different geologic formation processes.

Data treatment in this analysis aimed to eliminate non-diagnostic reflectance values using least absolute shrinkage and selection (LASSO) regression to see if prediction accuracy improved for each group. Next, the random forest classification method was used as a second line of analysis to see if predictions made by the first classification model, linear discriminant function analysis (LDA), were overly optimistic.

In the interformational tests, the initial number of labeled classes was nine, which was subsequently reduced to four following the application of LASSO to the LDA and random forest model data. For the intraformational tests, there were seven labeled classes, corresponding to the seven distinct collection locales for sandstone (including quartz sandstone) deposits within Alabama and Mississippi.

Partitioning of the data set for the LDA and random forest models used an 80/20 split between training and test populations based on the Pareto Principle (Pareto, 1896). The LDA and random forest models used on the data were created using a single random seed, without exploring different splits. Here, 80% is assigned to the training data set and 20% to the test data set. The Pareto principle states that 80% of outcomes come from 20% of causes and is optimal for assessing the accuracy and precision of big data sets. The training set is a subsection of the data set that learns or trains the data and makes the model learn the hidden patterns between features and target variables, which are labeled with known outcomes. The test set is a separate set of data used to evaluate how well the model performs and provides an unbiased final performance metric of the precision and accuracy of the model. Additional measures of accuracy assessment were used in this investigation such as cross-validation, in addition to balanced accuracy to determine the influence of imbalanced group sizes on the results of the classification.

4 | RESULTS

4.1 | Interformational variability and outlier manipulation for four raw material types

PCA was used to identify potential outliers among the 284 siliceous stone reference samples from eight sampling deposits representing four different raw material types located within the Southeastern Coastal Plain. Using PCA analysis, two samples were found to be outliers, positioned far outside the main sample clusters. This implies that, except for the two outliers, the remaining 282 samples form several identifiable cluster groups (Figure 3 and Figure S1). These samples were distinguished by a separation of two standard deviation cutoff values between PC1 and PC2, accounting for 77% of the total variability explained by the first two principal components (Figure S1). The outliers in the model are associated with two distinct sampling deposits, comprising samples from separate locations but originating from the same Tallahatta formation. Specifically, these outliers are from the Jackson sandstone source in Clarke County, Alabama, and the Meridian sandstone in Lauderdale County, Mississippi. Unfortunately, there are limited clues, whether compositional or mineralogical, to explain why these two samples deviate as outliers compared to others from the same collection locations. It is noteworthy that both sandstone sources were situated on an elevated part of the landscape and showed larger-sized cobbles compared to other sampling locations along the Tallahatta formation in Alabama and Mississippi. After the removal of the sample outliers, the PCA model was run again with 282 samples and the model's explained variance changed to PC1 (48%) and PC2 (11%), equaling 59% of the total variability encompassed by the first two principal components. In this initial phase, we aimed to pinpoint outliers beyond the calculated 6σ probability thresholds (not ellipses) within each examined raw material group or class (Figure 4). It is important to note that these processing treatments could potentially influence or distort the accuracy of source assignments when using supervised classification approaches used in the present study, such as LDA and random forest, but are worth exploring (Table 2).

4.1.1 | Interformational characterization model: LDA

The data set of 282 different samples represents four different material type groups, each consisting of 30 or more samples per

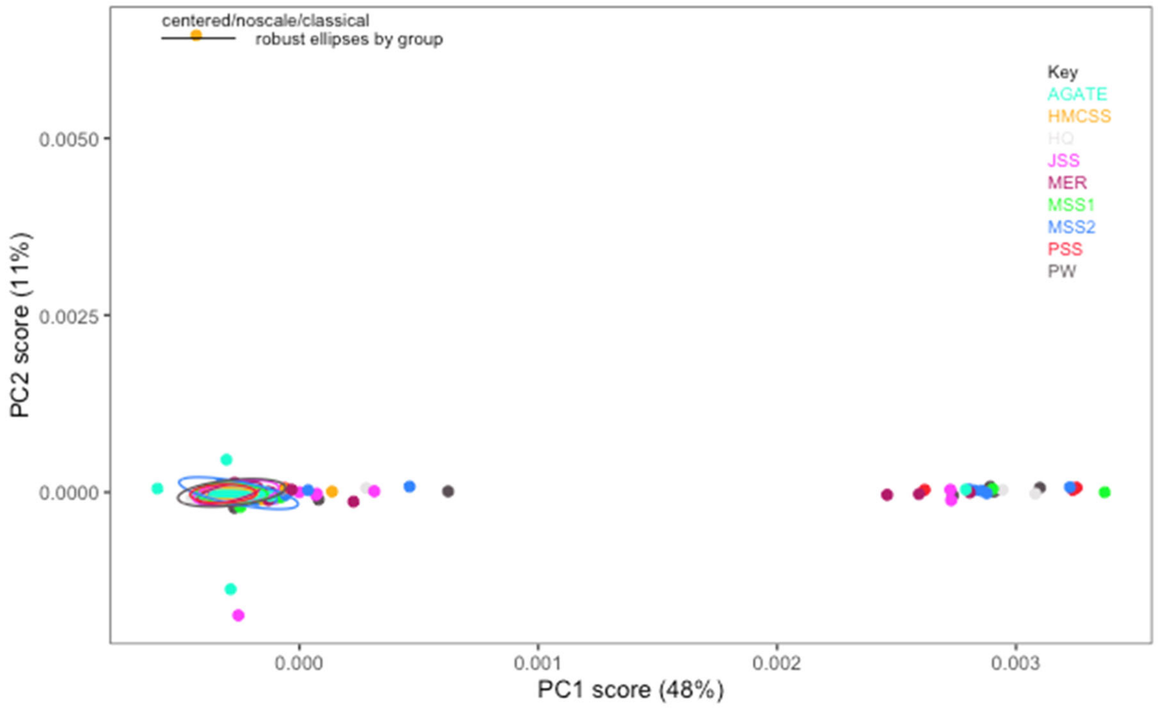

FIGURE 3 Principal component analyses (PCAs) of the eight coastal plain deposits, representing nine different types stone (AGATE, Tallahatta agate [AL]; HMCSS, Holly Mill Creek sandstone [AL]; HQ, Hattiesburg orthoquartzite [MS]; JSS, Jackson sandstone [AL]; MER, Meridian sandstone [MS]; MSS1 and MSS2, Monroe sandstone [AL]; PSS, Peterman sandstone [AL], PW, petrified wood [AL]) after the removal of two outlier samples. Note that the variability of the data is 59% encompassed by the first two principal components.

