Uncertainty, Subjectivity and Bias*

Andrew Curtis[1] and Matthew Walker[1]

Search and Discovery Article #120048 (2012)*
Posted December 31, 2012

*Adapted from extended abstract prepared in conjunction with poster presentation at AAPG Hedberg Conference, Fundamental Controls on Flow in Carbonates, July 8-13, 2012, Saint-Cyr Sur Mer, Provence, France, AAPG©2012

[1]University of Edinburgh, U.K. (Andrew.Curtis@ed.ac.uk)

Abstract

Lack of detailed knowledge about heterogeneities in subsurface carbonates renders the injection of expert-derived information important, yet uncertainties are seldom analyzed or quantified. We discuss the sources of expert uncertainty, and show how expert knowledge may be elicited to encapsulate uncertainties and reveal natural human bias.

Compared to many other areas of industrial endeavour, geologically related industrial problems are often characterised by large uncertainties. Success rates of exploration wells intersecting hydrocarbon reserves are generally lower than 60% in the USA (data: PetroStrategies, Inc.) showing that uncertainties in prospect evaluation are significant. Likewise, uncertainties in reservoir characterisation for industrial storage of CO_2 have been shown to be an order of magnitude higher than are acceptable in conventional business practice for the power companies who wish to store the CO_2 (Polson et al. 2011, 2012). Because these uncertainties stem from the lack of suitable physics to interrogate the Earth's subsurface at resolutions comparable to the reservoir heterogeneity, all methods applied to geological problems must account for intrinsic uncertainty if they are to provide robust results.

Introduction

Assessing the prospectivity of carbonate reservoirs is particular challenging due in large part to the large range of scales of variability (<1 μm to ~1 km) in pore space characteristics that contribute to flow properties of single or multi-phase fluids. Each of the individual geophysical, geological, well testing, well log and core analysis methods used to assess such heterogeneity is only sensitive to a relatively small range of scales, and homogenising or even calibrating information provided across different scale ranges is extremely difficult. Between wells, the range of methods available to interrogate reservoir properties directly reduces to geophysical remote sensing, well testing, and (post-reservoir development) production history matching techniques, all of which provide information only (significantly) above length scales of ~1 m. This results in an information scale gap illustrated in Figure 1, and huge uncertainties in properties across the majority of any subsurface carbonate reservoir. The only existing way to inject information into this gap is to use Geological Prior Information and interpretation (Curtis and Wood, 2004a; Wood and Curtis, 2004).

Discussion

Key to assessing and ultimately reducing these uncertainties is to understand their origins, and in this paper, we focus on geological uncertainties. These derive from a number of factors, but usually substantially from variations in hypotheses or interpretations of individual, or groups of Geologists, Geophysicists and Georeservoir Engineers (henceforth, Geologists). Such experts derive their knowledge from a wide range of experience, and this information is crucial to be able to fill in the information scale gap between wells. However, recent studies have shown that such experts are subject to well-known cognitive and motivational biases. Since the level of various biases varies with the individual, expert opinion is rendered subjective. Some biases reduce, while others increase in prevalence as more experts are added, hence such subjectivity cannot necessarily be removed simply by ‘averaging’ over experts (Polson and Curtis 2010).

We study methods to retrieve the information and experience of experts in a quantitative and probabilistic form, that is both useful for uncertainty and risk analysis for subsurface reservoirs, and which is as unbiased as possible.

Subjective biases in Geologists are expected: processes through which experts in any field of endeavour develop such biases are understood by cognitive psychologists. Particularly prevalent biases include over-confidence, anchoring and adjustment, availability, and motivational bias, and their definitions can be found in Kahneman et al., (1982) or O’Hagan et al., (2006). All such biases occur in situations of uncertainty (such as when forming geological hypotheses); in such situations, all humans employ heuristics (rules of thumb) subconsciously.

Allowing subjectivity within reservoir evaluation and characterisation is a positive aspect of the scientific method: it allows for leaps of faith, which occasionally lead to new proposals that prove to be valid. However, the effects of biases on expert judgment should also be subject to scientific investigation, and quantified where possible to reduce the possibility of biasing overall uncertainty and risk analysis (Curtis 2012).

