AEGC2019 Data to Discovery
Australasian Exploration Geoscience Conference
2-5 September 2019 • Perth, Western Australia
Incorporating the AIG, ASEG, PESA, and WABS

Information Theory Based Probabilistic Machine Learning And Wireline Inversion: Surat Basin Case Study

Oliver Gaede
Queensland University of Technology
2 George St
Brisbane, QLD 4000
oliver.gaede@qut.edu.au

Florian Wellmann
RWTH Aachen University
Computational Geoscience and Reservoir Engineering.
Wüllnerstrasse 2
52062 Aachen, Germany
florian.wellmann@cgre.rwth-aachen.de

Thomas Flottmann
Origin Energy
Brisbane, QLD 4000
thomas.flottmann@originenergy.com.au

Klaus Regenauer-Lieb
University of New South Wales
Minerals and Energy Resources Engineering
Sydney, NSW 2052
klaus@unsw.edu.au

SUMMARY

We present a probabilistic machine learning approach to determine lithologies for the wireline data from the Springbok Sandstone formation in the Surat Basin. Deterministic inversions of the data are compared to the new machine learning approach in order to develop a generic method for wireline log inversions. The approach is designed to combine the benefits of classical deterministic inversion with new machine learning algorithms hence providing an approach to a *statistical learning* method. In this approach model parameters have a physical meaning and therefore can be readily related to petrophysical standards.

We start by establishing a prior distribution of minerals within the Springbok Sandstone Formation using geochemical laboratory data. The prior distributions are used in a Bayesian framework to improve wireline-based predictions of the mineral assemblages within the formation. The uncertainties of the predictions are quantified using the associated information entropy measure. This approach allows us to choose between various model mineral assemblages and apply an information theory augmented statistical learning approach.

Key words: Wireline Data, Probabilistic Method, Machine Learning, Information Entropy, Coal Seam Gas

INTRODUCTION

For lithological end members with simple mineral assemblages (e.g. clean sandstone), well-established and straightforward petrophysical methods are available to estimate porosity and permeability (e.g. Ellis & Singer, 2007) For more complicated lithologies the mineral assemblage of the formation has to be considered explicitly. State of the art methods for determining mineral assemblages are based on deterministic methods, in particular matrix inversion (e.g. Doveton, 2014). When the number of mineral and fluid components exceeds n+1 the number of available wireline measurements the inversion problem is underdetermined. This is not an uncommon situation where only triple- or quad-combo wireline suites are available.

One approach to alleviate this shortcoming is to employ advanced borehole geophysical logging tools, such as neutron-gamma ray spectroscopy logs (e.g. Maliva et al., 2009), that will increase the number of measurements. These tools are not always available due to practical considerations like the costs of running advanced logging tools or the impetus to use legacy data.

In our approach we make use of the data rich environment present for Australia's major onshore hydrocarbon basins and cast the inversion problem in a probabilistic framework to address (1) under-determination and (2) assess model uncertainties with an information entropy measure. We use the hetero-lithic Springbok Sandstone in the Surat Basin as an example for our probabilistic machine learning approach (see **Figure 1**). The Surat Basin of Central Queensland contains substantial hydrocarbon resources and includes major aquifers of the Great Artesian Basin. The late Jurassic fluvial sediments of the Surat Basin were deposited during slow subsidence rates and are dominated by channel sandstones and overbank mudstones and coal swamps (Exon, 1976; Green et al., 1997).

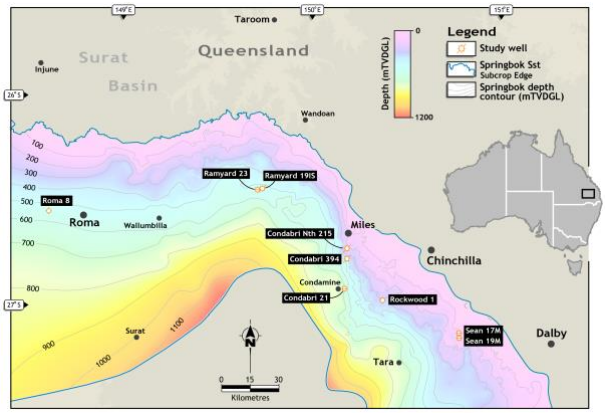

Figure 1 Map of study area and study well and the thickness of the Springbok Sandstone.

