SAND2011-7362C

MULTIVARIATE STATISTICAL ANALYSIS of MICRO-XRF SPECTRAL IMAGES FROM A BRUKER M4 TORNADO SYSTEM

Mark A. Rodriguez, Paul G. Kotula, James J. M. Griego,
Jason E. Heath, Stephen J. Bauer, and Daniel E. Wesolowski

Sandia National Laboratories, Albuquerque, NM 87185-1411

ABSTRACT

We employ multivariate statistical analysis (MSA) to the extraction of chemically relevant signals acquired with a micro-X-Ray Fluorescence (μ-XRF) mapping (full-spectral imaging) system. The separation of components into individual histograms enables separation of overlapping peaks, which is useful in qualitatively determining the presence of chemical species that have overlapping emission lines, and holds potential for quantitative analysis of constituent phases via these same histograms. We demonstrate the usefulness of MSA for μ-XRF analysis through application to a geological rock core obtained from a subsurface Compressed Air Energy Storage (CAES) site. Coupling of the μ-XRF results to those of quantitative powder X-Ray Diffraction (XRD) analysis enabled improved detection of trace phases present in the geologic specimen. The MSA indicates that the spatial distribution of Pyrite, a potentially reactive phase by oxidation, has low concentration and thus minimal impact on CAES operations.

INTRODUCTION

Compressed Air Energy Storage (CAES) is an important technology for renewable energy systems such as wind and solar power, due to the inherently intermittent nature of power-generation of these systems. To cope with the intermittent nature, it is ideal to store power that is generated "off-peak" in a temporary energy storage facility, until peak demand requires its release. CAES systems are designed to store energy as pressurized-air, pumped into large underground caverns or porous rock formations such as saline aquifers (Succar & Williams, 2008). When stored energy is required, the pressure is released to power electrical generators that restore electricity back to the power grid. The rock formations for CAES require detailed characterization regarding their geological constituents to predict and optimize performance. A concern is possible reaction of gases present in the compressed air, such as oxygen, with phases present in the formation. Pyrite oxidation could alter the pH, salinity, and mineralogy. Oxidation could potentially produce colloidal products, such as ferric hydroxide and Melanterite, and dissolution of carbonates with associated precipitation of Gypsum (Succar & Williams, 2008). Thus, oxidation can substantially impact CAES through alteration of porosity and permeability, which in turn govern storage capacity and flow rates of air into and out of the formation. This manuscript discusses the use of X-Ray Diffraction (XRD) and micro-X-Ray Fluorescence (μ-XRF) in the analysis of one such core-drilled specimen taken from a bore-hole

at a CAES facility. We shall discuss the use of Principal Component Analysis (PCA), a common form of Multivariate Statistical Analysis (MSA), and how PCA was employed to decompose the convoluted μ-XRF dataset containing elemental information regarding at least 8 distinct phases. These XRF results were then used to improve trace phase identification of the quantitative XRD results.

EXPERIMENTAL

The rock core was obtained from a borehole to support site characterization of a potential CAES facility in Iowa. A small piece (~24 mm x ~40 mm x ~15 mm) was cut from the original core block. From that piece, a small portion was removed and ground to a powder (~5 g) for routine powder XRD analysis. The remaining portion of sample was impregnated with a red fluorescent, low-viscosity resin under vacuum. The sample was then mounted on a glass slide and polished to a thickness of ~30 μm. This thin-section specimen was analyzed via optical petrography (using a Leitz Wetzlar Orthoplan-Pol polarizing microscope) as well as the subsequent μ-XRF experiment.

The μ-XRF analysis was performed using a Bruker M4 Tornado μ-XRF mapping system. The instrument was equipped with a micro-focused Rh source (50 kV, 600 μA) with a poly-capillary optic (~30 μm spot-size). The detector system employed two silicon-drift detectors to collect fluorescence spectra from the specimen. The specimen was secured to the x-y translation stage within the M4 chamber, and XRF spectra were collected under vacuum conditions (~10^{-3} Torr). The XRF-mapping dataset for the thin-section specimen was collected as a large datacube with full X-ray spectra (4096 channels, 0-40 keV range) collected at each pixel in a 2D array. The step-size employed for the μ-XRF spatial map was 50 μm. This resulted in a 502 x 860 matrix for the map which covered an area ~25 mm x 43 mm, thus incorporating the entire cross-sectioned specimen. Total data collection time was ~4 hours. The dimensionally of the datacube (502 x 860 x 4096) was greater than 1.7 billion elements, and encompassed a file size of 747 MB. Elemental maps for individual atomic species were generated within the M4 software package. PCA analysis was performed by reading the datacube into an in-house modified PCA software package written in *MATLAB* (The Mathworks, 2008). The multivariate statistical analysis procedure used, which seeks to maximize the mutual simplicity of the spatial components (Keenan, 2009) took 90 seconds on a PC equipped with four-core 3.8 GHz Intel Xeon processors and 8 Gbytes of memory. Full details on the use of PCA are beyond the scope of this paper and we refer the reader to existing references on this topic (e.g. Jollife, 2002; Keenan & Kotula, 2004; Rodriguez, *et al*., 2007; 2010).

