Paper 72

Rock Property Database System

Parsons, S.[1], McGaughey, J.[1]

1. Mira Geoscience Limited, Westmount, Québec, Canada

ABSTRACT

Rock property data are invaluable to an exploration program as they provide a quantitative link between geology and geophysics. This leads to improved characterization of ore deposit environments through integrated interpretation and modeling. We introduce here the Rock Property Database System, which brings together geological and geophysical information on a common integration platform, facilitating the interpretation of rock properties and corresponding geological description across geographic areas. Some of the main challenges involved in this type of integration are outlined along with the key database features that respond to these difficulties.

INTRODUCTION

Rock properties represent an important quantitative link between geology and geophysics because geophysical data is responsive only to physical rock properties. Physical property values can be correlated with geological description to characterize the rock property environment of specific ore deposits. Proper characterization of the physical property environment of ore deposits leads directly to significant exploration benefits through improved geophysical survey design, forward modelling, inversion, and interpretation. Advances in geophysical data acquisition and interpretation will yield higher quality drillhole targets.

Over the past eight years, an ongoing industry-government collaboration has resulted in the design and implementation of a reasonably comprehensive rock property database. It is called the *Rock Property Database System* (RPDS). RPDS brings together geological and geophysical information and facilitates interpretation of rock properties and corresponding geological description across geographic areas. This permits statistical and spatial characterization of the rock property environment for various ore deposit types in different geological settings. The significance of RPDS is that it provides a single repository for rock property data, as opposed to many disparate sources, thus allowing large-scale aggregation of data and in-depth analysis of rock property relationships.

CHALLENGES

The vision of an extensive, authoritative compilation of rock property data that could underpin many avenues of quantitative interpretation is appealing. It has proven to be difficult to achieve in practice due to a number of significant challenges. The challenges have all been met through a series of individual project initiatives and collaborative efforts over the years. Some of the key challenges are described below.

Lack of Value Recognition

Lack of recognition of the value of rock property knowledge stems from historical reliance on dominantly non-quantitative assessment and interpretation practices. This is, in part, due to a lack of available tools for quantitative data analysis, and slow adoption of quantitative geophysical modelling and inversion technologies. It is our observation that unconstrained geophysical inversion, which does not require *a priori* physical property knowledge, has taken about a decade to advance from proven technology to widespread industry adoption. On the other hand, forward modelling of geophysical surveys in complex, 3D geological settings for ore detectability or survey design applications does require *a priori* physical property knowledge. The same is true of constrained inversion, in which geological data is used to influence the computation. This kind of analysis is still in its relative commercial infancy. Physical rock property analysis is still not part of conventional geophysical processing and interpretation workflows in mineral exploration. This stands in marked contrast to geophysics in the oil industry, where petrophysical analysis is a standard, recognized, and critical component of interpretation. The benefits of a rock property database to serve as a foundation for quantitative approaches should accelerate their adoption.

In "Proceedings of Exploration 07: Fifth Decennial International Conference on Mineral Exploration" edited by B. Milkereit, 2007, p. 933-938

Data Procurement

Database design and implementation presented challenges easily met by software engineering (or re-engineering). The ongoing compilation exercise necessary to populate the database has proven a more difficult challenge. Physical properties of rocks have been measured by companies, universities, and government geological surveys the world over. The data, for the most part, reside in reports, papers, electronic spreadsheets on the computers of individual workers, old floppy disks on dusty shelves, etc. Acquiring and understanding data from a multitude of sources, and from institutions often reluctant to release data for reasons of lack of confidence in its completeness or quality, or sometimes having issues of control of its use, is difficult. We have found that reluctance to release physical property data for reasons of confidentiality is rare. To date we have focused primarily on compilation of data from cooperating government geological surveys.

Data Formats

Data provided from various data sources are inevitably in many different formats. This gives rise to a multitude of data formatting and integration issues, which are compounded by the fact that some data exist only in hardcopy paper format. Integration issues, such as digitizing and conversions, need to be overcome in order to accurately correlate different datasets. Although this is a labour-intensive and time-consuming task, we have made major strides in both ease of import and quality control by the use of intermediate spreadsheet files in standardized formats from which data can be mass-uploaded. The database itself then provides a common integration platform on which all datasets can be manipulated and stored. RPDS also allows for hardcopy legacy data and metadata to be properly archived preventing the loss of information over time.

