Geosci. Instrum. Method. Data Syst. Discuss., https://doi.org/10.5194/gi-2018-46
Manuscript under review for journal Geosci. Instrum. Method. Data Syst.
Discussion started: 8 May 2019

Geoscientific Instrumentation Methods and Data Systems Discussions
Open Access
EGU

A Universal and Multi-dimensional Model for Analytical Data on Geological Samples

Yutong He[1], Di Tian[1*], Hongxia Wang[1], Li Yao[2], Pengfei Chen[3]

[1]College of Instrumentation & Electrical Engineering, Jilin University, 130021, Jilin, China

[2]College of Earth Sciences, Jilin University, 130061, Jilin, China

[3]The 41st research institute CETC, 266000, Qingdao, China

[*]Corresponding author, email: tiandi@jlu.edu.cn, 18686648159

Abstract. To promote the sharing and reutilization of geoanalytical data, various geoanalytical databases have been established over the last 30 years. Data models, which form the core of a database, are themselves the subjects of intensive studies. Data models determine the contents stored in the databases and applications of the databases. However, most geoanalytical data models have been designed for specific geological applications, which has led to strong heterogeneity between databases. It is therefore difficult for researchers to communicate and integrate geoanalytical data between databases. In particular, every time a new database is constructed, the time-consuming process of redesigning a data model significantly increases the development cycle. This study introduces a new data model that is universally applicable and highly efficient. The data model is applied to various geoanalytical methods and corresponding applications, and comprehensive analytical data contents together with associated background metadata are summarized and catalogued. Universal data attributes are then designed based on these metadata, which means that the model can be used for any geoanalytical database. Additionally, a multi-dimensional data mode is adopted, providing geological researchers with the ability to analyze geoanalytical data from six or more dimensions with high efficiency. Part of the model is implemented with the typical database system-MySQL and comprehensive comparison experiments with existing geoanalytical data model are presented. The result unambiguously proves that the data model developed in this paper exceeds existing models in efficiency.

Keywords: geoanalysis, multi-dimensional data model, conceptual data model, logical data model, data structures

Geosci. Instrum. Method. Data Syst. Discuss., https://doi.org/10.5194/gi-2018-46
Manuscript under review for journal Geosci. Instrum. Method. Data Syst.
Discussion started: 8 May 2019

Geoscientific Instrumentation Methods and Data Systems Discussions

1 Introduction

Geoanalytical data includes measurements of major and trace elements, rare earth elements (REEs), isotopes, and structures and morphology, of geological samples analyzed by various analytical instruments such as ICP (LA-ICP-MS, ICP-MS), EPMA, SIMS, SEM, TEM, XRF, and XRD. Geoanalytical data effectively reflect the material composition, internal structure, external characteristics, interaction and evolution history of the Earth and represent the most important support for geological researchers in their aim to understand the Earth and exploit its resources for the survival and development of human society. Enormous financial, material, and human resources have been invested into the geological surveys and geoanalysis required to acquire more comprehensive and abundant geoanalytical data. Over time, tremendous volumes of geoanalytical data have been created, and these volumes continue to increase at a high rate. It is of paramount importance that this data is curated effectively and that adequate background information, such as sample description, sampling information, and analysis information, is included, so that geological researchers can utilize the data according to their requirements. This will also facilitate the reutilization of the precious geoanalytical data. In addition, with the accumulation of large volumes of data, statistical analysis and data mining can be conducted on these data to provide a more comprehensive and advanced scientific understanding of the Earth. Hence, a variety of geoanalytical databases aimed at managing, sharing and reutilizing geoanalytical information have been constructed and are used as advanced tools in geological studies. The analysis and comparison of existing geoanalytical data models, as well as the development of improved models, is therefore a worthy and significant study to be conducted.

Over the last decades, several studies of geoanalytical data models have been conducted. As early as 1977, Jeorge Van Trump and colleagues described a data model for environmental geochemical surveying and mineral resource exploration in the United States of America (Jr and Miesch, 1977)(. Lehnert et al. (2000) suggested a data model for the storage of global geochemical data of rocks. Their data model provides a complete summary of essential geochemical data contents and a robust structure with relational database management system (RDBMS). Numerous databases such as GEOROC, NAVDAT, and PetDB have since been constructed based on this model, and it is used by geological researches worldwide. In particular PetDB has been used for a considerable amount of high-impact research such as *Nature* (Brandl et al., 2013; Carbotte et al., 2013; Cheng et al., 2016; Dick and Zhou, 2014; Helo et al., 2011; Hoernle et al., 2011; Kamenov et al., 2011; Kelley, 2014; Samuel and King, 2014; Schlindwein and Schmid, 2016; Straub et al., 2009) and *Science* (Cottrell and Kelley, 2013; Greber et al., 2017; Joy et al., 2012; Kelley and Cottrell, 2009; Mcnutt et al., 2016). A limitation of existing geoanalytical data models is their specificity to particular applications or geological

