Assessing Uncertainty in Rock Physics Interpretations: The Pitfalls of Ignoring Variability

Tapan Mukerji*, and Gary Mavko.
Stanford Rock Physics Laboratory, Stanford University, USA
E-mail: mukerji@stanford.edu

Summary

Ignoring the variability of rock properties in quantitative computations can cause critical errors in interpretation decisions. We show how calculations based on a single "best guess" input may not give the "best guess" output. Monte Carlo simulations, by taking in to account distributions of values instead of single average values, help to avoid the flaw of averages. Monte Carlo simulations also give us confidence intervals and other measures of uncertainty. Computations using averages and average trends alone do not give any indication of the uncertainty due to the variability in the properties. We also show how estimating the variability in attributes can be difficult from just visual examination of clusters in cross-plots. The density of points gets obscured in cross-plots. What appears to be the best choice in terms of visual separability of clusters may not be the best quantitative discriminants.

The flaw of averages

Ignoring the variability of rock properties in quantitative computations can lead to distortions in interpretations. Consider the following calculation as an example of the pitfall of ignoring distributions. We would like to compute the normal incidence reflectivity between an overlying shale layer and a packet of very thinly bedded, sub-seismic resolution sand/shale layers. The effective properties (velocity, density, and impedance) of the thin sand/shale layers are computed using Backus average (e.g. Mavko et al., 1998). The contrast between the impedance of the shale and the effective impedance of the sand/shale packet then gives us the normal incidence reflectivity. The Backus average elastic modulus depends on the volumetric sand fraction. For normal incidence propagation perpendicular to the layers, the effective P-wave modulus is given by a volumetric harmonic average of the sand and shale moduli. The effective density is given by the usual volumetric arithmetic average of sand and shale densities. By doing the calculations for all sand/shale ratios (0 to 1) we get a relation between sand ratio and normal incidence reflectivity. We could do the computation in two different ways. We could take an average value for sand V_p, density, and impedance, and another average value for the shale V_p, density, and impedance, and use them in the equations for Backus average and reflectivity. The average values could be estimated from blocked well logs, for example. This average computation, ignoring the variability, gives us the single line shown in Figure 1, relating the computed reflectivity to the sand/shale fraction. One might use this line for interpreting observed reflectivities to estimate sand/shale ratio in the thinly layered packet.

Instead of taking average sand and shale properties, another way to do the computation would be by Monte Carlo simulations, taking in to account the natural variability in the properties of the sand and shales (Avesth et al., 2005).

We draw a bivariate sample of sand (V_p, density) and another sample of shale (V_p, density) from the distributions observed in the well log. The simulated values are then used in the equations. A number of realizations are

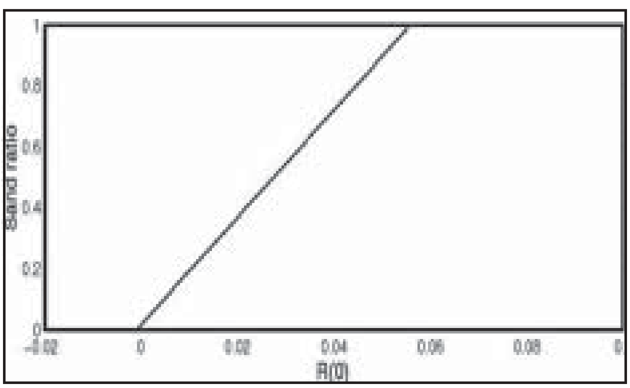

Fig. 1: Relation between normal-incidence reflectivity and sand/shale ratio in very thin-bedded sand-shale layers. The line is obtained using Backus average for thin layers, with blocked average values of sand and shale properties from well log. In this computation we ignored the variability of sand and shale velocity, density, and impedance.

drawn to get the full distribution of reflectivity for every sand fraction. In other words we bombard the equations with the whole range of possible inputs in accordance with their probability of occurrence. Now we see something very different (Figure 2). The average of the Monte Carlo simulations (blue curve) is not the same as the result from the average computation (black line). The black line would over predict considerably the sand ratio for large values of reflectivity, and slightly under predict the sand ratio at small values of reflectivity. The average computation also completely misses the negative branch of the relation between sand ratio and reflectivity. As we see from the distributions of the shale and sand impedances in Figure 3(a), there is considerable overlap, leading to a small finite chance of getting negative reflectivity. The histogram of reflectivity from the Monte Carlo Backus calculation is shown in Figure 3(b).