Data Quality

In many cases, there are concerns involving the integrity or quality (accuracy) of the data. For example, if data were collected over a number of years, there may be measurement accuracy concerns as measurement methods or equipment changes. Calibration is a critical issue. Our approach is to initially import data as they are presented to us, while recording as much metadata as there is available, and ranking the data by a quality indicator. The quality indicator is based on the source of the data. For example, government laboratories with traceable calibration standards get a high rating, while measurements of magnetic susceptibility made in the field with a hand-held probe get a lower rating. The quality indicator can be used in queries and the results tested or filtered as necessary. By housing the data in RPDS no vital information is lost and one now has the opportunity to test hypotheses about data reliability and calibration on a large population sample. Without the availability of a very large store of rock property data in a single repository, it is difficult to test hypotheses about the relationships between different measurement methods for the same property. The database administrator has the capability to make simple or quite complex quality tests and rank, or even delete, data based on the result.

Incomplete Metadata

Legacy data is often missing pertinent metadata information, such as geological logs, acquisition characteristics, and processing or calibration information. A significant effort is required to recover this missing information, including communicating with the various data holders who may or may not have been involved with the original acquisition of the data. In some cases, this metadata may be impossible to find. Although we regard the metadata as extremely important, it is not vital for storing rock property data. By using a single data repository and defining critical metadata, the problem is minimized because the user has background information on the metadata, or lack thereof. The availability of metadata can be used itself as an indirect quality indicator.

Data Confidentiality

Data owners are sometimes concerned with distributing confidential information, particularly in exploration where data is often regarded as highly confidential. We generally argue that not making the data available prevents valuable information from being added to a common database, limiting the potential of the system to everyone’s disadvantage. Also, if a data owner has an abundance of confidential data but does not have a proper integration database to house the information, proper analysis is difficult. We have assigned a confidentiality indicator to data so that private information can be stored, inaccessible to unauthorized users. This may be implemented in the publicly-accessible web interface by means of a log-in system so that confidential information is accessible only to the appropriate end-user.

Data Type

Rock property data comes from various data types, the most challenging being borehole wireline data. Data measured by borehole probes are by nature heavily sampled and associated with a wide diversity of acquisition, calibration, and processing history metadata. The dense sampling is essential for characterizing physical property variability, for which laboratory samples on their own are insufficient. The quantity of data makes it effectively impossible to link metadata with actual data without the use of a database. A database system can store metadata separately from the actual physical property data while easily conserving the link between the two. In addition, by generating pre-rendered summary statistics of the thousands of records per borehole, it allows for efficient query and analysis, leading to greater interpretation potential.

Classification

Data classification is a requirement for any database system in order to facilitate data organization and queries. The most significant classification challenge in RPDS was the geological rock type classification. Each physical property measurement is made on some "rock type" described in the original data source. Some geological mapping databases use the North American Commission on Stratigraphic Nomenclature (1983) for classifying rock units, and provincial and federal governments have been implementing more standards for their geological classifications (Quebec Ministry of Natural Resources-Sigéom system). However, each organization uses its own local classification scheme resulting in discrepancies when data from the various sources are merged. Many present locally idiosyncratic rock names, abbreviations to be deciphered, or even simple misspellings. We devised an automated system of classifying rock type for the purpose of RPDS queries, while retaining availability of the original rock type name.

RPDS DESIGN AND IMPLEMENTATION

The challenges in design and implementation of a large-scale repository for physical rock property data for explorationists were met through the creation of RPDS, whose key features are described below.

RPDS overview

RPDS is an Oracle-based relational data management system for borehole and surface sample rock property data. It is designed to store, manage and query physical, geochemical and geotechnical property data and metadata in correlation with geological information. RPDS provides a traceable archive of data as well as mechanisms for data quality assessment, editing, and meta-classification.