Geosci. Instrum. Method. Data Syst. Discuss., https://doi.org/10.5194/gi-2018-46
Manuscript under review for journal Geosci. Instrum. Method. Data Syst.
Discussion started: 8 May 2019

domains and their focus on the description and curation of only a certain portion of geoanalytical data. For example, RU_CAGeochem is specifically focused on major and trace element concentrations and Sr, Nd, and Pb isotopic ratios of American volcanic rocks (Carr et al., 2014). Another database is focused on lead-isotopes of copper ores from the south-eastern Alps (Artioli et al., 2016). Many other examples of similarly specific geoanalytical databases and associated models exist (e.g. Artioli et al., 2016; Hellström, 2016; Lopes et al., 2014; Siegel et al., 2012; Strong et al., 2016). The consequence of this development is that each database exists as a separate island and it is difficult for researchers to communicate and integrate geoanalytical data between databases. In particular, every time a database is constructed, a data model has to be redesigned. This consumes considerable amounts of time and prolongs the development cycle. In addition, the vast majority of models are designed based on relational models, which focus on the construction of relations between different data categories. When users query and utilize the geoanalytical data from different dimensions, these types of models utilize complicated joints between different tables to query the target data, which decreases efficiency as the amount of data increases. However, the exploration of such data models including the background items, have laid a solid foundation for later study of advanced geoanalytical data models.

At present, the development of various new techniques provides us with the opportunities to design more comprehensive and advanced geoanalytical data models. In this study, we introduce a novel, universal and efficient geoanalytical data model. First, we provide an overview of geoanalytical methods and applications to summarize the geoanalytical data available. Then, we design universal data attributes based on this data and develop a multi-dimensional data model. Finally, we evaluate the model to validate its efficiency.

2 Overview of geoanalytical data contents

In recent years, many new geoanalytical methods and instruments have been developed, creating novel kinds of data (Linge et al., 2017). A truly universal data model should have the ability to accommodate all kinds of geoanalytical data. In addition, the data model should be capable of making all stored data readily available for reutilization by geological researchers. In order to develop a model with such capabilities, a comprehensive set of geoanalytical data, together with related background information required for reutilization of the data, was summarized and categorized, as outlined below.

First, analytical techniques and their applications were studied to comprehensively summarize geoanalytical measurement data. This process is outlined in Figure 1. Because of the great diversity of analytical methods and geological applications, Figure 1 only shows a few examples to indicate the method adopted in this paper. The five

Geosci. Instrum. Method. Data Syst. Discuss., https://doi.org/10.5194/gi-2018-46
Manuscript under review for journal Geosci. Instrum. Method. Data Syst.
Discussion started: 8 May 2019

Geoscientific Instrumentation Methods and Data Systems Discussions

categories, namely, bulk analysis; microanalysis; isotope analysis; morphology, structure, and valence analysis; and organic analysis, were divided according to the analytical technique used. In this way, data from each category was categorized according to analytical instruments (e.g., SEM, SNM, and EPMA for microanalysis). In the next step, the data were grouped according to geological applications. The comprehensive list of geoanalytical measurement data items used in the present study, compiled from a thorough literature review, is presented in Figure 2. In the case of bulk analysis, most measurements ultimately provided major, trace and ultra-trace element concentration data. Microanalysis can yield data of elemental concentrations in a micro-region, as well as structural information of geological samples acquired by secondary electron and backscattered electron techniques, commonly stored as image files. For geochronology and stable isotopic analysis (GSI analysis), most measurement data are isotopic ratios. For morphology, structure and valence analysis (MSV analysis), the most common measurement data are image files such as XPS spectra or XRD patterns. Organic analysis is a new analytical method, which is used for the analysis of environmental geological samples. The most common application of this method in the geological literature is the analysis of the sixteen kinds of Polycyclic Aromatic Hydrocarbons PAHs in soils.

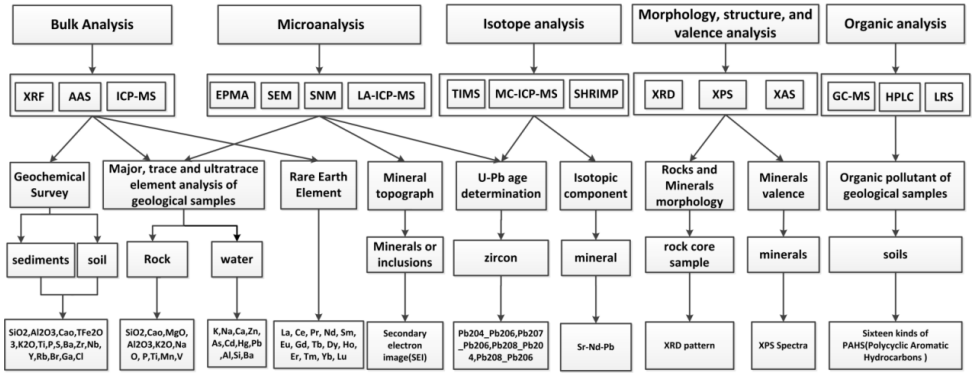


Figure 1: Process of summarizing the geoanalytical data contents.

