USING DECLINE CURVE ANALYSIS, VOLUMETRIC ANALYSIS, AND BAYESIAN METHODOLOGY TO QUANTIFY UNCERTAINTY IN SHALE GAS RESERVES ESTIMATES

A Thesis

by

RAUL ALBERTO GONZALEZ JIMENEZ

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Co-Chairs of Advisory Committee,	Duane McVay
	John Lee
Committee Member	Thomas Blasingame
Head of Department,	Dan Hill

December 2012

Major Subject: Petroleum Engineering

ABSTRACT

Probabilistic decline curve analysis (PDCA) methods have been developed to quantify uncertainty in production forecasts and reserves estimates. However, the application of PDCA in shale gas reservoirs is relatively new. Limited work has been done on the performance of PDCA methods when the available production data are limited. In addition, PDCA methods have often been coupled with Arp's equations, which might not be the optimum decline curve analysis model (DCA) to use, as new DCA models for shale reservoirs have been developed. Also, decline curve methods are based on production data only and do not by themselves incorporate other types of information, such as volumetric data. My research objective was to integrate volumetric information with PDCA methods and DCA models to reliably quantify the uncertainty in production forecasts from hydraulically fractured horizontal shale gas wells, regardless of the stage of depletion.

In this work, hindcasts of multiple DCA models coupled to different probabilistic methods were performed to determine the reliability of the probabilistic DCA methods. In a hindcast, only a portion of the historical data is matched; predictions are made for the remainder of the historical period and compared to the actual historical production. Most of the DCA models were well calibrated visually when used with an appropriate probabilistic method, regardless of the amount of production data available to match. Volumetric assessments, used as prior information, were incorporated to further enhance

the calibration of production forecasts and reserves estimates when using the Markov Chain Monte Carlo (MCMC) as the PDCA method and the logistic growth DCA model.

The proposed combination of the MCMC PDCA method, the logistic growth DCA model, and use of volumetric data provides an integrated procedure to reliably quantify the uncertainty in production forecasts and reserves estimates in shale gas reservoirs. Reliable quantification of uncertainty should yield more reliable expected values of reserves estimates, as well as more reliable assessment of upside and downside potential. This can be particularly valuable early in the development of a play, because decisions regarding continued development are based to a large degree on production forecasts and reserves estimates for early wells in the play.

DEDICATION

This thesis is dedicated to my mother, father, and sister for all the love and support they have provided me throughout my life.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. McVay for his teachings and supervision, my committee co-chair, Dr. Lee and my committee member, Dr. Blasingame, for their guidance and support throughout the course of this research.

I would like to thank the Harold Vance Department of Petroleum Engineering at Texas A&M University for giving me the opportunity to pursue a Master's degree. Also, I also want to extend my gratitude to the Crisman Institute for Petroleum Research for providing the funding for this research.

Thanks also go to my friends, colleagues, faculty and staff of the Harold Vance Department of Petroleum Engineering for making my time at Texas A&M University a great experience. Special thanks go to Mr. Xinglai Gong for his patience, support and endless discussion on statistics. I would also like to thank Mr. Jose Romero and Dr. Juan Monge for supporting my goals.

Finally, thanks to my mother, Eneida my father, Raul, and my sister Hazel for their encouragement, patience and love.

TABLE OF CONTENTS

LIST OF FIGURES

Page

LIST OF TABLES

1. INTRODUCTION

1.1 Statement and Significance of the Problem

Reserves estimates and production forecasts in hydraulically fractured horizontal shale gas wells have considerable uncertainty. Major sources of uncertainty in shale gas production forecasts and reserves estimates arise from complex flow geometry, large variability in reservoir and completion properties, and lack of long-term production data. Shale gas reservoirs possess extremely low matrix permeability. For this reason, shale reservoirs require massive hydraulic fracture treatments to become economical (Agarwal et al., 2012). In addition, the desorption dynamics of adsorbed gas are uncertain (Mengal and Wattenbarger, 2011). All of these phenomena result in complex flow geometry, which contributes to uncertainty in production forecasting and reserves estimation.

