

## Quantagenetics® and Machine Learning: A New Paradigm in Material Identification

Nancy McMillan<sup>1</sup> and Catherine McManus<sup>2</sup>

<sup>1</sup>Geological Sciences, New Mexico State University, Las Cruces, New Mexico, United States,

<sup>2</sup>Materialytics, Harker Heights, Texas, United States

Quantagenetics® is a term coined by McManus, Dowe, and McMillan (2018, *Spectrochimica Acta Part B*, 145: 79-85). It stems from *quanta*, referring to quantum photons of light emitted from laser ablation plasmas, and *genesis*, referring to the origin of materials. Fundamentally, the Quantagenetics® concept is that all materials, synthetic and natural, contain within them information that records both origin and history, and that this information can be accessed, in large part, using Laser-Induced Breakdown Spectroscopy (LIBS). By partnering the enormous amount of information in LIBS spectra with Machine Learning and Multivariate Statistics techniques, it is possible to understand the origin of materials and the processes that have affected them in more holistic ways compared to traditional technologies.

LIBS is a laser ablation technique that diffracts light emitted from a cooling laser ablation plasma. The resulting spectrum contains information about the concentration of all elements, some isotopic ratios, and structural information of the material. The method is very fast and little sample preparation is required, allowing analysis of hundreds to thousands of samples in the same time as tens of samples using traditional technologies. This approach allows one to analyze a sample set that is truly representative of complex geologic, biologic, and other systems.

Machine Learning employs a variety of multivariate statistical methods to sort through the information in LIBS spectra; the NMSU LIBS lab primarily uses Principal Component Analysis and Partial Least Squares Regression. Models are trained and validated to predict specific parameters of the system; spectra of unknown samples are compared to the models to predict the desired parameter.

The fundamental difference between the Quantagenetics® approach and the traditional approach of quantification of the concentration of elements using univariate or multivariate techniques is that no elemental concentrations are calculated in Quantagenetics®. Instead, LIBS spectra are used as extremely detailed compositional and structural diagnostic signatures of materials; all the information in the spectra is used simultaneously to discern relationships between samples. The approach does not require calibration standards for every element present in the spectra; this would be prohibitively expensive. Instead, the materials used to calibrate a model are defined by the problem being investigated. For instance, provenance studies are calibrated on suites of samples from each location or group. Multivariate analysis reveals the relationships inherent in the data, independent of human hypotheses, assumptions, or hunches. In effect, the materials themselves are telling their compositional and structural stories.

Examples of projects that utilize the compositional information in LIBS spectra include: 1) the provenance (country or mine of origin) of gems such as ruby, sapphire, and emerald (success rates > 95%); 2) the provenance of industrial conflict minerals such as wolframite, gold, tantalite, and cassiterite (success rates > 93%); 3) discrimination between biologically-influenced and inorganically-precipitated calcites in caves (success rates > 91%); and 4) changes in sediment source through time detected by changes in sand grains composed of the mineral tourmaline (based on models with success rates > 92%). A study of stainless

steel demonstrates that it is possible to detect both the source and heat treatments, showing that LIBS data are recording both compositional and structural information. Finally, LIBS data have been used to predict the values of engineering properties of highway aggregate to improve quality control and detect instances of the use of fraudulent, non-approved aggregates (McMillan et al., 2019). The highway aggregate project is significant because one LIBS spectrum that takes about 30 minutes to produce can potentially replace 5-6 tests for a variety of engineering properties that each take days to months.

#### References

- McManus, Catherine E., Dowe, James, and McMillan, Nancy J., 2018, Quantagenetics® analysis of laser-induced breakdown spectroscopic data: Rapid and accurate authentication of materials: *Spectrochimica Acta Part B: Atomic Spectroscopy*, 145: 79-85.
- McMillan, Nancy J., Chesner, Warren H., and O'Neil, Eileen, 2019, Laser Scanning of Highway Aggregates: Results of a Transportation Pooled Fund Study: 70<sup>th</sup> Highway Geology Symposium, 13 p.