SPE 68591

Problems Reconciling Probabilistic and Deterministic Reserve Classifications and Evaluations

Dwayne C. Purvis, SPE, and Richard F. Strickland, SPE, Cawley, Gillespie & Associates

This paper was prepared for presentation at the SPE Hydrocarbon Economics and Evaluation Symposium held in Dallas, Texas, 2–3 April 2001.

Abstract
The current SPE/WPC definitions for reserves quantified by probabilistic methods are ambiguous and largely inconsistent with deterministic criteria. Probabilities are least well defined and most subjective at the extremes of the ranges of outcomes. Since reserve definitions are meant to provide consistency across evaluators and evaluations, the current definitions are not workable. When defining reserves, volumes should first be classified into a category. A probabilistic analysis can then be made of the uncertainty in each category, and a central estimate of Proved reported. When probabilistic analyses are used to determine value, the same restraints are not necessary, but the same wariness of the extremes of the distributions should be applied.

Introduction
The debate on whether and how to include probabilistic criteria in reserve definitions has been going on for many years. Despite the inclusion of criteria in the 1997 version of the SPE/WPC reserve definitions (ref. 1), the meanings of the definitions remain unclear.

The definition of Proved is ambiguous as to whether the quantification of the volume should be 'estimated with reasonable certainty' or whether the presence of some amount of volume should be reasonably certain. Common deterministic practice is to offer a best, or central estimate, of the quantity of hydrocarbons after the producibility of some quantity has been established with reasonable certainty by largely non-quantitative criteria.

However, the probabilistic terms in the definition of Proved suggest that the specific quantity of Proved should be reasonably certain. The definitions neglect to define what types of volumes that are to be included in a probabilistic analysis from which the P10[1] value is chosen, i.e. how to apply non-quantitative definitional criteria to a probabilistic analysis. It also fails to specify the level of aggregation.

The following discussion explores problems with probabilistic analyses that can lead to artificial discrepancies with deterministic analyses and the implications for the use of probabilistic analyses in defining reserves and value. Probabilistic analyses can be easily constructed, either by indifference, ignorance or manipulation, to result in estimates of Proved reserves higher than deterministic estimates. Given the current state of the art and the SPE/WPC reserve definitions, users should beware of excess reliance on probabilistic analyses and should continue to regard reserves and value in light of deterministic measures.

Despite problems placing probabilistic estimates into the qualitative taxonomy of reserves, the value of the method exists apart from reserve labels. The rules of probabilistic analysis for classification of reserves do not apply when the object is not compliance with SPE/WPC definitions, such as in the case of determining economic value of a project from a probabilistic analysis. However, observations about the implications of probabilistic analyses do have importance when determining value.

Principles for resolution of reserve definitions
Reserve definitions are a convention not based on scientific truth but on the need for a consistent terminology to communicate between people and compare across diverse situations. Dichotomous and incongruent standards for reserves do not promote communication; they provide an avenue for chicanery. In addition, no one has suggested that the well-accepted and explicit criteria for reserve classifications have ceased to be good rules or that they should be discarded. In this state of ambiguity, the guiding principle must be one of backwards compatibility.

[1] P10 is taken to mean the 10th percentile of a cumulative frequency, i.e. sorted low to high.

Implications of reserve definitions and probabilistic analyses

The shortcomings of probabilistic models are most often expressed as excessively narrow distributions. The most often-cited imperatives for good model-building (correlation, aggregation, ranges of uncertainty) all pertain primarily to the impact on the range of outcomes. Selection of what types of volumes to include in an analysis is also a major driver of both the mean and the range of a resultant distribution. Consequently, the construction of a probabilistic model can have a significant impact on the implied reserves per the current SPE/WPC definitions and for the determination of value. None of these issues are addressed by the SPE/WPC definitions.

Given a static central value for a distribution, a narrower range means that the low side of the distribution is higher and that the high side is lower. A narrower range also means that the difference between a high percentile and a central estimate decreases as does the difference between a low percentile and a central estimate. Though obvious, these observations have profound implications.

A narrower distribution means higher estimates of Proved reserves and attendingly lower estimates of Probable and Possible reserves. In the limit, Proved reserves of the aggregate approach the sum of Proved and Probable for each input distribution. Also in the limit, Probable and Possible reserves approach zero. This suggests that Probable and Possible reserves are converted to Proved reserves by any one of a number of mechanisms that artificially reduces the range of a distribution.

