DOI:10.1190/segam2021-3582154.1

Bayesian inversion for rock composition and petrophysical endpoints in multimineral analysis

Liwei Cheng, and Ge Jin, Colorado School of Mines; Reinaldo J. Michelena, SeisPetro Geosoftware, LLC.*

Summary

Multimineral analysis that quantifies the volume fractions of minerals and fluids from a set of well logs has been used for reservoir characterization in complex geological settings. However, due to the data errors and the similarity between petrophysical endpoints, the solutions of multimineral analysis are non-unique. Furthermore, defining the petrophysical endpoints is challenging in complex geological settings because standard endpoint values may not be optimal. All the uncertainties must be evaluated but cannot be achieved by standard linear solvers. Stochastic Bayesian inversion methods have been developed to assess the uncertainties, but the high computational time and the need for detailed prior information hinder their practical use. We employ a Markov chain Monte Carlo with ensemble samplers (MCMCES) in the Bayesian framework, which is more efficient in convergence than the conventional random walk methods in high dimensional problems. We apply the new method in two different applications. First, we evaluate the uncertainty of constituent volume fractions resulting from the data errors and the similarity of endpoints on a conventional carbonate reservoir. In our second implementation of MCMCES, we assess the uncertainty of key endpoints that are difficult to estimate and optimize multimineral analysis using a synthetic dataset and field data from the Bakken Formation. Our proposed method provides different realizations in volume fractions or in petrophysical endpoints for interpreters to better evaluate multimineral results.

Introduction

Multimineral analysis is an analytical tool that relates the well logs to the rock composition through a matrix of petrophysical endpoints. The rock composition can be represented by the volume fractions of minerals and fluids (collectively called constituents) which are jointly inverted at every depth by minimizing the difference between the theoretical tool responses and a set of well logs (Mayer and Sibbit, 1980; Quirein et al., 1986). Due to the uncertainties in both data errors and petrophysical endpoints, the solutions of multimineral analyses are non-unique. Independent data, like core measurements, are generally required to calibrate the multimineral results, but they may not be available. The conventional linear solvers can only optimize the inverse problem for limited solutions and do not provide uncertainties of volume fractions and the endpoints used in the analysis.

Bayesian inversion methods have been used in the past to assess uncertainty in log evaluation workflows. For instance, Spalburg (2004) applies the random walk Markov chain Monte Carlo simulation to evaluate the uncertainty, but the proposed implementation only focuses on net-to-gross ratio, porosity, and saturations. Yang and Torres-Verdin (2015) and Deng et al. (2019 and 2020) formulate Bayesian inference methods for the uncertainty of mineral and fluid volume fractions. However, their methods generally require normal distributions as prior information to help convergence and require large computing times (approximately 1 minute per sample). The slow convergence and the large computational time of the method in high dimensions prevent its use in practical applications.

In this abstract, we show two implementations of the Markov chain Monte Carlo with ensemble samplers (MCMCES) to the problem of multimineral analysis. First, we estimate volume fractions. Given a set of constant petrophysical endpoints, we assess the uncertainty of fractions resulting from the data errors and the similarity between endpoints of constituents. The new implementation estimates the posterior distribution of each constituent volume fraction more efficiently than the previously published methods and does not require any specific prior distributions except the lower and the upper limits of volume fractions.

Besides the uncertainty in volume fractions, another common challenge in multimineral analysis is the selection of the constituent endpoints. Most commercially available applications use default endpoints for the tool responses of constituents. However, these predefined values are not always optimal, especially for the unconventional reservoirs, because the number of constituents is commonly greater than the number of well logs. In practice, multimineral analysis is an iterative process where petrophysicists tune these endpoints within reasonable ranges by manual trial-and-error to minimize the difference between the theoretical predictions and the well logs as well as core measurements. To automate this time-consuming process of adjusting the endpoints, Michelena et al. (2020) use a genetic algorithm to solve simultaneously for fractions and endpoints. However, a statistical method to assess the uncertainty of the petrophysical endpoints is still needed.