---

*Part of this chapter (portions of pages 1-2, 14-15) is reprinted with permission from Gonzalez, R.A., Gong, X., and McVay, D.A. 2012. Probabilistic Decline Curve Analysis Reliably Quantifies Uncertainty in Shale Gas Reserves Regardless of Stage of Depletion. Paper SPE 161300 presented at the SPE Eastern Regional Meeting, Lexington, Kentucky, USA, 3-5 October. DOI: 10.2118/161300-MS.

Another challenge to production forecasting and reserves estimation in shale gas reservoirs is lack of long-term production data. The history of drilling horizontal wells with multiple hydraulic fractures in shale gas reservoirs is relatively short, only a few years. Because of their low permeability, shale gas reservoirs require years to reach boundary dominated flow. To the best of our knowledge, only a small number of hydraulically fractured horizontal shale gas wells have experienced boundary-dominated flow. Conventional decline curve analysis (DCA) using Arps' equation requires the analysis of production data from stabilized flow (Arps, 1945). Despite the lack of a stabilized flow period, DCA using Arps' equations coupled with a minimum terminal decline is the preferred methodology to estimate reserves and forecast production in shale gas wells (Lee and Sidle, 2010).

Uncertainty will always be present in reserves estimates and it can be quite large early in the producing lives of hydraulically fractured shale gas wells. Early in the development of a field, producing wells are used as analogous wells for wells that have not been drilled, yet. Hence, while taking development decisions it is important to have accurate production forecast and reserves estimates. To ignore the quantification of uncertainty or to do a poor while quantifying the uncertainty in production forecast and reserves estimates can yield to overconfidence. And if we are trying to identify the most profitable fields to develop, overconfidence could yield to the selection of fields that might not be the most profitable. McVay and Dossary (2012) quantified the cost of underestimating the uncertainty. McVay and Dossary (2012) concluded that by reducing overconfidence other biases and decision error are reduced as well.

1.2 Status of the Question

To analyze production data from hydraulically fractured horizontal shale gas wells, several DCA models have been developed. In addition, PDCA methods have been proposed to quantify uncertainty in production forecasts and reserves estimates.

1.2.1 Decline Curve Analysis Models

DCA using Arps' equations (1945), or Arps with an imposed minimum decline are the preferred methodologies to estimate reserves and forecast production in shale gas wells (Lee and Sidle 2010). However, it might not be the optimal DCA model to forecast production and estimate reserves as several DCA models have been developed to analyze transient production data from shale gas reservoirs, such as: Power Law (Ilk et al., 2008) model, Stretched Exponential (Valko and Lee, 2010) model, Rate-Decline analysis for fracture-dominated shale gas reservoirs (Duong, 2011), and the logistic growth (Clark et al., 2011) model.

Arps' Equations

Arps' equations (Eq. 1) are the most commonly used DCA models to forecast production in oil and gas reservoirs. The most common two forms of the equations are exponential and hyperbolic decline. The exponential case is observed when b is set to be zero and the hyperbolic for b values between zero and one. When b is one, the decline is harmonic.

$$q(t) = \begin{cases} \frac{q_i}{(1+D_i bt)^{\frac{1}{b}}}, & b \neq 0 \\ q_i \exp(-Dt), & b = 0 \end{cases} \quad \text{......} (1)$$

In Eq. 1, $q(t)$ is the production rate as a function of time, Mcf/month, q_i is initial gas production rate, Mcf/month, b is Arp's dimensionless hyperbolic decline constant, and D_i is Arps' initial decline rate, 1/month, and t is the time, month.