Some scientific studies have analysed how subjectivity contributes to the progression of ideas, and some of those studies are in the Geological sciences (Aspinall 2010). Bond et al. (2012) showed a computer-generated seismic cross-section created from an underlying (invented) geological model to several hundred individual geologists. The model included structural deformation, inversion of faults, with pre-, syn- and post-deformational stratigraphic development. Each geologist interpreted the cross-section to hypothesize a geological model; they also provided information about their academic and professional background. Concepts employed by each geologist were categorised (e.g., as dominantly diapirism, thrusting, extension, inversion, etc.) and analysed statistically. Importantly, geologists' background and experience correlated significantly with their likelihood of having invoked the correct concepts. Those with M. Sc. or Ph. D. degrees were most likely to make a successful interpretation. Analysing techniques employed (e.g., feature identification, horizon picking, annotation, evolutionary sketches), successful interpretations were most often obtained from using multiple techniques, particularly if they included evolutionary sketches: academic staff were notably successful because they tended to use multiple techniques. Thus, variations in prior experience are shown to bias the formation of evidence-based geological hypotheses.

Bond et al. (2012) use large numbers of geologists to identify such biases, which is not usually practical in interpretational settings. An alternative approach to analysing bias is to use the theory of Elicitation – how to interrogate people in a manner designed to obtain the most reliable information (O'Hagan et al., 2006). Elicitation theory is cross-disciplinary, combining elements of statistics, cognitive psychology, and the field under investigation (here, Geology). Structured elicitation methods, and even real-time optimisation of questions posed during elicitation, have been used to assess and reduce uncertainty and bias in expert opinions (e.g., Rankey and Mitchell, 2003; Curtis and Wood, 2004b; Polson et al. 2009; Aspinall 2010). In all cases, a facilitator manages the process of elicitation, and the entire system of facilitator, experts and information flow may be analysed using probabilistic statistical techniques (e.g., Lindley et al., 1979).

To reduce biases associated with individual experts, information is often elicited from groups of experts simultaneously. However, groups of experts are subject to additional biases caused by social influence, resulting in convergence, divergence or herding behaviour (Kahneman et al., 1982; Baddeley et al., 2004). Polson and Curtis (2010) conducted an experiment representing an asset-team environment in the hydrocarbon industry, in which a range of experts was asked to assess the potential of a prospective reservoir stratum (in their case for CO_2 storage). Four experts were asked to interpret existing geological and geophysical data to assess the likelihood of existence of a particular fault, a specific reservoir stratum, and a sealing cap-rock. The experts' individual levels of certainty were quantified three times: days before the group meeting, near the beginning of the meeting, and approximately five minutes after the end of the meeting. During the meeting, the geologists were asked to reach a consensus position on their joint level of certainty through reasoned discussion.

Figure 2 shows the range of group-averaged individual experts' uncertainties in whether each of a potential reservoir (A), seal (B), and fault (C) exist, and the respective group consensus positions. Expert opinion changed significantly during the process, even in the absence of new information. For case C, the group consensus position combined both the extreme position and the highest degree of certainty (the narrowest range). The final positions shown in Figure 2 were obtained ten minutes after the final consensus position had been agreed, and at this point one particular geologist was even shown to disagree with the consensus to which he had just agreed. In this case, there is a clear lack of objectivity in hypothesis formation due to group dynamics. However, subjectivity is also shown to be important: the consensus position in case C was not adopted by any geologist before the meeting – without group dynamics, it might not have been considered.

Similar dynamism of opinion was observed in a study by Phillips (1999): two sets of 10 experts estimated corrosion timescales of containers to be used for geological storage of nuclear waste. Even accounting for the experts' own uncertainty estimates, final results from the two groups were barely consistent.

A portable computerised laboratory has been developed at the University of Edinburgh to explore these phenomena. This laboratory can be used in many debate, discussion or board-meeting type forums to track dynamic opinions of individuals throughout group elicitation or decision-making sessions. All studies to-date have exhibited similar dynamism in opinions of geoscientific experts during discursive information exchange.

The above studies significantly influence the way one should interpret consensus-driven results. Consensus positions clearly only represent the group opinion at one instant in time, and may not represent the true range of uncertainty about the issue at hand (e.g., Figure 2C). This is disturbing because consensus is often used in the Geosciences, and in particular (within asset teams) to make decisions about reservoir development.

