

Coupling Probabilistic Methods and Finite Difference Simulation: Three Case Histories

Dwayne C. Purvis, SPE, Cawley, Gillespie and Associates, Inc. (CG&A); Richard F. Strickland, SPE, CG&A; Richard A. Alexander, SPE, CG&A; and M. Anthony Quinn, SPE, CG&A

Copyright 1997, Society of Petroleum Engineers, Inc.

This paper was prepared for presentation at the 1997 SPE Annual Technical Conference and Exhibition held in San Antonio, Texas, 5-8 October 1997.

This paper was selected for presentation by an SPE Program Committee following review of information contained in an abstract submitted by the author(s). Contents of the paper, as presented, have not been reviewed by the Society of Petroleum Engineers and are subject to correction by the author(s). The material, as presented, does not necessarily reflect any position of the Society of Petroleum Engineers, its officers, or members. Papers presented at SPE meetings are subject to publication review by Editorial Committees of the Society of Petroleum Engineers. Electronic reproduction, distribution, or storage of any part of this paper for commercial purposes without the written consent of the Society of Petroleum Engineers is prohibited. Permission to reproduce in print is restricted to an abstract of not more than 300 words; illustrations may not be copied. The abstract must contain conspicuous acknowledgment of where and by whom the paper was presented. Write Librarian, SPE, P.O. Box 833836, Richardson, TX 75083-3836, U.S.A., fax 01-972-952-9435.

Abstract

Advances in computing power have made possible the combination of probabilistic methods with the conventional practice of finite difference simulation. This paper presents and demonstrates a methodology for probabilistic finite difference simulation to determine and examine a range of potential production rate profiles and ultimate recoveries by simulating hundreds of scenarios.

Due to the high degree of uncertainty in the three case histories discussed, finite difference simulation coupled with probabilistic methods offered the best tool to predict production profiles given a wide variety of assumptions about reservoir character and producing conditions. A Monte Carlo simulation was run in each case to create several hundred possible combinations of the uncertain variables. These variables became inputs to the same number of finite difference simulations. The resulting forecasts were ranked and P10, P50 and P90 values were determined depending on the needs of those using the forecasts.

When probabilistic simulation is used to predict a range of results, care should be taken when combining the forecasts to properly honor their history-dependent nature. Discounting techniques can be applied to the forecasts before ranking in order to make them useful for economic decisions. In the one case in which a comparison was made, it was also observed that, depending on the skewness of the input distributions, the median forecast was similar to the deterministic, "most likely" forecast.

Introduction

Monte Carlo analysis has been frequently applied in the petroleum industry as it permits a quantitative analysis of the associated risks. In traditional Monte Carlo applications the calculations are rather simple, such as a summation to obtain total cost or multiplication to obtain oil-in-place. Distributions of the calculated values are analyzed to ascertain the total and combined uncertainty implied by the input values.

Advances in computing power have made it feasible to extend Monte Carlo techniques to complex calculations such as finite difference simulation. The coupling of these powerful tools allows the modeler to quantify the uncertainty in the generated rate forecasts. In the three examples presented, the uncertainty in the production forecast was particularly important to the development decision, making them good candidates for a probabilistic approach.

All three studies involved large uncertainties in reservoir parameters due to limited information. A single simulation forecast was not deemed adequate because it could not describe the associated uncertainty. In each case, an appropriate finite difference model was created. Probability distributions were developed for the major variables that controlled the production profile and ultimate recovery. The most important parameters controlling the production profile varied between the cases. Among the variables considered were porosity, water saturation, net thickness, gas-water contact, permeability and areal extent.

Monte Carlo techniques were used to create several hundred combinations of uncertain reservoir parameters. Each combination was then simulated producing a unique production profile. The resulting, simulation-based production forecasts can be interpreted for technical risk. Such characterization of technical risk is becoming critically important to the increasingly sophisticated strategies used in the petroleum industry.¹

Statement of Theory and Definitions

Traditional Monte Carlo applications address uncertainty of data inputs to simple calculations such as oil-in-place. Finite difference simulation, on the other hand, has been used to assess uncertainty on a larger scale, such as that between qualitatively different reservoir characterizations. In combination, the Monte Carlo techniques and finite difference simulation

can be used to test both kinds of uncertainty. Naturally, meaningful analysis requires that the drivers of the uncertainty be adequately reflected in the construction of the Monte Carlo simulation and the design of the finite difference model.

Discussion of Uncertainty. In general, there are three types of uncertainty: data or input uncertainty, model uncertainty and compound uncertainty. Data or input uncertainty refers to the uncertainty associated with the values which are input to the model. Model uncertainty refers both to the applicability and the accuracy of the conceptual or mathematical model used. Compound uncertainty describes the use of an uncertain result in a second calculation or third calculation.

Data Uncertainty. Data uncertainty pertains to the range of raw data values which can be assumed as input to a calculation. For example, when porosity varies throughout a field, the distribution of porosity can become an input to a probabilistic volumetric calculation. The distributions are usually based on the data available. In the case where hard data is not available, assumptions must be made to supply the necessary inputs to the calculations.

Data uncertainty can also arise from questions about the accuracy or relevance of data. Of course, data can be inaccurate due to the limited ability to measure data as in the interpretation of pressure build-ups, or data can simply be wrong due to measurement error.

Model Uncertainty. Model uncertainty describes the uncertainty associated with how adequately any procedure, technique, algorithm or formula accurately represents the reality of physical processes it is meant to describe. There are two types of model uncertainty, one deals with the applicability of the model, the other with the accuracy of the model.

Of course, different conceptual and mathematical models exist to describe different types of situations, and the choice of models has a significant impact on the final result. For example, Darcy's law is accurate and precise only for Darcy flow. If it is not known whether flow is Darcy or turbulent, then applying Darcy's law introduces model uncertainty, since the applicability of the model is in question. Often, the correct choice of a model is not straightforward. Since a change in the model used to describe the system can lead to a quantum change in the result, model uncertainty can be the most important type of uncertainty.