Monte Carlo simulations also give us confidence intervals, such as the 10 percentile and 90 percentile curves shown in Figure 2. Computations using averages alone do not give any indication of the uncertainty due to the variability in the properties. But the percentile curves obtained from Monte Carlo simulations are a measure of uncertainty in the relation, subject to the assumed model being correct. The example described above applies where identification of thin sand/shale packets is a practical problem.

The surprising outcome is an example of the flaw of averages (Savage, 2002, 2003), or equivalently, Jensen's

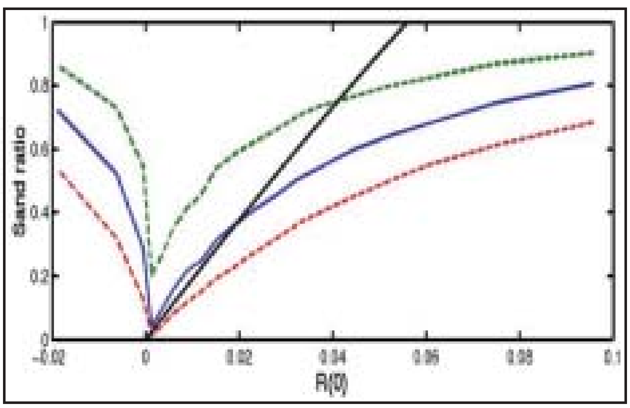

Fig. 2: Relation between normal-incidence reflectivity and sand/shale ratio in very thin bedded sand-shale layers The back line is obtained using Backus average for thin layers, with blocked average values of sand and shale properties from well log. Curves show the result of computations using Monte Carlo simulation to incorporate the variability of sand and shale velocity, density, and impedance. The blue curve is the mean of the distributions obtained from the simulations, and the red and green dotted curves are the 10 percentile and 90 percentile curves, respectively. We see clearly the pitfalls of ignoring variability.

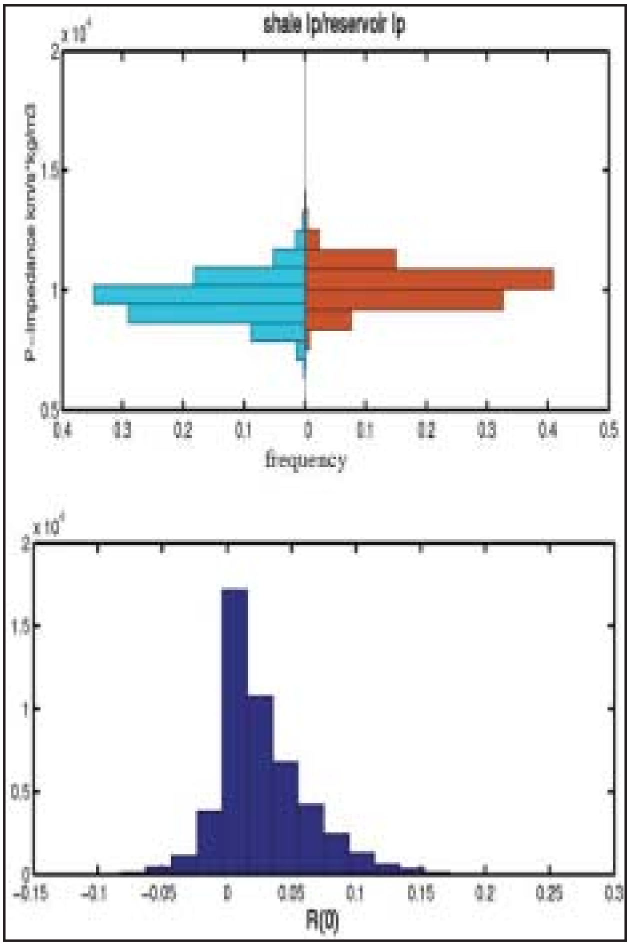

Fig. 3: a) Top: Distribution of shale and sand impedances used in the computation for Figure 2. On average the sand has slightly higher impedance than the shale, but there is considerable overlap. **b)** Bottom: Histogram of reflectivity from Monte Carlo simulations. Because of the overlap in shale and sand impedances, there is a small but finite chance of getting negative reflectivity values

inequality (e.g. Claerbout, 1992). One should not expect to get even correct average results using average values of inputs. Plugging in a single "best guess" input does not result in the "best guess" output. Mathematically, for a function g(x), in general, $\langle g(x)\rangle \neq g(\langle x\rangle)$, unless g(x) is linear. The symbol $\langle\ \rangle$ denotes the expectation operator. Other common non-linear rock physics models where one would have to be vary of Jensen's inequality include the Yin-Marion binary mixture model for shaly sands (e.g. Avseth et al., 2005) and of course the ubiquitous Gassmann model for low frequency fluid effects. When there is a lot of variability, calculations using single point values are almost worthless. Simple simulations to take in to account the variability can be easily performed using commonly available software such as Excel, Matlab, S, and others. Thereis hardly any excuse for falling prey to the flaw of averages.