METHOD AND RESULTS

Deterministic Inversion

The deterministic inversion (and the associated forward model) is based on sound physical principles such as the petrophysical properties of possible solid rock components (e.g. minerals) and possible pore fluids. If the physical response of a sufficiently

large number of rock forming minerals is known for the suite of wireline tools used (e.g. the GR response for all major rock forming minerals for a specific tool is tabulated) then the generalisation of the deterministic inversion to arbitrary rock formation is possible. Further, "unphysical" model parameters can be identified given basic understanding of the system (i.e. response of give measurement platform to lithology).
As an example of a well determined system we will consider a hypothetical formation with quartz, feldspar, clay and coal as the mineral assemblage and water as the sole pore fluid. The available wireline logs in our example are natural gamma ray (GR), bulk density (RHOB), neutron porosity (NPHI) and sonic slowness (Δt). The log-response equations are simply defined as the sums of the log response of the individual formation components:

$$GR = \gamma_{Qz} \cdot Qz + \gamma_F \cdot F + \gamma_{Cl} \cdot Cl + \gamma_C \cdot C + \gamma_{fl} \cdot \phi$$

$$RHOB = \rho_{Qz} \cdot Qz + \rho_F \cdot F + \rho_{Cl} \cdot Cl + \rho_C \cdot C + \rho_{fl} \cdot \phi$$

$$NPHI = n_{Qz} \cdot Qz + n_F \cdot F + n_{Cl} \cdot Cl + n_C \cdot C + n_{fl} \cdot \phi$$

$$\Delta t = \Delta t_{Qz} \cdot Qz + \Delta t_F \cdot F + \Delta t_{Cl} \cdot Cl + \Delta t_C \cdot C + \Delta t_{fl} \cdot \phi$$

Where *Qz*, *F*, *Cl*, *C* and Φ are the component fractions for quartz, feldspar, clay, coal and water, respectively. Hence, we have four measurements and need to solve for five component fractions. The system is determined if the formation can indeed be modelled with these five components. In this instance, using the closure assumption can extend the set of equations:

$$1 = Qz + F + Cl + C + \phi$$

This set of equations can be expressed conveniently using a matrix algebra formulation:

$$\mathrm{L} = \mathrm{CV}$$

where *C* is the matrix of log responses of the individual minerals and water; extended by a line of unit values. *V* is a vector containing the component fractions for quartz, feldspar, clay, coal and water; again, extended by a unit value. *L* is then the vector of log measurements and a unit value. Hence the above equation is the forward model of the formation-log response. In order to obtain the unknown component fractions, the inverse of the C matrix, C^{-1} is used:

$$\mathrm{V} = C^{-1}L\,.$$

The application of this inversion is easy to implement with modern programming languages, although it does have significant shortcomings when the underlying assumptions are not met. Firstly, the mathematical inversion is likely to be underdetermined meaning that there is a greater number of unknown component fractions then log measurements. Further, the underlying closure model is subjective because the components of the mineral assemblage have to be pre-assumed and the relative proportions are then calculated by the petrophysical algorithm. This is particularly problematic when a mineral is present in the formation that was not included in the closure model. Finally, no convenient measure of the uncertainty of the prediction is obtained.