When reserves are aggregated across fields or projects, the mean of the resultant distribution is the sum of the means, but the total range of the sum is diminished. The P10 of the sum is greater (closer to the mean) than the sum of the P10s from each distribution. Exactly the same conversion of high uncertainty volume estimates into high certainty estimates can occur as the result of the construction inside a model for a single project. Modeling of too many parameters or modeling the uncertainty in too many separate distributions can cause the model results to show an excessively narrow range.

Other modeling pitfalls such as the exclusion of a correlation between parameters or narrow ranges for key parameters reduce the range of calculated outcomes. All else being equal, both of these pitfalls result in an increase of Proved and a decrease in Probable and Possible volumes by the current definitions.

Perhaps most importantly, the reserve definitions do not draw boundaries on the classifications of volumes that may be included in a probabilistic analysis for the purpose of determining reserves. If there is no limit on what might be included in a probabilistic analysis of reserves, then a portion of Undiscovered Resources could make the leap to Proved reserves by virtue only of statistical rules. Though these rules may change the probability that some portion of Possible volumes may be realized, it is clear that statistical rules of aggregation do not change the reality of the state of knowledge about uncertain nonproved reserves.

The extremes of a resultant distribution are the least well-defined portion of the distribution and the most subject to change with changes in the model or input ranges. Consequently, high confidence Proved reserves are proposed to be defined in such a way that the estimate could change significantly with a modest change in the calculation model. These extreme values also have implication for the determination of value.

Classify first

If an evaluator performs the probabilistic analysis first and determines reserves from the prescribed probability levels as implied by the definitions, then the evaluator must first decide what should be included in the analysis, e.g. what recovery processes and which potential volumes.

A looser threshold for inclusion in a probabilistic analysis causes a higher estimate of P10. Besides the issue of aggregation narrowing the range, the addition of more considerations raises the mean. Furthermore, a higher degree of aggregation or the inclusion of more processes would raise the lower end of the resultant distribution. In all cases, the result is growth in "proved" reserves. The solution is to classify first then perform an appropriate probabilistic analysis.

Some have proposed that speculative processes like reserves growth or advancing technology may be included in probabilistic analyses. Of course, one of the strengths of probabilistic analyses is improved understanding of the impact of high uncertainty processes such as pre-drill prospect evaluation. While probabilistic methods are suitable to estimating in these situations, it is not suitable to use the results for determination of reserves.

Reserves should not be mathematically manufactured from resources, and volumes that fail the most fundamental criteria of the reserve definitions ('known accumulations') should not sneak in through the back door. One should not include volumes in an evaluation that do not meet the explicit criteria for that classification. Resources should not be included in a probabilistic analysis of Reserves.

Similarly, nonproved reserves should not be included in a probabilistic analysis of Proved reserves. To do either one creates a second and very poorly defined set of reserves definitions.

The use of risk factors with lower confidence volumes in a probabilistic analysis is necessary to modeling the uncertainty of the situation. Depending on the number of such risked volumes and the risk factors applied, the P10 of such a distribution could include some volumes directly attributable to lower probability occurrences. For example, inclusion of four volumes of independent, 50% probability means that 94% of the cases will include success of at least one. Moreover, creating more probability of recovering greater volumes, either by adding volumes of any risk or by adding the possibility of more optimistic outcomes, moves the 10th percentile figure of the distribution. Both of these cases are the equivalent of applying risk factors to deterministic summaries of non-proved reserves and reporting a risked total equivalent Proved.

In the same way that "Proved Undiscovered" reserves should not be created by the inclusion of Speculative Resources in a probabilistic reserve determination, one should not create Proved reserves from a collection of Probable or Possible. Reserves should first be classified, most importantly to the level of Proved, and then quantified with probabilistic analysis, to exclude the possibility of volumes outside the specific reserve criteria. (ref. 2)

Alternative Solution

There is an alternative, though it has limited application. It is possible to integrate the explicit criteria for reserve classification, e.g. the rules that define Proved reserves, into a probabilistic analysis. That is, it is possible to construct input distributions to comply with the apparent meaning of reserve definitions but only when there is no aggregation and limited consideration of uncertainty. Input distributions may be limited or otherwise defined to reconcile with reserve criteria. Though this approach works relatively easily when there are few relevant criteria, inclusion of multiple such considerations can cause the results to exceed deterministic measures.

For example, a probabilistic analysis may set a distribution such that the water contact is equal to the lowest known hydrocarbon (LKH) 10% of the time. All else being equal, a P10 of this distribution would appear to comply with SPE/WPC definitions as stated. The results, however, could be significantly less or more than a deterministic Proved evaluation as a result of the range of the LKH and other variables. (ref. 3) All else being equal, a more optimistic high side assumption of LKH raises the P10 of the distribution.