For XRD analysis, the ground powder specimen (<100 μm particle size) was loaded into a side-drifted specimen holder against filter paper (to reduce preferred orientation effects). Standard powder XRD analysis was performed using a Siemens model D500 θ–θ diffractometer equipped with a sealed-tube (Cu Kα) source, fixed (1^o) slits, a diffracted-beam graphite monochromator, and a scintillation detector. Generator settings were 40kV and 30mA. Scan parameters were 10-65^o 2θ angular range, a 0.04^o step-size, and 4 second count-time. Phase identification was performed within the program Jade (ver. 9.3, Materials Data, Inc., Livermore, CA). Quantitative analysis of the resulting XRD dataset was performed via Rietveld refinement within the program *GSAS* (Larson & Von Dreele, 1986; Toby, 2005).

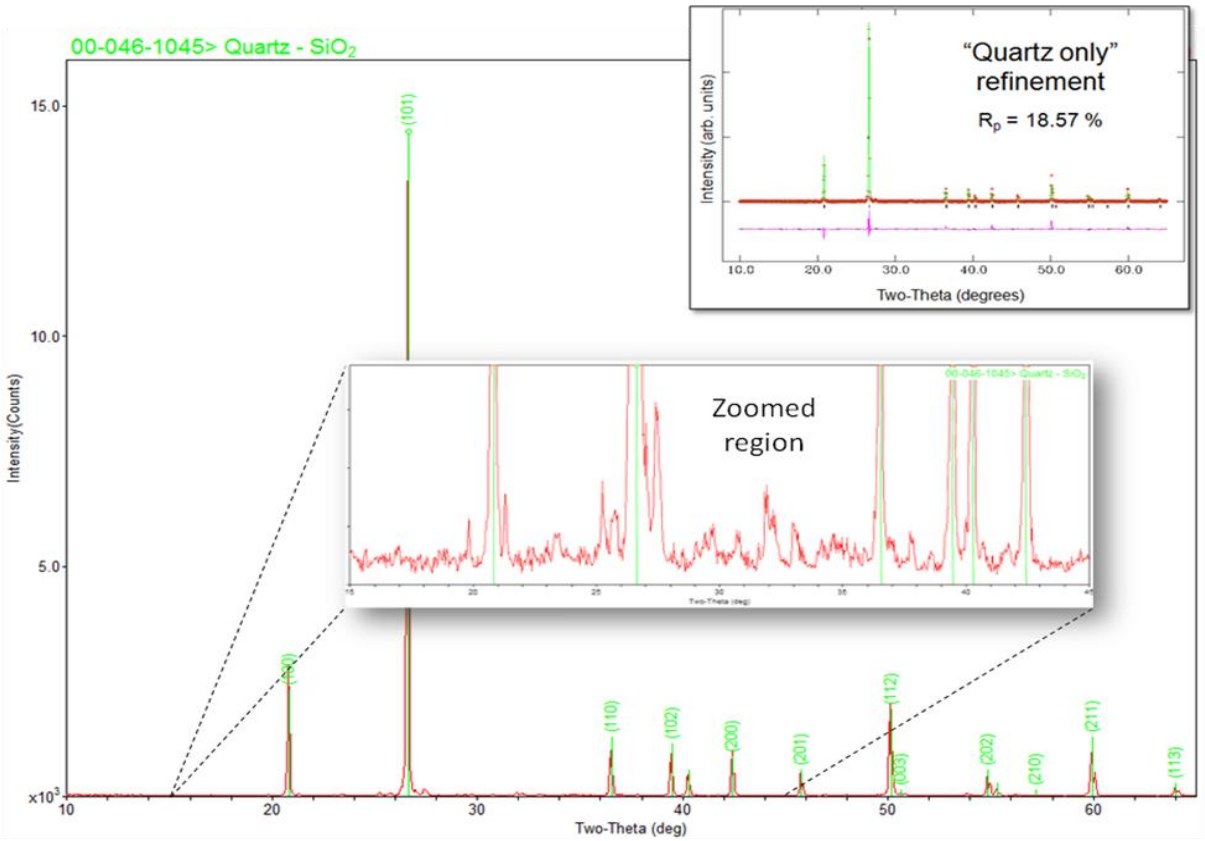

Figure 1. Initial XRD pattern showing domination of Quartz content, with only trace peaks of additional phases present in the background.