Even when using an appropriate model, there may be uncertainty associated with the accuracy of the model. Empirical models, for example, suffer this limitation, since actual values do not correlate perfectly to predicted values. When significant, the error can be quantified statistically and used in uncertainty analysis.

Model uncertainty offers a different challenge than data uncertainty since it involves a quantum change in methodology used to calculate an answer. Uncertainty analysis can describe model uncertainty by reporting different ranges under different models or assumptions. Model uncertainty can also be addressed by applying a likelihood to the occurrence of a certain

model coupled with an uncertainty of the value implied by that model. For example, there may be a 50/50 chance of a concept being valid. This 50/50 probability can be modeled, as well as the uncertainty of the result given each assumption.

Compound Uncertainty. Finally, compound uncertainty results when an uncertain result is assumed to be exact and becomes input for another process. This, of course, is the natural flow of data in engineering calculations, and it must be considered in designing the probabilistic analysis. Compound uncertainty can generally be avoided by modeling the uncertainty at the stage where it can best be described. In one case, for example, the log analysis criteria are varied to determine the range of uncertainty associated with net thickness. It is important to identify the key uncertainties and to adequately reflect those uncertainties in the next level of computation.

Distributions of Unknown Variables

The single most important step of Monte Carlo simulation is defining the input distributions. Often, however, certain critical data is limited or simply not available, making design of distributions somewhat abstract. In the absence of complete data, distributions must be based on judgments made by the estimator.

Data uncertainty and model uncertainty can often be described based on available data. Used judiciously, statistics can be a powerful tool. The estimator should exercise judgment and look for bias, particularly when the sample populations are small. Additionally, statistics should generally be applied at the same level at which they were measured. The lack of data sometimes requires that statistics be applied in less than ideal ways. For example, an estimator may be forced to assume that the statistics of a distribution in one area applies to another area.

Frequently, factual evidence does not exist to permit statistical analysis, and the estimator must rely on his experience and judgment to select a reasonable range of input values and a representative distribution type. Many reservoir parameters tend to distribute themselves in normal or log normal forms.² In practice, triangular distributions are often used to approximate these distributions or represent values for parameters whose distribution type is unknown.

An estimator should understand the implications of the distribution he selects. Triangular distributions, for example, imply that it is impossible to attain a value equal to or beyond the extremes of the distribution. Additionally, when distributions are not symmetrical, the distribution favors either a high side or a low side value. The most likely value (mode) located at the peak of the distribution will not be the same as the median (P50) value or the expected value (average) of the distribution. In the absence of data and when the estimator does not have a good sense about relative probability, a uniform distribution may be a good option. However, the estimator may limit the range to maintain reasonableness, not over-emphasizing unlikely results.

It has been pointed out that estimators tend to have more confidence in their estimates than they should, i.e. they choose uncertainty ranges that are too narrow. After constructing a distribution, the estimator should test his selection for reasonableness.^{2, 3} The estimator may confirm, for example, that there is about a 1 in 3 chance of the actual values being lower/higher than the P30 and P70 values of the distribution. However, the estimator should also be careful not to overestimate or to skew inadvertently the bulk of the distribution.

When limited data is available, distributions should be defined to include some probability at the values which are the extremes of the data sets. This can mean assuming that a given percentage, usually 5% or 10%, of the total population will fall above/below the most extreme values encountered. The more data points available, the less likely the actual distribution of values will include many values outside the range encountered in the sample. If the extreme value of the data set is used as the extreme of the distribution, then a conventional triangular distribution should not be used. The endpoint values of a triangular distribution have a zero probability. Alternatively, a truncated triangular distribution as demonstrated in **Figure 1** can be used. Truncated distributions are also useful in situations which do not allow values above or below a certain point.

Process and Tools

In order to perform Monte Carlo-type finite difference simulation, tools and procedures were developed to incorporate the relevant types of uncertainty and to process the results. Simple FORTRAN and spreadsheet-based programs can accomplish these tasks. Software to perform the Monte Carlo and finite difference simulations are readily available on the market.

The following list describes the generalized steps in the process:

1. **Identify the major uncertainties.** The engineer must define which variables have the most impact on the final result and the source of the uncertainty in those values. This may involve a limited number of simulation runs or analytical calculations.

2. **Define the uncertainties.** The uncertainties must be quantified into probability distributions, including appropriate correlations between variables.

3. **Choose the number of Monte Carlo trials necessary and create multiple realizations.** Although thousands of trials can be run, it is only necessary to simulate a few hundred scenarios in order to adequately define the range of values when the input distributions have been sampled using Latin Hypercube Sampling (LHS). A Monte Carlo simulation must be run to create the chosen number of combinations of variables.

4. **Run the finite difference simulations.** Using a program which will automatically substitute the next combination of parameters in the simulation input deck and rerun the simulation, sequential simulations can be run in batch mode and the rate forecasts for each well or combination of wells saved.

5. **Combine the results, if necessary.** Depending on the situation, the results of the simulations may be combined with others in order to define a range of total forecasts for an entire project. The process of combining multiple sets of probabilistic forecasts should be tailored to the specific project. In many instances, Monte Carlo techniques can also be applied here.

6. **Statistically examine the final set of simulation outputs.** In our work, we selected a sort criteria and determined the P10, P50 and P90 profiles. The forecasts were ranked according to a sort criteria related to the end use of the values. For example, if the forecasts are used to determine the likelihood of a given rate in a given year, then each year's forecasted rates can be ranked separately. If the main issue is economics of a project, then each forecast should be ranked related to economic return. Based on the resulting range of predictions, an informed decision can be made given the potential rewards of a project and the decision-makers' tolerance for risk.

Data and Results

Case #1: Stacked Gas Sands. The first example of the Monte Carlo simulation technique presented here involved a series of stacked gas sands in an offshore location. Based on seismic data, two wells plus a sidetrack had been drilled, establishing the presence of 11 potentially productive zones. Many of the zones were cored, and numerous RFT measurements were taken. Among the 11 zones, there was significant uncertainty in five parameters: net thickness, gas/water contact, porosity, permeability and water saturation.