Our second MCMCES application aims to estimate the uncertainty of key endpoints by simply setting ranges as prior information, instead of constant values, in the endpoint's matrix. MCMCES explores the model space of endpoints and provides the posterior distributions needed to optimize the multimineral results. We demonstrate the new implementation with a synthetic model and a field dataset from the Bakken Formation.

Bayesian inversion in multimineral analysis

Review of Multimineral Analysis

For simplicity, we use a linear mixing model to predict the log response (*d*) for each depth as follows

$$d = Gm, \quad (1)$$

where *G* is the matrix that contains the endpoints of constituents for each tool (e.g., bulk density, neutron porosity, etc.) and *m* is the vector of constituent volume fractions. Among the simulated logs (*d*), resistivity is modeled through the Archie's equation and converted to conductivity (assuming both cementation and saturation exponents equal to two).

Multimineral analysis implements a joint inversion to minimize the difference between the theoretical tool responses and the well log data. The L-2 norm of equation (1) is given by

$$\min_m \|Gm - D\|_2^2, \quad (2)$$

where *D* is a vector of well log data. Additional constraints are that the volume fraction of each constituent should be between 0 and 1 and that the sum of all fractions should be equal to 1 (the unity constraint). Due to the errors in the data (*D*) and the endpoint matrix (*G*), the solutions of multimineral analysis are non-unique. Moreover, constituents with similar endpoint values may result in a high condition number of the matrix *G*, and solutions become unstable. All the uncertainties must be assessed with different realizations.

Markov Chain Monte Carlo with Ensemble Samplers

Monte Carlo simulations in the Bayesian framework provide a means by which uncertainties in the data can be translated into uncertainties in the simulated results. For a given dataset (*D*), the posterior probability distribution $p(\theta|D)$ of the model parameter θ can be sampled from

$$p(\theta|D) \propto p(D|\theta)p(\theta), \quad (3)$$

where $p(D|\theta)$ is the likelihood function, and $p(\theta)$ is the prior distribution of the model parameter.

The implementation of Bayesian inference methods may be straightforward, but the difficulty lies in the parameter tuning as well as the slow convergence for the high dimensional problems. To optimize this problem for multimineral analysis, we employ the Markov chain Monte Carlo with ensemble samplers (MCMCES), which uses a simultaneously evolving ensemble of walkers where the proposal distribution for one walker is based on the current positions of the ensemble walkers. More details about MCMCES can be found in Goodman and Weare (2010).

Depending on the goal of the Bayesian inference application, the model parameter (θ) in equation 3 can be the volume fractions (*m*), which is a linear simulation (for a given *G*), or the endpoint matrix (*G*), which is a non-linear simulation. In the following sections, we set up different examples for each application.

Estimation of Uncertainty in Volume Fractions (*m*)

Our first application of MCMCES is to assess the uncertainty of volume fractions. In this case, the likelihood function $p(D|m)$ can be described by

$$p(D|m) \propto \exp\left[-\frac{1}{2}(Gm - D)^T W^{-1}(Gm - D)\right], \quad (4)$$

where *W* is a diagonal matrix containing terms related to the noise of well logs (Deng et al., 2019). The primary uncertainty in this case comes from the data errors. After determining the acceptable data misfits, MCMCES aims to produce the posterior results within the targeted data misfit. The secondary uncertainty results from the similarity in constituent endpoints. As shown in equation 2, the inverse problem is simulated through the endpoint matrix (*G*) in which the values of individual endpoints play a significant role. Even when the inverse problem is well determined (the number of logs plus the unity constraint is equal or greater than the number of unknown volume fractions), the uncertainty of each constituent volume fraction is different. The ambiguity exists when the endpoints of constituents are similar, such as anhydrite, dolomite, and calcite. Furthermore, such an ambiguity changes with the weighting function of input logs defined by the interpreters.