RESULTS AND DISCUSSION

Figure 1 reveals that the specimen is almost entirely composed of the mineral Quartz (PDF entry 00-046-1045, also see Table 1). An initial single-phase refinement of the powder XRD data within GSAS revealed a reasonable fit, as shown in the inset labeled "Quartz only" (upper right corner of Figure 1). However, careful evaluation of the low intensity peaks near the background (see zoomed region; inset in Figure 1) reveals small peaks from additional phases. Detailed evaluation of the low intensity "trace" phases was desirable, since initial optical petrography analysis (not shown) suggested additional phases besides the dominant Quartz. In particular, the presence of Pyrite was confirmed by optical microscopy. It is worth pointing out here that petrographic microcopy techniques such as point-counting (Dikenson, 1970; Gazzi, 1966) can require hours of detailed measurements. The PCA augmented μ-XRF analysis generates, in minutes, the same spatially-distributed chemical information as obtained through these time-consuming microscopy methods. Interestingly, the powder XRD analysis, employed specifically to identify these trace phases, could not confirm the presence of Pyrite within the sample based on the obtained pattern as shown in Figure 1. It was for this reason that μ-XRF mapping was undertaken, i.e. to test if the chemical signature for the suspected Pyrite could be isolated spatially by employing μ-XRF.

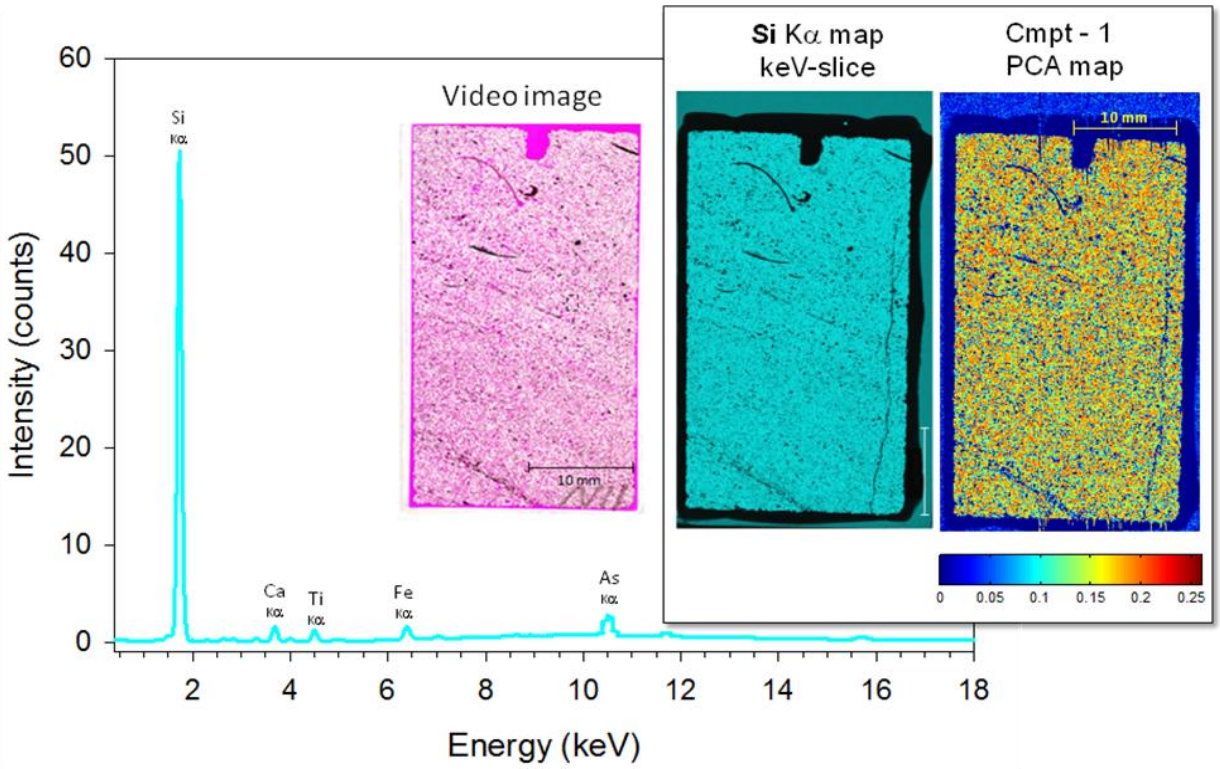

Figure 2. This figure shows the Intensity vs. Energy spectrum for the 1st isolated PCA component (Cmpt-1), along with the corresponding PCA map for Cmpt-1 (shown as a color contour map; upper right). A chemical map for the Si Kα keV-slice is shown for comparison, along with a video image from optical microscopy. See text for detailed discussion.

Typical PCA analysis employs algorithms that search large datasets for commonalities within the dimensionally of the data matrix. The resulting PCA outputs are displayed in generic terms as "loadings" and their corresponding "scores" (see Rodriguez, *et al*., 2007). For the case of this datacube the loadings (also called "components") are extracted as isolated histograms, along with the corresponding scores which extract as x-y spatial maps. The typical PCA map output for a component is a (rainbow) color contour image (red = strong, blue = weak) to indicate how that component is distributed spatially. The histograms derived from the PCA look like typical XRF spectra, and can be treated as such. The combination of component histograms (abbreviated as "Cmpt-#" throughout the text) and their corresponding contour maps taken together can adequately be recombined to fully describe the measured dataset, while removing background noise. The PCA analysis of the μ-XRF datacube yielded as many as 13 possible components necessary to describe the dataset. Some had to do with specific chemical signatures that could easily be assigned to phases within the sample. Others were more subtle effects. Table 1 at the conclusion of the manuscript, outlines all the assigned components from the PCA analysis as well as the quantified XRD results from the powder XRD analysis. The comparative results of XRD and μ-XRF analysis will be discussed in detail later.