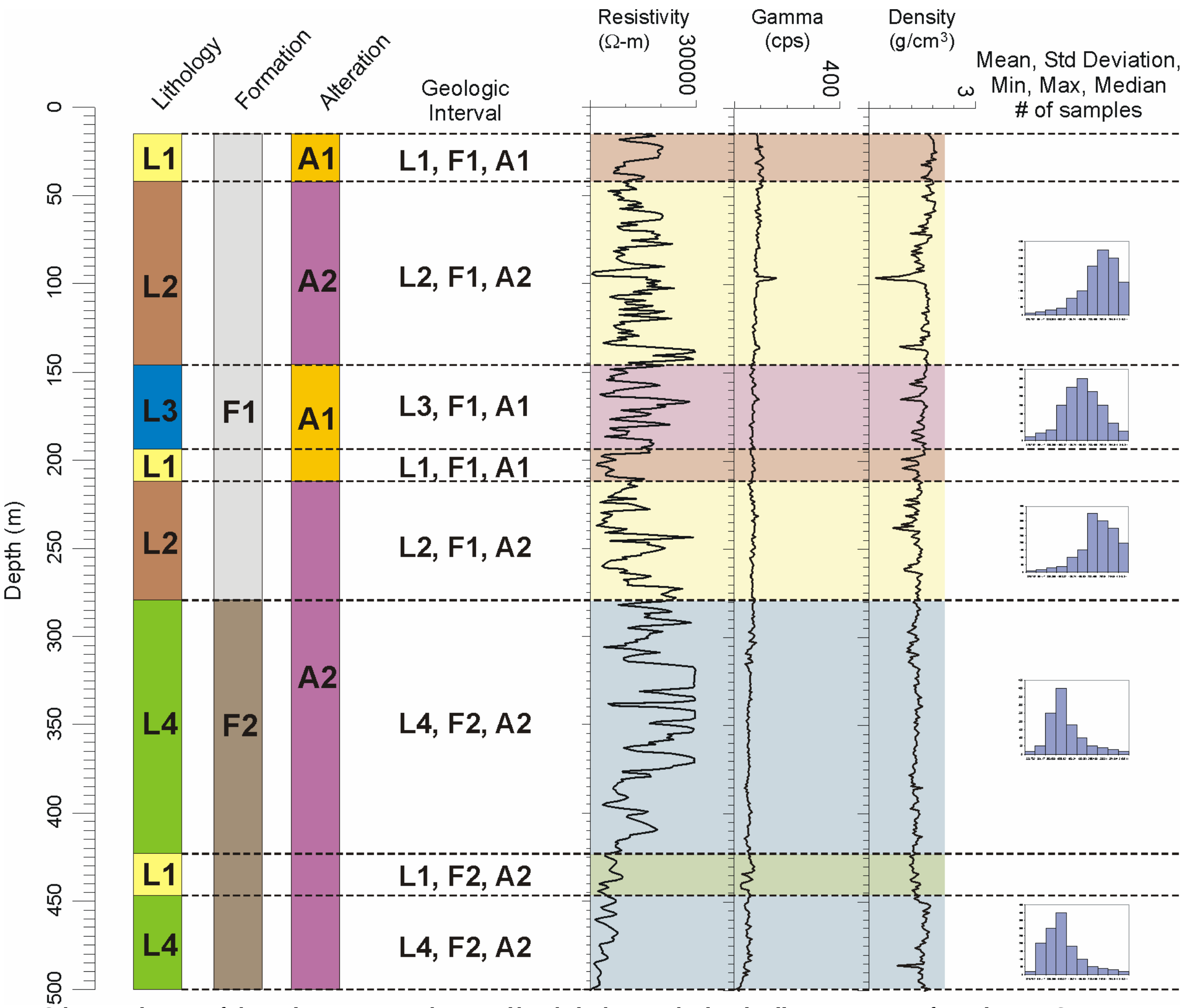


Figure 1: Schematic diagram of physical properties in a theoretical borehole showing the data distillation process performed in RPDS.

The storage of wireline and sample data requires an advanced relational database system capable of storing various types of information with gigabytes of storage space. Due to the large amount of data collected with borehole logging systems, and the difficulty involved in correlating multiple parameters per borehole, a data amalgamation and distillation process is required.