Geosci. Instrum. Method. Data Syst. Discuss., https://doi.org/10.5194/gi-2018-46
Manuscript under review for journal Geosci. Instrum. Method. Data Syst.
Discussion started: 8 May 2019

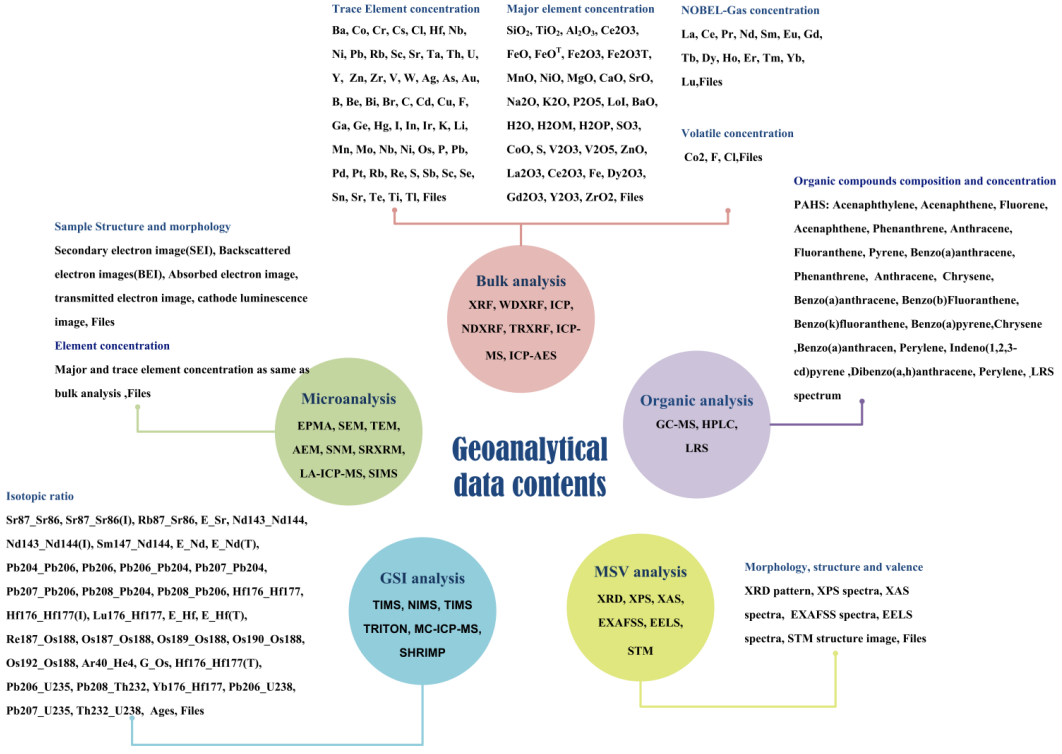


Figure 2 lists of geoanalytical methods and measurement data items

Background information describing the analyzed samples and data quality has to be incorporated, because it is indispensable for proper evaluation, efficient recovery, and sorting of the compiled data. Hence, background metadata are summarized based on the investigations of the geological researchers and the contents of existing databases (Adcock et al., 2003; Lehnert et al., 2000). Table 1 lists details of the background metadata used during the present study. In this study, the Background metadata are divided into three parts: sample metadata provide geological researchers with information about geological materials, sampling metadata provides information about environmental conditions in the field, and quality metadata allows geological researchers to make an assessment of data quality and usability (Table 1). The background metadata items listed in Table 1 are the most essential information required for every kind of geoanalytical measurement data. More specific attributes are not included in our model.

Table 1 Background metadata of geoanalytical data

Background Metadata

Geosci. Instrum. Method. Data Syst. Discuss., https://doi.org/10.5194/gi-2018-46
Manuscript under review for journal Geosci. Instrum. Method. Data Syst.
Discussion started: 8 May 2019

Geoscientific Instrumentation Methods and Data Systems Discussions
Open Access
EGU

Sample metadata	sample type, sample name, sample description
Sampling metadata	sampling site description, longitude, latitude, sampling methods, sampling depth, sampling institution
Quality metadata	laboratory name, project name, publish link, analytical method, analysts