Classical Machine Learning Approaches

A possible solution to underdetermined inversion problems is the use of supervised machine learning algorithms (i.e. multi-layer artificial neural networks). The forward model is usually not based on physics and rather relies on a multi-layered parameter fitting. This is advantage over deterministic inversion algorithms, as the artificial neural networks are not model based and hence do not rely on pre-assumption of the mineral assemblage of the formation. At the same time, this is a drawback, as the fitting parameters cannot be physically interpreted. Further, over-fitting to the training data set is prevalent, which means caution needs to be exercised when generalising results from a particular well to other areas of a basin. Similarly, to the deterministic inversion, no convenient measure of the uncertainty of the prediction is obtained. As the classical machine learning approaches are not the focus of this contribution the reader is referred to following published application examples (e.g. Ma, 2011 and He et al., 2019).

Probabilistic Machine Learning Approaches

The probabilistic machine learning approach overcomes the issue of lack of uncertainty measure. However, since the nature of the algorithms usually relies on multi-layered parameter fitting without adhering to basic physics principles the interpretation of the uncertainty is not straight forward. Results can be interpreted in terms of randomness of pre-assumed mineral assemblages but without additional information no objective measure can be obtained for the likely prior probability.
The advantage of the method lies in the possibility to develop a hybrid between deterministic and classical machine learning algorithms opening the opportunity to alleviate the under-determination by incorporation of prior knowledge. Similar approaches have been used for reservoir property inversion (e.g. Gunning & Glinsky, 2007; Grana, 2018).

We propose an extension of probabilistic ML that includes information entropy as an objective and quantifiable uncertainty measure. We show here how information entropy can be used to incorporate prior knowledge and derive an objective measure of prior probability distributions.

Probabilistic Machine Learning Approaches with Information Entropy

The concept of information entropy was first defined and applied by Claude Shannon (1948), as a measure of uncertainty due to attenuation in phone-line signals. Information entropy encapsulates an overarching concept of thermodynamic entropy as the statistical thermodynamic entropy can be seen as just a particular application of Shannon's information entropy. Shannon defined the entropy H as:

$$\mathrm{H} = -\sum_{\mathrm{i}}^{\mathrm{N}} \mathrm{p_i} \log_{\mathrm{b}} \mathrm{p_i}$$

Where p_i is the probability mass function, or the information content of the *i-th* constituent. This stochastic source of data contributes information to the information entropy H at an average rate. Common values of the base of the logarithm b are 2 for a 1 bit system, Euler's number e and 10.

In our application we interpret the information entropy in a spatial context (with reference to a space vector $\boldsymbol{x}$) and perform a statistical sampling over infinite time (ignoring time dependence).

$$H(x) = -\sum_{i=1}^{n} p_i(\boldsymbol{x}) \log_b\left(p_i(\boldsymbol{x})\right)$$

We use the principle of maximum entropy for the most likely probability function. This principle can be illustrated by a classic example. When tossing a coin, the probabilities of coming up heads or numbers, is defined by a 1-bit system with two possible outcomes where each side should have a probability of 0.5 if sufficient tosses of coins were made. If the coin is fair the maximum information entropy on a $\log_2$-base is 1. A coin that has heads or numbers on both sides has an information entropy of 0. Values between 1 and 0 are expected for coins that have been tampered with (see **Figure *2***).

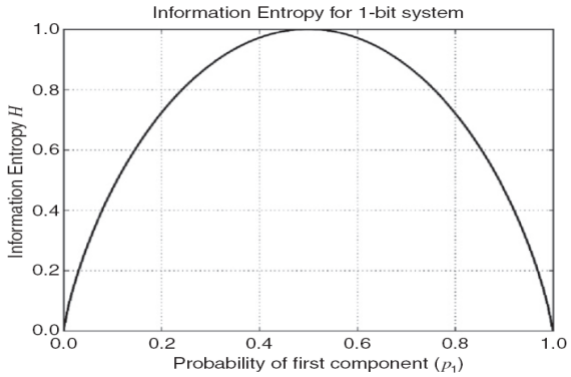

Figure 2 Information entropy for a fair coin has a maximum entropy of $H = -0.5 \log_2(0.5) - 0.5 \log_2(0.5) = 1$

When we have prior knowledge, such as knowing that the only minerals in the logs are Quartz and Feldspar, we have the equivalent of a coin flip and the distribution function is determined by the maximum entropy. If Quartz and Feldspar are present in the formation in equal proportions then the maximum entropy is 1 and lower values are expected for uneven distributions. The same principles can be extended to *n*-composition models and this gives the necessary extension to give meaning to uncertainties obtained from probabilistic inversions (Wellmann and Regenauer-Lieb, 2012).