An important feature of RPDS is the process it uses to generate physical property summary statistics. The process is schematically illustrated in Figure 1. A "geologic interval" is defined, which represents a combination of the lithology, formation, and alteration information at depth along a borehole (e.g. L2-F1-A2). For each of these intervals, a mean value, standard deviation, minimum, maximum, median, and number of samples is calculated for all physical property parameters measured in the hole. This results in each "geologic interval" having a unique physical property signature. The final step in the data distillation process is to combine the physical property statistics for similar "geologic intervals" in a borehole, and to amalgamate this information with sample data possessing the same combination of geologic information in similar geographic areas. This provides a pre-computed physical property summary for all unique occurrences of lithology, formation, and alteration (e.g. L2-F1-A2) in the same geologic setting.

RPDS data model

Various tables in RPDS store the metadata pertaining to all borehole and sample data entered into the database. This metadata includes information related to the entire logging/sampling process (location, equipment, personnel, project descriptions, laboratory methods, and processing/calibration history), as well as information related to geological units, and associated geochemical and geotechnical data. RPDS takes the physical property data, linked with geological description, and calculates physical property summary statistics tables. The simplified data model showing the sequence of tables used to generate the summary statistics is shown in Figure 2.

The storage of borehole wireline physical property data in RPDS is based on the concept of logging runs. Logging run data is stored in the Process Log Table, which contains the calibrated and processed logging run data for each borehole. This data is considered the "live data" in RPDS and is used for calculating the population statistics. Raw data is stored elsewhere in the database for archival purposes only.

The Process Log Table stores the physical property values from various depths as measured along the borehole. Since the depth intervals for each measurement may vary per logging run, it is important to normalize these values to a constant depth interval in order to correlate each of the parameters for different logging runs. This is performed in the Forced Interval Table of RPDS.

The Forced Interval Table interpolates the Process Log data for each physical property to a common reference sampling interval of 10cm. Physical properties from the Forced Interval

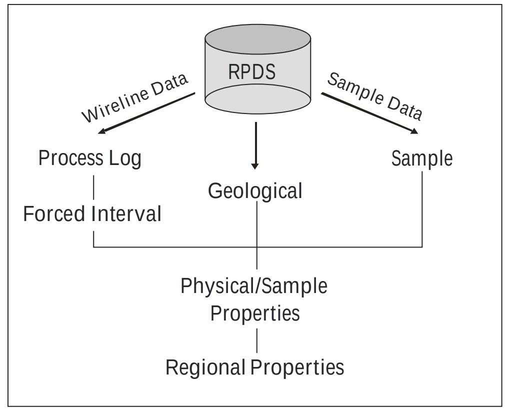


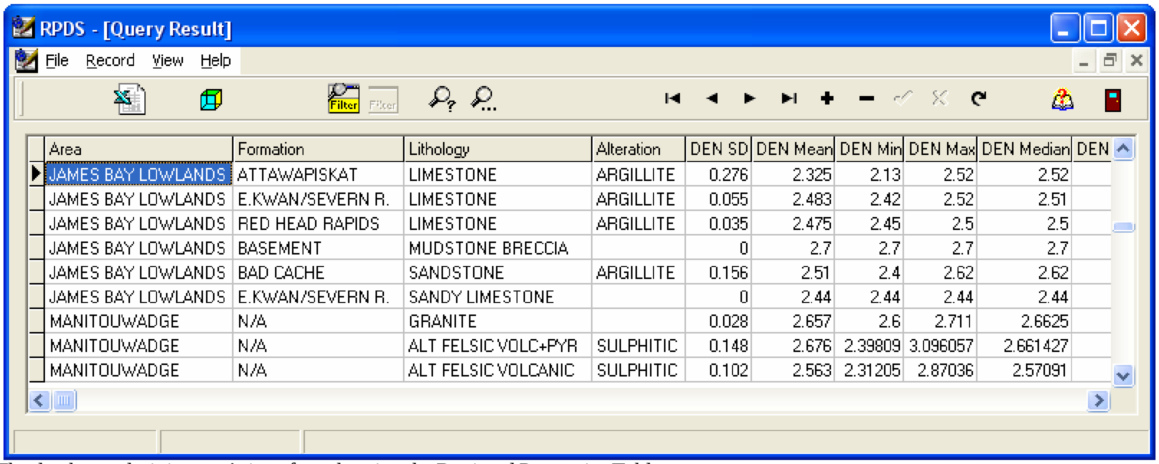
Figure 2: Schematic diagram of the simplified RPDS data model, leading to the generation of the Regional Properties summary table.