Geoanalytical data modelling

This section outlines how the novel geoanalytical data model was designed, utilizing the data summarized above. Despite their limitations, currently relational data mode is the most commonly used pattern for geoanalytical data models. Rational data mode construct relations between each group of data within the database. This means that more data categories inevitably lead to much more data relations, increasing storage demands and the time required to query the database. Compared to such conventional relational data models (Beynon-Davies, 2004), multi-dimensional models (MDM), which are widely utilized during the development of big data science and data mining, are single subject-oriented sources for analyzing data based on various dimensions (Niemi and Hirvonen, 2003). Multi-dimensional modelling approaches share characteristics with Fast Analysis of Shared Multidimensional Information (FASMI). In particular, MDM offers the advantage of a relatively simple and straight-forward database design, which nevertheless supports powerful analyses, and is relatively well understood by the end users (Hoberman, 2005). As a modelling framework, MDM has a conceptual and a logical phase of design, composed of a fact table and several dimension tables (Höpken et al., 2013). Facts comprise numeric and additive characteristics of the data, which can be accumulated along multiple dimensions. Frequently, researchers are interested in analyzing geoanalytical measurement data from different metadata perspectives. Hence, the MDM approach is ideally suited for the design of geoanalytical data models. Here, the geoanalytical data are the fact data, and other background information are dimensions data. The use of the MDM modelling framework applied in the present study will allow geological researchers to rapidly analyze geoanalytical data based on numerous meta-data criteria.

2.1 Conceptual data model (CDM)

A conceptual data model (CDM) includes the definition of its universal attributes and a rough design of its structure. It represents the primary phase of data model design, independent from the detailed techniques of computer systems. Figure 3 presents the multi-dimensional CDM we developed for geoanalytical data. Here, the abstraction

Geosci. Instrum. Method. Data Syst. Discuss., https://doi.org/10.5194/gi-2018-46
Manuscript under review for journal Geosci. Instrum. Method. Data Syst.
Discussion started: 8 May 2019

Geoscientific Instrumentation Methods and Data Systems Discussions

of universal concepts present in geoanalytical data, the model becomes more flexible and universally applicable. The geoanalytical data are placed in the center of the model, in the form of a facts table. The associated background information is categorized and abstracted as various dimensions which are represented by different axes in Figure 3. The six dimensions of our CDM are sample, analysis type, analytical methods, location, time and quality. This arrangement allows geological researchers to analyze geoanalytical data from six different dimensions or any combination thereof. The marks in each dimension represent the detailed measurement conditions. The "n" dimension is an expansible dimension, which can be added according to the specific model application.

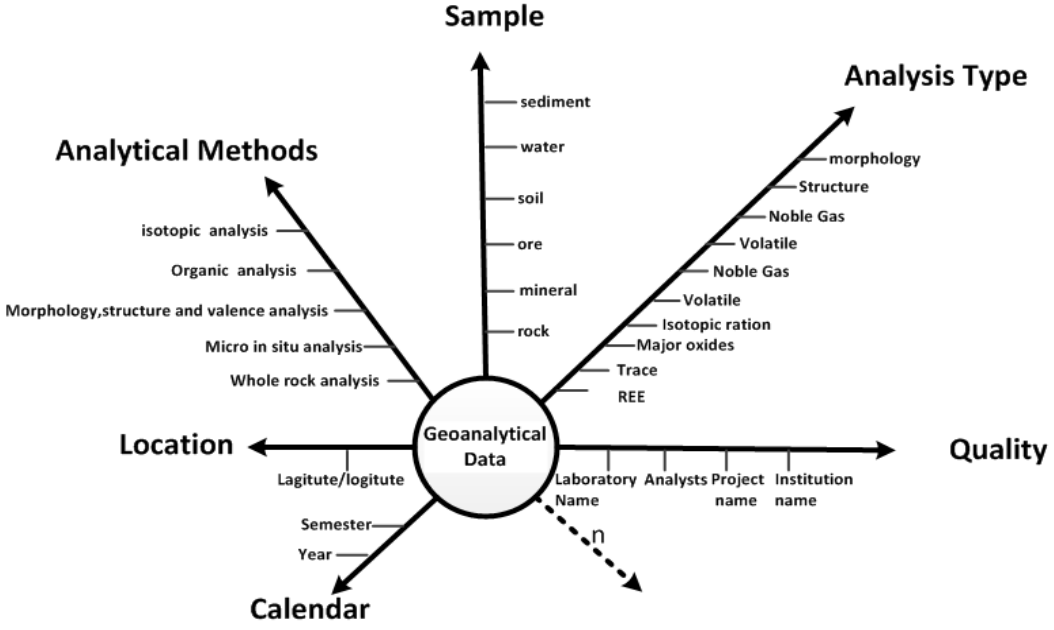


Figure 3: Multi-dimensional CDM for geoanalytical data.

