CPMG Categorization of Pottery Sherds with t-SNE and UMAP

Kelly Ann Sutterfield, Shinjer Li, Tyler Meldrum
William & Mary Department of Chemistry

Motivation

There are many different styles, time periods, and firing methods connected to pottery sherds found by archaeologists. Complex trade routes are considered when sherds are found in areas like Virginia. With a rich history and many different cultures, styles, and relationships, there are many factors that make it difficult to connect a sherd to one origin from just visual identification. Additionally, breaking sherds to visually identify porosity, while beneficial in examining the firing method, is extremely invasive and not a conservative method. Single-sided NMR allows for non-invasive quantification of pore size in sherds, and therefore an additional way to group similar sherds to one origin.

Figure 1: Three pottery samples, one from each category

Jamestown Sherds

The sample of sherds in the study totaled 14 different sherds of three different varieties. Two of these groups were sherds that were found in Kecoughtan, Virginia (A and B). There are two different styles of the Kecoughtan sherd from different time periods. RS (A) corresponds to Roanoke simple stamp and TWN (B) corresponds to Townsend Creek ware. The other type of sherd was from Nansemond, Virginia (C). The sherd was found on Dumpling Island in the Nansemond River, hence the notation "DI."

Figure 2: Map of Hampton Roads and Norfolk Area demonstrating where the sherds originated and where they were analyzed at Jamestown and the College of William & Mary

Methods

We were given 14 different sherds that were classified into three different categories(see Figure 1). These sherds were then soaked in de-ionized water to saturate pores in the sample. The water-soaked samples are placed on the PM5 single-sided magnet (Figure 3) for NMR measurements through the Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence. We collect the decays from each sample keeping as many of the parameters set as possible:
Number of Echos: 512
Number of Scans: 512
These parameters were kept the same in each of the runs in order to ensure that each decay has the same number of time points (echos) and the same number of scans.

Limitations

Only having 14 sherds to work with is a huge limitation in finding a true correlation in pore size and origin. We can only begin to see that there might be some sort of correlation with this sample size.

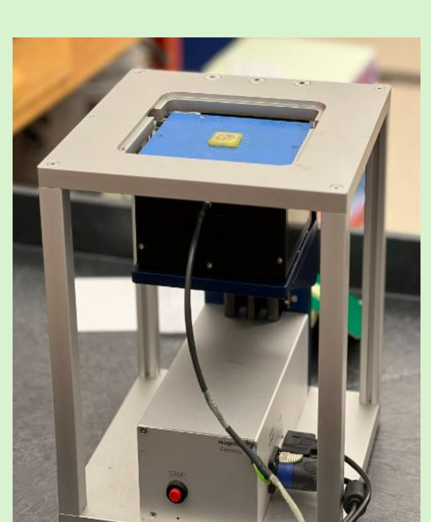


Figure 3: The PM5 magnet used for the sherd CPMG measurements

Inverse Laplace Transform

To analyze T_2 visually, the Inverse Laplace Transform (ILT) is performed on the decay that is collected from the CPMG sequence.

$$S(t) = \int_0^\infty P(T_2) e^{-t/T_2} \, dT_2$$

Figure 4: The formula for the Inverse Laplace Transform

In doing this, we integrate to find the probability of the T_2 distribution from the decay. This also requires setting a smoothing factor (α). The different smoothing factors are shown in the key. From these plots, sherds were manually grouped based on T_2 appearance. These groupings were purely visual names, such as 3 decreasing bumps, 3 equal bumps, 2 decreasing bumps, etc.