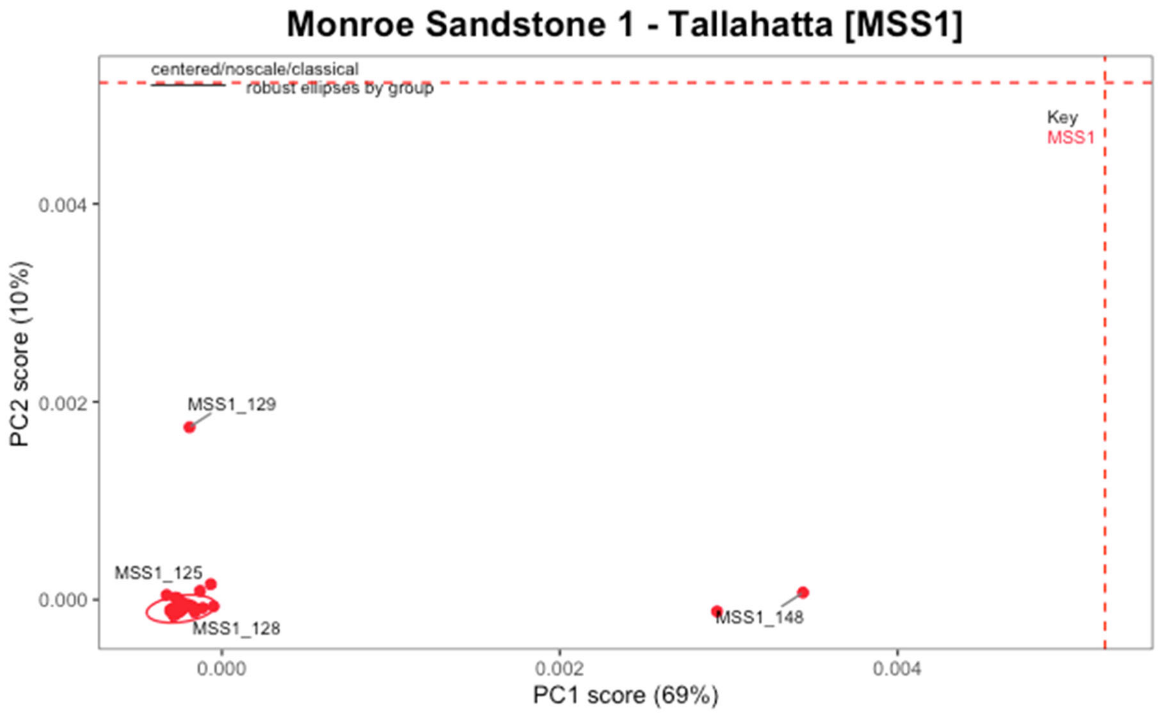

FIGURE 4 Outlier identification using principal component analyses (PCAs) was conducted for each specific group under investigation; in this figure, Monroe sandstone 1 (AL) serves as the example. The threshold values for the x- and y-axis are indicated by red dashed lines, with 6σ serving as the initial threshold. It is important to observe that 79% of the data variability is captured by the first two principal components.

group and containing more than 853,440 individual reflectance values, each potentially diagnostic for material type and/or deposit. Partitioning of the data sets into training and test groups created 228 samples for training data and 54 samples for test data. These numbers for the partitioned data sets will remain consistent throughout the study. The accuracy assessments for the data set containing all 3048 reflectance values resulted in an accuracy of 67% for the test prediction group. In addressing the model's accuracy, specifically in the context of class imbalances, cross-validation was executed on the data set, yielding an accuracy of 13% and a κ value of 0.236. After the use of LASSO regression to identify and isolate only the most diagnostic reflectance variables, 190 spectral bands from the original 3048 were retained in the model as being the most diagnostic (Figure 5).

TABLE 2 Each siliceous stone group and the variability encompassed by the first two principal component scores.

Raw material name	6σ percentage
1. Petrified wood	83%
2. Hattiesburg orthoquartzite	81.8%
3. Peterman sandstone	81.8%
4. Jackson sandstone	83%
5. Monroe sandstone 1	79%
6. Monroe sandstone 2	83.7%
7. Holly Mill Creek sandstone	86%
8. Meridian sandstone	80.8%
9. Tallahatta agate	64%

After executing the model, only using the retained 190 diagnostic wavelengths, the model's prediction accuracy improved to 98% (Table 3). The LDA model showed that the first two components (discriminant scores) account for 68% of the variability, with LD1 accounting for most of that total (~50%). Conducting a cluster analysis on all the sampled deposits using LDA scores reveals distinct separation among the nine different raw material classes into four separate cluster groups. There is a noticeable overlap between the sandstone and orthoquartzite sources, with the latter being predominantly represented in the largest LDA cluster. This overlap is expected due to higher compositional and spectral reflectance similarities between these two classes compared to agate or petrified wood samples. It is worth noting that both Hattiesburg orthoquartzite and Tallahatta sandstone are, in fact, quartz sandstones. The true relationship of the group clusters was observed in three dimensions to view the multivariant cluster in a multidimensional reality (Figure 6).

4.2 | Interformational characterization model: Random forest

The random forest classification model implemented an 80/20 split for the training and testing data sets. The number of trees in