If the evaluator considers uncertainty in prices in addition and specifies a 10% chance of current economic conditions continuing, then the evaluation would be more likely to result in a dollar value of P10 greater than a deterministic calculation considering all the definitional criteria. It is possible to define multiple such uncertainties so that the P10 of the distribution is consistent with Proved criteria, but establishing equivalence will very likely require a deterministic evaluation for comparison.

For the defined probability levels to be consistent between different entities, the reserve definitions would need to explicitly define specific guidelines for model building, as has been done inside some companies. (ref. 4, 5) Since considerations like input ranges, correlations, levels of aggregation, and chance of success factors all can significantly affect the P10 of the outcome, consistency in model construction is essential to comparability. This is neither practical nor consistent with the stated purpose of the definitions: "The proposed terminology is not intended as a precise system of definitions and evaluation procedures to satisfy all situations."

For the sake of backward compatibility and universal acceptance of reserves by numerous entities, the solution is not to build probabilistic models by a rigid set of rules to conform to the meaning and intent of the definitions. Instead, classifying first and reporting a measure of central tendency solves the problem consistently and clearly. Probabilistic methods are most suitable to aid the user of the reserve calculation in understanding the uncertainty associated with the reserve volume.

Level of aggregation and probability level

It has been argued that even deterministic Proved reserves are high confidence figures despite the lack of a 'P10' label. (ref. 6) Only the means of distributions can be added to get a figure that applies to the distribution of the sum. Percentile figures cannot be added to get a corresponding percentile in the summed distribution. Since the laws of statistics are immutable, so the argument goes, the addition of Proved reserves is improper. On the other hand, evaluators are frequently warned about (or against) aggregation of reserves since the act of adding distributions probabilistically raises the combined P10, i.e. 'Proved', value. The P10 of a distribution summed from two independent distributions is greater than the sum of P10s for each distribution.

The two competing perspectives are based on different premises, one probabilistic and one deterministic. The probabilistic premise posits that Proved is a low side estimate and that Proved reserves are, at least in part, completely independent of other evaluations.

Despite statistical theory, empirical research has not demonstrated trends which might be attributable to the benefits of aggregation. Specifically, research shows no "economy of scale" for the market value of reserves as measured by stock price. (ref. 7) That is to say, the stock market does not recognize incremental value to Proved reserves when they are part of a larger reserve base. Additionally, data on the purchase of producing assets has not shown a higher price for reserves purchased in larger volumes. (ref. 8) It is possible that the lack of empirical evidence draws from the lack of recognition of the benefits of aggregation[2] or other cause. Nevertheless, the data indicates no empirical benefit from aggregation.

Moreover, anecdotal evidence (ref. 9 *inter alia*) indicates that Proved figures appear to be median figures since they seem to be revised down almost as often as are revised up. This appears to be the case at the company level and at the project level for fields early and late in life.[3]

If the probabilistic perspective on aggregation fails, it might be because Proved reserve figures are not P10 figures or because they are not independent or both.

Probability level of Proved. An estimate of Proved reserves should be a best estimate of the volumes that fall into

[2] If this were the case, the research would still support not adding reserves due only to aggregation since it would violate the principle of backwards compatibility.

[3] The reserves growth phenomenon seems to occur after the field is well understood and before it is fully developed. Even during this period, reserve estimates may decrease.

the class of Proved. As a best estimate, it should lie near the center of a range of uncertainty.[4] It is not an ultra-conservative (exceeded 9 of ten times) figure for the volumes in the class. Proved usually represents a high level of confidence, i.e. reasonably certain, that the volumes will be recovered not a high level of confidence that the estimates will be exceeded.

Proved may be a low side estimate of total reserves for an entire project, but only for projects with large upside development potential or a good deal of uncertainty. Large fields or fields early in development, for example, may have a large portfolio of Probable and Possible reserves. In this case, the best estimate of the Proved class may lie at the low end of the range of outcomes. If, however, the field is more mature and the majority of reserves are Proved, then a deterministic Proved would lie much closer to the middle of the distribution of total reserves.

At one extreme, a project with all Proved reserves would find a deterministic Proved close to the median of the distribution and a P10 value much less. References 10 and 11 both posit a group of Proved reserves and then observe that the P10 value of reserves is much less than a deterministic estimate. At the other end of the extreme, a single well in a massive discovery could qualify only a very small portion of the reservoir for the category of Proved. In a more typical example of a field in ongoing development, the probability level of Proved reserves will depend on the proportion of Proved reserves to nonproved reserves.