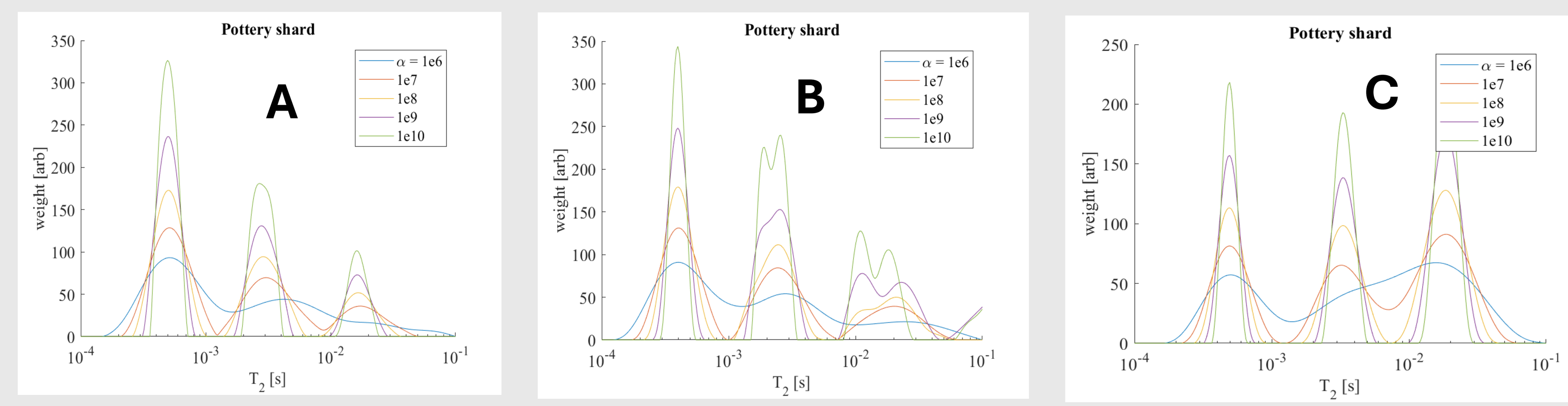


Figure 5: The ILT plots of the corresponding sherds from Figure 1.

Data Analysis with t-SNE and UMAP

T-distributed Stochastic Neighbor Embedding (t-SNE) maps high-dimensional data into a lower-dimensional space while preserving local relationships. It calculates distances between points and uses probability distributions to identify close neighbors. This ensures that points close together in the reduced space reflect similar patterns from the original data.
Uniform Manifold Approximation and Projection (UMAP) like t-SNE, preserves local patterns when mapping high-dimensional data into lower-dimensional space but also captures global structure. It's typically faster than t-SNE, using a weighted graph instead of probability distributions to connect points and optimize the final projection.

Figure 6: t-SNE using multiple values for perplexity on a sherd (A) and the t-SNE graph that gave the best grouping with this sherd (B) at perplexity=3