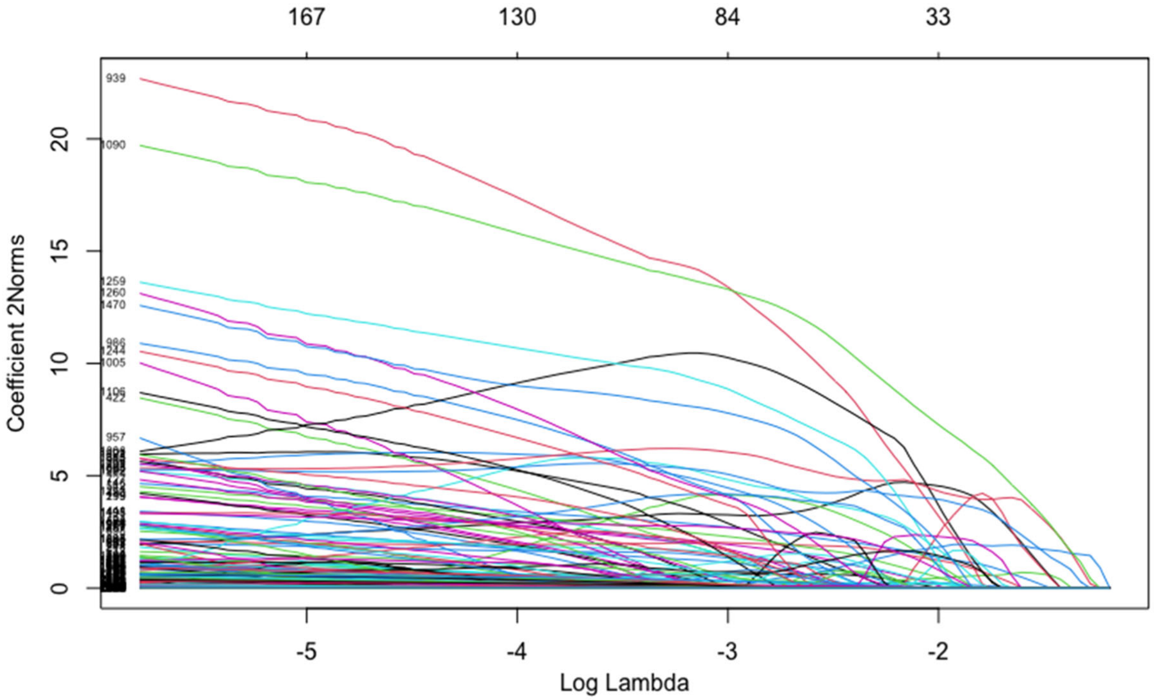

FIGURE 5 Example of the LASSO regression model for seven different sandstone deposits representing two different raw material groups from this study. Each line is one of the 179 most diagnostic spectral bands. Lambda min, the value of λ that yields the minimum mean cross-validated error, is 0.0030515.

TABLE 3 Linear discriminant function analysis (LDA) and random forest model accuracy assessments for the 282 primary analysis group samples of nine separate raw material groups.

Count	Raw material name	LDA prior probabilities	LDA prediction	LDA balanced accuracy	Random forest prediction	Random forest ICP 1	Random forest ICP 2	Random forest ICP 3	Random forest ICP 4	Random forest balanced accuracy
1	Petrified wood (1)	0.1132	2	0.5195	1	0.656	0.184	0.116	0.044	1
2	Hattiesburg orthoquartzite (2)	0.1132	2	0.4787	2	0.154	0.424	0.392	0.03	0.857
3	Peterman sandstone (3)	0.1132	3	0.4149	3	0.016	0.236	0.712	0.036	0.941
4	Jackson sandstone (3)	0.0943	3	0.5375	3	0.018	0.23	0.74	0.012	0.941
5	Monroe sandstone 1 (3)	0.1132	3	0.5301	3	0.046	0.288	0.646	0.02	0.941
6	Monroe sandstone 2 (3)	0.1132	3	0.6454	3	0.036	0.256	0.678	0.03	0.941
7	Holly Mill Creek sandstone (3)	0.1321	2	0.5171	3	0.018	0.228	0.742	0.012	0.941
8	Meridian sandstone (3)	0.0943	3	0.5583	3	0.01	0.228	0.744	0.018	0.941
9	Tallahatta agate (4)	0.1132	4	0.4149	4	0.174	0.026	0.09	0.71	1

Note: The first accuracy assessment (second column) is the actual prediction of the LDA model. The second accuracy assessment is the actual prediction accuracy when testing the model for unknown samples, which coincides with the number of samples correctly identified by raw material type (inter- and intra-raw material deposit and outcrop variation). The third column shows the prior probabilities of the classes in LDA, followed by the balanced accuracy of the LDA model. The sixth column shows the random forest prediction. The seventh through tenth columns show the individual class prediction(s) per raw material group (RF ICP 1–4). The final column shows the balanced accuracy of the random forest model.

Abbreviations: ICP, individual class predictions; LDA, Linear discriminant function analysis; RF, random forest.

this model was 500, while each split was performed at every seven samples. The prediction accuracy for the test data with the confusion matrix had a prediction accuracy of 95%, with a κ of 0.917 (Table 3). When applying random forest models to the same sample of unknowns, the prediction was just as accurate as the LDA models. Additionally, the benefits of random forest classification are that each prediction will be accompanied by a percentage of confidence per each group within the model (Table 3).

The accuracy results in interformational analysis across four raw material groups demonstrated superior performance for random forest compared to LDA. However, both were successful in discriminating between otherwise visually characteristic resource groups. Following LASSO regression and subsequent application of LDA, real-time predictions of stone achieved a total accuracy of 78%. The misclassification was petrified wood and Holly Mill Creek sandstone being misclassified as Hattiesburg orthoquartzite. In contrast, when using random forest models for interformational classification tests on a set of unknown samples, the prediction accuracy surpassed that of LDA. The random forest model accurately predicted all four types of materials in the data set (100%) (Table 3).

4.3 | Intraformational tests for Tallahatta and Hattiesburg formation

4.3.1 | Intraformational characterization model: LDA

An accuracy assessment of seven individual deposits of both Hattiesburg orthoquartzite and Tallahatta sandstone materials was conducted for all the wavelength variables before LASSO regression to measure the ability of reflectance spectroscopy to characterize and differentiate individual deposits of Tallahatta sandstone. Therefore, the data set had 3048 potentially diagnostic wavelength bands. After partitioning the samples into an 80/20 training and test data set into sample sizes of 173 for the training data set and 41 for the test data, an accuracy of 80% was achieved. Again, to account for the accuracy of the sandstone LDA model, particularly class imbalances, cross-validation was performed on the data, with an accuracy of 13% and a κ of 0.1435. After the use of LASSO regression, 179 of the most parsimonious wavelengths from the original 3048 spectra were retained. The predictive accuracy of the intraformational LDA model test data was 85%. To assess the accuracy of the sandstone LDA model, especially considering class imbalances, cross-validation was conducted on the data set, revealing an accuracy of 17% and a κ value of 0.308. This is impressive and potentially identifies intraformational differences within the Tallahatta sandstone. The LDA model showed that the first two components account for 74% of the variability among the first two discriminant function score groups, with the first discriminant function score accounting for the majority (50%) of that total (Figure 7).