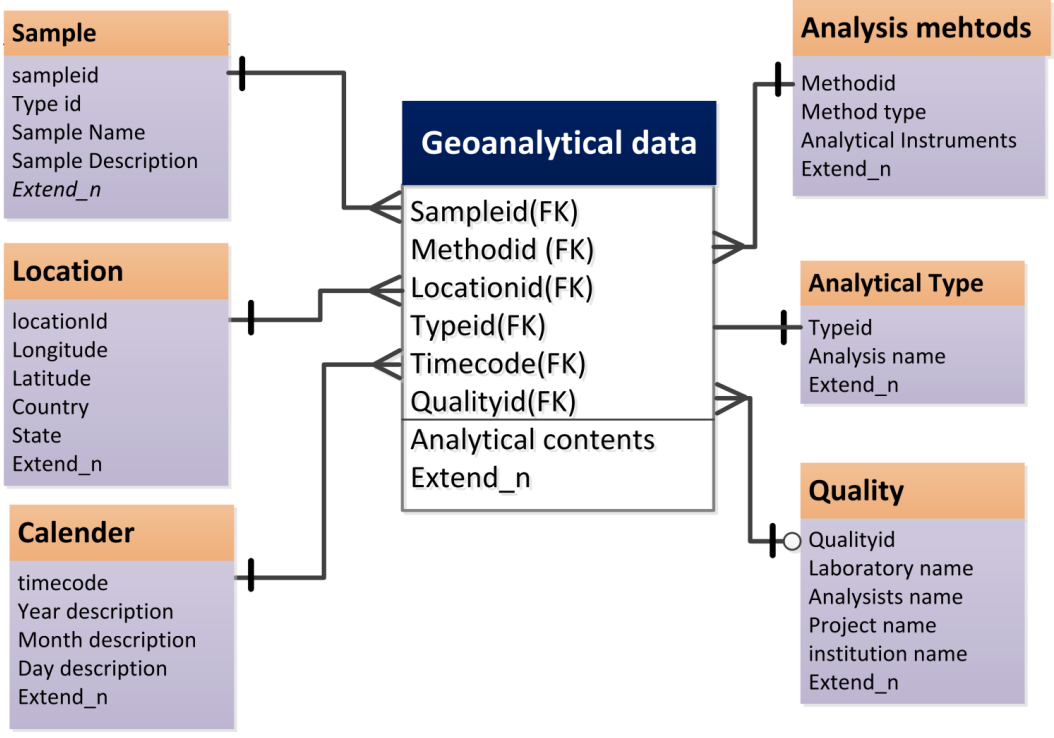
2.2 Logical data model (LDM)

A logical data model (LDM) is a CDM written in unified modelling language (UML) (Evans et al., 2014). Logical model design leads to a logical scheme, defining objects, attributes, and relationships (Chmura and Heumann, 2005). The LDM scheme can be easily implemented by any DBMS. Figure 4 shows the LDM scheme designed for geoanalytical data. Each box in the LDM represents an object, and items in the box are its attributes. The relations between object are represented with lines. There are three kinds of symbols associated with the lines. The short line

Geosci. Instrum. Method. Data Syst. Discuss., https://doi.org/10.5194/gi-2018-46
Manuscript under review for journal Geosci. Instrum. Method. Data Syst.
Discussion started: 8 May 2019

Geoscientific Instrumentation Methods and Data Systems Discussions

150 denotes "1", the circle denotes "0" which means "maybe", and the triangle denotes "many". Lines and symbols
151 define the relations between objects. The additional notation foreign key (FK) is added if the attribute in one object
152 uniquely identifies an attribute in another object. For example, the sample ID in the geoanalytical data object is a
153 foreign key of Sample_id in the sample object, because they have the same value. By means of this foreign key, the
154 data contents of the two objects are connected. For each object, a few extended attributes are added (*Extend_n* in
155 Figure 4). This feature allows developers to add database specific attributes to this model, increasing its flexibility
156 and universal applicability.