Arps' equations were derived empirically. Nevertheless, Fetkovich et al. (1996) derived Arps' equations using the following assumptions: the reservoir is in boundary-dominated flow, there is constant bottomhole pressure, the reservoir fluid is slightly compressible, and the skin factor does not change. An additional constraint is that the b value should remain constant through the life of the well. Nevertheless, because of shale's low permeability, shale gas reservoirs require long times, often years, to reach boundary-dominated flow. Hence, the only available production data to be matched is still on transient flow. When transient data is analyzed using Arps' equations, it can yield b values greater than 1. When a b value greater than 1 is used, it could overestimate reserves estimates (Ilk et al., 2008; Lee and Sidle, 2010). To fix this problem, a minimum decline is often imposed on Arps' equations.

Arps' Hyperbolic Equation with an Imposed Minimum Decline

Arps' hyperbolic equation with a minimum decline is an empirical relation used to account for wells in which a b factor greater than 1 is used to match the production data. It can be proved that if a b factor greater than 1 is used, the cumulative production goes to infinity as times goes to infinity. In the other hand, if a b factor less than or equal to 1 is used to match the production data, the forecasted production rate eventually goes to zero and the cumulative production reaches a limit.

To solve the problem of infinite cumulative production, an exponential tail is commonly imposed on the hyperbolic fitting by imposing a minimum decline. By doing this, it can ensured that a finite cumulative production is forecasted. Nevertheless, the correct minimum decline rate is unknown, as most of the reservoirs that need a b factor greater than 1 to match the data, such as shale gas reservoirs, are relatively new and little production data have been acquired. The limitations of this model are the same as with Arps' equations.

Modified Arps' Equation

The modified Arps' equation (Eq. 2) is based on Arps' equations and a fourth parameter defined as the time in which the production rate goes to an exponential decline. At time T_o, the model calculates the instantaneous decline, and the model an exponential decline after T_o.

$$q(t)=\begin{cases}\dfrac{q_i}{(1+D_ibt)^{\frac{1}{b}}},\ t\le T_0\\[2ex]\dfrac{q_i}{(1+D_ibT_0)^{\frac{1}{b}}}\exp\left[\dfrac{-D_i(t-T_0)}{1+bD_iT_0}\right],\ t>T_0\end{cases}\quad\text{.......................................}(2)$$

In Eq. (2), T_o is the modified Arps' time required to go into exponential decline, months.

The modified Arp's equation describes both the early and the latest trend from the production data. For example, if the data initially displays a hyperbolic decline, but the latest trend is an exponential decline, the model will provide a model that fits the early hyperbolic trend and the latest exponential decline. On the other hand, if the data follows an exponential decline the model will use T_o as zero and it will present a match based on an exponential decline. Another scenario is that the production data does not exhibit an exponential decline. In this case, T_o will be extremely large and the behavior fitted by the model will be only hyperbolic. This scenario can be modified by manually setting T_o, as the last month of the production data. In this case the model will follow and exponential.

The modified Arps' model can accommodate transient data, provided latest trend form the production data exhibits an exponential decline. If the latest trend is still in hyperbolic decline, the model by itself does not solve the problem of the cumulative production going to infinity when a b greater than 1 is used to match the data. The user can set T_o, at the end of the data to solve this problem and provide a conservative estimate for cumulative production.

Arps' equations and its variants have worked well for the industry for years. Nevertheless, with the boom of shale gas newer models have been proposed to analyze the transient data. Several deterministic decline curve models have been developed specifically for shale gas reservoirs: Power Law (Ilk et al., 2008) model, Stretched Exponential (Valko and Lee, 2010) model, Rate-Decline analysis for fractured-dominated shale gas reservoirs (Duong, 2011), and logistic growth (Clark et al., 2011) model.

Power Law Model

The Power Law model (Eq. 3) was developed by Ilk et al. (2008). The model is based on a power law loss ratio. The loss ratio was modeled "by a decaying power law function with a constant behavior at large times" Ilk et al. (2008). The constant behavior at large times is described by constant decline parameter, D_∞. By having a four parameter model, the model can describe transient and boundary-dominated flow.

$$q(t) = \hat{q}_i \exp\left(-D_\infty t - \hat{D} t^n\right) \qquad (3)$$

In Eq. (3), $\hat{D}$ is the power law decline constant, 1/month, D_∞ is the power law decline at infinite time constant, 1/month, and n is dimensionless time exponent.