Table may be correlated since, as they are interpolated to the same depth, they represent measurements of the same rock sample.

In parallel, a significant amount of available laboratory measurements are stored in the Sample Table. This table accommodates the physical property data and all associated metadata from laboratory measurements of both borehole core samples ("borecore") originating from boreholes, and surface samples of varying origin.

Geological information for borehole wireline, borecore, and surface samples are stored separately in the database in the Geological Property Table. This table includes information on lithology, alteration, formation, geologic age, assay analyses, as well as space for storing core photos which are rapidly visible on-the-fly. Lithology is stored as the specific lithological unit name using the local nomenclature from the data source. However, in addition to this naming, a geological "Master Lithology Classification" scheme has been developed to provide a more general hierarchical description of the unit. This allows for consistent and more practical data querying within the RPDS environment. The geological data is combined with the borehole and sample data to produce the comprehensive Physical/Sample Properties Table.

The Physical/Sample Properties Table is a composite table where logging run data taken from the Forced Interval Table and sample data taken from the Sample Table are correlated with geological information. This is also where population statistics of physical properties as a function of geological classification are pre-stored for rapid query. This table lists, for each borehole, the mean values, standard deviations, and sample counts for physical properties per unique lithologic interval encountered in the borehole. For the sample data, this table simply combines each physical property measurement with the associated geologic information, since sample data only have one measurement per sample for each physical property parameter. At present, population statistics are calculated on the following 16 parameters, although others can be added to this list: gamma-ray, potassium, uranium, thorium, density, magnetic susceptibility, conductivity, temperature, temperature gradient, IP, resistivity, self potential, self potential gradient, velocity, neutron porosity, and caliper. This table is further summarized in the Regional Properties Table.



Area	Formation	Lithology	Alteration	DEN SD	DEN Mean	DEN Min	DEN Max	DEN Median	DEN
JAMES BAY LOWLANDS	ATTAWAPISKAT	LIMESTONE	ARGILLITE	0.276	2.325	2.13	2.52	2.52	
JAMES BAY LOWLANDS	E.KWAN/SEVERN R.	LIMESTONE	ARGILLITE	0.055	2.483	2.42	2.52	2.51	
JAMES BAY LOWLANDS	RED HEAD RAPIDS	LIMESTONE	ARGILLITE	0.035	2.475	2.45	2.5	2.5	
JAMES BAY LOWLANDS	BASEMENT	MUDSTONE BRECCIA		0	2.7	2.7	2.7	2.7	
JAMES BAY LOWLANDS	BAD CACHE	SANDSTONE	ARGILLITE	0.156	2.51	2.4	2.62	2.62	
JAMES BAY LOWLANDS	E.KWAN/SEVERN R.	SANDY LIMESTONE		0	2.44	2.44	2.44	2.44	
MANITOUWADGE	N/A	GRANITE		0.028	2.657	2.6	2.711	2.6625	
MANITOUWADGE	N/A	ALT FELSIC VOLC+PYR	SULPHITIC	0.148	2.676	2.39809	3.096057	2.661427	
MANITOUWADGE	N/A	ALT FELSIC VOLCANIC	SULPHITIC	0.102	2.563	2.31205	2.87036	2.57091	

Figure 3: The database administrator's interface showing the Regional Properties Table.