It is interesting to note that Bayesian methods now enjoy widespread acceptance within the geological and other sciences (Tarantola 2005). These inference methods are objective, being governed by mathematical laws, yet they explicitly represent (often subjective) prior information. From the prior position, inferences are formed by assimilating new data in a quantitative, probabilistic manner. What is seldom acknowledged is that mathematically the prior information is given identical weight to information present in any newly available data (Figure 3 - top): distributions representing these two states of information (the prior probability distribution and the so-called likelihood function, respectively) are simply multiplied together to obtain the final (or posterior) state of combined information. While much time and money has been and continues to be invested into improvements in data acquisition, processing and interpretation (i.e., into the likelihood function), relatively little is invested into improving the quality of how we obtain and parameterise prior information. This imbalance in investment seems paradoxical given that in many situations the prior information is mathematically orthogonal to information in the likelihood function, potentially resulting in significantly reduced overall uncertainties (Figure 3 - bottom). Since relatively little has been invested in this to-date, these gains may also be relatively cheap to achieve.

Some non-Bayesian methods also allow, for example, qualitative geologists to influence quantitative inferences about the Earth by interacting intuitively with the optimisation of Earth model parameter values (Boschetti and Moresi, 2001; Curtis and Wood 2004b). Thus, a range of methods explicitly facilitates the use of subjective geological information within otherwise objective scientific inference.

As an example, Figure 4 shows three results from an experiment designed to assess the intrinsic human uncertainty (in this case, that of a Geologist) in determining the spatial multi-point statistics that best represent the character of a rock pore space observed on a thin-section. In this experiment, artificial binarised thin-sections (referred to as target images) are produced using a geostatistical model. As such, the target images were defined by a set of statistics called transition probabilities. A Geologist is asked to estimate these transition probabilities by finding an image, produced using the same geostatistical model, which had the same texture as the target image. To do this, a particular optimisation method called a genetic algorithm (GA) is employed. This algorithm iteratively generates a 'population' of six sets of transition probabilities and six corresponding artificial thin section images. The GA then prompts the Geologist to rank each image (and hence, each set of transition probabilities) in the current population by the similarity of its texture to that of the target image. This is sufficient information for the GA to update each set of transition probabilities in the population such that the next population of images generated is (usually) improved relative to the previous iteration.

Usually a genetic algorithm is used to perform optimisations where a particular quantitative objective function is to be minimised: in this problem, the objective function would normally quantitatively measure aspects of the difference in character between the current population of images and the target image. By minimising this objective function, the population of images would converge to have similar characteristics to the target image (and hence such convergence may correspond to convergence of the model parameters). In our case, however, the objective function is replaced by the Geologist's subjective opinion about the extent of these differences (henceforth referred to as a subjective function); by interacting with the algorithm their subjective function is minimised (see, e.g., Boschetti and Moresi 2001).

The Geologist continues to interact with (iterate) the GA until they judge that at least one member of the population of images has a texture indistinguishable to the target image, at which point they stop. This RMS measure is not visible to the Geologist during the experiment. Nevertheless, the RMS difference value at the final iteration allows us to quantify the uncertainty that is intrinsic to the Geologist's ability to discriminate between different spatial statistics (those of the population, and those of the target image) given that the Geologist has determined that the final population contains an image with the same texture as that of the target.

Conclusions

The geostatistical parameters used here (transition probabilities) are often used to describe Geological analogue information on a variety of length scales ranging from pore to reservoir scales (Wu et al., 2006; Elfeki and Dekking, 2001). Usually such statistics are simply those of explicit digital analogues when these are available, but often they are obtained from analogues that are constructed artificially from geological process or object-based models (Zhang, 2008). In the latter case, the above method would provide an indication of the intrinsic uncertainty in the geologists' ability to discriminate a good artificial analogue to use. This uncertainty is subjective (individual-dependent), but may thus be assessed using quantitative methods.