Depending on the specific case, deterministic reserves for a single project could fall at a probability level anywhere from near P0 to P50 or higher (depending upon the skewness and the measure of central tendency chosen) for the entire project. The one constant is that an estimate of Proved reserves represents a central estimate of the volumes that fit the criteria. Consequently, it is not a low side estimate whose addition violates statistical rules.

Independence. As described above, the structure of a model, including dependencies, can be the most important driver of the outcome of a probabilistic model. The same is true of a model for aggregating distributions of reserves. The problem is that not all the drivers of correlation are quantifiable, or even fully conceived.

Besides the obvious correlation of economic thresholds, significant correlations/dependencies may exist as a result of the state of the art, limited capital budgets or manpower, or corporate culture. More importantly, systematic bias of an estimator or a company can cause correlation between outcomes. Not including any kind of correlation in an aggregation model seriously risks understating volatility and overstating the low side of possible outcomes. On the other hand, including a correlation mechanism would be, at best, qualitative.

When ‘proved’ reserves are created by aggregation, the volumes cannot be attributed to any single well, field or recovery mechanism. It is increased by the fact that, given enough attempts at lower probability events, there are bound to be some successes. The number of events necessary to guarantee success at a P10 level is a function of the dependence between the attempts (subjectively estimated), the chance of success (subjectively estimated), and the number of attempts postulated (vs. actually made). The reserve definitions do not offer guidelines to incorporate these considerations nor even make allowance for such considerations.

The proposal is to classify reserves first and then take a central estimate of the Proved reserves from a range of outcomes created by excluding possibilities outside the criteria of reserve classifications. This is most consistent with deterministic practices and least ambiguous. This proposal further leaves no meaningful problem with aggregation of reserves. If an explicit proposal for aggregation across evaluations must be made, then the creation of reserves should be avoided. Alternatively, both probabilistic and deterministic may be performed and the two reconciled. (ref. 4, 12)

Economic Value

The limitations on a probabilistic analysis for the determination of reserves according to the SPE/WPC definitions do not necessarily carry over to probabilistic analyses for the determination of value. Since value determination is more subjective and less standardized and less transferable between entities, there is more room for flexibility and customization in a probabilistic analysis. However, there are still some principles that should be acknowledged.

Once reserve and economic uncertainties have been described in a probabilistic model, a user must consider at least two topics to arrive at a final value[5]:

- what discount rate or other economic yardstick to use
- what consideration to make for the range of outcomes

In our industry, there are many yardsticks commonly used to examine the economics of a project. For example, some consider one or more of the following indicators:

- present value discounted at a hurdle rate
- present value discounted at a cost of capital rate
- undiscounted cash flow
- rate of return
- return on investment
- payout

The same yardsticks or combinations of yardsticks used in deterministic analysis may be used with a probabilistic analysis. A prescription for value to every decision-maker can no more be established for a probabilistic analysis than for a deterministic one as situations are unique and different considerations are important to different decision-makers. The following discussion assumes that the decision-maker has an understanding of how to arrive at value from a deterministic analysis. For comparison purposes with

[4] For this discussion, it is irrelevant whether the central value is a mode, median or mean.

[5] It is assumed that any application of chance of success factors have been included in the model.

probabilistic valuations, let us call the appropriate deterministic metrics the "normal" values.

Central measures. It has been argued that the mean of the distribution is the best, most unbiased measure from the distribution and should be preferred as a central estimate. In many cases, the mean, median and mode do not differ significantly and the point is moot. In cases where they do differ significantly, the median or mode might be more useful to a decision-maker. Not all users of the distribution have risk profiles sufficient to allow them to make decisions based on means when there is significant risk of much lower figures. By analogy, the casino makes money by playing the averages but individuals at the casino can only win by considering probabilities. The casino can endure the losses necessary to play enough times to realize the mean. Those with lesser resources cannot endure the swings.

High & low values. In some cases, the range shows a significant risk of either a loss or an insufficient return. In cases like these, one may not need to arrive at a specific value since the project is not acceptable.

One might look at the range to understand the volatility of the best estimate and qualitatively factor that into the economic metrics used to determine value. A very wide range of possible outcomes may cause a risk-averse person to regard the project with more skepticism and value the opportunity less. Greater ranges of outcomes generally mean more volatility and less value. Conversely, a narrow range of outcomes might be used to justify a greater value due to lesser risk than otherwise estimated. For example, the discount rate that would be used with a deterministic evaluation might be revised upward for very uncertain projects and revised downward for projects with more certainty. An evaluator runs the risk, however, of overcompensating in value metrics when he/she considers the range of outcomes.