157

158 **Figure 4: LDM of the geoanalytical data model.**

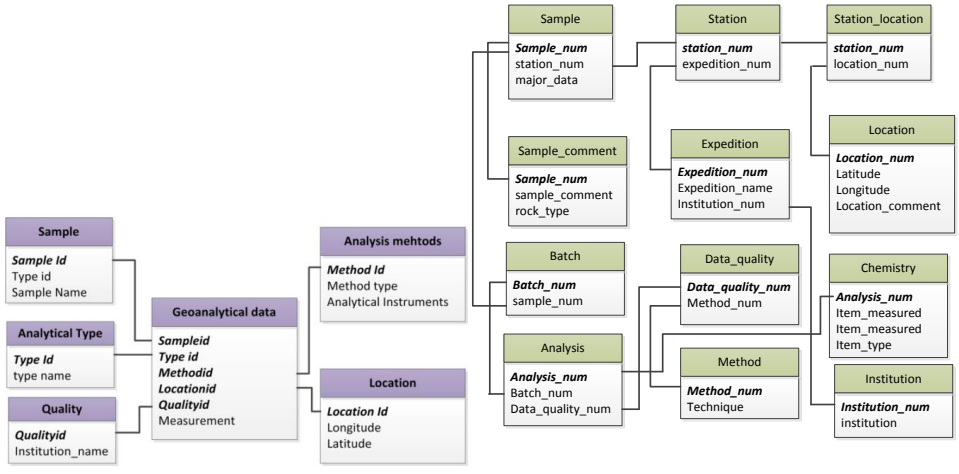
159 **3 Implementation and Evaluation**

160 In order to evaluate the performance of our model, we carried out a comparison experiment with the widely used
161 Lehnert rock geochemical data model (Lehnert et al., 2000). In order to conduct the experiment, a physical data
162 model (PDM) needed to be created with a database management system. As RDBMS is the most common
163 techniques used in geoanalytical databases, MySQL which is a widely used RDBMS was adopted to implement the

Geosci. Instrum. Method. Data Syst. Discuss., https://doi.org/10.5194/gi-2018-46
Manuscript under review for journal Geosci. Instrum. Method. Data Syst.
Discussion started: 8 May 2019

Geoscientific Instrumentation Methods and Data Systems Discussions

two models. A specific data item (rocktype: andesite; location: sycamore Hall; latitude: 36.27.12N, longitude: 83.34.12W; Institution: Jilin University; Method: ICP-MS; SiO2:58.9; FiO2:1.13) was used as test data and tables related to the data contents were implemented. We analyzed the two models from two perspectives: developers and users. For developer, the comparison of the PDM structure is shown in Figure 5, and query operation descriptions are presented Figure 6. The comparison clearly indicates that the geoanalytical data model is more succinct than rock data model, and saving time and computer resources. Three model performance indicators (insert time, storage space usage, and retrieval time) were evaluated with the increasing of amounts of data. The results are shown in Figures 7, 8, and 9, respectively. Figure 7 shows clearly that the process of data insertion is considerably faster for the geoanalytical data model, when compared to the rock data model. Figure 9 shows clearly that the storage space usage is relatively less than rock data model. In the case of data query (Figure 8), the difference in time consumption is even more striking. With increasing amount of data items, the query time of the geoanalytical data model remains very fast and efficient. In contrast, for the rock data model query time costs increased exponentially with the increasing amount of data items.



(a) PDM of geoanalytical data model (b) PDM of Lehnert rock data model

Figure 5: PDM structure comparison with the Lehnert rock data model.

Geosci. Instrum. Method. Data Syst. Discuss., https://doi.org/10.5194/gi-2018-46
Manuscript under review for journal Geosci. Instrum. Method. Data Syst.
Discussion started: 8 May 2019

SQL comparison

Insert operation

Geoanalysis model	Rock model
• insert into sample(sample_num,station_num) values(10, 50) • insert into analysis(analysis_num,batch_num,data_quality_num) values(30,10,90) • insert into batch(batch_num,sample_num) values(20, 10) • insert into station(station_num) values(50) • insert into institution(institution_num,institution_name) values(70,Jilin university) • insert into station_location(station_num,location_num) values(50, 80) • insert into location_num(station_num, longitude, latitude) values(30,83.34.10,36.27.12) • insert into data_quality(data_quality_num,method_num) values(90, 40) • insert into method(method_num,technique) values(40, ICPMS) • insert into chemistry(analysis_num,,item_measured,item_type) values(30, ,Sio2,59)	• insert into sample (sampleid, sample_name) values(10,sulphate) • insert into quality(qualityid,institution_name) values(20, Jilin university) • insert into location(locationid, longitude, latitude) values(30,83.34.10,36.27.12) • insert into analysis_methods(methodid,analytical_instruments) values(40, ICPMS) • insert into measurement(sample_id,methodid,Locationid,qualityid,Item,value) values(10,40,30,20,Sio2,59)

Query operation

Geoanalysis model	Rock model
select l.location_num, l.latitute, l.logitute, s.station_num, s.location_num, t.station_num, a.sample_num, a.station_num, b.batch_num, b.sample_num, y.batch_num, y.analysis_num, c.analysis_num, c.item_measured, c.item_type from location l, station_location s, station t, sample a, batch b, analysis y, chemistry c where l.latitute= '36.27.12' and l.logitute='83.34.10'and l.location_num=s.location_num and s.station_num=t.station_num and t.station_num=a.station_num and a.sample_num=b.sample_num and b.batch_num=y.batch_num and y.analysis_num=c.analysis_num;	select l.locationid, l.longitude, l.latitude, m.locationid, m.item, m.value from location l, measurement m where l.latitude= '83.34.10' and l.longitude='36.27.12'and l.locationid=m.locationid;

Figure 6: Comparison of insert and query operations in SQLs.