While Polson and Curtis (2009) and Bond et al. (2012) show clearly that subjectivity affects Geologists' interpretations either as individuals or in groups, the existence of subjectivity in forming hypotheses does not necessarily imply a lack of scientific rigor. When recognised explicitly, subjectivity may properly influence inferences, and can lead to novel hypotheses. Geoscientists should therefore not be ashamed of subjectivity, but we should strive to develop methods to quantify, and sometimes to reduce its effects.

References

Aspinall, W., 2010, A route to more tractable expert advice: Nature, v. 463, p. 294-295.

Baddeley, M.A., A. Curtis, and R.A. Wood, 2004, An introduction to prior information derived from probabilistic judgments; elicitation of knowledge, cognitive bias and herding, Geological Prior Information: Informing Science and Engineering: Royal Geological Society, London, Special Publications, v. 239, p. 15–27.

Bond, C.E., Lunn, R.J., Shipton, Z.K., and Lunn, A.D., 2012, What makes an expert effective at interpreting seismic images?: Geology, v. 40, p. 75–78.

Boschetti, F., and L. Moresi, 2001, Interactive inversion in geosciences: Geophysics, v. 66/4, p. 1226-1234

Curtis, A., 2012, The science of subjectivity: Geology, v. 40, p. 95-96.

Curtis, A., and R. Wood, 2004, Optimal elicitation of probabilistic information from experts: Geological Society, London, Special Publications, v. 239, p. 127-145.

Descartes, R., 1637, Discourse on the method of rightly conducting the reason and seeking for truth in the sciences, *in* R. Descartes (1596-1650) Descartes: Selected philosophical writings: English translation, republished 1988, Cambridge University Press, Oxford, England, 249 p.

Elfeki, A.M.M., and F.M. Dekking, 2001, A Markov chain model for subsurface characterization: Theory and applications: Mathematical Geology, v. 33/5, p. 569-589.

Kahneman, D., P. Slovic, and A. Tversky, 1982, Judgment under Uncertainty: Heuristics and biases: Cambridge University Press, New York, 555 p.

Lindley, D V., A. Tversky, and R.V. Brown, 1979, On the reconciliation of probability assessments: Journal of the Royal Statistical Society, v. 142, p. 146-180.

O'Hagan, A., C.E. Buck, A. Daneshkhah, J.R. Eiser, P.H. Garthwaite, D.J. Jenkinson, J.E. Oakley, and T. Rakow, 2006, Uncertain Judgments: Eliciting Experts' Probabilities: Wiley, Hoboken, N.J., 338 p.

Phillips, L.D., 1999, Group elicitation of probability distributions: are many heads better than one?, *in* J. Shantaeu, B. Mellors, and D. Schum, (eds.), Decision Science and Technology: Reflections on the Contributions of Ward Edwards: Kluwer Academic Publishers, Norwell, MA., p. 313-330.

Polson D., and A. Curtis, 2010, Dynamics of uncertainty in geological interpretation: Journal of the Geological Society, London, v. 167, p. 5-10.

Polson D., A. Curtis, C. Vivalda, and S. Saunier, 2009, Process for tracking the evolving perception of risk during CO2 storage projects: SPE 124703, 13 p

Polson D., A. Curtis, and C. Vivalda, 2011, Risk and Uncertainty, *in* M. Smith, D. Campbell, E. Mackay, and D. Polson, CO_2 Aquifer Storage Site Evaluation and Monitoring: Heriot-Watt University, Edinburgh, Chapter 6, p. 121-140.

Polson D., A. Curtis, and C. Vivalda, 2012, The evolving perception of risk during reservoir evaluation projects for geological storage of CO_2: International Journal Greenhouse Gas Control, v. 9, p. 10-23.

Rankey, E.C., and J.C. Mitchell, 2003, That's why it's called interpretation: Impact of horizon uncertainty on seismic attribute analysis: The Leading Edge, v. 22, p. 820-828.

Tarantola, A., 2005, Inverse Problem Theory and Methods for Model Parameter Estimation: Society for Industrial and Applied Mathematics, Philadelphia, Pa., 352 p.

Wood, R., and A. Curtis, 2004, Solving Geoscientific Problems Using Geological Prior Information, *in* A. Curtis, and R. Wood, (eds.), Geological Prior Information: Geological Society of London, Special Publication, v. 239, p. 1-14.

