Digitized Database Quantification, DDBQ, Analysis of Complex Mixtures using Fully Digitized Patterns

Deane K. Smith and Gerald G. Johnson, Jr., Departments of Geosciences and Computer Science and Materials Research Laboratory, The Pennsylvania State University, University Park, PA 16802, USA

Abstract

Use of the fully digitized X-ray diffraction pattern for phase quantification has proven to be the most effective approach to achieving the theoretical accuracy potential of 1% absolute. The digitized database quantification technique, DDBQ, was developed in the 1980s. Appropriate reference patterns are stored in small databases and selected for fitting to an experimental pattern of a mixture. The reference patterns may be experimental, calculated from crystal structure information, or simulated from d-I lists. A least-squares fitting procedure of the selected patterns to the experimental data provides a reference pattern scale factor that is converted to weight fractions of the phases by employing the reference intensity ratio. Comparisons with the Rietveld method on the same experimental mixture patterns verify that the same levels of accuracy are achieved by both methods.

Introduction

Concepts of Digitized Database Quantification, DDBQ

The principle that diffraction patterns of mixtures are composed of a weighted sum of the patterns of the component phases applies to the diffraction traces as well as to the numerical intensity data. The weighting is the pattern scale factor, PSF, times the reference intensity ratio, RIR, and the weight fraction, $x_j$, of each phase. Because the reference patterns are stored with the same peak maximum intensity in the databases, the PSF is the factor required to scale the database reference pattern to the intensity level in the pattern of the mixture. If one has a database with digitized diffraction patterns of the component phases in a mixture, an experimental pattern of the mixture may be matched by a best-fit sum of such patterns to yield the quantitative analysis.

History of Digitized Database Quantification

The initial studies using the profiles of diffraction peaks as the basis for quantitative analysis was applied to oil shale phase quantification (Smith et al., 1983). This study was a modification of the procedure first applied by Copeland and Bragg (1958). For the analysis of up to 12 mineral species, 20 different scan ranges were defined using a programmable Rigaku system where each range included one or more peaks from one or more of the 12 minerals. This 12 by 20 matrix of integrated intensities was solved by a least-squares procedure based on the ratio method described by Karlak and Burnett (1966) using the condition $\sum x_j = 1$. 
This document was presented at the Denver X-ray Conference (DXC) on Applications of X-ray Analysis.

Sponsored by the International Centre for Diffraction Data (ICDD).

This document is provided by ICDD in cooperation with the authors and presenters of the DXC for the express purpose of educating the scientific community.

All copyrights for the document are retained by ICDD.

Usage is restricted for the purposes of education and scientific research.

DXC Website – www.dxcicdd.com

ICDD Website – www.icdd.com
The next stage in the development of using digitized diffraction patterns was applied to clay mineral analysis (Smith, Johnson, and Ruud, 1986). This project was supported by a consortium of oil and gas companies. Because of the broad profile nature of the diffraction patterns produced by most clay minerals, the application of digitized patterns for phase identification was very appropriate. It was also coincident with the acquisition of the first fully automated powder diffractometer system at the Materials Research Laboratory of The Pennsylvania State University, MRL/PSU. This project concentrated on the phase identification aspects of the use of the digitized traces and lead to the development of the database building routines and their early use for phase identification. Within this same time frame, General Motors Technical Center sponsored the development of the use of similar patterns of corrosion products for phase quantification that lead to the preparation of the routine GMQUANT (Smith et al., 1987, 1988). Later, Arco Oil and Gas Company sponsored the modification of the quantification routine to include chemical restraints that lead to ARCOQUANT (Smith et al., 1989). The clay consortium evolved into a consortium devoted to the characterization of zeolites known as ZEOPAK that lead to the development of new programs and databases including MATCHDB, a program that used the databases for phase identification (Smith, Johnson, and Hoyle, 1991).