Cross plot pitfall in discrimination and classification

Once we realize the importance of accounting for the natural variability in the rock properties the next step is to do Monte Carlo simulations of various seismic signatures. Often cross-plots of Monte Carlo simulated seismic attributes are used as an aid in seismic interpretation. For example, one might make a cross-plot of AVO attributes R(0) and G (Figure 4), where R(0) is the intercept or normal-incidence reflectivity and G is the AVO gradient. Other cross-plots may include acoustic impedance vs. elastic impedance (AI-EI) cross plot, LMR cross-plots (λρ–μρ, or λ–μ), etc.

From the scatter-plot of different clusters (say oil sands vs. background, red cluster vs. black cluster, etc.) we try to answer questions such as a) whether the clusters are well separated in a particular cross-plot domain, b) which pair of seismic attributes would be most useful in separating the clusters, and (c) what cut-off values should be used for discriminating the clusters. Bivariate cross plots are definitely an improvement over trying to find a single linear combination of the two attributes for discriminating lithologies or pore fluids. A cross-plot of R(0)-G is better than using only a linear combination of R(0) and G. However, a pitfall of the cross-plot is that it obscures the third dimension – the *density* of the points on the scatter-plot.

Figures 5 and 6 show two cross plots in each case with two clusters. Visually it is apparent that there is some separation between the clusters in Figure 5, while in Figure 6 the situation is almost hopeless because the clusters almost completely overlap. Or do they?

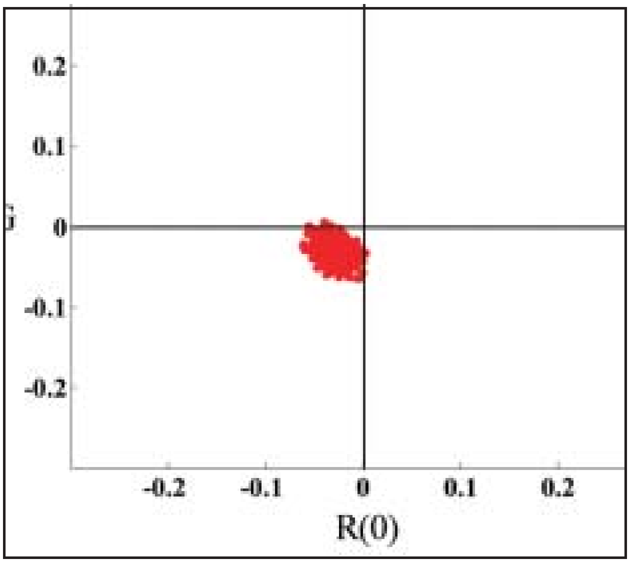

Fig. 4: Monte-Carlo simulated values of R(0) and G plotted on a scatter plot.

To quantify the separation of the clusters we do a discriminant analysis (Hastie et al., 2001) taking in to account the bivariate density of the clusters as well as their covariance matrix. Using either Bayesian classification or quadratic discriminant analysis (QDA) the misclassification error is about 12% for Figure 5, while the misclassification is only about 4% for Figure 6. In other words *the two clusters in*

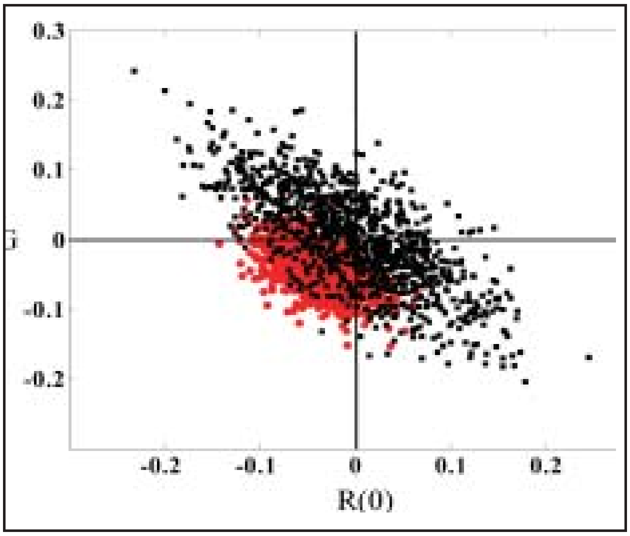

Fig.5: Monte-Carlo simulated synthetic R(0)-G cross plot showing separation between two clusters; background in black and oil sands in red.