Wu, K. , V. Dijke, I.J. Marinus, G.D. Couples, Z. Jiang, J. Ma, K.S. Sorbie, J. Crawford, I. Young, and X. Zhang, 2006, 3D stochastic modeling of heterogeneous porous media; applications to reservoir rocks: Transport in Porous Media, v. 65/3, p. 443-467.

Zhang, T., 2008, Incorporating geological conceptual models and interpretations into reservoir modeling using multiple-point geostatistics: Earth Science Frontiers, v. 15/1, p. 26-35.

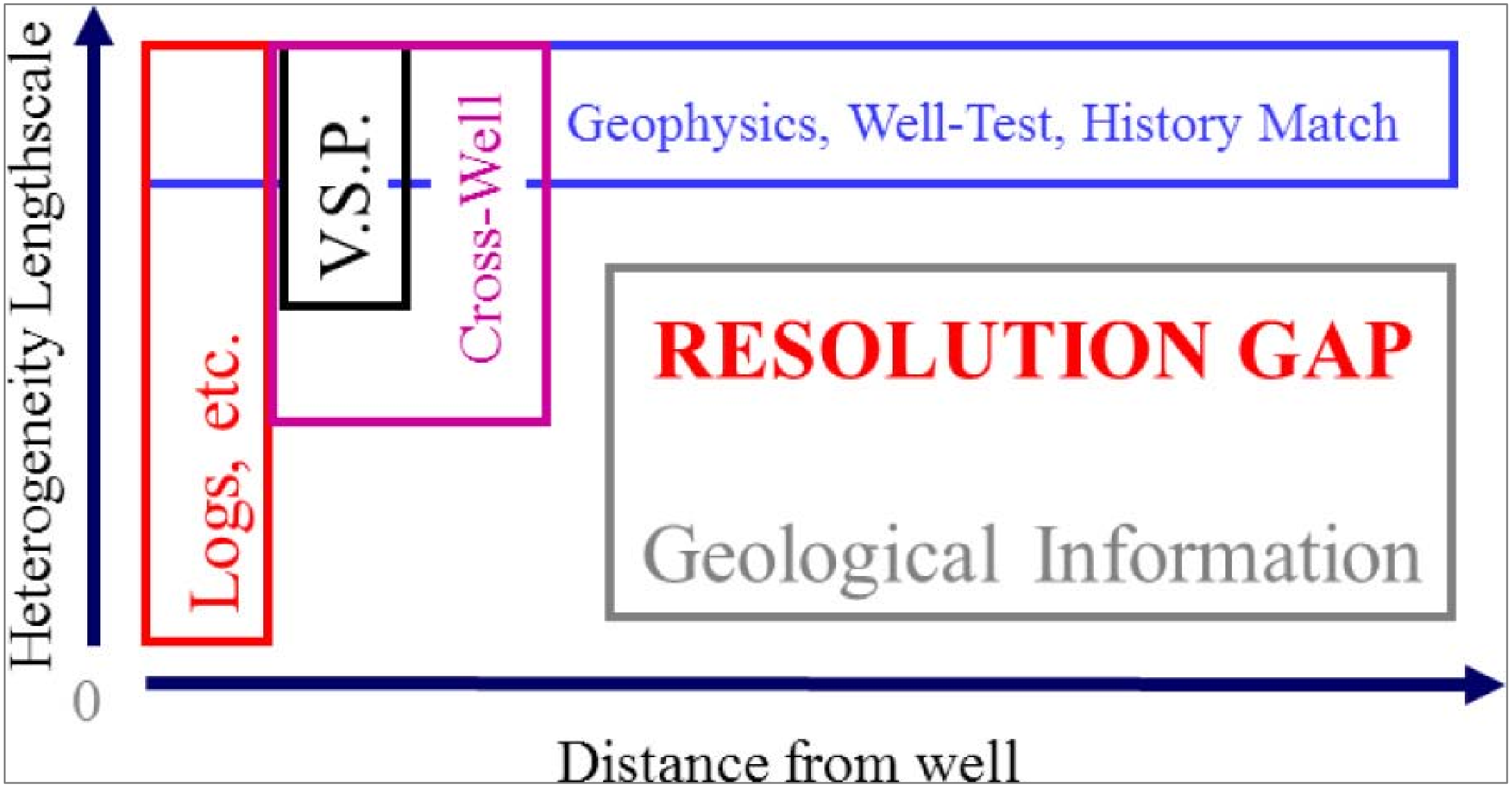

Figure 1. Techniques used to derive information of different length-scales as a function of distance from wells. The “Resolution Gap” is at small length-scales far from wells; must be filled by Geological information.