Mattar (2008) analyzed public monthly production data of gas wells from the Barnett Shale with the Power Law model. Five wells were analyzed with available production data from three to seven years. It was found that the model fit the data and that it was able to model the different observed flow regimes. They also found that the Power Law model worked well with simulated data from horizontal wells with multiple fractures. Also, they found that for their limited scenarios the inclusion of D_∞ did not affect the matches. For all the scenarios the D_∞ value was on the order of 10^{-5}. Proper estimation of D_∞ would require the matching of longer periods of production data.

Stretched Exponential Production Decline

The stretched exponential production decline (SEPD) model (Eq. 4) was introduced by Valko and Lee (2010). The SEPD model assumes a large number of volumes individually decaying exponentially. It can be proved, by rearranging the variables and the elimination of the D_∞, that it is equivalent to the Power Law (Ilk et al., 2008). The difference is the approach taken and the objectives of the models.

$$q(t) = q_i \exp\left[-\left(\frac{t}{\tau}\right)^{\eta}\right] \qquad (4)$$

In Eq. (4), η is a dimensionless exponent parameter and τ is the characteristic time parameter, months.

Can and Kabir (2012) analyzed production data from 820 wells from three different shale formations (220 wells in the Bakken oil shale, 100 wells in the Marcellus shale and 500 wells in the Barnett Shale). The wells analyzed were both oil and gas wells. The amount of data being analyzed was seven years of synthetic data and at least a year for real production data. They concluded that the SEPD model is better than Arps' model to estimate reserves in unconventional reservoirs, because it is a bounded model, and did not overestimate reserves.

Rate-Decline Analysis for Fracture Dominated Shale Reservoirs

The rate-decline analysis for fractured dominated shale reservoirs model (Eq. 5) was developed by Duong (2011). The driving assumption for Duong's model is long-term linear flow. Matrix contribution to the EUR is negligible when compared to fracture contribution to production forecast and EUR estimations (Duong, 2011). The model is based on three variables, from which two of them are strongly correlated. Therefore, we can estimate one of them, making it a two parameter model. One of the key features is that the model is bounded as the production rate eventually goes to zero.

$$q(t) = q_i t^{-m} \exp\left[\frac{a}{1-m}\left(t^{1-m} - 1\right)\right] \qquad (5)$$

In Eq. (5), a is the intercept constant for Duong's model, 1/month and m is the dimensionless slope for Duong's model.

Duong (2011) used different types of wells to validate his model: tight/shallow, dry shale and wet shale gas. He tested for dry shale gas a sample of 25 wells from the Barnett shale. The amount of data used and the other sources were not described in the paper. Duong (2011) found that for various shale plays a varied from 0 to 3 and m from 0.9 to 1.3. Duong (2011) concluded that hi model provided more conservative estimates for cumulative production than the Power Law (Ilk et al., 2008) and Arps' hyperbolic model.

Logistic Growth Model

The logistic growth model was developed by Clark et al. (2011). The model is based on the logistic growth curves with are used to forecast growth (for example cumulative oil or gas production). The logistic model is based on 2 or 3 parameters. The third parameter depends on the knowledge of the estimated ultimate recovery (EUR), which can be obtained independently from a volumetric estimate (recoverable amount of production without economic constraints). The model is bounded by the EUR. However, when the EUR is not known, the solution is non-unique, as multiple good matches can be fitted.

$$q(t) = \frac{K n_L a_L t^{n_L - 1}}{(a_L + t^{n_L})^2} \quad (6)$$

In Eq. (6), K is the carrying capacity (EUR) Mcf, n_L is the dimensionless decline exponent, and a_L is the time to the power n at which half of the carrying capacity has been produced, months.