4.3.2 | Intraformational characterization model: Random forest

The random forest model implemented an 80/20 split for the training and testing data sets. The number of trees in this model was 500, while each split was performed at every 11 samples for the intraformational random forest model. The prediction accuracy for

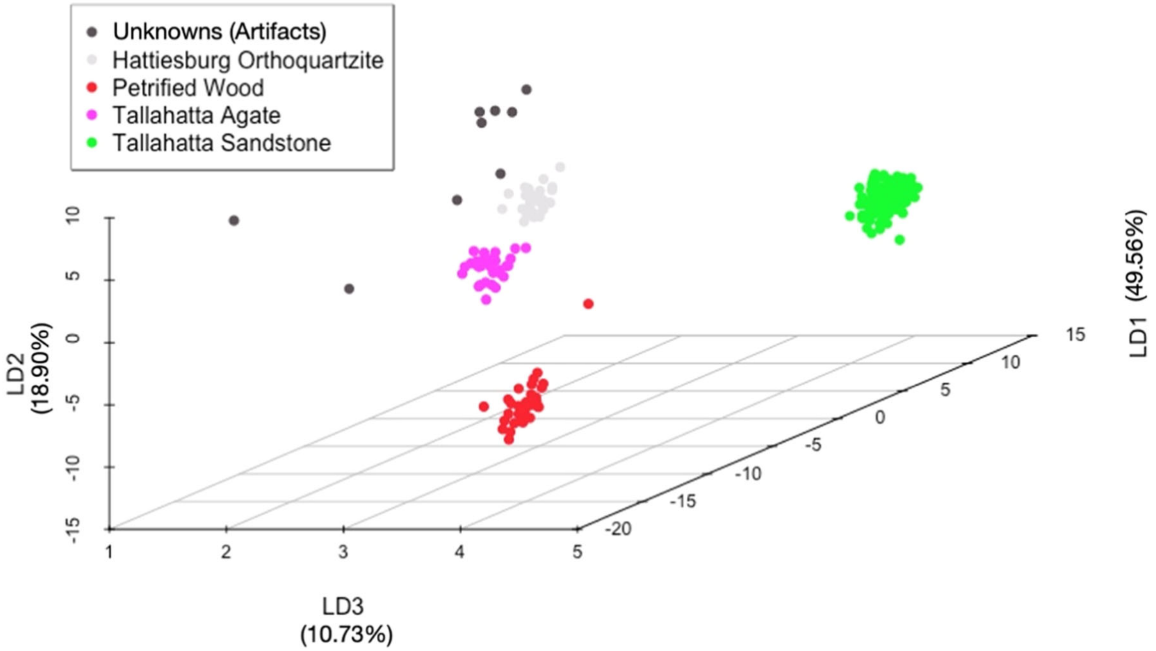

FIGURE 6 Linear discriminant function (LDA) analysis scores and 95% confidence intervals for interformational analysis of four individual material source groups (petrified wood, Hattiesburg orthoquartzite, Tallahatta sandstone, and Tallahatta agate) in three dimensions.

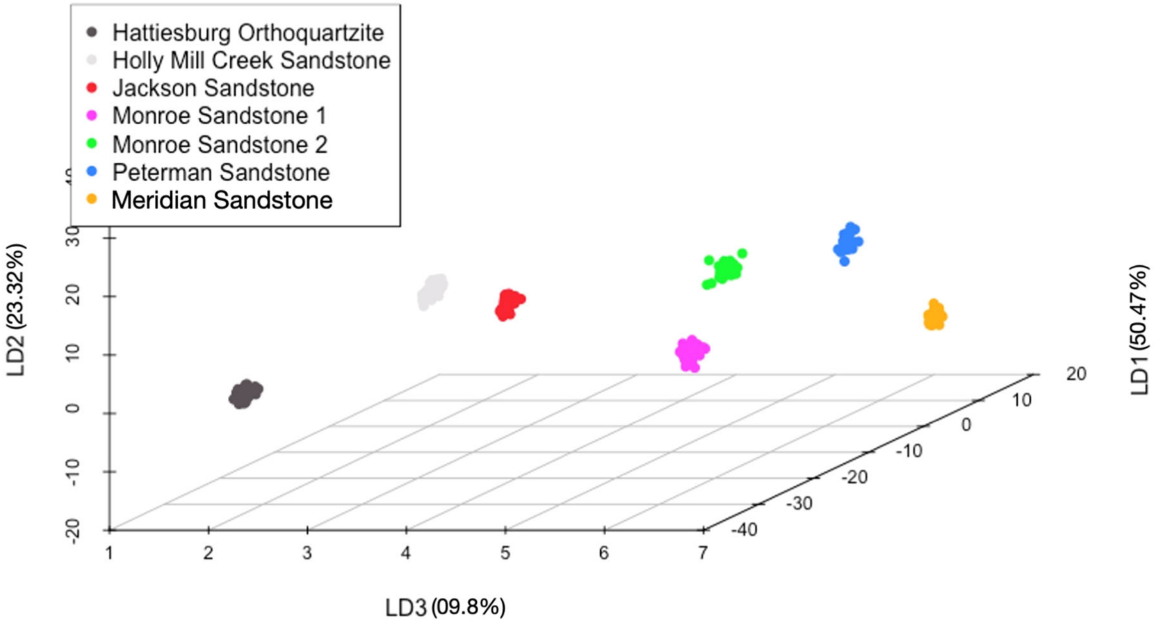

FIGURE 7 Linear discriminant function (LDA) analysis scores and 95% confidence intervals for intraformational variation tests of the seven Tallahatta and Hattiesburg sandstone deposits sampled in this study in three dimensions.

the test data with the confusion matrix had a prediction accuracy of 84%, with a κ of 0.8142. Raw material collected in southwestern Alabama classified as Monroe sandstone 1 as well as Hattiesburg orthoquartzite from southwestern Mississippi had individual class predictions exceeding 50% and outperformed other groups in accurately assigning test group samples to their respective source deposits when utilizing the random forest classification method (Table 4).

The intraformational accuracy results, encompassing seven groups including an individual Hattiesburg quartz sandstone deposit in Mississippi and six Tallahatta sandstone groups, show lower precision in the case of the LDA model. Only two real-time correct material predictions were made, constituting a total accuracy of 29%. When applying random forest models to the same sample of unknowns for the intraformational classification tests, the prediction accuracy was more accurate than those of the LDA models, with four out of the seven sandstone types being predicted accurately (57%) by the random forest model (Table 4).

5 | DISCUSSION

The major research objective in the present study was to both characterize and differentiate four major toolstone types for future artifact sourcing programs. A secondary but equally important research objective was whether individual deposits (prehistoric procurement sites) of Tallahatta sandstone could be differentiated. These objectives are necessary first steps in sourcing unknown archaeological stone artifacts such as projectile points and/or knives. The eight sampling locales, representing four raw material groups discussed in the study, capture meaningful variability for sandstones and orthoquartzites, as well as for additional stone resource materials, agate, and petrified wood.

Depending on the chosen classification approach, accurate prediction of material classes ranges from 78% to 100% across four distinct material types when applying flakes from the same deposits and treated as unknowns. The sole misclassification occurred with petrified wood using LDA, where it was erroneously identified as Hattiesburg orthoquartzite. Accuracy tests across different formations indicate reliable predictions between sandstones from the Hattiesburg and Tallahatta formations. The structural and mineralogical similarities between Tallahatta sandstone and Hattiesburg orthoquartzite pose challenges for discrimination; however, reflectance spectroscopy proved to be successful in discriminating between the two different formations. It is crucial to recognize the proximity in individual class predictions for interformational tests concerning Hattiesburg orthoquartzite and Tallahatta sandstone. The random forest prediction percentages for Hattiesburg and Tallahatta sandstone were 42% and 39%, respectively (Table 3). The results also suggest that both classification approaches are warranted for siliceous material characterization: LDA for visualizing the data in two to three dimensions and random forest for viewing the individual class prediction percentages.