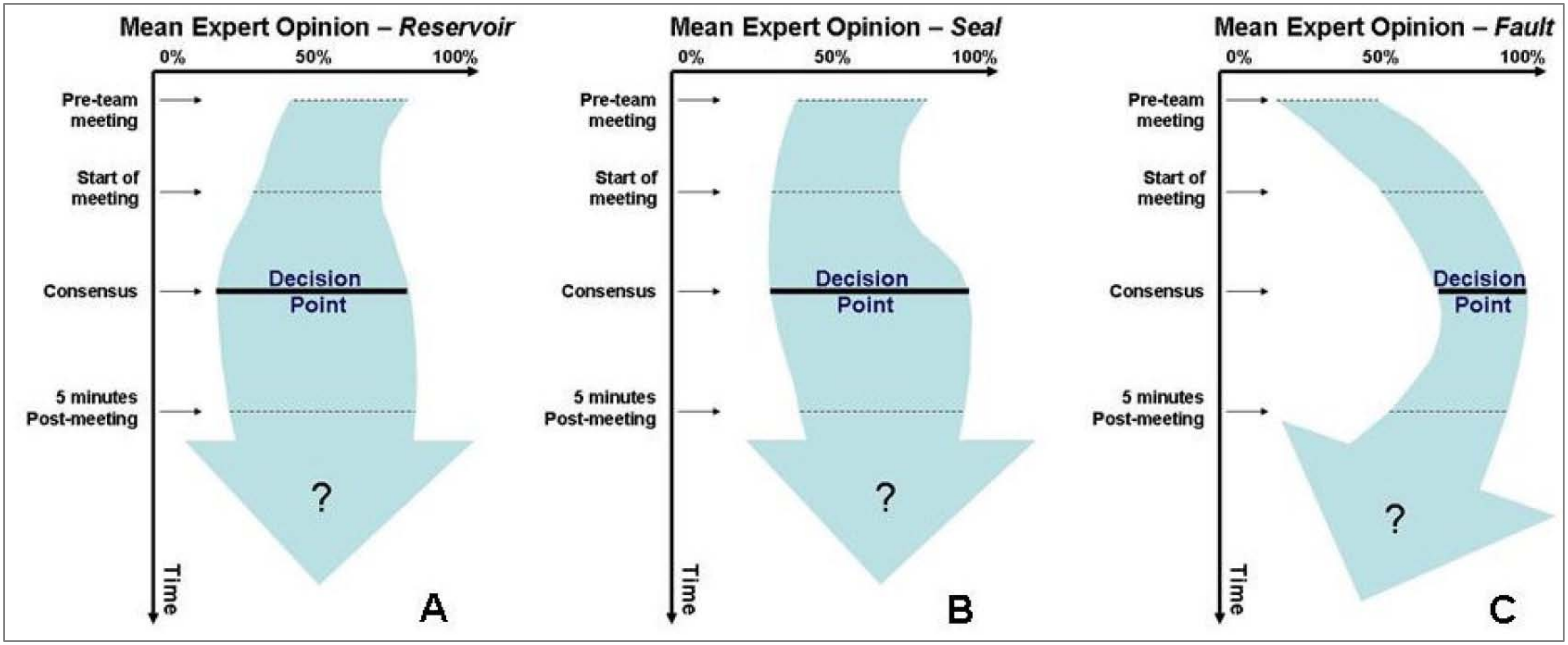

Figure 2. Evolution of expert opinion during the structured group elicitation process of Polson and Curtis (2010). Horizontal axis: estimated probability of the existence of a specific reservoir (A), cap rock (B), and fault (C). Vertical axis shows expert opinion at four points in time. Thin-dashed lines show average range of experts' opinions. Bold-solid line shows the group consensus on the range of probabilities, representing the decision point in a usual committee of experts.