Acceptance of the DDBQ method by outside users was limited by two factors. One was a contractual limitation on the distribution of the routines in the package of programs developed under the Consortium. The second was the VAX-based aspect of the program package at a time when the PC was becoming the computer systems of choice for laboratory analyses. The contracts contained clauses that restricted all the routines developed for use only by Consortium members for two years beyond the termination of the project. This time range overlapped the adaptation and perfection of quantification by the Rietveld method (Hill and Howard, 1987; and Bish and Howard, 1988). The success and availability of the Rietveld routines precluded widespread application of the DDBQ method. The two methods will be contrasted later in this paper.

Advantages and disadvantages of digitized traces for quantification

Smith (1992) re-emphasized the role of crystallite statistics (de Wolff, 1958) in diffractometer specimens as the most limiting factor in the accuracy of phase quantification. He showed that the small volume of active crystallites in a specimen limits the attainable accuracy to 1% absolute whether the phase is abundant or sparse in the sample. In order to achieve this accuracy, the specimen has to have crystallites at least below 5µm, and the specimen must be prepared so that all the crystallites are oriented randomly. Even with these specimen conditions, peak quantification methods rarely exceed 5% accuracy. In order to achieve 1% accuracy, the entire diffraction trace has been shown to be required. Hill and Howard (1987) and Bish and Howard (1988) showed that the Rietveld method using the full trace approached the 1% accuracy. Howard and Smith (1990) compared the DDBQ and Rietveld methods on the same data sets and showed that 2% absolute accuracy was obtained by both methods on specimens comprised of mineral particles that were larger than the ideal 5µm crystallite size.
Digitized diffraction patterns appear to be desirable to achieving the theoretical accuracy with real samples.

The main advantage of the digitized pattern methods is that they use all the data that is available in the pattern. This use of the full information leads to the partial compensation of some pattern aberrations such as preferred orientation and profile shape problems. Problems associated with use of single peak methods are overcome, and the level of potential accuracy can approach the theoretical limit. The main disadvantage of the digitized pattern methods is the time required to collect the diffraction data and prepare the reference information. In the case of DDBQ methods, the initial database of patterns requires considerable effort obtaining appropriate samples and then building the database. Reference intensity ratios, RIRs, also must be determined for each reference sample. In addition, the data for the sample to be analyzed does require some preprocessing before it can be quantified. However, once the database is established, the analyses of multiple related samples can be accomplished in a routine and very rapid manner. In the Rietveld method, the structure information has to be entered into a control file, and then, each run is a research project to control the flow of the calculation.

The background problem

All experimentally collected data contains background that must be considered in the calculations. In the Rietveld method, the background is retained in the raw data and fit analytically during the refinement procedure. In the DDBQ procedure, there is background in all the experimental data including the reference patterns and the mixture pattern, and it is necessary to remove it prior to any use of these patterns. Determining and stripping a realistic background is not easy. There is also the problem of any amorphous content in either the reference materials or in the mixture and whether it should be retained or stripped. In general, the amorphous component should be removed from all reference patterns, but it may not be desirable to remove it from the mixture pattern where it indicates a percent crystallinity less than 100%. In fact, if a pattern is available in the database for the amorphous phase in a given sample, it can be used to quantify the percent crystallinity. An example is the starting gel phase used for syntheses of some materials such as zeolites. A pattern of the uncryrstallized gel can be recorded, and a reference intensity ratio value can be determined. Amorphous patterns may be obtained from mixtures by subtracting the crystalline component and stripping a flat background. Such patterns do not usually have a RIR, but one may be assigned if the percent crystallinity of the mixture used to obtain the trace is known.

Equations of the DDBQ method

The basic equation

The diffraction pattern of a mixture is a weighted sum of the diffraction patterns of each component phase. The function can be expressed as:

$$I_i(2\theta) = \sum_j W_j I_j(2\theta)$$
where the sum is over the number of phases, $W_j$ is the PSF, and $I_{ij}(2\theta)$ is the digitized reference pattern. Because each reference pattern is stored in the database with the same maximum intensity value, the PSF is multiplier that changes its scale to the relative scale appropriate to the mixture pattern. The weight fraction $x_j$ is related to the PSF through the relation $W_j = x_jC_{JS}$ where $C_{JS}$ is the RIR of phase J to standard S.