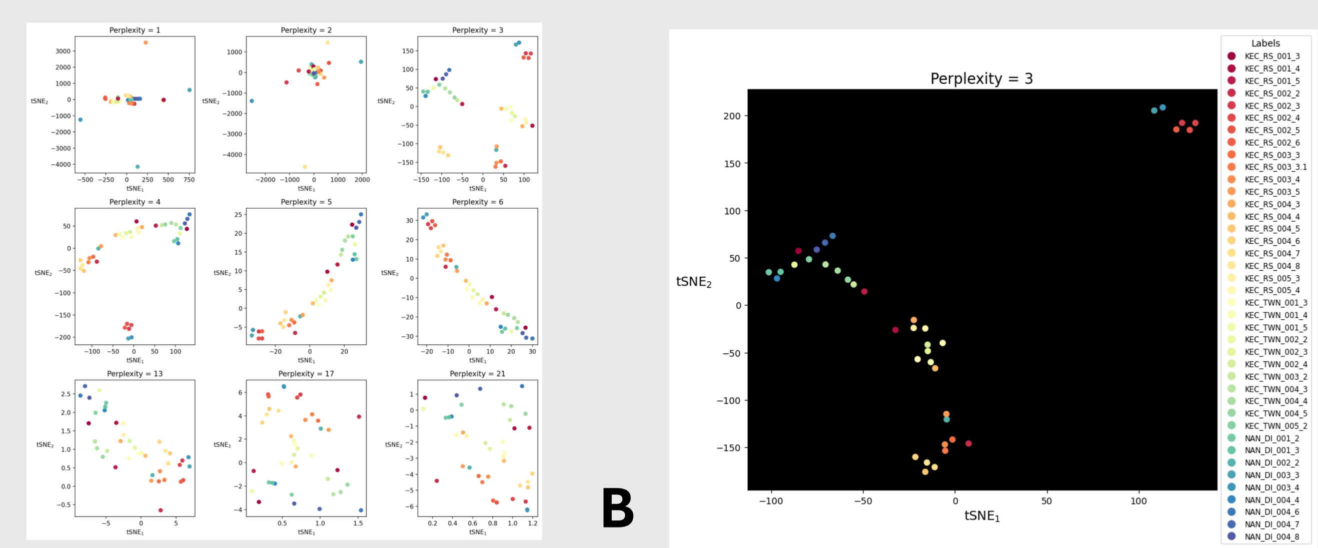
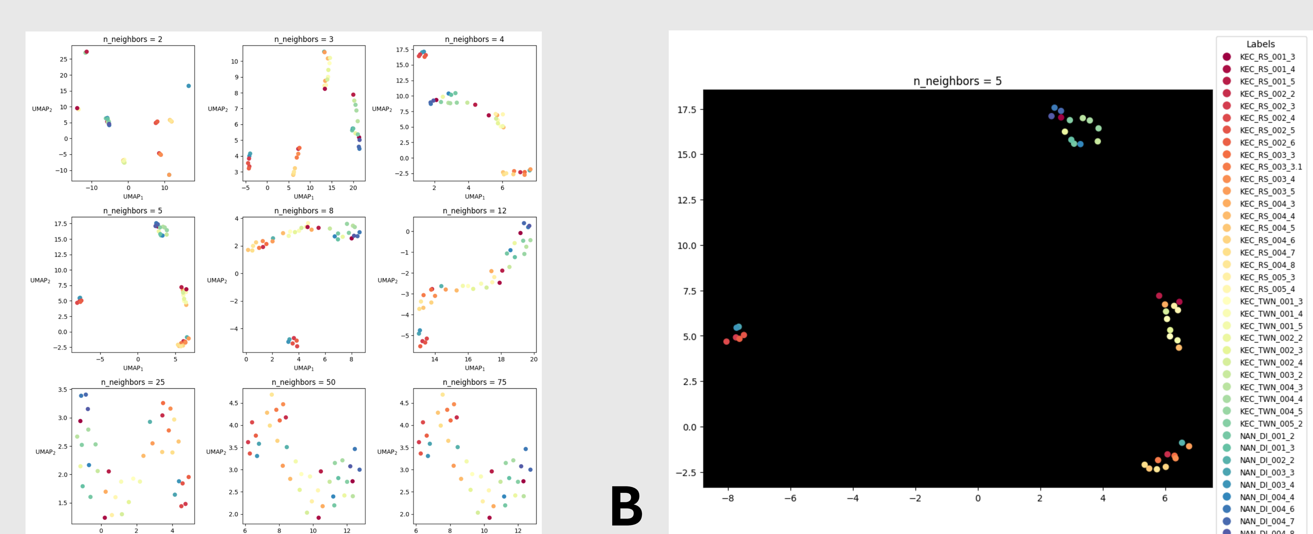


Figure 7: UMAP using multiple number of neighbors on a sherd (A) and the UMAP graph that gave the best grouping with this sherd (B) at n_numbers = 5



Results

From the information that we gathered from the visual ILT grouping, the t-SNE groupings, and from the UMAP groupings, we were able to make a Venn Diagram of the data (**Figure 8**.) The visual ILT groups were marked by coloring the text. The Venn Diagram was based on the groupings from the t-SNE and UMAP. If a sherd appeared near or in both groupings, it was placed in the overlap of the Venn Diagram.