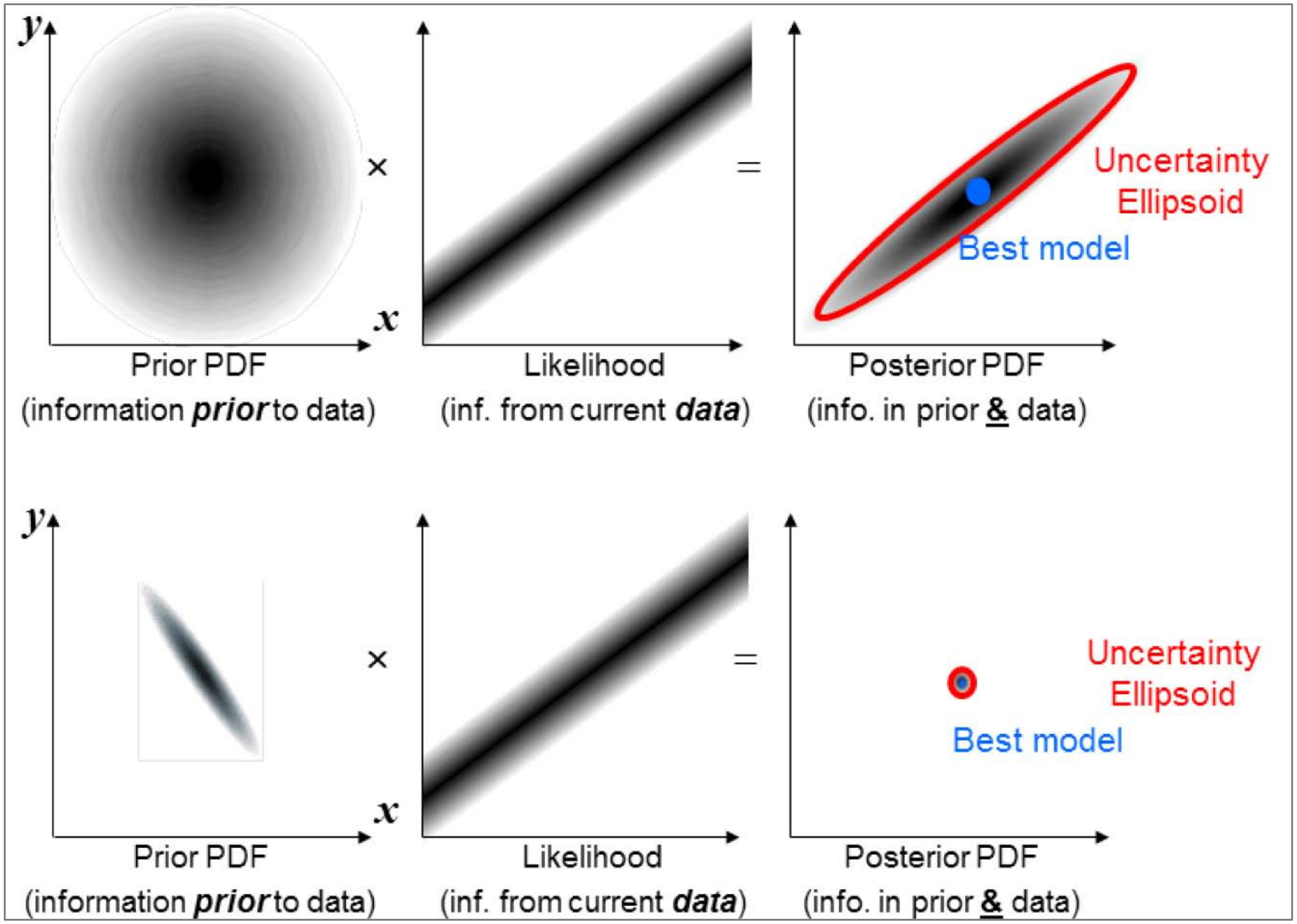

Figure 3. Top: Application of Bayes Rule to combine prior probabilities (e.g., Geological information) with likelihood functions (e.g., new Geophysical data). The result is the 'posterior' distribution obtained by multiplication. Bottom: same, but with improved prior information → improvement in posterior information.