The PSFs are determined by a least-squares calculation that minimizes

$$\sum_i |\delta I_i(2\theta)|^2$$

where

$$\delta I_i(2\theta) = I_i^{exp}(2\theta) - \sum_j W_j I_{ij}(2\theta).$$

Once the best-fit set of $W_j$ values are determined, the weight fractions may be obtained applying $W_j = x_jC_{JS}$ along with the condition $\sum_j x_j = 1$. Where the crystallinity is not 100% or where all the phases are not included in the analysis, an internal standard must be added in a known amount and used as one of the phases.

The Reference Intensity Ratio Concept

The RIR concept was first introduced by de Wolff and Visser (1964) to allow intensities measured on a relative scale as used in the Powder Diffraction File to be placed on a relative absolute scale for quantitative comparisons. The RIR was defined as the ratio of the intensity of the strongest peak for phase J to the strongest peak for a standard S in a 50/50 weight fraction mixture. The ratio measurement eliminates the complication imposed by the differential attenuation of phases J and S. The most common standard was $\alpha$-Al$_2$O$_3$, so the RIR is also referred to as $I/I_{corundum}$, but any standard may be used. Using RIR values is essential in the DDBQ method because the reference patterns are all stored on the same relative scale. If the reference patterns were to be stored on a relative-absolute scale, RIR values would not be necessary.

Obtaining accurate RIR values requires considerable care. It is always best to measure the RIR for a sample experimentally using the same diffractometer and conditions as will be used for the analyses. Alternate sources exist for RIR values, but they are usually less appropriate to attaining accuracy. The PDF does contain RIR values for many entries. Pattern calculations provide a theoretical RIR value. These values are close to the correct RIR, but because the conditions are not always appropriate, the accuracy is seldom better than two significant figures on these alternative sources. It should be noted that RIR values may be obtained using the DDBQ procedure on any sample whose phase composition is known with high accuracy. During the running of the analysis, the RIRs are adjusted so that the final weight fractions yield the known values. If one of the phases is the standard, then the RIR scale is equivalent to that which is usually employed, but it is not necessary to use a standard scale unless a large database is involved. A review of the use of RIRs in X-ray diffraction has been presented by Johnson and Zhou (1999). The RIR measured against any standard, $S_2$, may be referred to the corundum scale using the following relation:

$$RIR = \frac{I_j/I_{corundum}}{I_j/I_{S_2}} = \frac{I_j/I_{S_2}}{I_{S_2}/I_{corundum}}.$$
Quartz is a commonly used standard when analyzing rock suites (e.g. Davis and Walawender, 1982) because of its abundance in nature. Using the relation above, it is easy to convert from the Al₂O₃ base to the SiO₂ base or vice versa.

Building the Database

The databases for the quantification package developed at MRL/PSU, use a standard file structure for the reference patterns and a slightly different format for the unknown mixture patterns. The decision was made early to fix the range of the data to 70°2θ using a step of 0.02°2θ yielding 3501 data points. The range was considered adequate to cover the pertinent portion of any pattern used for quantification, and the step was based on having at least five data points within the FWHM on a standard diffractometer from a well-crystallized sample. This choice has proved to be good, but it was necessary to allow the range to slide along the angle scale. Current procedures allow the start angle to vary from 0°2θ to 20°2θ depending on the nature of the sample.

Ideally the reference diffraction pattern is collected experimentally on an appropriate phase-pure specimen of the reference phase. Alternatively, the pattern may be obtained by calculation from crystal structure parameters using a program such as POWD12+, or the pattern may be simulated using a program in the MRL/PSU package named SIMUL from a set of d-I values such as in the Powder Diffraction File, PDF, (McClune, 1998). Smith et al., (1988) have shown that calculated patterns do yield useful results, but the results are still best considered semiquantitative compared to using reference data measured experimentally.