We do not present spectra and PCA maps for all 13 components within this manuscript. Instead, we select a few important components to illustrate the functionality of the PCA technique. For the purposes of orienting the reader to PCA output, Figure 2 shows the 1st derived component,

(Cmpt-1) along with other images for comparison purposes. The histogram in Figure 2, shown as Intensity vs. Energy, is a straightforward XRF spectrum with the Si Kα peak as the main observed emission line, as well as some small peaks from other elements near the background level. This histogram can easily be assigned to the Quartz phase. Since the μ-XRF system is unable to detect elements below sodium, it is expected that the only peak we would observe from the Quartz phase is the Si Kα line. The contour map for the 1st component, shown in the upper right corner of Figure 2 and labeled as "Cmpt-1 PCA map", shows this Si signal coming from nearly all the locations across the specimen. There are only small regions that indicate any absence of this signal. For comparison purposes, an optical image (labeled "Video image") is included in Figure 2 so one can observe what the specimen looked like under typical optical conditions. One can see the pink-staining within some areas of the specimen and around the edges; this is from the added epoxy resin. Also, there are regions of darker coloring, suggesting different phases beside the dominant Quartz. In addition to the video image and the PCA map, the Si map generated via the M4 software package is also included, which simply isolates signal from the Si Kα emission line and plots its relative magnitude as a function of sample position. This is dubbed the "keV-slice" as it is an energy cross-section through the datacube. The Si Kα keV-slice map looks very much like the PCA map, indicating the strong presence of Si across the specimen, with small, isolated regions indicating absence of Si-containing minerals. It is clear from both Figure 1 and Figure 2 that any additional phases present in this sample will be minor or trace constituents, making this sample particularly challenging regarding phase identification.

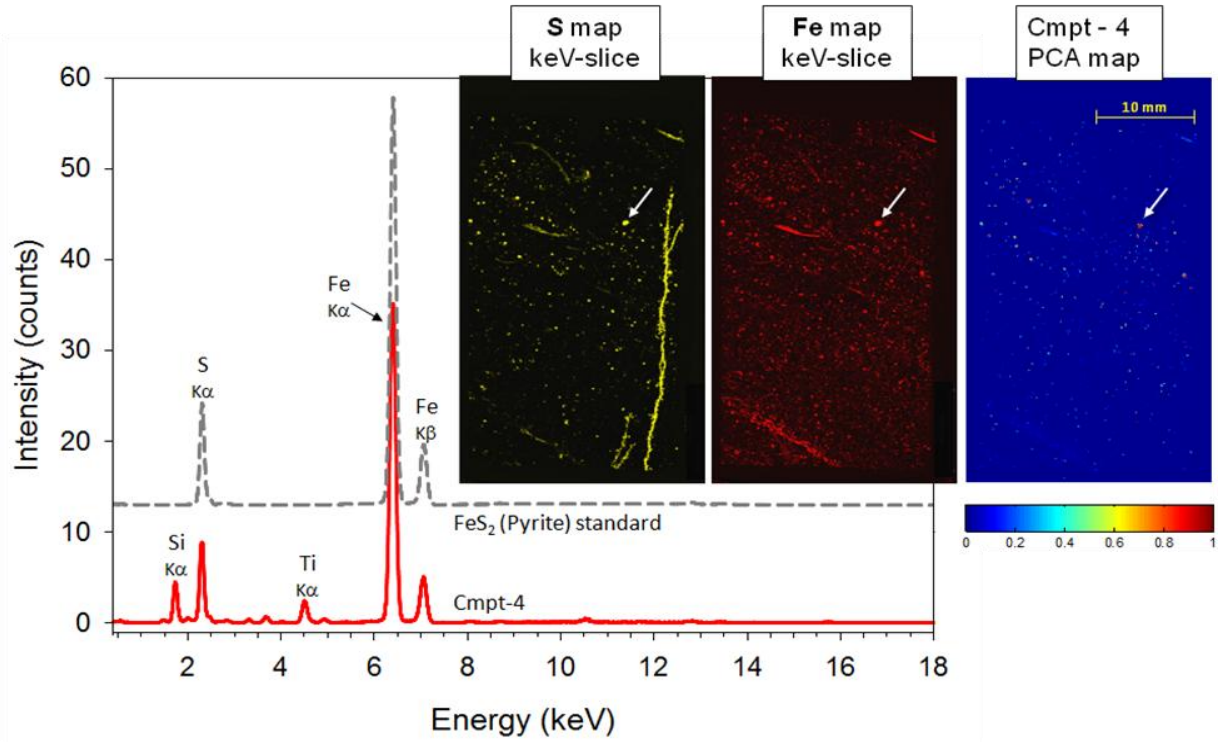

Figure 3. PCA-derived Cmpt-4. The Intensity vs. Energy plot shows the obtained Cmpt-4 spectrum (solid line) along with a spectrum from a Pyrite standard (offset dashed line). keV-slice maps for S and Fe are shown along with the PCA derived contour map for Cmpt-4. A small arrow on map images indicates the location of a ~500 μm Pyrite grain. See text for details.