According to geological descriptions by Starnes (2016) and Starnes and Leard (2020, p. 1), the Hattiesburg formation in southwestern Mississippi is considered orthoquartzite, while sandstones from the Tallahatta formation in this study are sedimentary. Strictly speaking, quartzite is a metamorphic rock and should not be used interchangeably with orthoquartzite, which is merely a sedimentary quartz sandstone.

An assumption is that the low sample size for the petrified wood sample is the primary reason for discrepancies in the prediction accuracy in the LDA classification model between the training and test groups, possibly warranting the need for more sampling of petrified wood from this location. It is also important to note that petrified palmwood was noticed during collection. The Tallahatta agate analyzed can be part of chert, and petrified wood could also constitute chert nodules.

Accurately characterizing intraformational tests for classification is challenging, regardless of the chosen classification approach. The random forest model's accuracy (57%) can vary due to its random attribute selection, necessitating an ample number of base learners for consistency. Although LDA has lower accuracy (29%) and faces challenges with inseparable data, such as sandstone deposits within a two-state formation, both random forest and LDA prove valuable for processing separable data efficiently and swiftly. Petrological characterization provides useful information and can refer to diverse features and characteristics of the rock because of their genesis, and transformation (metamorphic), sometimes being affected by physical, chemical, and biological alteration processes in subaerial weathering conditions (Prieto et al., 2019). The accuracies of the predictions made by LDA and random forest should be taken lightly in the absence of real-world data such as stone tool implements used and treated as unknowns.

In summary, the results of both LDA and random forest classification models show that the spectral signatures are unique to individual sandstone sources in the Southeastern United States. The same holds true for agate, petrified wood, and orthoquartzite (excluding metamorphic quartzite), as they are easily distinguishable on a macroscopic level. Consequently, it is expected that the reflectance data could be effectively clustered. Secondarily, the results demonstrate that the agate, petrified wood, and orthoquartzite samples separate successfully, which highlights the potential for future archaeological lithic sourcing studies using these materials. When isolating individual sandstone sources, the task of differentiation becomes more challenging, necessitating a larger number of samples from various sandstone deposits to achieve a precise identification and comprehensive characterization of distinct sandstone sources.

6 | CONCLUSION

The methodological study presented here highlights the potential of reflectance spectroscopy to characterize and differentiate various siliceous material types in the Southeastern United States and abroad. The non-destructive characterization of sandstones and orthoquartzite supports the application of a series of research

TABLE 4 Intraformational variation accuracy assessments of the Hattiesburg orthoquartzite and Tallahatta sandstone formations.

Count	Raw material name	LDA prior probabilities	LDA prediction	LDA balanced accuracy	Random forest prediction	Random forest ICP 1	Random forest ICP 2	Random forest ICP 3	Random forest ICP 4	Random forest ICP 5	Random forest ICP 6	Random forest ICP 7	Random forest balanced accuracy
1	Hattiesburg orthoquartzite (1)	0.1402	1	0.6786	1	0.518	0.082	0.092	0.064	0.112	0.092	0.04	1
2	Peterman sandstone (2)	0.1402	1	0.4429	2	0.058	0.342	0.256	0.016	0.088	0.098	0.142	0.8021
3	Jackson sandstone (3)	0.1355	5	0.5722	2	0.094	0.326	0.278	0.02	0.118	0.12	0.044	0.9839
4	Monroe sandstone 1 (4)	0.1402	4	0.5119	4	0.174	0.03	0.01	0.57	0.02	0.074	0.12	1
5	Monroe sandstone 2 (5)	0.1402	2	0.5405	2	0.09	0.358	0.15	0.004	0.162	0.154	0.082	0.8333
6	Holly Mill Creek sandstone (6)	0.1682	5	0.7122	6	0.066	0.136	0.028	0.072	0.172	0.358	0.168	0.9559
7	Meridian sandstone (7)	0.1355	5	0.5167	6	0.036	0.164	0.016	0.134	0.094	0.334	0.222	0.75

Note: The first accuracy assessment (second column) is the actual prediction of the LDA model. The second accuracy assessment is the actual prediction accuracy when testing the model for unknown samples, which coincides with the number of samples correctly identified by raw material type (intra-raw material deposit and outcrop variation). The third column shows the prior probabilities of the classes in LDA, followed by the class prediction and balanced accuracy of the LDA model. The sixth column shows the random forest prediction. The seventh through thirteenth columns show the individual class prediction(s) per raw material group (RF ICP 1–7). The final column shows the balanced accuracy of the random forest model.

Abbreviations: ICP, individual class prediction; LDA, Linear discriminant function analysis; RF, random forest.

protocols and a combination of unsupervised, supervised, and ensemble classification models that provide a more robust source assignment of lithic raw materials. Non-destructive provenance applications of agate, petrified wood, orthoquartzite, and siliceous sandstone provide a mechanism to study entire artifact assemblages, and the reasons for prehistoric use of the raw material can be better understood. The current study contributes data in two ways. The first is that it describes in detail the methodology for characterizing sandstone, orthoquartzite, petrified wood, and agate found in primary and secondary depositional settings, as well as archaeological procurement sites. The experimental data analyses and their subsequent success in discriminating from one another. Their final grouping according to the classification models based on diagnostic spectral features illustrates the characterization and differentiation of four non-chert siliceous raw material types. Second, the study presents the importance of establishing a systematic sampling methodology that obtains representative reference collections. The successful differentiation of quartzites and siliceous sandstones may require continued sampling efforts of these structurally similar materials.

The prehistoric use of sandstones and orthoquartzites, in addition to petrified woods and agates as toolstone materials, encompass broad regional and temporal spans, making the application of sandstone and quartzite source programs of paramount interest in archaeological investigations. The range of human behavioral questions potentially addressed through non-destructive source methods illustrates the need for accurate and replicable techniques. The capability of reflectance spectroscopy to accurately assign unknown artifacts is illustrated by the study's use of a “test” sample group. Both LDA and random forest are statistical multivariate methods that are more representative than traditional scatterplot graphs with ellipsoids. The goal of provenance and provenience studies is to track the movement and interaction of prehistoric people via their procurement, consumption, use, and discard of toolstone resources. Future investigations will implement these methods on archaeological materials to better gauge the range of variability present between the toolstone types of interest. While these experimental data analyses are statistically robust, the study is certainly not multitudinous or multilayered (Brandl, 2016) in its approach. The preliminary examination of sandstone, orthoquartzite, agate, and petrified wood shows that the accuracy of the classifications using non-destructive reflectance spectroscopy is encouraging, and the result of the study showcases the feasibility of future applications on these toolstone materials.