The MRL/PSU Program Package

The software package in use at MRL/PSU for quantification contains many routines for preparing the databases and running the quantification. There are routines for transferring and processing files from different diffractometers and building databases from these files. Other routines read files from the Inorganic Crystal Structure Database, ICSD, (Bergerhoff and Brown, 1987) for calculating reference patterns or read d-I lists in the PDF for simulating reference patterns. There are also routines for displaying patterns and interacting with the data primarily to strip the background. There is a routine, MATCHDB, that employ the databases for the identification of the phases in the mixture and to setup the control file for quantification in addition to the two main quantification routines, GMQUANT and ARCOQUANT. The major difference between these two routines is that ARCOQUANT allows the use of chemical restraints during the least-squares pattern-fitting calculation to take advantage of analytical data that is available on the mixture. GMQUANT is described in detail in Smith et al., (1987), ARCOQUANT in Smith et al., (1989), and MATCHDB in Smith et al., (1991, 1993). SIMUL has not been described in the open literature but uses the profile technique of POWD12. All the PSU routines are in VAX/VMS FORTRAN compiled on a MicroVAX II and DEC ALPHA using standard computer options. A complete GMQUANT package is available from the authors.
An Example Using Full Patterns for Phase Quantification

An example involving Jamaican bauxite will illustrate the power of the DDBQ method. Bauxite is a mixture of aluminum and iron oxyhydroxides along with residual silicates and oxides that form as a very deep soil in a tropical climate. It is the principal source of aluminum throughout the world. The Bayer process is used to concentrate the aluminum, but a major problem is the presence of iron in some of the minerals. It is important to quantify the mineral composition as the material enters the recovery plant. The sub micrometer particle size of all the components makes the DDBQ method very suitable as the means for phase quantification. The very small crystallite size in these samples meets the criteria described in Smith (1993) for achieving potential accuracy below 1% absolute.