Figure 1 shows posterior volume fraction models from a carbonate reservoir using MCMCES. The input well logs are gamma ray, resistivity, bulk density, neutron porosity, acoustic slowness, and volumetric cross-section. The rock composition consists of anhydrite, dolomite, calcite, quartz, water, and oil. The posterior distributions are displayed as color schemes at every depth whereas the deterministic solutions are shown in the black dashed lines as reference.

In this application, we use 1,000 walkers in volume fraction with a random prior distribution between [0, 1] for minerals and [0, 0.5] for fluids. Each walker takes 140 steps. Figure 2 shows the detailed trace plots where all the steps of the walkers are recorded for one depth sample at 11,624 ft. The realizations are only acceptable when their weighted data misfits are acceptable (> -15), which defines the burn-in section and the posterior distribution of each constituent volume fraction. The posterior distributions are all plausible realizations that fit the data within the acceptable misfit.

After 90 steps, most walkers converge and the data misfits (top panel of Figure 2) stabilize within the target range. Both Figures 1 and 2 demonstrate the greater uncertainties (lower resolutions) for the volume fractions of anhydrite, calcite, and dolomite than those of clay, water, and oil, mainly due to the similarity in petrophysical endpoints of those minerals. Note that the new method does not use any Gaussian prior distributions, which is best for applications where only upper and lower limits of the volume fractions are known. Regarding computing time, a single MCMCES simulation at one depth sample (shown in Figure 2) takes 3.5 seconds on a desktop PC (i7 CPU at 3.47 GHz with 24-GB memory) using a MATLAB 2018b platform.

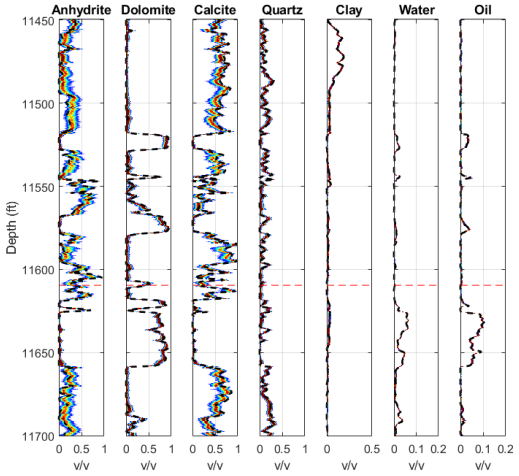

Figure 1. The posterior distributions for constituent volume fractions. Anhydrite, dolomite, and calcite have greater uncertainties than those of clay, water, and oil, due to the similar endpoint values. The solutions from a conventional linear solver are shown as black dashed lines for reference.

Estimation of Uncertainty in Endpoints (*G*)

Our second application of MCMCES aims to explore the model space of endpoints that are typically optimized by the interpreter's manual, trial-and-error process. The standard values of endpoints are commonly not optimal in a complex geological setting, which increases the uncertainty of multimineral results. The reasons why the endpoints may depart from the standard values are (Michelena et al., 2020): 1) more unknown minerals than logs available which forces the use of mixed "pseudo-minerals"; 2) lack of pure minerals that exhibit standard values interval of interest; 3) spatial variation of endpoints; 4) errors in tool calibration. Therefore, tuning the petrophysical endpoints is necessary but may also be a painstaking process before reasonable multimineral results can be obtained.

Assuming that the endpoints do not change in the interval of interest, we explore the model space of selected endpoints using MCMCES and the likelihood function $p(D|G)$ is now given by

$$p(D|G) \propto \exp\left[-\frac{1}{2}\sum(Gm-D)^T W^{-1}(Gm-D)\right], \quad (5)$$

where the summation is over depth. The key differences between equation 4 and 5 are: (1) the endpoint matrix (G) is not constant as selected endpoints perturb within reasonable ranges; (2) the volume fractions (m) are solved by a gradient-based optimizer at every depth and the misfits are then summed along depth to compute the errors. The results of equation 5 are the posterior distributions of selected endpoints.