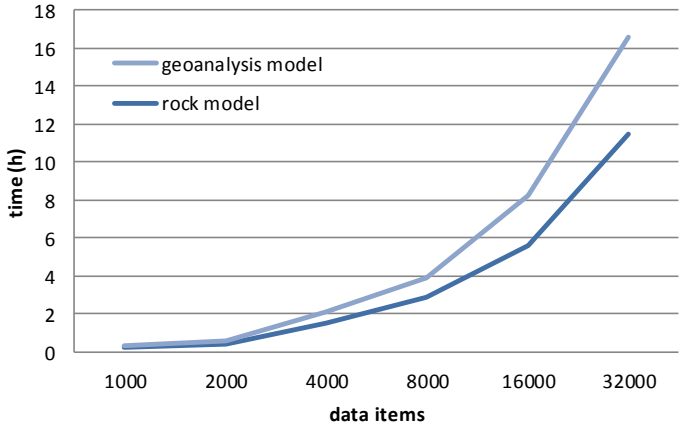


Figure 7: Time spent on data insert operations with increasing amounts of data.

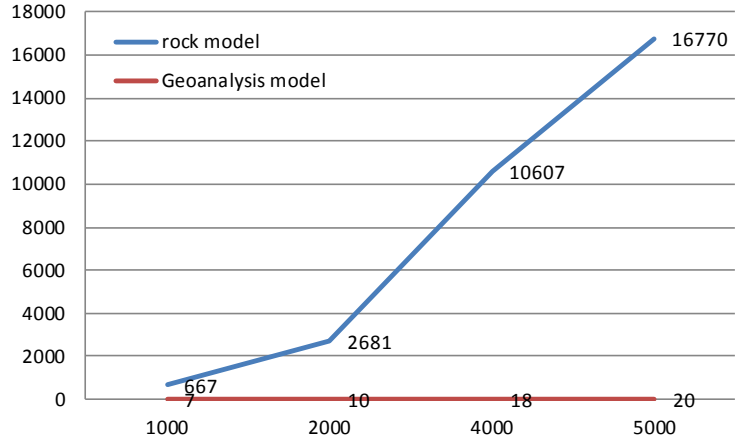


Figure 8: Time requirement for data queries (latitude and longitude).

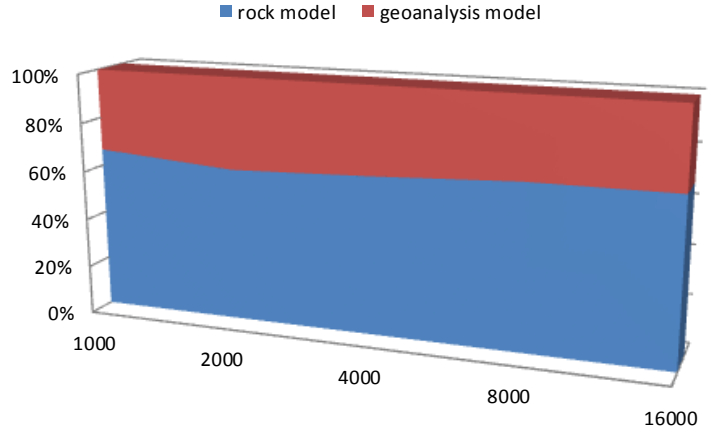


Figure 9: Space usages of the two data models with increasing amount of data items.

4 Conclusions

The geoanalytical data model presented herein is flexible and appropriate for a broad range of applications to geoanalytical data. The model has the following general characteristics:

[1] Its universality allows the model to accommodate any type of geoanalytical data for various geological

materials, as well as all significant metadata.

[2] The adoption of a multi-dimensional data model framework provides geological researchers with the ability to analyze geoanalytical data from different dimensions. In addition to the sample description and location criteria commonly used in existed databases, this model provides four additional query criteria (method, quality, time, and analysis).

[3] There are minimum data relations between different objects. Relations between different background metadata objects have been avoided in order to construct robust relations between background metadata and measurement data. This increases the model's efficiency when geoanalytical data are inserted or queried while simultaneously decreasing its space usage.

It is hoped that the design of this model will allow for the unified construction of geoanalytical databases. The model enables the accumulation and integration of significant amounts of diverse geoanalytical data. By utilization of the big data analysis techniques described in our study, geological researchers could analyze geoanalytical data with high efficiency and develop novel methods to conduct Earth science studies.

Acknowledgments:

This work was financially supported by National Major Scientific Instruments and Equipment Development Special Funds [2016YFF0103303, 2011YQ050069] and CGS research fund [JYYWF20181702].

Reference

Adcock, S.W., Grunsky, E. and Laframboise, R., 2003. A Universal Geochemical Survey Data Model.

Beynon-Davies, P., 2004. Relational Data Model.

Brandl, P.A., Regelous, M., Beier, C. and Haase, K.M., 2013. High mantle temperatures following rifting caused by continental insulation. Nature Geoscience, 6(5): 391-394.

