EPSC Abstracts,
Vol. 4, EPSC2009-479, 2009
European Planetary Science Congress,

Compositional classes of the IDPs in the "Cosmic Dust Catalog"

J. Lasue, T. Stepinski and S. Bell

Lunar and Planetary Institute, Houston, TX USA
(lasue@lpi.usra.edu / Fax: +1-281-4862162)

Abstract

Thousands of interplanetary dust particles (IDPs) have been collected by the NASA Johnson Space Center (JSC); their properties have been published in the 'Cosmic Dust Catalog'. We use clustering technique applied to particles energy dispersive X-ray spectra (EDS) to classify them into compositionally distinct classes. These classes are likely to correspond to particles of different origin.

Background

Recent studies of Stardust samples from comet 81P/Wild 2 have shown that cometary dust has been mixed over the whole solar system and presents a wider range of composition than previously thought [1-3]. Very peculiar compositions including Na, Cr and K-rich particles were found by [4-5] showing the limits of the current IDPs classification. A number of publications [6-9] have shown the importance of composition for the classification of IDPs. We apply multivariate clustering technique to EDS spectra of particles collected in the volume #15 of the `Cosmic Dust Catalog.' The clustering groups particles into classes with similar spectra and thus having similar composition.

Methodology

Only images of particles EDS are available in the catalog. For each particle an image of EDS is processed by a pixels-to-graph software that outputs a list of points (counts vs energy) constituting a digital graph of the spectrum. The electrons accelerating voltage is 20 kV and the spectrum is recorded over the 0-10 keV range [10]. Spectra of standard materials are also available for comparison.

We will classify the particles using a *projection technique* that maps a structure of N-dimensional vectors (spectra) into 2-dimensional vectors (points in the plane) in a way that best preserves the inherent structure of the data. Thus, a projection assigns two particles with similar spectra to two nearby points on a 2-dimensional graph. The presence of clusters of points indicates existence of classes each having a distinct type of spectrum and thus distinct composition.

The most widely used projection technique is the Principal Component Analysis or PCA [11]. The PCA assumes that many of the original N variables are correlated and projects the N-dimensional vectors into 2-3 uncorrelated variables that account for as much of the variability in the data as possible. PCA has been used for clustering LIBS optical emission calibration spectra for ChemCam [12-13]. Fig. 1a shows a graph of 2 principal components calculated for the spectra of 103 samples of 21 different materials (see the legend). The PCA is a *linear* projection technique and could result in an inaccurate view of the data structure. In order to illustrate this point Fig. 1b shows a graph of 3 principal components calculated for the aforementioned 103 spectra. Note that the two graphs show *different* grouping of the spectra; although many clusters depicted by the 2-components PCA remain valid in the 3-components representation, other clusters become widely separated, e.g. the GBW 07105 basalt and the nontronite samples. Thus, popular 2-components PCA cannot be relied upon to reflect faithfully the inherent structure of the original spectra.

In order to avoid such ambiguities we will use the Sammon's map technique [14] instead of the PCA. This is a *nonlinear*, iterative projection technique designed to calculate inter-point distances ratio in a 2-dimensional graph as close as possible to the inter-vector distances ratio in the original N-dimensional space. The advantage of the Sammon's map is that its application to the

EPSC Abstracts,
Vol. 4, EPSC2009-479, 2009
European Planetary Science Congress,

database will result in a 2-dimensional graph that faithfully reflects the structure of the N-dimensional space enabling an accurate classification. Fig.1c shows the Sammon's map calculated for the spectra of the 103 samples [13] as a 'proof-of-concept'. It correctly resolves the spectral difference between nontronites and other samples.

Expected results

The classes determined by the multivariate clustering of IDPs will be tentatively assigned a specific origin depending on previous analyses. This study will pave the way for further development of pattern recognition and clustering tools for application to different datasets relevant to cosmochemistry.

References

[1] Brownlee, D.E., et al. (2006) *Sci.,* 314, 1711.

[2] Zolensky, M., et al. (2006) *Sci.,* 314, 1735.

[3] Flynn, G.J. (2008) *EMP,* 102, 447.

[4] Joswiak, D.J., et al. (2007) *LPS XXXVIII*, 2142.

[5] Joswiak, D.J., et al. (2008) *LPS XXXIX*, 2177.

[6] Mackinnon, I.D.R, et al. (1982) *JGR,* 87, A413.

[7] Rietmeijer, F.J.M. (1998) *Reviews in Mineralogy and Geochemistry* 36, 2.1.

[8] Rietmeijer, F.J.M. (2004) *EMP,* 95, 321.

[9] Joswiak, D.J., et al. (2007) *Proc. 'Dust in Planetary Systems'* ESA-SP-643, 141.

[10] Warren, J.L., et al. (1997) Cosmic Dust Catalog Volume 15. Johnson Space Center, Houston.

[11] Pearson, K. (1901) *Philosophical Magazine*, 2, 559.

[12] Wiens, R.C., et al. (2009) *LPS XXXX*, 1461.

[13] Forni, O., et al. (2009) *LPS XXXX*, 1523.

[14] Sammon Jr., J.W. (1969) *IEEE Transactions on computers,* C-18, 5, 401.

Figure 1: a) projection of the data with PCA-2D method; b) projection with PCA-3D method; c) projection with Sammon's map method