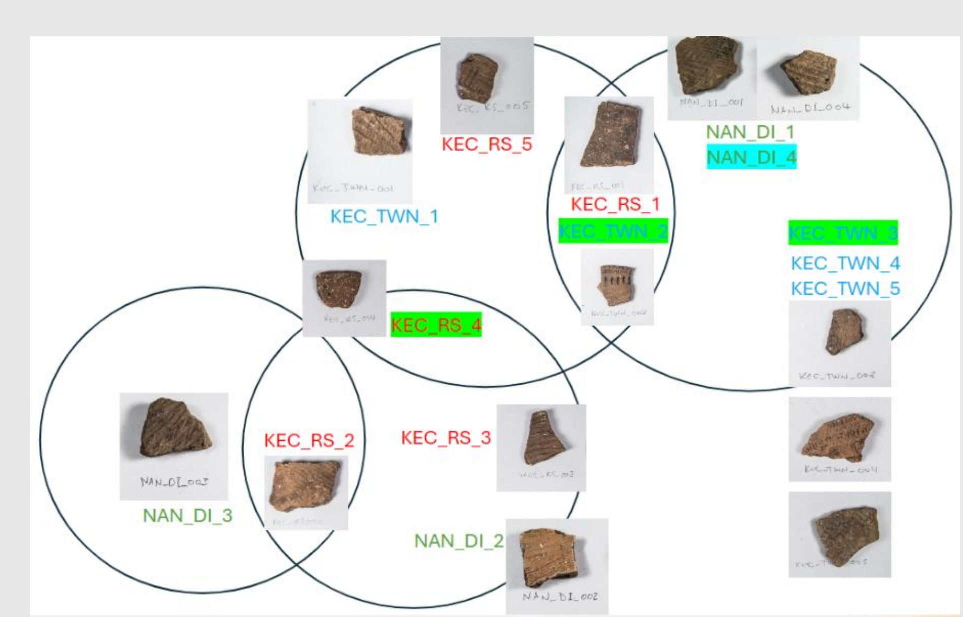


Figure 8: The Venn Diagrams of our Results

Single-Sided NMR

Single Sided NMR is a non-invasive and portable technique used to analyze samples. Because of the single-sidedness, the sample does not have to be modified to fit into the traditional NMR machine. The device generates a magnetic field on the surface of the platform, allowing for measurements of certain relaxation times in the molecule.

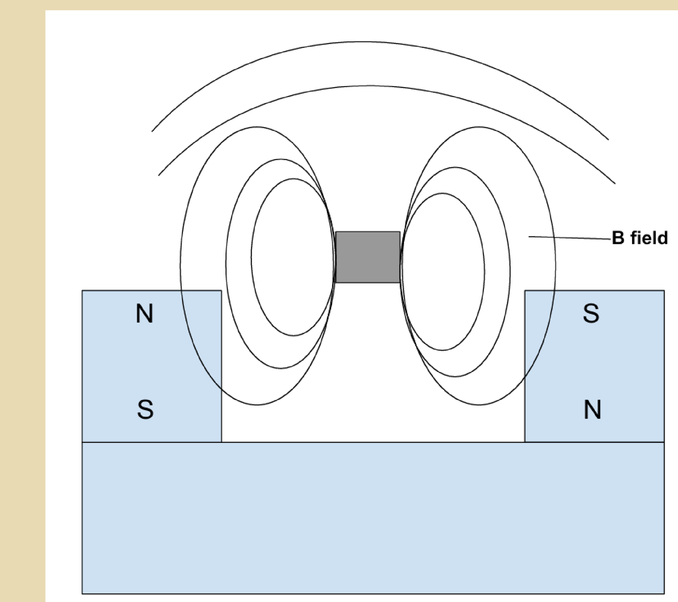


Figure 9: Schematic of how the magnetic field exists relative to the magnet. The grey square is where the sample is placed

CPMG THEORY

The transverse relaxation time (T_2) is how quickly spins in the nuclei lose their signal in the transverse plane. This measurement is how we gather most of our information regarding the sherds. To measure T_2, a Carr-Purcell-Meiboom-Gill (CPMG) sequence is applied to a sample. This aligns the magnetic spins of the molecules in the transverse plane so we can measure the decay of their magnitude (T_2) over time. A short T_2 in our sample correlates to a small pore size, and a long T_2 correlates to a larger pore. These are displayed in the Inverse Laplace Transform.