Table 1. Database for the analysis of Jamaican bauxite

<table>
<thead>
<tr>
<th>Number</th>
<th>Mineral Phase</th>
<th>Source</th>
<th>RIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Gibbsite, Al(OH)₃</td>
<td>PDF-7-324</td>
<td>0.73</td>
</tr>
<tr>
<td>2</td>
<td>Nordstrondite, Al(OH)₃</td>
<td>Calculated, ICSD</td>
<td>0.72</td>
</tr>
<tr>
<td>3</td>
<td>Boehmite, AlO(OH)</td>
<td>Calculated, ICSD</td>
<td>0.76</td>
</tr>
<tr>
<td>4</td>
<td>Bayerite, Al(OH)</td>
<td>Calculated, ICSD</td>
<td>0.51</td>
</tr>
<tr>
<td>5</td>
<td>Goethite, FeO(OH)</td>
<td>PDF-29-713</td>
<td>1.10</td>
</tr>
<tr>
<td>6</td>
<td>Goethite-o, Fe₀.₅Al₀.₅O(OH)</td>
<td>Calculated, modeled</td>
<td>0.81</td>
</tr>
<tr>
<td>7</td>
<td>Goethite-d, (Fe,Al)O(OH)</td>
<td>Calculated, modeled</td>
<td>1.04</td>
</tr>
<tr>
<td>8</td>
<td>Hematite, Fe₂O₃</td>
<td>Calculated, ICSD</td>
<td>1.18</td>
</tr>
<tr>
<td>9</td>
<td>Anatase, TiO₂</td>
<td>PDF-21-1276</td>
<td>2.93</td>
</tr>
<tr>
<td>10</td>
<td>Rutile, TiO₂</td>
<td>PDF-21-1272</td>
<td>2.36</td>
</tr>
<tr>
<td>11</td>
<td>Quartz, SiO₂</td>
<td>PDF-33-1161</td>
<td>2.33</td>
</tr>
<tr>
<td>12</td>
<td>Crandallite, Ca₃Al₅H(PO₄)₆(OH)₆</td>
<td>PDF-33-257</td>
<td>2.06</td>
</tr>
<tr>
<td>13</td>
<td>Lithiophorite, (Al₅Li)MnO₆(OH)₂</td>
<td>PDF-16-364</td>
<td>2.04</td>
</tr>
<tr>
<td>14</td>
<td>Calcite, CaCO₃</td>
<td>PDF-5-586</td>
<td>2.08</td>
</tr>
<tr>
<td>15</td>
<td>Kaolinite, Al₃Si₂O₅(OH)₄</td>
<td>Med. crystallinity, expt.</td>
<td>1.08</td>
</tr>
<tr>
<td>16</td>
<td>Siderite, FeCO₃</td>
<td>PDF-29-696</td>
<td>0.69</td>
</tr>
<tr>
<td>17</td>
<td>Sodalite, Na₄Al₅Si₅O₁₂(CO₃)</td>
<td>PDF-15-469</td>
<td>0.69</td>
</tr>
<tr>
<td>18</td>
<td>Hydroxylapatite, Ca₅(PO₄)₅(OH)</td>
<td>PDF-9-432</td>
<td>0.49</td>
</tr>
<tr>
<td>19</td>
<td>Cancriinit, Na₄Ca₃Al₆Si₅O₄(CO₃)₂H₂O</td>
<td>PDF-25-776</td>
<td>0.68</td>
</tr>
<tr>
<td>20</td>
<td>Dicalcium phosphate hydrate – 1</td>
<td>PDF-12-8</td>
<td>2.04</td>
</tr>
<tr>
<td>21</td>
<td>Dicalcium phosphate hydrate – 2</td>
<td>PDF-11-205</td>
<td>2.04</td>
</tr>
<tr>
<td>22</td>
<td>Tricalcium phosphate hydrate</td>
<td>PDF-2-83</td>
<td>2.04</td>
</tr>
<tr>
<td>23</td>
<td>Goethite-d, (Fe₀.₇₅,Al₀.₂₅)O(OH)</td>
<td>Calculated, modeled</td>
<td>0.82</td>
</tr>
<tr>
<td>24</td>
<td>Si</td>
<td>PDF-27-402</td>
<td>4.12</td>
</tr>
</tbody>
</table>

A database of 24 selected mineral patterns was established to cover the phases expected in the raw and processed bauxite. These patterns were mostly simulated from the PDF or calculated from fundamental crystal structure information obtained through the ICSD. A list of the phases is given in Table 1 along with the reference-intensity-ratio, RIR, used in this study. Many of the patterns were adjusted for small crystallite-size.
effects. The effective size used for most of the phases was assumed to be 400Å. The RIRs were appropriately modified for this size. Up to 18 phases were quantified using this database and DDBQ.

Although many DDBQ analyses were performed on bauxite samples, the most illustrative results show the consistency obtained on pairs of samples from each of three different mines. It is not possible to evaluate the accuracy of these analyses because the true values are not known, but the precision is on the order of 0.5% absolute. The analyses for 13 phases are given in Table 2. All analyses are normalized to 100%. Where negative values are reported, the mineral would be removed from the control file, and the analyses would be rerun.

The results of the bauxite analyses confirm the potential of the DDBQ method. When the specimens are prepared so as to meet the three fundamental criteria, sampling representation, counting statistics, and particle statistics, results are possible to achieve better than 1% precision. The barrier to accuracy in the DDBQ method is the RIR value. To achieve the potential accuracy, the RIRs have to be measured experimentally on real representative samples using very careful procedures, and the same samples must be used for the reference patterns in the database.