Clark et al. (2011) tested the model on a sample of 600 wells. Their original sample included 1000 wells. Nevertheless, 400 wells exhibit unreasonable values for any of the parameters. The authors did not provide any criteria to describe values that were considered unreasonable. The wells they used were completed through January 2004 and December 2006. For the given sample they estimated *K* to be log normal distributed with a minimum value of 1.3 Mcf, a maximum value of 9 Bcf and a mean value of 1.7 Bcf. They also estimated a_L to be normally distributed with a minimum value of 7 months, a maximum value of 153 months and a mean of 33 months. Finally, they estimated n_L to be log normal distributed with a minimum value of 0.1, a maximum value of 1.3 and a mean of 0.9. They concluded that the logistic model estimates for cumulative production are more modest that the ones provided by Arps' equations.

1.2.2 Probabilistic Decline Curve Analysis Models

All the DCA models described above are deterministic and do not quantify the uncertainty in production forecasts and reserves estimates. Capen (1976) pointed out that by performing a probabilistic analysis, we should obtain better estimates for the mean than following a purely deterministic methodology. Probabilistic decline curve analysis (PDCA) methods have been proposed to quantify the uncertainty in production forecasts and reserves estimates in conventional and unconventional reservoirs. Three PDCA methods published in the literature are the bootstrap method (JSM) developed by Jochen and Spivey (1996), the modified bootstrap method (MBM) developed by Cheng et al. (2010), and the Markov Chain Monte Carlo method (MCMC) developed by Gong et al. (2011).

The Bootstrap Method or Jochen and Spivey Method (JSM)

The JSM relies only on the DCA of synthetic data sets. The synthetic data sets are created by resampling the data. A data set is created by picking production data points at a random time to replace others at a random time. From the DCA of the synthetic data sets a distribution for reserves estimates and production forecast is created.

Bootstrapping as a sampling technique requires no time dependency between the data. However, this is not true for production data; production data is time dependent and should not be treated as independent events. This problem was addressed by Cheng et al (2010).

Modified Bootstrap Methodology

The MBM relies only on DCA of synthetic data sets. The MBM creates synthetic data sets based on data that is not time dependent. The creation of the synthetic data sets is based on blocks of residuals obtained from production data and the best fit from any DCA model. The residuals are blocked to take into account the log difference in production data.

A three stage backward approach was introduced to take into account the analysis of transient data. The backward approach eliminates a certain amount of data, which is considered to be in transient flow, from the synthetic data sets. The first stage of the backward approach involves performing DCA on the most recent 50% of the synthetic data sets. The second and third stage analyzes the most recent 30% and the 20% of the synthetic data sets respectively. The percentages used for the backward approach were estimated, calibrated and tested on conventional mature wells. Cheng et al. (2010) data set consisted of 100 oil and gas wells with no apparent changes in development strategy or production operations.

From the DCA of the three stages, for a total of 360 synthetic data sets, three distributions for reserves estimates and production forecasts are created. The reserves' distribution P_{90} is the minimum P_{90} from the three distributions from the stage analysis. The P_{50} is the mean of the three P_{50} from the distributions from the stage analysis, and the P_{10} is the maximum P_{50} from the three distributions from the stage analysis. The MBM method was demonstrated to be well calibrated for conventional reservoirs and unconventional reservoirs (Gong et al., 2011).

Markov Chain Monte Carlo Methodology

The MCMC method is based on Bayes' theorem (Eq. 7), more specifically on Bayesian inference.

$$\pi(\theta_j|y) = \frac{f(y|\theta_j)\pi(\theta_j)}{\int f(y|\theta)\pi(\theta)d\theta} \quad (7)$$

In Eq. (7), $\pi(\theta|y)$ represents the posterior distribution, $f(y|\theta)$ a likelihood function, $\pi(\theta)$ the prior distribution of parameters. Gong et al. (2011) allowed θ_j to be a candidate of the DCA parameter (j = 3 to 4 depending on the DCA model) while y is the available production data for matching. A Bayesian study's goal is to estimate the posterior distribution, after a certain amount of data has been observed (Gong et al., 2010). Gong et al. (2010) set the DCA parameters to be random variables. Therefore, they wanted to evaluate the distribution of DCA parameters after all the available production data was observed.