Net Thickness. Due to log analysis limitations, there were significant uncertainties in the log-calculated net thickness values. As the net pay cut-off criteria was the major driver in the log analysis, it became the basis of a Monte Carlo analysis. High side and low side net thicknesses were determined by changing the two net pay cut-off criteria up and down by 20%. The ratios of the resulting net pay thicknesses to the deterministic value defined the endpoints of a triangular distribution of net thickness ratios used in the Monte Carlo analysis.

Rather than defining a range of net thickness multipliers based on an amorphous sense of uncertainty in the resulting value, basing the range on the inputs to the calculation of net thickness eliminated compound uncertainty and yielded a more meaningful range of net thickness multipliers unique to each zone. For example, in one zone, the net thickness multipliers ranged from 0.47 to 3.00 while another zone varied only from 0.84 to 1.16.

Gas/Water Contact. The first well was drilled near the crest of the structure and did not encounter any water zones. The second well was drilled off structure, and though most of the sands were present, many were wet. The presence of gas in one well and water in the other made it possible to approximate the gas-water contact (GWC) directly from RFT data. For zones with the contact between the wells, the uncertainty in the GWC was the uncertainty in the pressures measured by

RFT. Like the method used in net pay, this method quantifies the uncertainty in the inputs to the calculations rather than the uncertainty of the result.

For layers of limited areal extent which did not correlate between the wells, a uniform distribution was chosen to represent a range of reasonable assumptions for GWC. In absolute terms, the contact could be no higher than the lowest known gas in the upstructure well, and it could be no lower than the top of the correlation spot in the lower well. However, the contact, and hence the areal extent, is not likely to occur over the whole range of theoretical possibilities. The highest reasonable contact was chosen to be one sand thickness below the lowest known gas. The lowest reasonable contact was chosen to be the same distance below the deterministic case contact as the highest reasonable contact was above it. This algorithm defined the possible range of contacts as approximately the upper two-thirds of the displacement between the wells. This distribution is consistent with geologic limitations, and it captures the downside possibility without exaggerating the upside. A similar uniform distribution was used to define the contact in the zones which tested gas in both wells.

Since the Monte Carlo results were to be used for internal company decision-making and not public reserves reports, standard US classifications of reserves did not constrain the evaluation of potential gas-in-place. It would be possible, however, to define the probabilities to align with the risks associated with a deterministic classification of reserves.

Porosity, Water Saturation and Permeability. Statistical analysis was used to define the uncertainty in porosity and water saturation values. Core analysis results were correlated against log-based calculations in order to determine an empirical relationship between the two. The resulting correlations supplied values of porosity and water saturation in zones without direct measurements, and the scatter in the correlation was the basis of the Monte Carlo analysis. Assuming the error would form a normal distribution centered at zero, a distribution was defined using the standard deviation of the percent error in the correlation.

A similar analysis was attempted to quantify the uncertainty in permeability. However, lab tests indicated that permeability was a strong function of overburden stress, casting doubt on the core analysis run at low overburden stress. Instead of core data, permeability determined by pressure-buildup analysis formed the basis of a non-linear, empirical relationship to log readings. Although most predicted values are close to measured values, the predicted permeability was off by a factor as great as 3.

The small sample of data made it unreasonable to calculate an average percent error. Instead, a triangular distribution of uncertainty as shown in **Figure 1** was used. If a conventional triangular distribution had been used, then most of the cases would have fallen above the most likely value. Rather than exaggerate the upside in this manner, the distribution was chosen so that fifty percent of cases will be high by up to a factor

of 3, and fifty percent of cases will be low by a factor down to one third.

Analysis and Conclusions. The number of trials necessary to define P10 and P90 forecasts depends in part on how the Monte Carlo simulation is run. If the simulation uses traditional sampling techniques to choose values for each trial, then thousands of trials may be necessary to adequately define the extremes of the outcome distribution. Latin Hypercube Sampling (LHS) divides each input distribution into "bins" and selects randomly in each bin before sampling from a bin a second time. This methodology provides random sampling but permits the outcome distribution to be adequately represented using significantly fewer trials.⁴

The number of trials necessary depends upon the relationships between the uncertain variables and the number of bins used for LHS. In cases where P5 and P95 will be used for analysis instead, more trials may be necessary. In this example, LHS was used to create the combinations of parameters to be used in each simulation, and examination of cross-plots of variables confirmed that there was sufficient coverage of the possible range of variables and combinations of variables.

To minimize the number of simulation runs necessary, a prototype finite difference model was used to simulate 100, 400, and 1,000 combinations of variables. Comparison of the resulting distribution of forecasts showed only a minor difference between the cases with 400 and 1,000 trials. Subsequent comparison of the final set of 400 simulation runs to a set of 3,000 simulations runs confirmed that the 400 simulations adequately captured the uncertainty in the forecasts. **Figure 2** illustrates that comparison.

For each of the 11 zones, 400 finite difference simulations were run based on the 400 combinations of variables produced by the Monte Carlo simulation. The simulation model was a two-phase, two-dimensional model with approximately 600 cells for each zone. A small FORTRAN driver program sequentially started the runs and saved the forecasts. The computation time to run 400 simulations for each zone was approximately 13 hours, totaling about 150 hours to simulate all the zones.

The result was 4,400 unique production forecasts representing 11 zones which would, in fact, produce simultaneously. The forecasts had to be combined meaningfully to create a forecast for the total field. Each zone was assumed to be an independent event, i.e. there was no correlation between the results for the different zones. A FORTRAN program randomly selected one of the 400 forecasts for each zone and summed the forecasts for all zones.

Since there was some uncertainty as to whether the permeability reduction exhibited in core analysis would in fact occur in the field, a second set of simulations was run without permeability reduction due to increases in net overburden stress. This analysis quantified the model uncertainty associated with permeability reduction due to net stress effects.