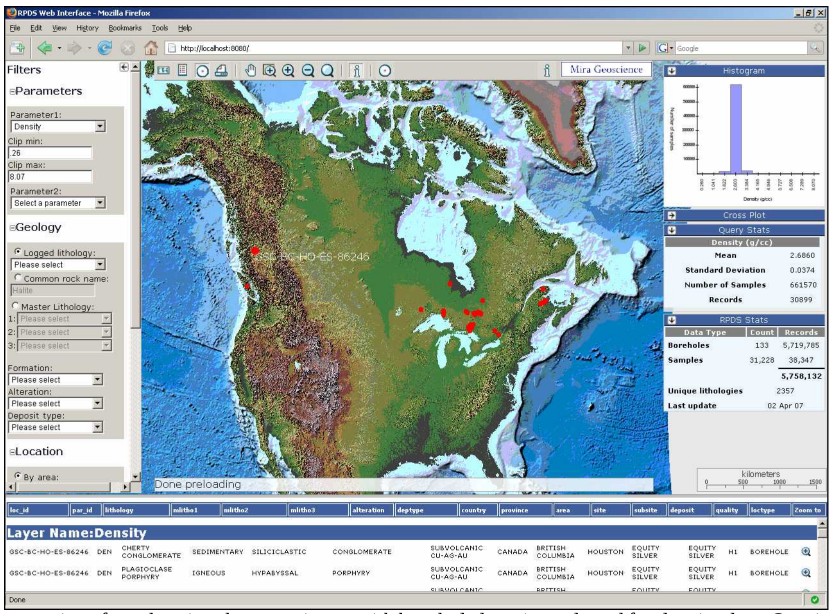


Figure 4: The RPDS web query interface showing the map viewer with borehole locations plotted for density data. Queries are defined in the pane to the left of the map, a histogram and query statistics are shown on the right, and metadata for selected data locations are shown in the table at the bottom of the screen.

The Regional Properties Table (Figure 3) is the final step in the data distillation process where physical property data is summarized and stored by combining mean physical property values from the same regional area that possess a common geological fingerprint, i.e. the same formation/lithology/alteration combination. Therefore, the physical properties of all occurrences of one geological unit in a borehole are averaged and combined with any other occurrences of that geological combination in the same area. As mentioned above, this provides one series of statistical summary values (mean, min, max, standard deviation, median, number of samples) for each physical property, for each unique geological combination in the same regional geographic area (e.g. one mean value for the density of basalt in Sudbury). The web query interface interacts directly with this table.

RPDS web interface

The prototype RPDS web interface is a publicly accessible web-based query tool available at: www.mirageoscience.com/rpds (Figure 4). It communicates with the Regional Properties Table to provide rapid query results on population statistics, including histograms and multiparameter cross-plots. Queries can be refined by geological parameters, location information, location type (borehole or sample) and data quality. The map interface also includes a series of pre-rendered map layers for rapid visualization. These layers include base maps and various symbolized layers showing the data distribution per physical property parameter. Some typical queries might be:

> "What is the average density of Paleozoic basalt in the Abitibi greenstone belt?" or,

"What is the correlation between velocity and fracture frequency for Sudbury granites?"

CURRENT STATUS AND LOOKING AHEAD

RPDS currently stores 5.7 million physical property records, distributed as follows:

- 5,719,795 measurements from 621 logging runs performed on 133 boreholes.
- 2,268 sample measurements from 1,965 borecore samples taken from 29 boreholes.
- 36,079 sample measurements from 29,271 surface samples.

Measurements were recorded on 2,357 unique lithologies representing 38 deposits from across Canada. Continued data entry is proceeding with funding from various government and industry sources. Although the RPDS data model is complete and the web query interface is fully functional, continued software enhancement is expected.

Initially, we hope that the RPDS technology will be widely used through the web interface as a tool for querying data and for integration of rock properties with ore deposit models. Eventually, a direct link between RPDS and Gocad to facilitate direct assignment of physical properties to 3D earth models will be developed.

ACKNOWLEDGEMENTS

The development of RPDS has been an 8 year collaboration of Mira Geoscience, Noranda/Falconbridge (now Xstrata), the Geological Survey of Canada, the Government of Ontario, McMaster University, Quantec Geoscience, Geoscience BC and other industry sponsors through CAMIRO projects. Many individuals have played a role in the design and implementation of RPDS, but we particularly acknowledge the key contributions of the following individuals: Jonathan Mwenifumbo and Barbara Elliott of the GSC, Bill Morris and Heather Slavinski of McMaster University, Gaétan Léonard of Flairbase Inc., and Tom Lane of CAMIRO.

REFERENCES

North American Commission on Stratigraphic Nomenclature, 1983, North American Stratigraphic Code. American Association of Petroleum Geologists, v.67, p. 841-875.