The prior distribution is the distribution of DCA parameters before any production data have been analyzed. The likelihood function is the conditional probability of the available production data given the DCA parameters. The posterior distribution is the distribution of the DCA parameters after all the available data have been taken into account. Production forecast and reserves estimates can be obtained from the distribution of DCA parameters. Further details regarding the formal definition of these distributions can be found in Gong et al. (2011).

Gong et al. (2011) used a random walk algorithm for MCMC sampling. In this algorithm samples are drawn from a proposal distribution. The proposal distribution is not necessarily related to what we know before we draw the samples; it is simply a starting point. The proposal distribution should not be confused with the prior distribution. The prior distribution is what we believe the distribution should be before any additional information, and it is part of the calculation of the posterior distribution. In other words, a proposal distribution is just a starting distribution to sample from, while the prior distribution is part of the posterior distribution calculation and is more important. Properties of the proposal distribution can be found in Gong et al. (2011). The standard deviation of the proposal distribution for each DCA model parameter is the only parameter needed to be specified.

Gong et al. (2011) created a likelihood function which honors the quality of the fit provided by each sample of decline curve parameters. Gong et al. (2011) assumed that the difference between the log of the available production data and the log of the predicted production from any given DCA model follows a standard normal distribution (Eq. 8).

$$f(y|\delta_i \dots \delta_n) = \frac{1}{2\pi\sigma} \exp\left(-\frac{\sigma^2_{\text{proposal}}}{\sigma^2}\right) \quad (8)$$

In Eq. (8), δ_i is the candidate for DCA parameter (δ_i for i = 1 to 3 or 4 depending on the DCA). It is fairly simple to integrate information from independent assessments when using the MCM. Gong et al. (2011) used a uniform prior distribution for DCA parameters. The prior distribution can be enhanced to accommodate other types of information, such as volumetric.

Of the three described PDCA methods, according to the literature only the MBM and the MCMC have been tested on shale gas wells (Gong et al., 2011). The MCMC method performed faster and provided narrower uncertainty ranges.

There is a gap in the literature regarding how well the MBM and the MCMC methods perform when limited production data are available. In the literature, the MBM and the MCMC have undergone only one test on a dataset from hydraulically fractured horizontal shale gas wells. The dataset included Barnett shale wells with producing times ranging from 50 to 120 months. Most of the newly developed shale fields do not possess wells with such long histories.

One of the weaknesses of MBM and the MCMC is that both were developed and tested in the literature using Arps' equations. In the literature, neither method has been tested with other DCA models. Arps' equations might not be the optimal DCA model as new DCA models have been developed for unconventional reservoirs. Using PDCA methods with DCA models developed specifically for shale gas wells could enhance reserves estimates and provide better quantification of uncertainty.

The advantage of performing PDCA is that it relies solely on the analysis of production data. Some authors have combined estimation methods to decrease the uncertainty and provide more accurate reserves estimates. Typical Bayesian applications in the oil industry involve updating probability estimates for reserves from volumetric data (static data) as additional information is gathered from production data (dynamic data). Ogele et al. (2006) and Aprilia et al. (2006) coupled volumetric assessments with material balance equations to better quantify the uncertainty and produce more accurate estimations of oil initially in place (OIIP) and gas initially in place (GIIP) respectively. Both MBM and the MCMC rely on the analysis of production data only and do not take advantage of all the known information regarding the reservoir. A more robust probabilistic method should integrate other sources of information to reduce uncertainty.

In summary, the main limitations of performing PDCA on hydraulically fractured shale gas wells are:

- PDCA methods have been tested using fairly large amounts of production data (5 to 10 years), and limited work is available on their performance outside those time ranges.
- PDCA methods have been tested using one DCA model (Arps' equations) that might not be the best available for analysis of production data from hydraulically fractured horizontal shale gas wells
- Currently there is no published method to integrate other types of information into PDCA production forecasts and reserves estimates.