Figure 3 illustrates the 4th derived component of the analysis. The solid line in the Intensity vs. Energy plot (labeled Cmpt-4) shows the spectrum that is very similar to that obtained from an independent Pyrite mineral sample, shown as the offset (dashed-line) in the plot. The intensity ratios for the S Kα, Fe Kα, and Fe Kβ lines are all nearly identical between the two spectra. In essence, the derived component histogram Cmpt-4 is the chemical spectrum for Pyrite. The signature is *quantitative*, and can be employed to characterize phase presence based on species concentration. There are some additional elemental lines present in the component spectrum, namely the Si Kα and the Ti Kα peaks. These peaks are not likely part of Pyrite, but there origin is likely "bleed-in" of the 1st component (silicon from Quartz) and a 2nd identified component (Titanium from Anatase, TiO_2). See Table 1. The presence of the extra Si Kα and the Ti Kα peaks shows that the PCA analysis is not always perfect for the generation of pure components, but that it sometimes suffers from "component mixing". This is especially true when a datacube contains one or two components that dominate (e.g. Quartz), while other components are indicative of more subtle effects (i.e. Pyrite). In fact nearly all the 13 derived components had some bleed-in from the Quartz component as indicated by a small Si peak present in each spectrum. This can make quantification of compounds containing silicates more difficult. However, this issue had little impact on the relative intensity ratios of the S Kα and Fe Kα,Kβ peaks observed in Cmpt-4.

The keV-slice maps shown in Figure 3 for S and Fe are very interesting. They actually do not show a very high degree of correlation. In the case of the S keV-slice map, we see a long, vertical vein on the right side of the map, suggesting a sulfur containing mineral. This vein is not present in the Fe keV-slice map. Careful analysis of other keV-slice maps found that the Ca map (not shown) also displayed this vein. The material was ultimately identified as being from a calcium sulfate mineral Gypsum. This identification was confirmed based on weak peaks in the XRD pattern and the presence of other similar spatial mapping features from other bore-hole specimens obtained from nearby depths. It is worth mentioning that PCA *did* find a component for Gypsum (Cmpt-5) with proper chemical ratios of S and Ca (see Table 1). In truth, many of the features in the S and Fe keV-slice maps are not actually signal from Pyrite. This fact underscores the difficulty of isolating chemical signatures of trace phases when the atomic species of that desired constituent phase are also present in other phases in the specimen. What one really needs is to spatially identify the commonality of location for the Fe and S fluorescence signals that possess the proper emission-line intensity-ratio to be assigned to the FeS_2 chemical composition. This, in fact, is exactly what the Cmpt-4 PCA map reveals as seen in Figure 3 (upper right). This color contour map looks nearly featureless, with a blue background level across most of the map area. However, dispersed across the map are small regions, shown as red dots, which indicate strong "scores" for the Cmpt-4 histogram. These small dots indicate the presence of tiny Pyrite grains (100 to 500 μm) within the specimen matrix. One such dot is highlighted with an arrow, shown in the upper right portion of the Cmpt-4 PCA map. Arrows also highlight this same position on the Fe and S keV-slice maps; this is done to demonstrate that S and Fe signals are observed in this location, having approximately the same spatial dimensions as that of the PCA map. Such fine-tuned analysis would be very difficult without the use of PCA techniques. The obvious low-quantity of Pyrite in this specimen explains the difficulty in detecting this trace phase via conventional powder XRD analysis. In addition, the information derived from the Cmpt-4 PCA map regarding the small size, and spatially isolated nature of the Pyrite grains, increased confidence that the Pyrite

phase would not pose a major concern regarding acidic damage within the bore-hole. The spatial mapping via μ-XRF was critical in the diagnosis of Pyrite as isolated grains, and this information was of great use in assessing a minimal impact of Pyrite reaction during CAES operation.