When determining value, it is imperative to keep in mind the limitations of the probabilistic model. The range of results can be artificially inflated by poor model construction or aggregation. Even the range of a good model is sensitive to model construction. Moreover, the model very likely does not include all types of business uncertainty. As with deterministic analyses, the range of reserves is additionally subject to a paradigm shift such as a major change in reservoir description. Failure to recognize the limitations of a probabilistic analysis can lead to overconfidence in a range of results and an inflated evaluation.

Since those assigning value are likely most familiar with deterministic analyses, a conventional value determination from a central estimate should be the starting point and the paragon by which alternative conclusion of value are sense-checked. Suggesting that the "normal" measures cannot be adjusted denies the value of a probabilistic assessment, and that is not the conclusion. However, deviation from the well understood should be judicious and circumspect.

Conclusions

1. Reserve definitions are meant to promote clear communication. Any changes to the definitions should be backwards compatible.
2. The SPE/WPC definitions for reserves evaluated by probabilistic methods lead to a dichotomous and incongruent standard for reserves.
3. Classify first - When probabilistic methods are used in a reserves determination, volumes should first be grouped by classification and secondly evaluated for uncertainty.
4. Proved reserves should be a measure of central tendency, not a lower probability level, for a distribution of volumes which meet the criteria for the classification.
5. When using probabilistic methods for the determination of value, deviation from the well-understood should be judicious and circumspect.

Nomenclature

P10 = 10th percentile of cumulative frequency (sorted low to high)

P50 = median

P90 = 90th percentile of cumulative frequency (sorted low to high)

References

1. Society of Petroleum Engineers and World Petroleum Congresses: "Petroleum Reserves Definition" (March 1997). Available from SPE Book Orders and posted online at http://www.spe.org
2. Acuna, H.G. and Harrell, D.R.: "Adapting Probabilistic Methods to Conform to Regulatory Guidelines," paper SPE 63202 presented at the 2000 SPE Annual Technical Conference and Exhibition, Dallas, Oct. 1-4.
3. Jordan, D.G. and Collarini, C.R.: "Deterministic Versus Probabilistic Volume Estimates for Reservoirs Without a Known Downdip Limit," paper SPE 52945 presented at the 1999 SPE Hydrocarbon Economics and Evaluation Symposium, Dallas, Mar. 20-23.
4. Patricelli, J.A. and McMichael, C. L.: "An Integrated Deterministic/Probabilistic Approach to Reserve Estimations," paper SPE 28329 presented at the 1994 SPE Annual Technical Conference and Exhibition, New Orleans, Sept. 25-28.
5. Alexander, J.A. and Lohr, J.R.: "Risk Analysis: Lessons Learned," paper SPE 49030 presented at the 1998 SPE Annual Technical Conference and Exhibition, New Orleans, Sept. 27-30.
6. Capen, E.C.: "A Consistent Probabilistic Approach to Reserves Estimates," paper SPE 25830 presented at the 1993 SPE Hydrocarbon Economics and Evaluation Symposium, Dallas, Mar. 29-30.
7. Lohrenz, J. and Pederson, J.A.: "Buying and Selling Producible *In Situ* Oil and Gas on Wall Street and Other

Markets," *Proceedings*, 1998 North American Conference, Inter. Assoc. Energy Econ., Albuquerque, Oct. 18-21.
8. Lohrenz, J.: "Buying and Selling Producing Oil and Gas Properties: The Norms in the Data," paper SPE 52946 presented at the 1999 SPE Hydrocarbon Economics and Evaluation Symposium, Dallas, Mar. 20-23.
9. Keith, D.R., *et al*: "Reserve Definitions: An Attempt at Consistency," paper SPE 15865 presented at the 1986 SPE European Petroleum Conference, London, Oct. 20-22.
10. Cronquist, C.: "Reserves and Probabilities – Synergism or Anachronism?," *JPT* (Oct. 1991) 1258.
11. Caldwell, R.H. and Heather, D.I.: "How to Evaluate Hard-To-Evaluate Reserves," *JPT* (Aug. 1991) 998.
12. Nangea, A.G. and Hunt, E.J.: "An Integrated Deterministic/Probabilistic Approach to Reserve Estimation: An Update," paper SPE 38803 presented at the 1997 SPE Annual Technical Conference and Exhibition, San Antonio, Oct. 5-8.