Figure 3 illustrates the resulting P10, P50, P90 and deterministic rates for each year from three of the eleven zones and

for the combination of all zones in the permeability reduction case. In zones A and C, the deterministic values were near the center of the distributions for all the uncertain variables, and the P50 rates agree very well with the deterministic forecast. In zone B, however, the range of GWCs included a significant probability that the contact would be deeper than was estimated for the deterministic case. For this reason, the deterministic forecast is closer to the P90 rates, and the probabilistic analysis captures the upside potential.

The deterministic combination of forecasts exceeds the P50 forecast in earlier years but falls below it in later years. GWC assumptions in the deterministic case as described in zone B above may account for this behavior. It should also be noted that the P50 and deterministic may not be expected to be exactly the same since they are different types of forecasts. The P50 is a median, meaning that half of the forecasted rates are greater and half are less than that value. The deterministic forecast represents an opinion about the most likely, not the median, outcome.

Figure 4 shows the distribution of forecasts over time for all zones in the prospect combined. In early years, the distribution depends on the number of wells that came on line, but in later years, the distribution appears more normal, although skewed. Later in the forecast, the distribution becomes bimodal, apparently in response to the character of forecasts for certain important reservoirs.

Case #2: Vent Gas Project. In this second example, the owner of mineral rights in an area formerly mined for coal proposed converting a vent from the abandoned mine into a coalbed methane producer. Measurements of pressure and rate had been made as gas was vented to the atmosphere over a two month period. The large degree of uncertainty in critical reservoir parameters and operating conditions made the project a good candidate for probabilistic finite difference simulation.

Coal removal operations began at the mine around the turn of the century and concentrated on the Top Hard seam. Though the Top Hard is the largest seam, as many as 60 separate coal seams had been identified within about 1,000 feet above and below the Top Hard. Extensive removal of this zone resulted in the destressing of the seams located immediately above and below the main workings. This destressing caused both the non-coal rocks in between and the coal seams themselves to expand, creating fissures and conduits for flow into the tunnels of the main workings. Additionally, limited mining operations branched into a few of the surrounding seams. The proposed vent project is assumed to tap directly into this combined system of mine workings and fissure-connected coal seams.

Based on the available data, a numerical finite difference simulation model was developed to simultaneously characterize the interaction between 1) the flow of gas through the tunnel workings, 2) the desorption of gas from the remaining unmined coal, 3) the flow of gas from the surrounding seams

through fissures into the main tunnel workings, and 4) the productivity of the vent against a range of producing pressures.

The finite difference model was history matched to the existing measured data. The match quality was excellent, but the range of flowing pressures used in the match was limited to the variation in atmospheric pressure. Rate predictions against pressures lower than about 14 psia represented an extrapolation from the existing data. The limited duration of the test also made the history match insensitive to certain parameters important to long term production.

Relying on the history-tuned model, future gas rates were forecast for a variety of reservoir parameters and producing conditions. Probability distributions were established for effective drainage area, net coal thickness, effectiveness of the connecting fissures, and flowing pressure. In order to select the important parameters for the Monte Carlo simulation and to define the distributions of some of the parameters, simulations were run varying one parameter at a time. For example, areal permeability of the surrounding coal seams was found not to have a material impact on the future rates.

Drainage area was estimated to be very large based on the extensive mine workings, but the lack of information about the connectivity of the mine shafts caused uncertainty in the drainage area assumed. The range of flowing pressures was selected based on discussions with the project's designer and the type of compression technology envisioned. Flowing pressures were assumed constant over the life of the project.

Although important to the ultimate productivity of the well, the effectiveness of the stress-induced fissures to transmit gas from the surrounding coal seams into the main tunnel workings was unknown. The final distribution was based on simulation testing of the sensitivity to vertical connectivity. This analysis showed that vertical connectivity needed to vary over two orders of magnitude, from 0.01 to 1.0 millidarcies, in order to capture the effects of this variable.

Actual data was available from shaft logs to define the net thickness distribution. Detailed summaries of the logs were created and clean coal thickness was calculated for coal seams within 200 meters (about 650 feet) above and below the Top Hard workings. Published data from other mines estimate that coals this remote from the main workings can contribute gas into the main tunnels.

Our analysis of the uncertainty in this project did not directly address one issue of model uncertainty. Since the mine workings are not completely sealed from the surface, a risk exists that too low flowing pressures might cause air to be sucked into the mine, thereby contaminating the produced gas with high levels of nitrogen and oxygen. However, because of the large volume of gas-in-place and the modest assumptions for future flowing pressures, we have assumed that the produced composition will not change materially over time.

Like in the previous example, a finite difference simulator calculated forecasts of gas production rate for several hundred random combinations of variables chosen from the input distributions described above. For each year, the P10, P50 and

P90 were chosen based on the distribution of predicted rates for that year. Forecasted gas rates and cumulative gas produced for the P10, P50, and P90 cases are depicted in **Figure 6** and **Figure 7**.

In this example, there was less data to form a basis for the uncertainty distributions, so the distributions had to be based on subjective judgment. The example also demonstrates how probabilistic finite difference simulation can be used to bound forecasts after history-matching has helped to define some of the variables and/or their likely ranges.

Case #3: Coalbed Methane Project. The third example also addresses a coalbed methane project. In this case, it was proposed to use conventional vertical wells to develop unmined coal seams in an area with active mining operations. There had been no production in the area, but a large amount of data was available from nearly 200 core holes used to evaluate the coal resources for mining purposes. In this case, a single well finite difference model was chosen to simulate various combinations of ten variable parameters.

After a review of the available data, the development area was divided into three qualitatively different areas for the purpose of defining uncertainties and forecasts. Because the core holes which had been drilled had been intended to gather data for coal mining instead of gas production, there were still large uncertainties associated with many parameters. For each area, ten important and uncertain variables were identified and defined. They included ash content, gas saturation, reservoir pressure, number of coal seams, thickness of coal seams, permeability, porosity, net stress dependent permeability, and completion efficiency. In cases where good data was available from core hole data, statistics were used to define the probability distributions of the parameters. Where little or no factual data existed, distributions were based on our experience and judgment.