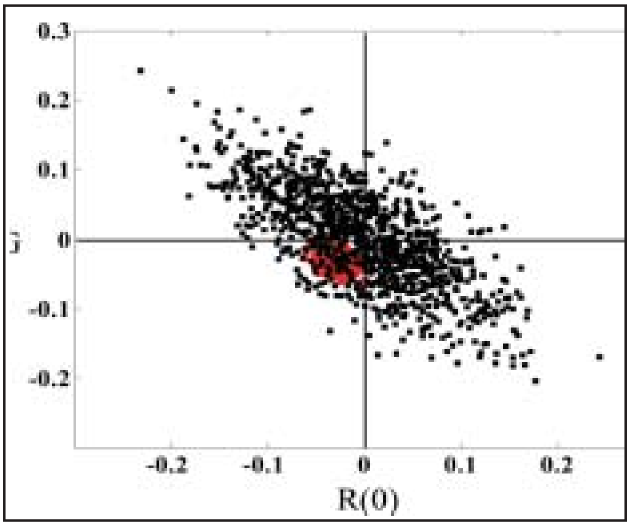

Fig. 6: Monte-Carlo simulated synthetic R(0)-G cross plot showing considerable overlap between two clusters; background in black and oil-sands in red.

Figure 6 are much better separable than the two clusters in Figure 5. If we use simple linear discriminant analysis (LDA) the clusters in both figures are equally well separated with a mis-classification error of about 15%. Visually trying to estimate separation of clusters in cross-plots may lead to erroneous interpretations. Quantitative discrimination between clusters depends not only on the classification method (e.g. QDA, LDA, Bayesian etc.) but also on the bivariate density of the points and their variance-covariance structure (Figure 7).

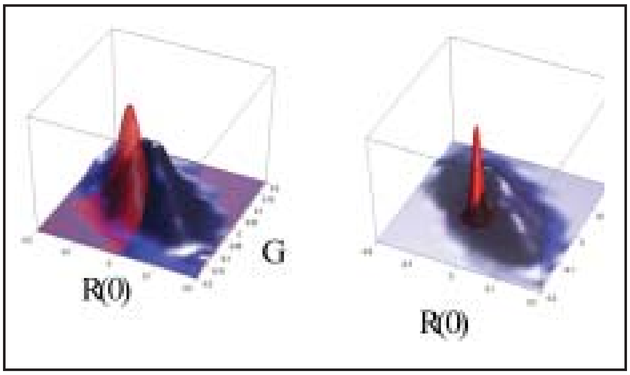

Fig. 7: Bivariate probability density functions corresponding to Figures 5 (left pdfs) and 6 (right pdfs). The volume of overlap between the pdfs of the two clusters is much smaller for the pdfs on the right. Therefore, contrary to visual appearance, the clusters in Figure 6 are quantitatively about three times better separable than the clusters in Figure 5.

Conclusions

Monte Carlo simulations, by taking in to account the full distributions of values instead of single average values, help to avoid the flaw of averages. Calculations based on a single "best guess" input may not give the "best guess" output. An estimate of a best "average" value should be derived from a simulation not from a point estimate. One should be vary of visually estimating discrimination of clusters in cross-plots. The density of points gets obscured in cross-plots. What appears to be the best choice of attributes in terms of visual separability of clusters may not be the best discriminants.

Acknowledgments

We thank the Stanford Rock Physics Project (SRB) and its affiliates.

References

Avseth, P., Mukerji, T., and Mavko, G., 2005, *Quantitative Seismic Interpretation: Applying Rock Physics Tools to Reduce Interpretation Risk*, Cambridge University Press, 359pp.

Claerbout, J. F., 1992, *Earth Sounding Analysis: Processing versus Inversion*, Blackwell Scientific Publications, Boston, 304pp.

Hastie, T., Tibshirani, R., and Freidman, J., 2001, *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*, Springer-Verlag, New York, 533p.

Mavko, G., Mukerji, T., and Dvorkin, J., 1998, *The Rock Physics Handbook*, Cambridge University Press.

Savage, S. L., 2002, The Flaw of Averages, Harvard Business Review, November, 20-21.

Savage, S. L., 2003, *Decision Making with Insight*, Brooks/Cole Publ. Co., Belmont, 360p.