We first benchmark our second application using a synthetic case. The synthetic model consists of anhydrite, dolomite, calcite, clay, water, and oil in a 100 ft interval. The simulated well logs (equation 1) are gamma ray, bulk density, neutron porosity, conductivity, acoustic slowness, and volumetric cross-section with additional 2% Gaussian noise. In the endpoint matrix, instead of constant values, we set up reasonable ranges with random distributions as prior information with 50 walkers for gamma ray (GR) and volumetric cross-section (U) endpoints of the solid constituents (anhydrite, dolomite, calcite, and clay) for a total 8 parameters. Figures 3a and 3b show the prior and posterior distributions of selected endpoints, respectively. The posterior distributions in Figure 3b correctly approximate the most probable endpoint values with uncertainties as compared with the true model parameters shown in black dashed lines.

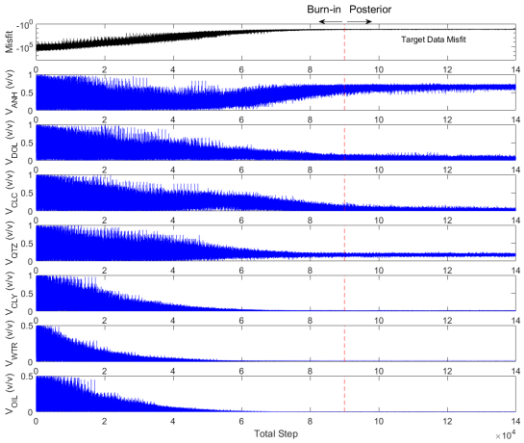

Figure 2. Trace plots at 11,624 ft after combining all 140 steps from 1,000 walkers. The top panel (black points) illustrates the change in data misfits along steps. Blue points are the perturbations of volume fractions of individual constituent along steps. The red dashed line defines the burn-in sections and the steps with the acceptable misfits for the posterior distributions.

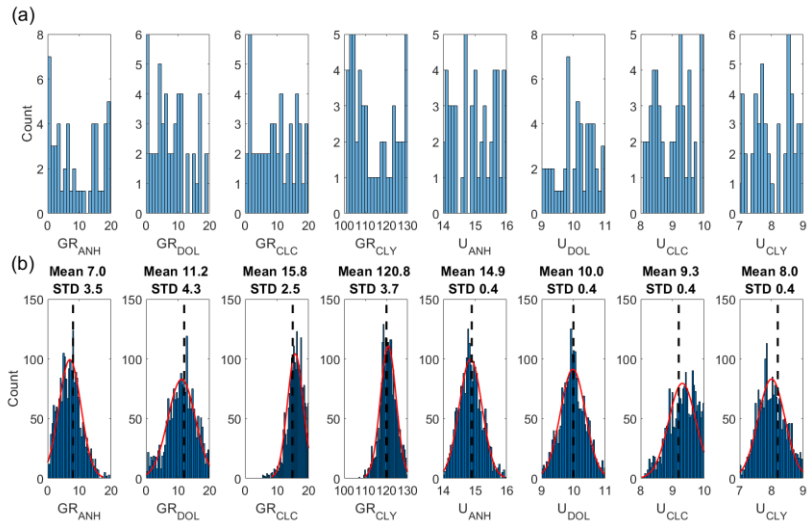

Figure 3. A synthetic model of MCMCES for selected endpoints in the matrix (*G*). (a) Top panels: random prior distributions of gamma ray (GR) and volumetric cross-section (U) of solid constituents (anhydrite, dolomite, calcite, and clay). (b) Bottom panels: posterior distributions of the corresponding endpoints. Black dashed lines indicate the true answer values.

