

MATERIAL IDENTIFICATION: THE DESCRIPTIVE STATISTICS OF PHASE IDENTIFICATION

T.G. Fawcett, V. Bosnic, S. N. Kabbekodu, F. Needham, J. R. Blanton, D. M. Crane,
C.E. Crowder
International Centre for Diffraction Data

The ability to analyze unknown solid state materials by powder diffraction depends on a comprehensive reference collection and the accuracy in determining key parameters in both the unknown and the reference. The Powder Diffraction File (PDF), was created through contributions of the global scientific community, to provide an accurate, comprehensive reference database. The latest version of the PDF for organics and pharmaceutical analyses, PDF-4/Organics Release 2009, contains 370,844, standardized, quality reviewed and edited, material entries. The primary reference data in the database is a collection of peaks, interplanar d-spacings (d's), and their corresponding peak intensities (I's). Comparison of d,I pairs from the experimental data with those of reference materials, was described by Hanawalt and Rinn in 1936 (1) and has been the basis for phase identification methods for decades. PDF-4/Organics Release 2009 contains 66,275,578 d,I pairs for search and comparison to an unknown. With such large numbers of d-spacings and materials how can we be confident that we have correctly identified an unknown?

In recent years, especially with powerful PC's, total pattern analysis techniques are now used in material identification. These techniques require the use of "full or total patterns" for both the unknown and reference data. With PDF-4/Organics all reference materials can be represented as full digital patterns either through experimental measurement or by calculation, enabling the use of several types of full pattern methods (Reitveld, LeBail, pattern fitting) in phase identification and quantitation. Not as obvious is the ability to use other dependent variables in conjunction with diffraction or crystallographic data to facilitate material identification. Such variables might include elemental composition, physical properties (density, color, mp), functional groups, nomenclature, prototype structures, and chemistry classification (i.e. subfiles). PDF-4/Organics is a relational database, with a JAVA interface and embedded plotting and statistics software. The database has 48 types of searches and data can be sorted and displayed for 70 fields, in any permutation. We can now statistically examine methods and techniques that have been historically and empirically applied to phase identification and materials analysis. How much data is required for minor phase analysis, how many d-spacings are required to uniquely describe a material for patent purposes? What limits in detection are imposed by not using a standard, or expanded by use of a standard? In this presentation we will discuss some of the fundamental statistics of powder X-ray diffraction and answer basic questions on how this method works.

(1) J.D. Hanawalt, H. W. Rinn, *Ind. Eng. Chem., Analytical Edition*, **8**, 244 (1936).