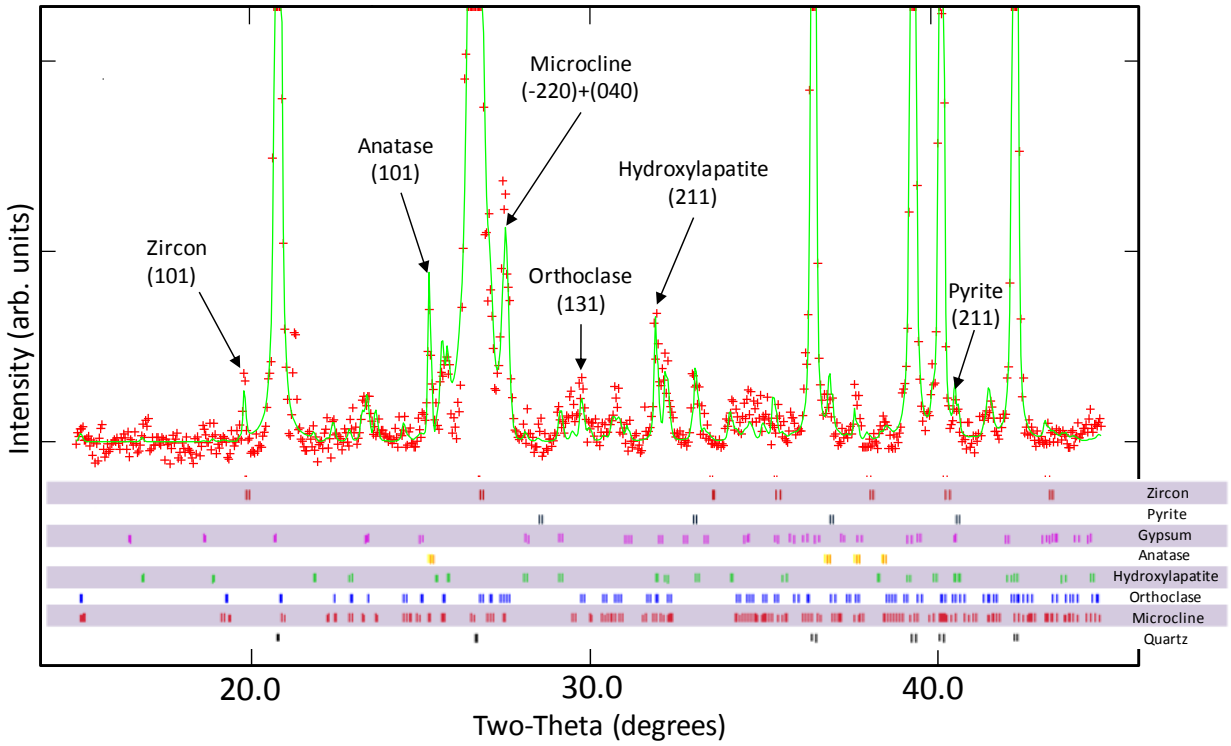

Figure 4. Zoomed range of final Rietveld refinement showing fitting of trace phases. Quartz peaks go off-scale. Reflection tick-marks underneath the fit pattern indicate locations of possible peaks from individual phases. Selected phases with assigned (hkl) values have been labeled for the trace phases (R_p = 17.16%). See text for details.

Figure 4 shows the final fit of the Rietveld refinement for the powder XRD data. The plot contains the same zoomed-range as shown in the inset of Figure 1. One can see in Figure 4 that the small impurity peaks are now accounted for in the refinement. The μ-XRF analysis coupled with PCA enabled straightforward clarification of the trace phases. Table 1 summarizes the synergistic results obtained by the merger of XRD and μ-XRF analyses. Along with Quartz, additional phases were confirmed to be present: Orthoclase, Hydroxylapatite, Microcline, Anatase, Gypsum, Zircon, and Pyrite. Note that the order in terms of PCA ranking of the derived components listed in Table 1 does not always follow the quantity present of a given phase (i.e. wt. %). This is because PCA rank depends on dataset variance and not concentration. Figure 4 shows specific indexed peaks for these trace phases. These labeled peaks were selected because they had significant relative intensities for the given phase (> 40% I_{rel} as documented in PDF entries for these phases listed in Table 1), and they did not overlap major peaks from the other phases. The residual error (R_p) from the Rietveld analysis reduced from 18.57% for the "Quartz only" refinement (Figure 1) to Rp = 17.16% after fitting the trace phases (Figure 4).

While this is not a large change, the reduction of the residual error by ~1.5%, along with the generally good fit of the observed to calculated patterns in Figure 4, indicate an accurate assessment of the trace phases in this specimen.

The PCA components were of great assistance in the phase identification. Specifically, isolation of Cmpt-6 as containing both phosphorous and calcium enabled the quick identification of Hydroxylapatite likely from fossilized remains within the specimen. Microcline and Orthoclase have the same chemical signature, and therefore were extracted as one component (Cmpt-8) in the PCA analysis. Phase identification searches of the XRD pattern with K, Al, and Si resulted in the identification of both these phases. The same process worked for the trace phases of Zircon and Pyrite. In fact, the quantitative results for Zircon and Pyrite trace phases via Rietveld refinement indicate their presence at ~0.2 wt%. This is typically at the limit of quantification for powder diffraction measurements, as the error associated with the phase fraction is of the same magnitude as the quantity refined to be present. Even with this limitation in quantification, qualitatively, the phases can be isolated spatially in terms of chemical signature, and confirmed present based on the location of diffraction peaks present in the bulk powder analysis. This was critical in the confirmation of Pyrite within this specimen.