Ash Content, Number of Seams and Seam Thickness. The data collected from core holes contained a good deal of data related to ash content, number of seams and seam thickness. For these parameters, distributions were fit to the available data. The analysis showed that a log normal distribution best characterized the distribution of ash content. The number of total seams present was chosen as a triangular distribution. Not enough data was available to fit a more sophisticated distribution type since a relatively small number of the core holes penetrated all zones.

The core hole data showed two distinct distributions for net thickness. In both cases, a log normal distribution fit the sample data. However, a minimum net pay thickness of 0.3 meters, approximately 1 ft, was chosen. Consequently, the distribution reflects a truncation at 0.3 m as shown in **Figure 8**.

Permeability, Porosity, Completion Efficiency, and Gas Saturation. The core hole data did not include any information about the permeability or porosity of the coals. Triangular distributions for these parameters were assumed based on our experience in coalbed methane reservoirs. Under the assump-

tion that the wells would be stimulated during completion, a range of hydraulic fracture half-lengths was assumed to represent the completion efficiency.

Gas saturation was based on a combination of available data and judgment. Gas saturation describes *in situ* gas content as a fraction of the theoretical maximum. Despite the possibility that the coals would be fully saturated with methane when drilled, available data led us to believe that there was a significant probability that the coals would be discovered in an undersaturated state. Three different saturation distributions were defined to describe the three regions, but the same methodology was used in each case.

A single probability was used to define whether or not the coal would be fully saturated. For each trial in which the coal was determined to be undersaturated, a second distribution predicted the degree of undersaturation. For example, Region I was considered to have only a fifty percent chance of being fully saturated, based on the available data. One distribution captured this either/or probability. A second distribution shaped as a triangle was used to define the degree of undersaturation. The minimum and most likely undersaturation was chosen to be zero with the probability declining to a maximum undersaturation of 20%.

Reservoir Pressure and Net Stress Dependent Permeability. Uniform distributions were defined for the possible range of reservoir pressure and overburden net stress. For each region, the initial reservoir pressure distribution included a range of pressures based on a normal pressure gradient and geologic structure.

Overburden net stress may have an impact on coal cleat permeability over the large vertical interval where coal seams are found. Because the area has not been developed, it is not known whether these coal seams would exhibit permeability net stress effects. However, coals in other basins have documented net stress relationships. The probabilistic simulation included the possible effect of net stress by defining a uniform distribution for a "net stress factor," ranging from zero to 1.0. The net stress factor was defined so that permeability was not a function of overburden net stress if the randomly selected net stress factor equals zero. However, if a net stress factor of 1.0 was randomly selected, then the permeability was 100% a function of overburden net stress. The adaptation of a net stress factor made it possible to describe a uniform distribution including all possibilities. The uniform distribution aptly described the situation since they are best suited to situations of large uncertainty with absolute limits to the range.

Combine the Results. Given these distributions, Monte Carlo simulations using Latin Hypercube Sampling created hundreds of combinations of parameters for each of the three regions. A single well simulator was used to predict the production for each combination of variables. As in the stacked gas sands project, a second set of simulations was run for each area to consider fundamental differences in assumptions. In this case, the second set of runs helped to defined the effects of doubling the well spacing.

Rate forecasts for this project had to be combined in a different way than in the stacked gas sands project. The three regions were treated as separate and independent development decisions and were not combined to create a total forecast. However, the simulations forecasted rates for individual wells rather than for an entire region, raising the issue of how to create forecasts for each region as a whole.

Unlike the first example, the forecasts which should be combined were not independent of each other. Many of the uncertain parameters would be highly correlated between wells. Random individual well forecasts should not be combined to create a total forecast for each region. Instead, each single well forecast was grossed up to describe the forecast for the region by multiplying the forecast by the number of wells to be drilled in the region.

Of course, there should have been some correlation between the forecasts of "independent" zones in the stacked gas sands project, and there should have been some variability among the "completely dependent" forecasts for each well in each region in the coalbed methane project. However, a methodology to incorporate the proper dependencies while permitting the proper independencies appears to require many more simulations than used for these cases.

Including the correlations between zones or wells in the Monte Carlo simulation of parameters was relatively simple, once the correlations were defined. For the combination of the resulting finite difference simulations to be meaningful, the finite difference simulations must be combined in the same way in which the inputs were generated. Although a few hundred combinations of parameters may be sufficient to describe the uncertainties in a single finite difference simulation, many more may be necessary to capture the distribution of total forecasts if that total is to accurately reflect the interdependencies between events.

Determine P10, P50, P90 Forecasts. In previous examples, the forecasted rates for each year were sorted and the P10, P50, and P90 values were identified from each year's distribution. This type of analysis produces the probability of a rate greater than a given value in a given year independent of the rate in any other year. Since the rates in a forecast are dependent on those of previous years, it is not proper to conglomerate the P10 or P90 rate for each year into a P10 or P90 forecast. If the forecasts never crossed, then the P10 forecast would be the same as the P10 rates for each year using many sort criteria. In these examples, it appears that there is little difference between the results of the two methods, although differences increased in the later years of the forecasts.

The gas production profiles were sorted on a discounted ultimate recovery basis for the purpose of selecting the P10, P50 and P90 probabilistic profiles. Assuming similar cost and capital requirements, discounting gas volumes approximates a sort based on net present value. Once the forecasts are ranked, the forecasts can be chosen at appropriate percentiles.

Since the selected probabilistic curve can look quite different when compared to its neighboring profiles, the final re-

ported profiles were based on arithmetic averages of the surrounding curves. For example, the final P50 curve is the arithmetic average of the actual P50 curve and the six surrounding curves (i.e., curves P50 minus 3 through P50 plus 3). **Figures 9 through 11** show the P10, P50 and P90 gross gas profiles for Regions 1, 2, and 3. The probabilistic water profiles were determined by averaging the companion water profiles associated with the gas profiles identified in the selection process. In all three regions, the predicted ultimate recovery formed a normal distribution. **Figure 12** is a probability graph of ultimate recovery for Region 2 demonstrating this observation. The forecasts for each region became input to full economic analyses.