### Table 2. Paired analyses of raw bauxite samples from three mines.

<table>
<thead>
<tr>
<th>Samples</th>
<th>Mine 1</th>
<th>Mine 2</th>
<th>Mine 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gibbsite</td>
<td>75.3</td>
<td>76.0</td>
<td>75.1</td>
</tr>
<tr>
<td>Nordstrandite</td>
<td>2.1</td>
<td>1.7</td>
<td>-0.7</td>
</tr>
<tr>
<td>Goethite</td>
<td>8.1</td>
<td>7.8</td>
<td>16.1</td>
</tr>
<tr>
<td>Hematite</td>
<td>6.0</td>
<td>6.1</td>
<td>0.3</td>
</tr>
<tr>
<td>Boehmite</td>
<td>1.5</td>
<td>1.4</td>
<td>2.2</td>
</tr>
<tr>
<td>Anatase</td>
<td>1.3</td>
<td>1.4</td>
<td>1.5</td>
</tr>
<tr>
<td>Rutile</td>
<td>1.3</td>
<td>1.1</td>
<td>1.2</td>
</tr>
<tr>
<td>Quartz</td>
<td>1.1</td>
<td>1.3</td>
<td>1.0</td>
</tr>
<tr>
<td>Crandallite</td>
<td>0.2</td>
<td>0.2</td>
<td>1.6</td>
</tr>
<tr>
<td>Lithiophorite</td>
<td>0.5</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td>Calcite</td>
<td>0.2</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>Siderite</td>
<td>2.1</td>
<td>2.0</td>
<td>1.2</td>
</tr>
</tbody>
</table>

Comparisons with Rietveld Quantification

A direct comparison between the DDBQ and Rietveld methods was carried out by Howard and Smith (1990). This test used five patterns obtained from experimental mixtures based on sedimentary minerals. The DDBQ database was based on patterns of the same samples as used for the mixtures. The Rietveld technique was applied to the
same experimental mixture patterns as were used for the DDBQ studies. Because the
mineral samples used for the test were natural, they were good but not absolutely phase
pure. They are sufficiently single phase for the Rietveld refinement structure modeling to
be successful. True sedimentary rocks are in physical equilibrium and may contain
grains of some minerals with highly variable chemical composition that will lead to
problems for both techniques.

The test used samples of quartz, albite, microcline, calcite, dolomite, siderite,
pyrite, kaolinite, and muscovite in varying proportions to make up the five mixtures
examined. Each mixture represented a different type of sedimentary rock and included
six to eight of the nine minerals. Figure 1 shows the average absolute errors obtained by
GMQUANT and Rietveld on each of the five samples in the test. The average errors are
1% to 2% for all samples compared to the nominal values and even closer when
comparing the agreement between the two methods on each sample. These results show
that when the two methods are compared on appropriate samples, that they do indeed
achieve very similar results.

Figure 1. Comparison of DDBQ and Rietveld analysis on five
mineral mixtures representing sedimentary rocks
Summary

The use of the whole diffraction pattern in quantitative analysis allows the user to approach the theoretical limit of accuracy. This result is partially due to the use of all measured experimental data. In addition, the whole pattern is insensitive to peak overlap and effectively compensates for small effects of preferred orientation (Bish, 1999). These statements apply to both the DDBQ and Rietveld methods.

With respect to the Rietveld method, the DDBQ method has an advantage in its speed once an appropriate database has been established. Such rapid analyses allow a high throughput for high-volume processing that may be important in an industrial laboratory. It is easily programmable for batch processing. The use of experimental traces as references allows low-crystallinity samples to be used and samples for which the crystal structure is not known. On the other hand, the Rietveld method can be applied to single samples where all the structures are known and it is standardless. It does not require the determination of a RIR for any phase. Because it is a refinement method, additional information on each phase often is produced.

With the widespread use of the Rietveld method now for quantification, the DDBQ method will see only limited application in the near future; however, its distribution as a PC Windows version would increase its use considerably. One application, however, that is being developed currently by ICDD is a clay minerals identification file. This file is composed of digitized patterns of many of the clay minerals that have been concentrated to near single-phase purity. The diffraction patterns for these samples have then been recorded in several physical states both natural and following specific treatments commonly used in clay mineral analysis. These patterns are mostly composed of broadened peaks and bands that require pattern recognition procedures to produce the most information. The DDBQ programs are very suitable for the analyses of these types of patterns. The pertinent VAX/VMS programs will be converted to a PC platform and adopted to the windows environment in the near future.

References