Table 1. Summary Table for XRD and μ-XRF Analysis

PCA Rank # from μ-XRF data	Key elements of component from μ-XRF	Identified material	Wt % from Rietveld refinement*	PDF[#] entry number
1	Si	Quartz	94.3(4)	00-046-1045
2	Ti	Anatase	0.5(1)	00-021-1272
3	Ca, Cl	Epoxy resin	-	
4	Fe, S	Pyrite	0.2	00-042-1340
5	Ca, S	Gypsum	0.3	00-033-0311
6	Ca, P	Hydroxylapatite	1.6(2)	04-011-6221
7	Cl	Epoxy in voids	-	
8	K, Al, Si, Fe	Orthoclase	1.9(2)	04-009-3700
8		Microcline	1.0(2)	04-008-1783
9	Si	Si Laue	-	
10	Ca, Si, As	Glass slide	-	
11	Zr, Hf	Zircon	0.2	00-006-0266
12	Si	Si Laue	-	
13	Si	Si Laue	-	

*****bracketed value refers to 1σ error on last significant digit
[#]Powder Diffraction File (PDF-4+, 2010): ICDD, Newtown Square, PA.

Notably, a defining (hkl) for the Gypsum phase is absent from Figure 4. For this phase, there was significant peak overlap for all the major reflections. The pattern showed consistency regarding possible intensity which could account for this phase, but the intensity always overlapped other major reflections from the other phases. Gypsum was confirmed based on several factors. First, the μ-XRF clearly showed a component with Ca and S in a 1:1 atomic ratio (see Cmpt-5 in Table 1) which suggested a sulfate. Secondly, other core-drill specimens taken from similar depths showed Gypsum, but with higher concentration, which made the

identification routine. Third, other calcium sulfate containing phases common in geological formations, such as Anhydrite and Bassanite were inconsistent with the XRD data. Lastly, Gypsum was accommodated within the Rietveld refinement without difficulty and refined to a realistic concentration.

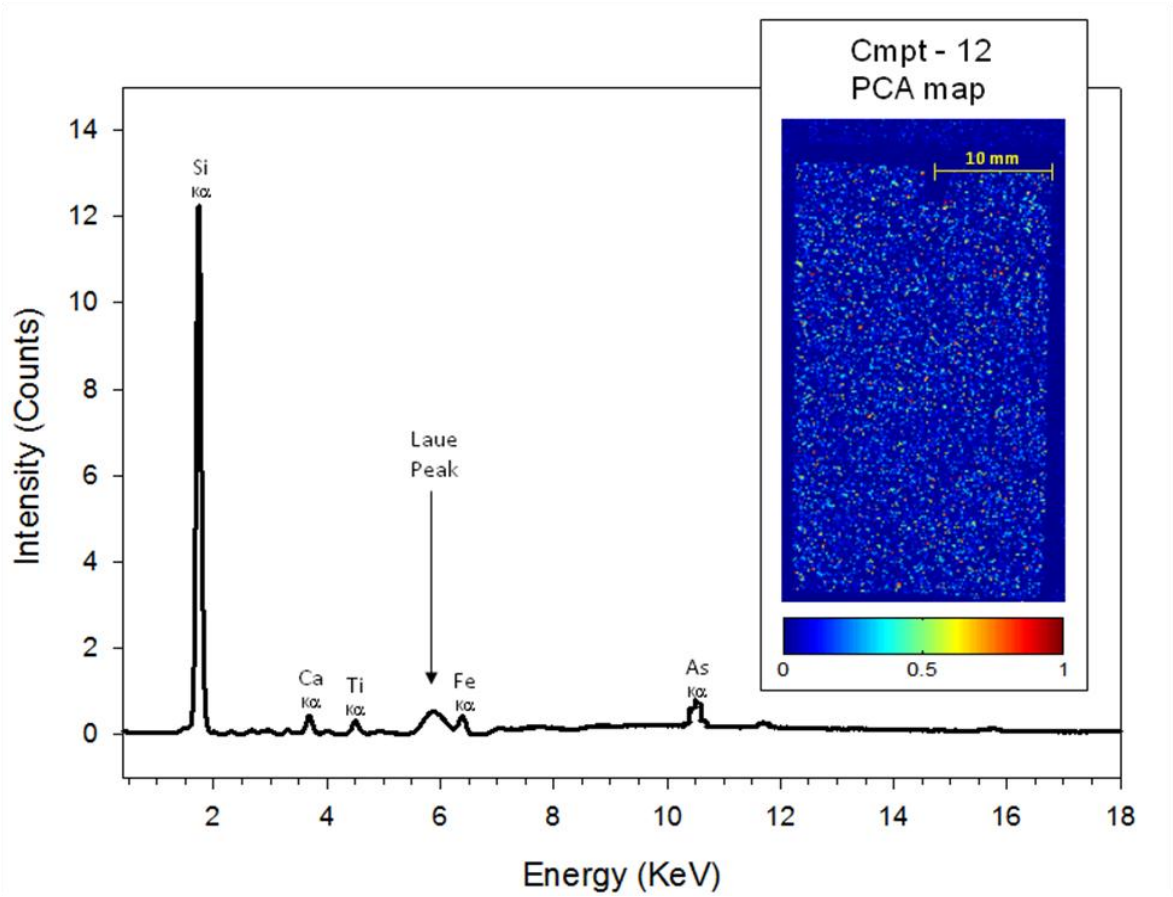

Figure 5. Derived PCA spectrum and corresponding PCA map for Cmpt-12. The Intensity vs. Energy plot looks identical to that of Figure 1, but with the addition of a broad Laue peak at ~5.8 keV. The PCA Map shows that the distribution of this component is correlated with large, individual grains distributed throughout the specimen.