On the basis of this study, the developer decided that the downside potential of two of the three regions was too great. The developer planned a pilot drilling program in the third region, expecting that the new wells might reduce the risk in that area and in the other two.

Conclusions

1. Monte Carlo techniques used with finite difference simulations can quantify the uncertainty in forecasts of production.
2. Selection of the P10, P50, P90 results should be based on the final use of the forecasts.
3. In the case for which a comparison was made, the relationship between the P50 forecast and the deterministic forecast seemed to depend on the relationship between the deterministic inputs and the range of possible values used in the probabilistic analysis.

Acknowledgments

We thank McCormick Energy and Energy Finance Limited for permission to publish this work.

Nomenclature

P10 = 90th percentile, i.e. 10% of the population exceeds this value

P50 = 50th percentile, i.e. the median of the population

P90 = 10th percentile, i.e. 90% of the population exceeds this value

References

1. Edwards, R.A. et al: "Quantification of Production Prediction Uncertainty and Its Impact on the Management of Oil and Gas Price Risk," paper SPE 28330 presented at the 1994 SPE Annual Technical Conference and Exhibition, New Orleans, Sept. 25-28.
2. Murtha, J.A.: "Incorporating Historical Data in Monte Carlo Simulation," paper 26245 presented at the 1993 SPE Petroleum Computer Conference, New Orleans, July 11-14.
3. Capen, E.C.: "The Difficulty of Assessing Uncertainty," *Journal of Petroleum Technology* (Aug. 1976) 843-850.
4. Startzman, R.A. et al: "An Improved Computation Procedure for Risk Analysis Problems with Unusual Probability Functions," paper 13772 presented at the 1985 SPE Hydrocarbon Economics and Evaluation Symposium, Dallas, March 14-15.

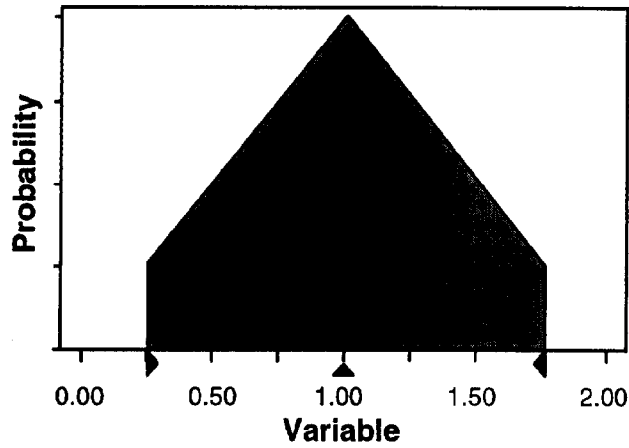


Figure 1—Example of truncated triangular distribution.

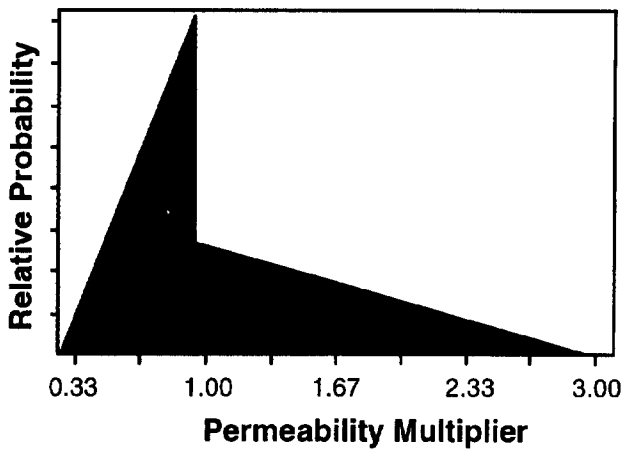


Figure 2—Modified triangular distribution to represent permeability variations.

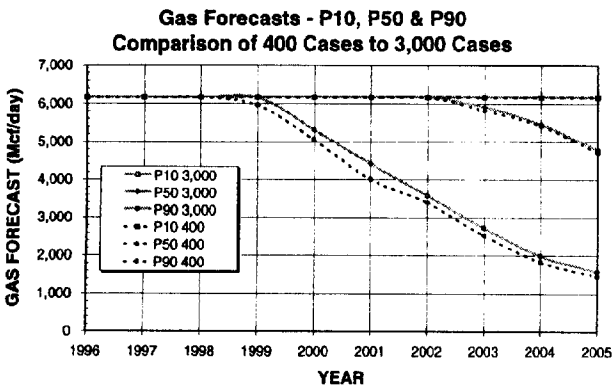
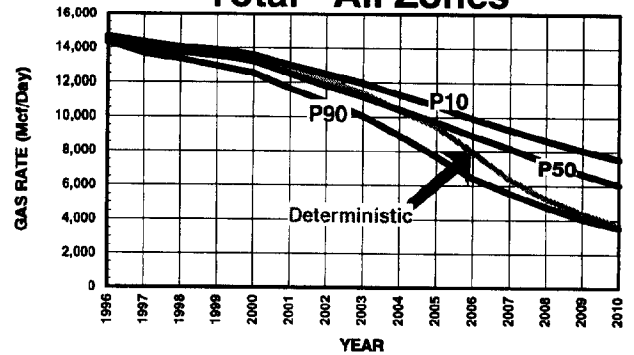
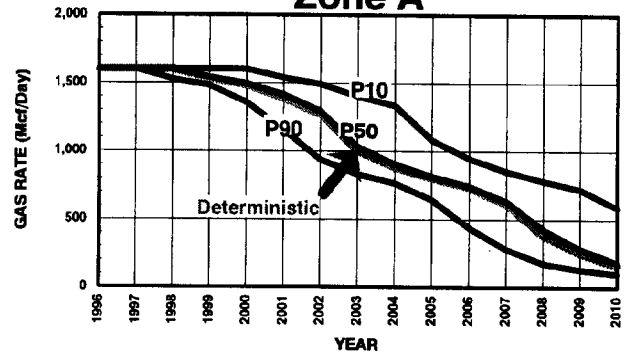


Figure 3—3,000 trials vs 400 trials, P10, P50 & P90.