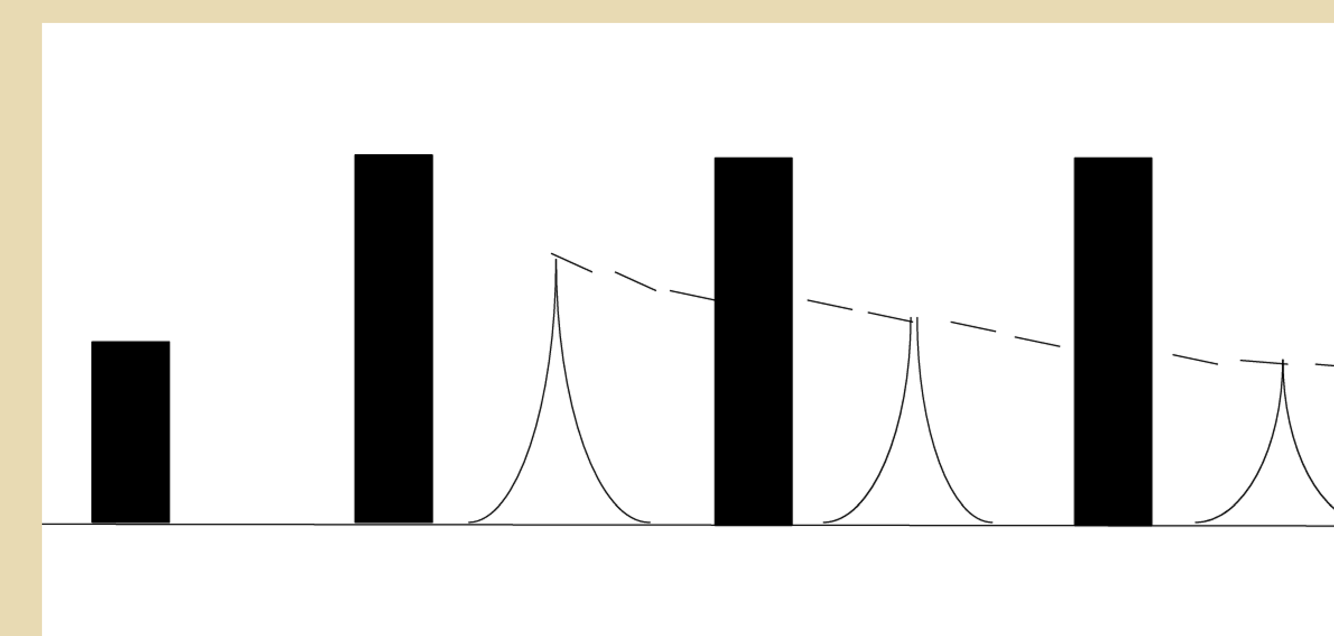


Figure 10: A sample CPMG pulse. This is a 90^o pulse followed by three 180^o pulses with a period to acquire between each. The signal decays over time and is measured, as seen by the dotted line.

Conclusion and Future Directions

As seen in the Venn Diagram, there is some correlation between the different groups of sherds and the visual and quantitative groupings we performed with the T_2 from the CPMG pulses. There were some severe limitations due to the sample size being very small. In the future, we wish to continue our collaboration with Jamestown and perform the experiment on more samples to have a stronger correlation and data base. We wish to continue using our techniques to advance the cultural research in the area and to further understanding the history that surrounds us.