AUTHOR CONTRIBUTIONS

Simon P. Sherman: Conceptualization; writing—original draft; visualization; software; data curation; supervision; investigation; methodology; formal analysis; project administration. **Ryan M. Parish**: Funding acquisition; conceptualization; investigation; methodology; validation; project administration; resources; supervision; writing—review and editing; formal analysis. **Youngsang Kwon**: Methodology; software; formal analysis; supervision; resources; data curation; writing—review and editing; validation. **Steven Meredith**: Validation; supervision; resources; investigation. **David Johnson**: Investigation; validation; supervision; resources.

ACKNOWLEDGMENTS

The authors would like to thank the Donovan Travel Enrichment Fund at the University of Memphis and Christine Smith for processing some of the samples used in spectroscopic analyses. Additional thanks are due to Philip J. Carr (University of South Alabama) for helping contact David Johnson. Special thanks are due to the property owners who granted access to deposits sampled in this study: Mr. Walter Davis and Robert Duke, and avocational archaeologists Dickie Andress and Maury Outlaw. Without their hospitality and shared appreciation for archaeology, these efforts would not be possible. We greatly appreciate the indispensable support provided by the Alabama Archaeological Society and the Alabama Department of Transportation in facilitating access to the deposit locations. Furthermore, we extend our gratitude to James Starnes at the Mississippi Department of Environmental Quality (MDEQ) and the Mississippi Department of Archives and History (MDAH) for their valuable insights into the intricacies of raw material geology.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in the Supporting Information of this article. The source code, implemented in the R programming language, is accessible in the supplementary materials, along with the analytically processed data stored in.csv format. Additionally, the supplementary materials encompass the data frame coefficients generated through the application of Linear Discriminant Analysis (LDA) and Least Absolute Shrinkage and Selection Operator (LASSO) regression statistics. Certain deliverables or data referenced in this research article are obtainable upon request, owing to privacy and ethical constraints. Specifically, information pertaining to individual outcrops is not publicly accessible, as constrained by privacy and ethical considerations.

ORCID

Simon P. Sherman https://orcid.org/0000-0002-5549-9502

REFERENCES

Barlow, R. A., Parish, R. M., & Kilborn, R. L. (2020). Using reflectance spectroscopy to assess Paleoindian through Early Archaic Tool stone selection at the Big Sandy Site (1LI573) in Limestone County, Alabama. *Journal of Alabama Archaeology*, *66*(2), 3–17.

Bense, J. A. (1983). *Archaeological investigations in the Upper Tombigbee Valley, Mississippi, phase I, volume III*. Report of Investigation No. 1. Office of Cultural and Archaeological Research, University of West Florida.

Blomme, A., Degryse, P., Van Peer, P., & Elsen, J. (2012). The characterization of sedimentary quartzite artefacts from Mesolithic sites, Belgium. *Geologica Belgica*, *15*(3), 193–199. https://popups.uliege.be/1374-8505/index.php?id=3690

Boulanger, M. T., Buchanan, B., O'Brien, M. J., Redmond, B. G., Glascock, M. D., & Eren, M. I. (2015). Neutron activation analysis of 12,900-year-old stone artifacts confirms 450–510+ km Clovis tool-stone acquisition at Paleo Crossing (33ME274), northeast Ohio, USA. *Journal of Archaeological Science*, *53*, 550–558. https://doi.org/10.1016/j.jas.2014.11.005

Brandl, M. (2016). The multi layered chert sourcing approach (mla) analytical provenance studies of silicate raw materials. *Archeometriai Muhely*, *13*, 145–156.

Bright, W. (2004). *Native American placenames of the United States*. University of Oklahoma Press.

Brooms, B. M. D. (1977). *An Archaeological Inventory of 1Pk2 within the Walnut Creek Reservoir*. Alabama Historical Commission.

Brown, I. W. (2008). *An archaeological survey in Clarke County*. University of Alabama.

Church, T. (1994). Ogalalla orthoquartzite: An updated description. *Plains Anthropologist*, *39*(147), 53–62. https://doi.org/10.1080/2052546.1994.11931708

Church, T. (1996). Lithic resources of the Bear Lodge Mountains, Wyoming: Description, distribution and implications. *Plains Anthropologist*, *41*(156), 135–164. https://doi.org/10.1080/2052546.1996.11931804

Cnudde, V., Dewanckele, J., De Kock, T., Boone, M., Baele, J. M., Crombé, P., & Robinson, E. (2013). Preliminary structural and chemical study of two quartzite varieties from the same geological formation: A first step in the sourcing of quartzites utilized during the Mesolithic in northwest Europe. *Geologica Belgica*, *16*(1–2), 27–34.

Crandell, O. (Host) (2021, November). Episode No. 2–17—Questions about quartzite: An interview with Alejandro Prieto [Audio podcast episode]. ArchaeoCafe. http://archaeocafe.kvasirpublishing.com/archaeocafe-podcast-ep-217-prieto

Dalpra, C. L. (2016). *Petrographic quartzite source discrimination in the Upper Gunnison Basin, Colorado: Implications for the archaeological study of prehistoric hunter-gatherers* [Unpublished Master's thesis, Department of Anthropology, University of Oklahoma].

Dalpra, C. L., & Pitblado, B. L. (2016). Discriminating quartzite sources petrographically in the Upper Gunnison Basin, Colorado: implications for Paleoamerican Lithic-Procurement studies. *PaleoAmerica*, *2*(1), 22–31. https://doi.org/10.1080/20555563.2015.1137684

Drennan, R. D. (1996). *Statistics for archaeologists: A common sense approach*. Plenum Press.

Dunning, A. B. (1964). The Tallahata formation in Clarke County, Alabama. *Journal of Alabama Archaeology*, *10*(2), 50–60.

Ebright, C. A. (1987). Quartzite petrography and its implications for prehistoric use and archaeological analysis. *Archaeology of Eastern North America*, *15*, 29–45.

Ensor, H. B. (1981). *Gainesville Lake area lithics: Chronology, technology, and use. Volume III of archeological investigation in the Gainesville Lake of the Tennessee–Tombigbee waterway*. Report of Investigation No. 13. Office of Archeological Research, University of Alabama.

Florea, L. J., & Vacher, H. L. (2009). The Southeastern Coastal Plain: An overview. In A. Palmer & M. V. Palmer (Eds.), *Caves and Karst in America* (pp. 181–184). National Speological Society. http://digitalcommons.wku.edu/geog_fac_pub/22

Hay, R. L. (1976). *Geology of Olduvai Gorge: A study of sedimentation in a semiarid basin*. University of California Press.