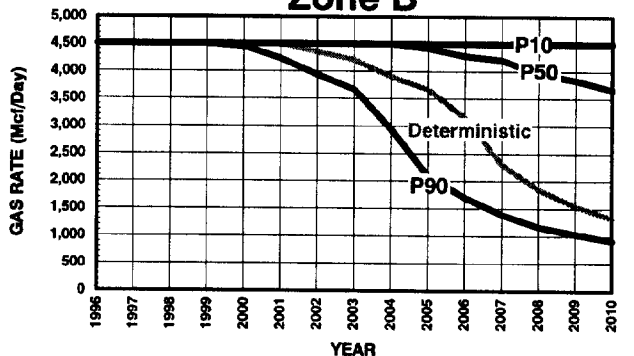
GAS RATE FORECASTS Total - All Zones



Zone A



Zone B



Zone C

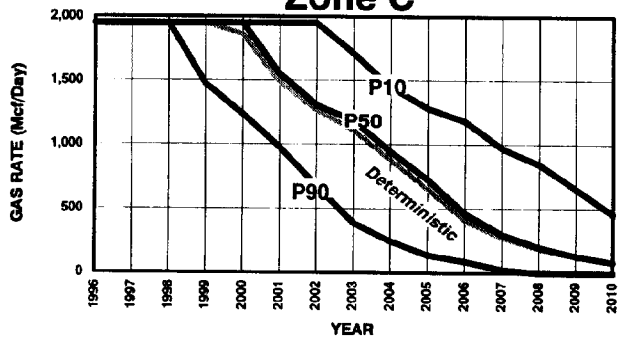


Figure 4—Comparison of deterministic & probabilistic forecasts.

GAS RATE HISTOGRAMS

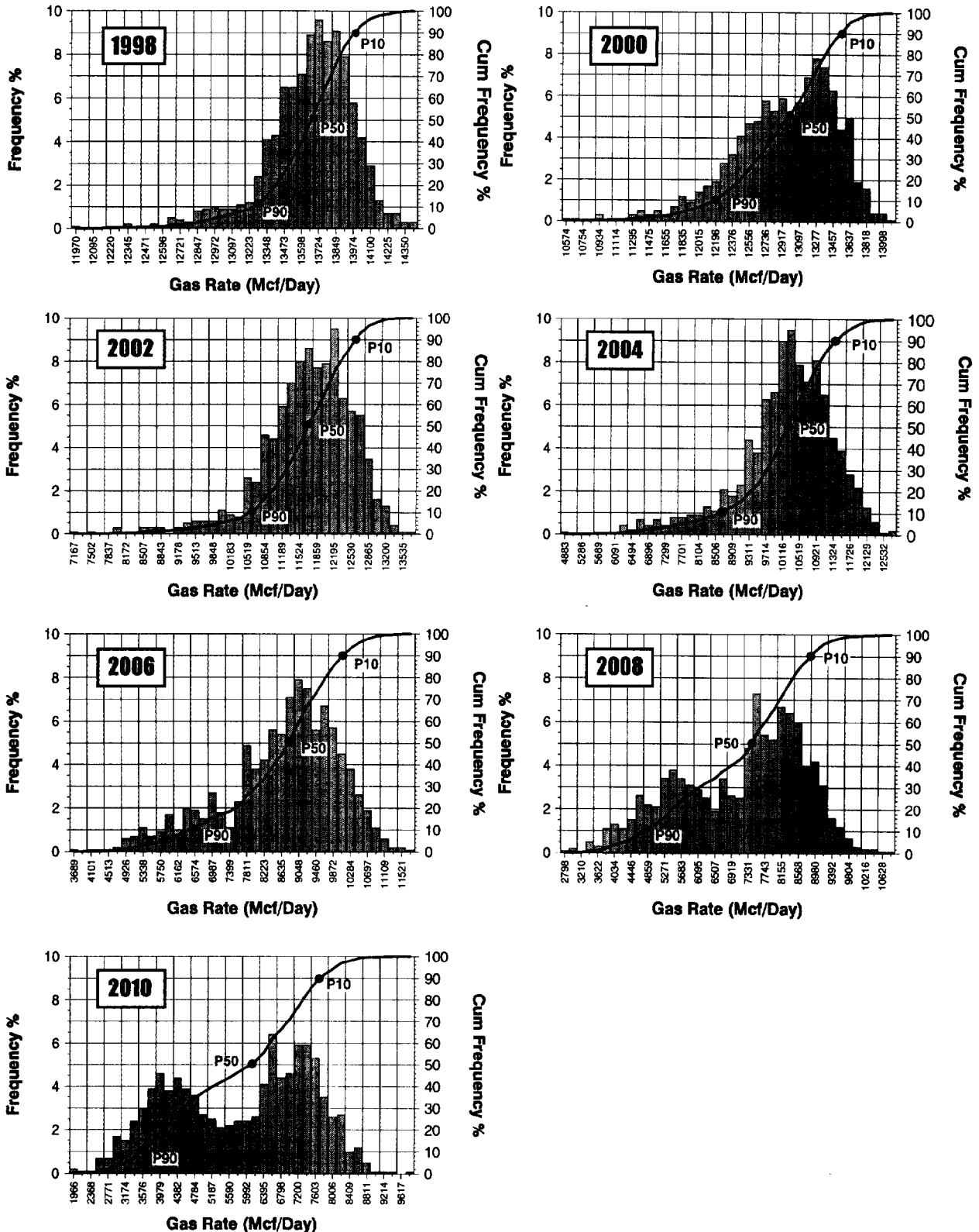


Figure 5—Histograms of total field forecasts for every other year.