Carbotte, S.M., Marjanović, M., Carton, H., Mutter, J.C., Canales, J.P., Nedimović, M.R., Han, S. and Perfit, M.R., 2013. Fine-scale segmentation of the crustal magma reservoir beneath the East Pacific Rise. Nature Geoscience, 6(10): 866-870.

Cheng, H., Zhou, H., Yang, Q., Zhang, L., Ji, F. and Henry, D., 2016. Jurassic zircons from the Southwest Indian Ridge.

226 Sci Rep, 6: 26260.

227 Chmura, A. and Heumann, J.M., 2005. Logical Data Modeling. Integrated, 5(2): 179-203.

228 Cottrell, E. and Kelley, K.A., 2013. Redox heterogeneity in mid-ocean ridge basalts as a function of mantle source.

229 Science, 340(6138): 1314.

230 Dick, H.J.B. and Zhou, H., 2014. Ocean rises are products of variable mantle composition, temperature and focused

231 melting. Nature Geoscience, 8(1): 68-74.

232 Evans, A., France, R., Lano, K. and Rumpe, B., 2014. The UML as a Formal Modeling Notation. Computer Standards &

233 Interfaces, 19(7): 325-334.

234 Greber, N.D., Dauphas, N., Bekker, A., Ptáček, M.P., Bindeman, I.N. and Hofmann, A., 2017. Titanium isotopic evidence

235 for felsic crust and plate tectonics 3.5 billion years ago. Science, 357(6357): 1271-1274.

236 Höpken, W., Fuchs, M., Höll, G., Keil, D. and Lexhagen, M., 2013. Multi-Dimensional Data Modelling for a Tourism

237 Destination Data Warehouse.

238 Helo, C., Longpré, M.A., Shimizu, N., Clague, D.A. and Stix, J., 2011. Explosive eruptions at mid-ocean ridges driven

239 by CO2-rich magmas. Nature Geoscience, 4(4): 260-263.

240 Hoberman, S., 2005. Data Modeling Essentials. Dm Review, 131(8): 654 - 660.

241 Hoernle, K., Hauff, F., Werner, R., Bogaard, P.V.D., Gibbons, A.D., Conrad, S. and Müller, R.D., 2011. Origin of Indian

242 Ocean Seamount Province by shallow recycling of continental lithosphere. Nature Geoscience, 4(12): 883–887.

243 Joy, K.H., Zolensky, M.E., Nagashima, K., Huss, G.R., Ross, D.K., Mckay, D.S. and Kring, D.A., 2012. Direct detection

244 of projectile relics from the end of the lunar basin-forming epoch. Science, 336(6087): 1426.

245 Jr, V.T. and Miesch, A.T., 1977. The U.S. geological survey rass-statpac system for management and statistical reduction

246 of geochemical data. Computers & Geosciences, 3(3): 475-488.

247 Kamenov, G.D., Perfit, M.R., Lewis, J.F., Goss, A.R., Jr, R.A. and Shuster, R.D., 2011. Ancient lithospheric source for

248 Quaternary lavas in Hispaniola. Nature Geoscience, 4(8): 554-557.

249 Kelley, K.A., 2014. Inside Earth Runs Hot and Cold. Science, 344(6179): 51-52.

250 Kelley, K.A. and Cottrell, E., 2009. Water and the oxidation state of subduction zone magmas. Science, 325(5940):

251 605-7.

252 Lehnert, K., Su, Y., Langmuir, C.H., Sarbas, B. and Nohl, U., 2000. A global geochemical database structure for rocks.

253 Geochemistry Geophysics Geosystems, 1(5): 179-188.

254 Linge, K.L., Bédard, L.P., Bugoi, R., Enzweiler, J., Jochum, K.P., Kilian, R., Liu, J., Marin-Carbonne, J., Merchel, S. and

255 Munnik, F., 2017. GGR Biennial Critical Review: Analytical Developments Since 2014. Geostandards &
256 Geoanalytical Research, 36(4): 337-398.
257 Mcnutt, M.K., Lehnert, K., Hanson, B. and Nosek, B.A., 2016. Liberating field science samples and data. Science,
258 351(6277): 1024.
259 Niemi, T. and Hirvonen, L., 2003. Multidimensional data model and query language for informetrics. John Wiley & Sons,
260 Inc., 939–951 pp.
261 Samuel, H. and King, S.D., 2014. Mixing at mid-ocean ridges controlled by small-scale convection and plate motion.
262 Nature Geoscience, 7(8): 602-605.
263 Schlindwein, V. and Schmid, F., 2016. Mid-ocean-ridge seismicity reveals extreme types of ocean lithosphere. Nature,
264 535(7611).
265 Straub, S.M., Goldstein, S.L., Class, C. and Schmidt, A., 2009. Mid-ocean-ridge basalt of Indian type in the northwest
266 Pacific Ocean basin. Nature Geoscience, 2(4): 286-289.
267