In a system where no uncertainty exists about the number *n* of constituents the root base of the logarithm b can be adjusted to be equal to *n*; in our application it is the number of expected compositions to normalize the maximum information entropy to 1, as in the toss for an *n*-sided dice. In our definition of the information entropy we cannot, however, expect the same compositions at different spatial positions $\boldsymbol{x}$ and therefore this is not an appropriate choice. We can either select the maximum number of compositions as a base (to obtain information entropies smaller than one) or allow entropy to become larger than 1. In order to normalize the values back to below 1 we use Shannon's approach and multiply the information entropy by K, which is the inverse of the maximum entropy of the largest *n* found in any of the wireline data sets. This approach works for any log root b chosen and particular rational numbers can be identified corresponding to a given *n*.

$$H(x) = -K\sum_{i=1}^{n} p_i(\boldsymbol{x}) \log_b\left(p_i(\boldsymbol{x})\right)$$

Information entropy in our context identifies the uncertainty of the composition at a particular location $\boldsymbol{x}$. Locations where no uncertainty exists about the composition are identified by 0 values in the information entropy. These locations are very distinct in wireline logs and should be used as marker beds for the construction of a 3-D static model. Values above 0 identify areas where uncertainty about the composition exists which can be parameterised by the maximum entropy distribution function.

Above 0 the information entropy encodes the uncertainty. If 2 compositions are equally likely an information entropy of 1 would define the prior distribution function for the probabilistic inversion. For locations where 3 compositions are equally likely and using a $\log_2$ base (without renormalisation) the information entropy of 1.58 would define areas where 3 constituents are equiprobable.

We have so far discussed information that can be added to reduce uncertainty by using data from geochemical analyses of core samples, where the distribution functions are based on discrete random variables. This analysis could provide a first prior information used in a Bayesian analysis.

The next step is to incorporate information from the log-response equations. The principle for deriving maximum entropy distribution functions remains the same. However, in this case we are adding information from continuous random variables. This can be accommodated by using the definition of relative entropy, which acknowledges the discreteness of the compositions as counting measures and can incorporate their probability distributions.

In summary, the addition of prior distribution functions based on information theory provides the important link to incorporate prior information from chemistry and physics and merges deterministic and artificial intelligence-based approaches.

CONCLUSIONS

We have presented a method of probabilistic inversion combining petrophysical theory with machine learning algorithms. The method relies on (1) utilizing prior knowledge from laboratory analysis of available core to determine prior distribution functions, (2) using an information theory-augmented Bayesian inversion method for testing the petrophysical rules, (3) interpreting the uncertainty of the results with respect to petrophysical insight. The method therefore provides a significant improvement of the classical deterministic approach with probabilistic measures. The objective uncertainty measure therefore allows to assess whether knowledge learned from an analysis, as presented here, can be transferred to far offset wells. This could then be further developed into a database of basin-wide petrophysical models.

REFERENCES

Doveton, J. (2014). Principles of Mathematical Petrophysics. Oxford: Oxford University Press.

Ellis, D. V., & Singer, J. M. (2007). Well Logging for Earth Scientists (Vol. 692). Dordrecht: Springer.

Exon, N. F. (1976). Geology of the Surat Basin in Queensland (Vol. 166). Australian Government Publishing Service.

Grana, D. (2018). Joint facies and reservoir properties inversion *Geophysics*, *83*(3), M15–M24.