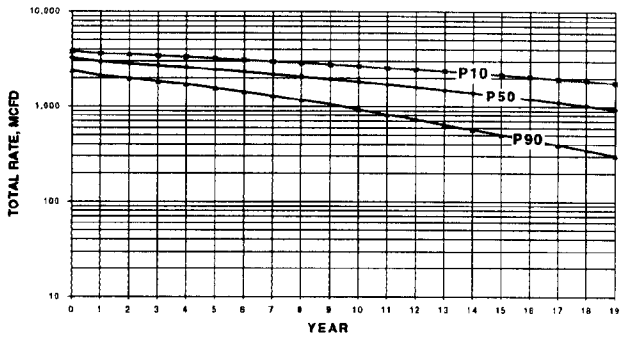


Figure 6—Probabilistic rate forecast, vent gas project.

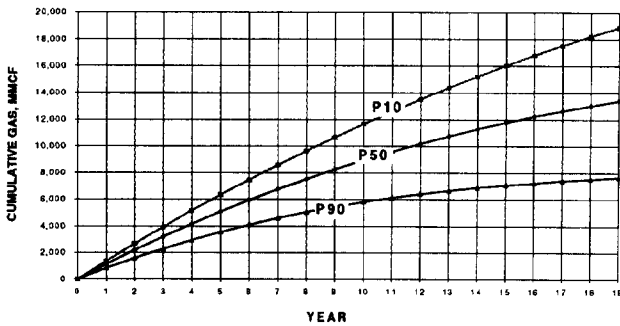


Figure 7—Probabilistic cumulative production forecast, vent gas project.

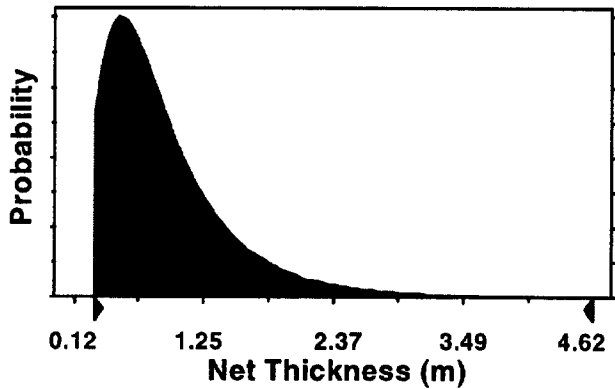


Figure 8—Truncated log normal distribution describing uncertainty in net thickness.

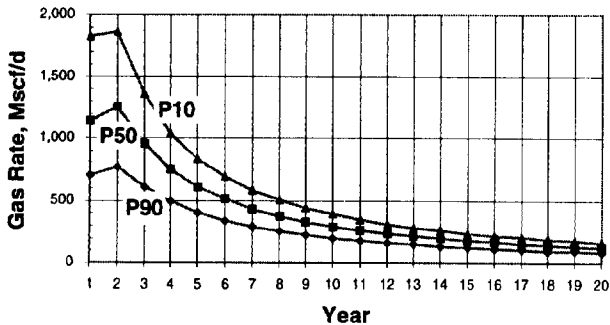


Figure 9—Probabilistic gas rate forecast, region 1.

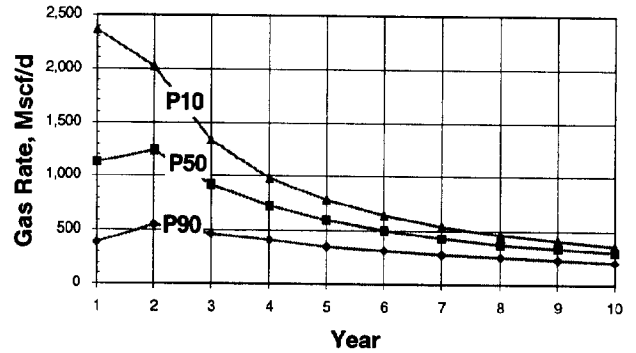


Figure 10—Probabilistic gas rate forecast, region 2.

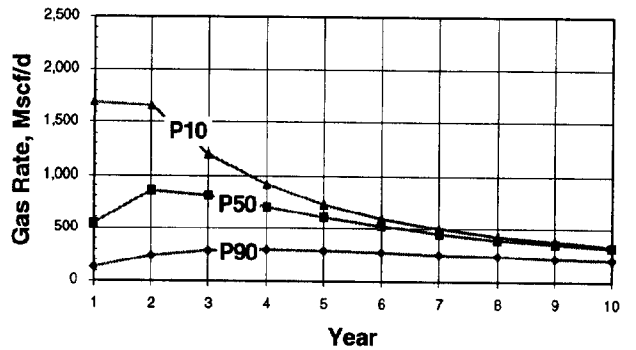


Figure 11—Probabilistic gas rate forecast, region 3.

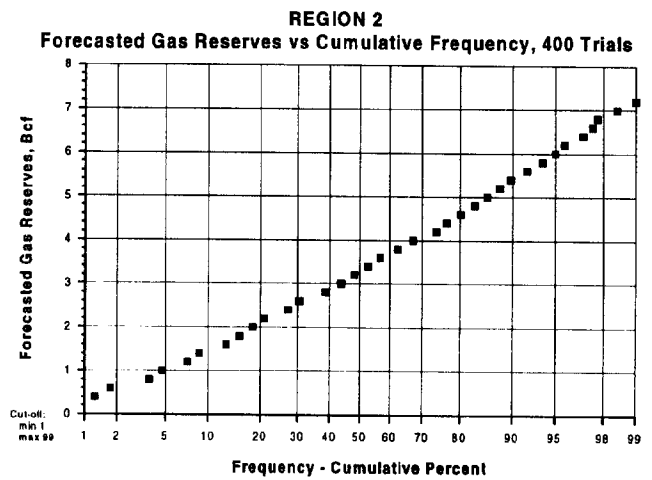


Figure 12—Normal distribution of predicted reserves, region 2.