Holdaway, S., & Douglass, M. (2015). Use beyond manufacture: Non-flint stone artifacts from Fowlers Gap, Australia. *Lithic Technology*, *40*(2), 94–111. https://doi.org/10.1179/2051618515Y.0000000003

Hu, J., Xiao, R., Shen, D., & Zhang, H. (2013). Structural analysis of lignin residue from black liquor and its thermal performance in thermogravimetric-Fourier transform infrared spectroscopy. *Bioresource Technology*, *128*, 633–639. https://doi.org/10.1016/j.biortech.2012.10.148

Jenkins, N. J. (1982). *Archaeology of the Gainesville Lake Area: Synthesis. archaeological investigations in the Gainesville Lake area of the Tennessee–Tombigbee waterway*. Office of Archaeological Research, Report of Investigation No. 23 (Vol. 5). University of Alabama.

Johnson, D., Meredith, S. M., & Dumas, A. A. (2017). *Handbook of Alabama's prehistoric Indians and artifacts* (2nd ed.). Borgo Publishing.

Lemire, R. (1975). South Bay quartzite quarry. *Bulletin of the Massachusetts Archaeological Society*, *36*, 17–21.

Lloyd, J. R., Bense, J. A., & Davis, Jr., J. L. (1983). Tallahatta quartzite quarries in the Escambia River Drainage. *Journal of Alabama Archaeology*, *29*(2), 125–142.

Longwell, C. R., & Flint, R. F. (1962). *Introduction to physical geology*. Wiley.

Luedtke, B. E. (1978). Chert sources and trace-element analysis. *American Antiquity*, *43*(3), 413–423.

Luedtke, B., & Meyers, J. T. (1984). Trace element variation in Burlington chert: A case study. In B. M. Butler & E. E. May (Eds.), *Prehistoric chert exploitation: Studies from the midcontinent*. Occasional Paper No. 2 (pp. 287–298). Center for Archaeological Investigations, Southern Illinois University.

Manthey, J. A. (2006). Fourier transform infrared spectroscopic analysis of the polymethoxylated flavone content of Orange oil residues. *Journal of Agricultural and Food Chemistry*, *54*(9), 3215–3218. https://doi.org/10.1021/jf053134a

McGahey, S. O. (1999). Use and avoidance of Kosciusko quartzite in prehistoric Mississippi flaked stone assemblages. In E. Peacock & S. O. Brookes (Eds.), *Raw materials and exchange in the mid-South: Proceedings of the 16th Annual Mid-South Conference, Jackson, Mississippi—June 3 and 4, 1995*. Archaeological Report No. 29 (pp. 1–11). Department of Archives and History.

McGahey, S. O., Dockery, D. T., & Ingram, S. L. (1992). Indian artifacts of Tallahatta quartzite from Tallahatta Creek Site 22-LD-645, East Central Mississippi. *Mississippi Geology*, *13*, 37–43.

Mehta, J. M., McCall, G., Marks, T., & Enloe, J. (2017). Geochemical source evaluation of archaeological chert from the Carson mounds site in northwestern Mississippi using portable X-ray fluorescence (pXRF). *Journal of Archaeological Science: Reports*, *11*, 381–389.

Meredith, S. M., & Grunewald, M. (2005). Projectile points from Porter Village, 1CK21, Clarke County, Alabama. *Journal of Alabama Archaeology*, *51*(1), 40.

Miller, P. E., Marine, J. T., & Vento, F. J. (2004). *Archaeological investigations, route 11/15 improvements (SR 0011, Section 008), Juniata and Perry Counties, Pennsylvania; volume III: Site 36Pe60*. Pennsylvania Department of Transportation, Engineering District 8-0. KCI Technologies.

National Park Service. (2022). Coastal Plain province [Electronic document]. https://www.nps.gov/articles/coastalplain.htm

Olivares, M., Tarriño, A., Murelaga, X., Baceta, J. I., Castro, K., & Etxebarria, N. (2009). Non-destructive spectrometry methods to study the distribution of archaeological and geological chert samples. *Spectrochimica Acta, Part A: Molecular and Biomolecular Spectroscopy*, *73*(3), 492–497. https://doi.org/10.1016/j.saa.2008.12.036

Pareto, V. (1896). La curva delle entrate e le osservazioni del Prof. Edgeworth. *Giornale degli economisti*, *13*, 439–448.

Parish, R., & Durham, E. (2015). The problems with visual identification of Dover and Fort Payne chert. *Southeastern Archaeology*, *34*(1), 71–83. https://doi.org/10.1179/2168472314Y.0000000003

Parish, R. M. (2011). The application of visible/near-infrared reflectance (VNIR) spectroscopy to chert: A case study from the Dover quarry sites, Tennessee. *Geoarchaeology*, *26*, 420–439. https://doi.org/10.1002/gea.20354

Parish, R. M. (2013). *The application of reflectance spectroscopy to chertprovenance of Mississippian symbolic weaponry* [Unpublished doctoral dissertation]. Department of Earth Sciences, University of Memphis.

Parish, R. M. (2016). Reflectance spectroscopy as a chert sourcing method. *Archaeologia Polona*, *54*, 15–128.

Parish, R. M., & Finn, A. (2016). Quantifying regional variation in terminal Pleistocene assemblages from the lower Tennessee river valley using chert sourcing. *Tennessee Archaeology*, *8*, 42–58.

Parish, R. M., & Werra, D. H. (2018). Characterizing “chocolate” flint using reflectance spectroscopy. *Archaeologia Polona*, *56*, 89–101. https://doi.org/10.23858/APa56.2018.007

Peacock, E. (1995). *A final report on ice-storm damaged timber tracts on the Holly Springs and Tombigbee National Forests, Yalobusha and Chickasaw Counties, Mississippi*. Mississippi Department of Archives and History.

Pitblado, B. L., Cannon, M. B., Neff, H., Dehler, C. M., & Nelson, S. T. (2013). LA-ICP-MS analysis of quartzite from the Upper Gunnison Basin, Colorado. *Journal of Archaeological Science*, *40*(4), 2196–2216. https://doi.org/10.1016/j.jas.2012.11.016

Price, S. E., Peacock, E., & Dumas, A. (2009). *Phase III data recovery at archaeological site 1CK56, Clarke County, Alabama*. Mobile District of the U.S. Army Corps of Engineers https://doi.org/10.13140/RG.2.2.24686.41285

Priddy, R. R. (1961). *Bulletin 91: Geologic study along Highway 80 from Alabama line to Jackson, Mississippi* (Vol. 91, p. 61). Mississippi State Geological Survey.