Upon investigation of Table 1, it is worth noting again that not all PCA components had origins in specific phases present in the core-drill cross-section. Actually, six of the thirteen components could be considered artifacts of the analysis. This includes components from the epoxy resin (i.e. Cmpt-3, Cmpt-7), a component from the underlying glass slide (Cmpt-10), and components from Si Laue peaks (Cmpt-9, Cmpt-12, and Cmpt-13). These artifacts prove very interesting because they demonstrate the ability of PCA to remove systematic errors that might otherwise contaminate the quantitative analysis of given phases. This is particularly true of Laue peaks. For example, the Cmpt-12 histogram, shown in Figure 5, looks nearly identical to that of Cmpt-1 (Quartz). However, there is one major difference: a broadened peak at ~5.8 keV. The peak at ~5.8 keV would not easily identify with any expected elemental species from the specimen, but could possibly be mistaken for Mn Kα (5.90 keV). The fact that this peak showed a broadened peak profile led to immediate suspicion regarding its origin. Evaluation of the PCA

map for Cmpt-12, as shown in Figure 5, gives support to the identification of this component as being from Laue diffraction effects. As seen in the Cmpt-12 PCA map, the appearance of individual grains, evenly distributed throughout the specimen in a similar fashion to that of the Quartz phase, is highly suggestive of Laue diffraction. In addition, the spectrum for Cmpt-12 tags the Laue peak with the Si Kα peak, assuring that this effect is from the Quartz phase. Geometrical considerations regarding the beam and detector angular orientation tentatively assign the Laue energy as being diffraction from the Quartz (102) and/or (110) planes. Here again, we see the effectiveness of the PCA technique. Not only does the method help to remove potentially misleading artifacts which may bias quantitative results, the process actually separates these artifacts for specific assignment, which often leads to additional information gleaned from the technique.

CONCLUSIONS

PCA analysis of μ-XRF datasets revealed a highly-detailed, spatially-constrained, set of components. Detection of trace phases in the core-drilled cross-section was greatly enhanced by PCA analysis when coupled with quantitative powder XRD results. Derived PCA components are quantitative for the purpose of chemical composition determination. PCA augmented μ-XRF analysis generates, within seconds, the same spatially-distributed chemical information as obtained from many hours of petrographic microscopy. Results support geological characterization of the renewable energy site for CAES. Pyrite oxidation is not a major concern during CAES operation because the mineral was detected at low quantities and it was well dispersed throughout the mineral matrix from the core-drilled specimen.

ACKNOWLEDGMENTS

Sandia is a multiprogram laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

REFERENCES

Dickinson, W. R. (**1970**). "Interpreting Detrital Modes of Graywacke and Arkose" *Journal of Sedimentary Petrology*, **40**, 695-707.

Gazzi, P. (**1966**). "Le Arenarie del Flysch Sopracretaceo dell'Appennino Modenese: Correlazioni con il Flysch di Monghidoro", *Mineralogica e Petrografica Acta*, **12**, 69-97.

Jollife, I. T. (**2002**). Principal Component Analysis, 2nd ed. New York: Springer-Verlag.

Keenan, M. R. & Kotula, P. G. (**2004**). "Accounting for Poisson noise in the multivariate analysis of ToF-SIMS spectrum images," *Surf. Interface Anal.*, **36**, 203-212.

Keenan, M. R. (**2009**). "Exploiting spatial-domain simplicity in spectral image analysis," *Surf. Interface Anal.*, **41**, 79-87.

Larson, A. C. and Von Dreele, R. B. (**2000**). General Structure Analysis System (*GSAS*) (Report LAUR No. 86-748). Los Alamos National Laboratory, Los Alamos, New Mexico.

Rodriguez, M. A., Keenan, M. R. and Nagasubramanian, G. (**2007**). "In situ X-ray diffraction analysis of $(CF_x)_n$ batteries: signal extraction by multivariate analysis," *J. Appl. Cryst.*, **40**, 1097-1104.

Rodriguez, M. A., Van Benthem, M. H., Ingersoll, D., Vogel, S. C. and Reiche, H. M. (**2010**). "*In situ* analysis of $LiFePO_4$ batteries: Signal extraction by multivariate analysis," *Powder Diffraction*, **25**, 143-148.

Succar, S. and Williams, R. H. (**2008**). "Compressed Air Energy Storage: Theory, Resources, and Applications for Wind Power," Energy Analysis Group, Princeton Environmental Institute, 81 pp.

The Mathworks (**2008**). *MATLAB*. Version 7.7.0.471. The Mathworks Inc., Natick, MA, USA.

Toby, B. H. (**2001**). "*EXPGUI* , a graphical user interface for *GSAS*," *J. Appl. Cryst.* **34**, 210-213.