We then apply the new method to a field dataset from the Bakken Shale Formation, where gamma ray endpoints of solid constituents generally are hard to estimate, especially for kerogen. The gamma ray log in this example exhibits extremely high readings (~ 1,000 API) at the Upper and the Lower Bakken intervals due to the high uranium content of kerogen in place. Typically, the optimal value of gamma ray for kerogen can only be tested in a trial-and-error fashion by manually checking the data misfits or core measurements. We implement the MCMCES for gamma ray endpoints of solid constituents (quartz, dolomite, calcite, illite, and kerogen) to estimate the optimal values as well as their uncertainties. Figures 4a and 4b show the prior and the posterior distributions of gamma ray endpoints after the simulation of 50 walkers with 200 steps. The gamma ray posterior distribution for kerogen displays a mean value of 3,979 API with a standard deviation of 106 API. The application can potentially be extended to all other endpoints of the matrix (*G*).

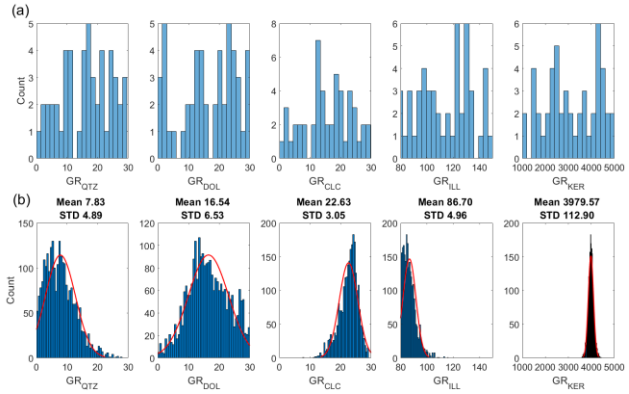

Figure 4. Gamma ray endpoints estimation for Bakken Shale Formation. (a) Top panels: random prior distributions for gamma ray of quartz, dolomite, calcite, and kerogen. (b) Bottom panels: posterior distributions.

Conclusions

To address the uncertainties in multimineral analysis, we have implemented the Markov chain Monte Carlo with ensemble samplers for posterior distributions in either the volume fractions or the petrophysical endpoints with only using lower and upper limits as prior information. By setting an acceptable data misfit, the new methods search the model space efficiently and provide plausible realizations to help the petrophysicist make proper interpretation decisions.

Acknowledgements

The authors would like to thank the sponsors of Reservoir Characterization Project (RCP) in the Geophysics Department of Colorado School of Mines and Patricia Rodrigues for finding the data of the Bakken Shale Formation.

REFERENCES

Deng, T., J. Ambia, and C. Torres-Verdin, 2019, Fast Bayesian inversion method for the generalized petrophysical and compositional interpretation of multiple well logs with uncertainty quantification: SPWLA 60th Annual Logging Symposium.

Deng, T., J. Ambia, and C. Torres-Verdin, 2020, Bayesian method for rapid multi-well interpretation of well logs and core data in unconventional formations: SPWLA 61st Annual Logging Symposium.

Goodman, J., and J. Weare, 2010, Ensemble samplers with affine invariance: Communications in Applied Mathematics and Computational Science, **5**, 65–80, doi: https://doi.org/10.2140/camcos.2010.5.65.

Mayer, C., and A. Sibbit, 1980, Global, a new approach to computer-processed log interpretation: Proceedings of the SPE Annual Technical Conference and Exhibition,.

Michelena, R. J., K. S. Godbey, M. J. Uland, and P. E. Rodrigues, 2020, Petrophysical multimineral analysis using genetic optimization to solve complex mineral composition in unconventional reservoirs: 90th Annual International Meeting, SEG, Expanded Abstracts, 2455–2463, doi: https://doi.org/10.1190/segam2020-3425780.1.

Mitchell, W. K., and R. J. Nelson, 1988, A practical approach to statistical log analysis: 29th SPWLA annual logging symposium.

Quirein, J., S. Kimminau, J. LaVigne, J. Singer, and F. Wendel, 1986, A coherent framework for developing and applying multiple formation evaluation models: 27th Annual Logging Symposium, Society of Petrophysicists and Well Log Analysts.

Spalburg, M. R., 2004, Bayesian uncertainty reduction for log evaluation: 11th Abu Dhabi International Petroleum Exhibition and Conference, SPE 88685.