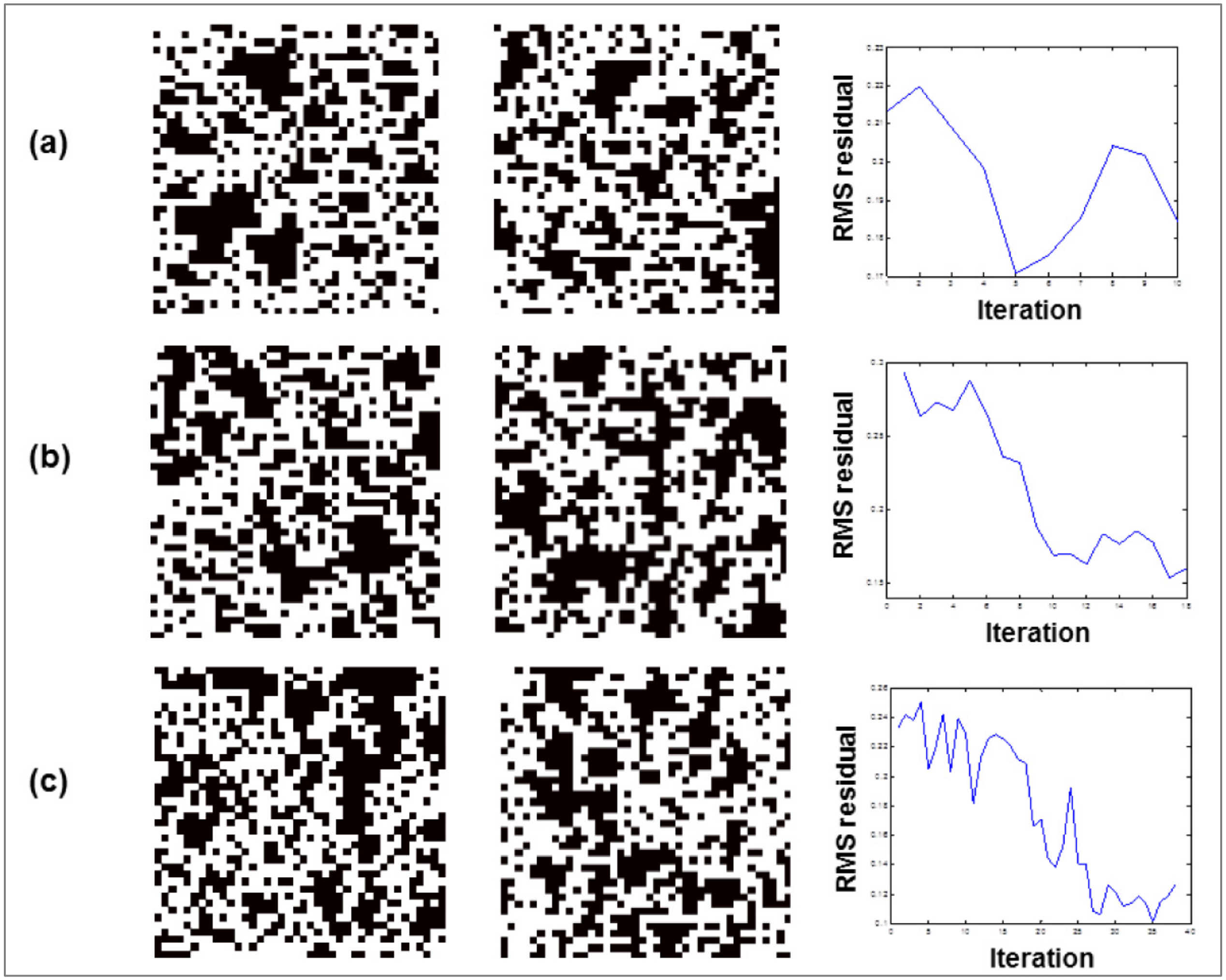

Figure 4. Rows (a), (b) and (c) each illustrate individual results of an experiment designed to quantify the intrinsic uncertainty of a Geologist determining the statistics used to model a rock pore space. The target image in each case is shown in the central pane in each row. The final image, which the Geologist deems to have the same texture as the target image, is given in the left pane of each row. The right pane in each case shows the root mean square (RMS) difference between the transition probabilities describing the statistics of the target image and the final image, throughout each trial.