Prieto, A., Arrizabalaga, A., & Yusta, I. (2021a). Lithic raw material in the Cantabrian region: Dialectical relationship between flint and quartzite in the Palaeolithic record. *Journal of Lithic Studies*, *8*(1), 32. https://doi.org/10.2218/jls.4334

Prieto, A., Yusta, I., & Arrizabalaga, A. (2019). Defining and characterizing archaeological quartzite: Sedimentary and metamorphic processes in the lithic assemblages of El Habario and El Arteu (Cantabrian Mountains, Northern Spain). *Archaeometry*, *61*, 14–30. https://doi.org/10.1111/arcm.12397

Prieto, A., Yusta, I., García-Rojas, M., Arrizabalaga, A., & Baena Preysler, J. (2021b). Quartzite procurement in conglomerates and deposits: Geoarchaeological characterization of potential catchment areas in the central part of the Cantabrian Region, Spain. *Geoarchaeology*, *2021*(36), 490–510. https://doi.org/10.1002/gea.21838510

Raber, P. A. (2021). The exploitation of quartzite in the Lower Juniata and Susquehanna Valleys: Outcrops and cobble sources. *North American Archaeologist*, *42*(3), 328–354. https://doi.org/10.1177/0098628320984354

Ramirez-Lopez, L., Stevens, A., Viscarra Rossel, R., Lobsey, C., Wadoux, A., & Breure, T. (2022). *resemble: Regression and similarity evaluation for memory-based learning in spectral chemometrics*. R package Vignette R package version 2.2.1.

Roy Sunyer, M., Mora Torcal, R., Plasencia Figueroa, F. J., Martínez-Moreno, J., & Benito-Calvo, A. (2017). Quartzite selection in fluvial deposits: The N12 level of Roca dels Bous (Middle Palaeolithic, southeastern Pyrenees). *Quaternary International*, *435*, 49–60. https://doi.org/10.1016/j.quaint.2015.09.010

Sánchez de la Torre, M., Le Bourdonnec, F. X., Dubernet, S., Gratuze, B., Mangado, X., & Fullola, J. M. (2017). The geochemical characterization of two long distance chert tracers by ED-XRF and LA-ICP-MS. Implications for Magdalenian human mobility in the Pyrenees (SW Europe). *STAR: Science & Technology of Archaeological Research*, *3*(2), 405–417. https://doi.org/10.1080/20548923.2017.1370842

Seale, L. L. (1939). *Indian place-names in Mississippi* (Historical Dissertations and Theses, 7812). Louisiana State University. https://repository.lsu.edu/gradschool_disstheses/7812

Shackley, M. S. (1998). Gamma rays, x-rays, and stone tools: Some recent advances in archaeological geochemistry. *Journal of Archaeological Science*, *25*(3), 259–270. https://doi.org/10.1006/jasc.1997.0247

Sherman, III, S. P., Parish, R. M., Greenlee, D. M., & Miller, D. S. (2023). Assessing raw material diversity at Poverty Point (16WC5) using non-destructive reflectance spectroscopy. *Geoarchaeology*, *38*(1), 76–88. https://doi.org/10.1002/gea.21940

Starnes, J. E. (2015a). Geology for archaeologists—Part I: The orthoquartzites of the Catahoula formation. *Mississippi Archaeological Association Newsletter*, *50*(2), 5–6.

Starnes, J. E. (2015b). Geology for archaeologists—Part II: Sioux quartzite of the pre-loess gravels in Mississippi. *Mississippi Archaeological Association Newsletter*, *50*(3), 4–5.

Starnes, J. E. (2016). Geology for archaeologists—Part III: Hattiesburg quartzite. *Mississippi Archaeological Association Newsletter*, *51*(1), 6–7.

Starnes, J. E. (2019a). Prehistoric utilization of ferrugenous orthoquartzite in Southeast Mississippi. *Mississippi Archaeological Association Newsletter*, *54*(2), 4–5.

Starnes, J. E. (2019b). Geologic resource utilization for mining with hammerstone tools at two distinct prehistoric orthoquartzite quarries in Mississippi. *Mississippi Archaeological Association Newsletter*, *54*(3), 5.

Starnes, J. E., & Leard, J. (2020). A newly discovered Hattiesburg Quartzite Quarry in Amite County, Mississippi. *Mississippi Archaeological Association Newsletter*, *55*(1), 1–2.

Stross, F. H., Hay, R. L., Asaro, F., Bowman, H. R., & Michel, H. V. (1988). Sources of the quartzite of some ancient Egyptian sculptures. *Archaeometry*, *30*(1), 109–119. https://doi.org/10.1111/j.1475-4754.1988.tb00439.x

Thomas, Jr., P. M., Campbell, L. J., & Weed, C. S. (1982). *The Caron Site (1Ba376): Archaeological excavation of a Fort Walton Site in Baldwin County, Alabama*. Report of Investigations No. 1. New World Research.

Veldeman, I., Baele, J. M., Goemaere, E., Deceukelaire, M., Dusar, M., & De Doncker, H. W. J. A. (2012). Characterizing the hyper siliceous rocks of Belgium used in (pre)history: A case study on sourcing sedimentary quartzites. *Journal of Geophysics and Engineering*, *9*, S118–S128. https://doi.org/10.1088/1742-2132/9/4/S118

Walthall, J. A. (1980). *Prehistoric Indians of the Southeast: Archaeology of Alabama and the Middle South*. The University of Alabama Press, University.

Wang, W., Bae, C. J., Huang, S., Huang, X., Tian, F., Mo, J., Huang, Z., Huang, C., Xie, S., & Li, D. (2014). Middle Pleistocene bifaces from Fengshudao (Bose Basin, Guangxi, China). *Journal of Human Evolution*, *69*, 110–122. https://doi.org/10.1016/j.jhevol.2013.11.002

White, N. M. (1981). *Archaeological survey at Lake Seminole*. Archaeological Research Report No. 29. Cleveland Museum of Natural History.

Willoughby, P. R. (2007). *The evolution of modern humans in Africa: A comprehensive guide* (pp. 243–318). Altamira Press.

Wise, Jr., R. (2016). The organization of technology in the pine hills of Mississippi (Order No. 10140494). ProQuest Dissertations & Theses Global. Retrieved from https://www.proquest.com/dissertations-theses/organization-technology-pine-hills-mississippi/docview/1821362099/se-2

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Sherman, S. P., Parish, R. M., Kwon, Y., Meredith, S., & Johnson, D. (2024). Non-destructively characterizing sandstones, orthoquartzites, agates, and petrified wood for provenance research: Perspectives from the Southeastern Coastal Plain, United States. *Geoarchaeology*, *39*, 628–644. https://doi.org/10.1002/gea.22018