Green, P. M. (1997). Lithostratigraphic units in the Bowen and Surat basins, Queensland. In The Surat and Bowen Basins, Southeast Queensland. Queensland Minerals and Energy Review Series (pp. 41-108). Brisbane.

Gunning, J., & Glinsky, M. E. (2007). Detection of reservoir quality using Bayesian seismic inversion. *Geophysics*, *72*(3), R37–R49.

He, J., La Croix, A. D., Wang, J., Ding, W., & Underschultz, J. R. (2019). Using neural networks and the Markov Chain approach for facies analysis and prediction from well logs in the Precipice Sandstone and Evergreen Formation, Surat Basin, Australia. Marine and Petroleum Geology, 101, 410-427.

Ma, Y. Z. (2011). Lithofacies clustering using principal component analysis and neural network: applications to wireline logs. Mathematical Geosciences, 43(4), 401-419.

Maliva, R. G., Clayton, E. A., & Missimer, T. M. (2009). Application of advanced borehole geophysical logging to managed aquifer recharge investigations. Hydrogeology Journal, 17(6), 1547-1556.

Shannon, C. E. (1948). "A Mathematical Theory of Communication". Bell System Technical Journal. 27 (3): 379–423.

Wellmann, J. F., & Regenauer-Lieb, K. (2012). Uncertainties have a meaning: Information entropy as a quality measure for 3-D geological models. Tectonophysics, 526, 207-216.

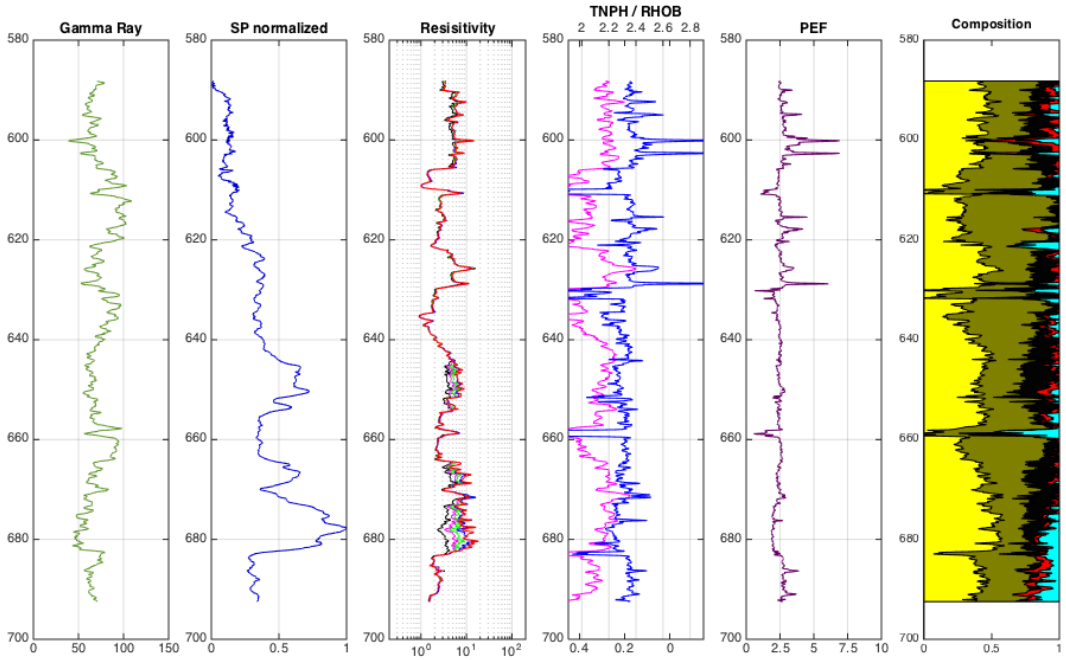

Figure 3 Example of deterministic inversion of formation components - Springbok Sandstone interval