1.4 Research Objectives

The objectives of this research are to:

1. Determine the probabilistic methods and decline curve models that reliably quantify the uncertainty in production forecasts from hydraulically fractured horizontal shale gas wells, regardless of the stage of depletion.

2. Develop a Bayesian method to integrate volumetric information with decline curve analysis to enhance quantification of uncertainty in production forecasts from hydraulically fractured horizontal shale gas wells.

1.5 Overview of Methodology

A systematic study of different combinations of PDCA methods, DCA models, amounts of production data available for matching, and availability of volumetric data was performed. The two best PDCA methods on unconventional wells were studied with all the DCA models. The best performing PDCA method from this analysis was evaluated with different amounts of production data available to match. The best combination overall of PDCA and DCA methods was enhanced by incorporating volumetric information.

2. VALIDATION OF PUBLISHED PROBABILISTIC DECLINE CURVE ANALYSIS METHODS

2.1 Methodology

Three published PDCA methods the JSM (Jochen and Spivey, 1996), the MBM (Cheng et al., 2010), and the MCMC (Gong et al., 2011) were studied. A validation study was performed on the implementation of the JSM and the MBM based on Cheng et al. (2010) dataset. The MCMC was also tested on a conventional data set of wells because there is no published work on its applications on conventional wells. The dataset studied contained mature wells with useable production data ranging from 48 to 335 months, with an average of 198 months.

A hindcast study was performed for each of the PDCA methods using Arps' equations as the decline curve model. In a hindcast study a portion of the known history is matched and the remainder of the known history is considered to be "future" production and is

---

*Part of this chapter (portions of pages 20-22) is reprinted with permission from Gonzalez, R.A., Gong, X., and McVay, D.A. 2012. Probabilistic Decline Curve Analysis Reliably Quantifies Uncertainty in Shale Gas Reserves Regardless of Stage of Depletion. Paper SPE 161300 presented at the SPE Eastern Regional Meeting, Lexington, Kentucky, USA, 3-5 October. DOI: 10.2118/161300-MS. Copyright [2012] by Society of Petroleum Engineers.

compared to the model's prediction to the same point in time (i.e., hindcast).50% of the known history hindcast were used, in which 50% of the known data was matched and the rest 50% of it was compared against the predictions from each of the PDCA models.

The probabilistic methodology was conducted independently for each well in the data set. The individual wells' P_{90}, P_{50}, and P_{10} predictions for the JSM and the MBM were determined from the distribution of each well's cumulative production at the end of the hindcast period (CPEOH). The individual wells' P_{90}, P_{50}, and P_{10} predictions for the MCM were based on the CPEOH period determined from the distribution of DCA parameters. However, a single well prediction cannot be used to evaluate the reliability of a PDCA method. Hence, a sample of wells must be analyzed to determine if a PDCA method is well calibrated. We believe there is likely correlation among the wells. Nevertheless, for simplicity we assumed that the wells were perfectly correlated. Therefore, the P_{90}, P_{50}, and P_{10} values for the set of wells, were calculated by adding the individual-well P_{90}, P_{50}, and P_{10} estimates, respectively, assuming the wells' individual estimates are perfectly correlated (r=1).

The coverage rate (CR) was used to assess the calibration of the PDCA methods. The CR is the number of wells in which the actual production falls within the P_{90}-P_{10} range divided by the total number of wells. For a well calibrated methodology generating 80% confidence intervals (C.I), approximately 80% percent of the wells would bracket the true cumulative production within its C.I.

A calibration curve was also used to assess the calibration of the PDCA methods (**Fig. 2.1)**. In a calibration curve, the x-axis represents the probability associated with each forecast, such as the P_{90}, P_{50}, or P_{10} estimates. The y-axis represents the fraction of wells that comply with the definition. For example, the well actual production at the end of hindcast should be greater than the P_{10} estimate for approximately 10% of the wells being analyzed. The fraction of wells in which the actual production exceeded the P_{10} estimate was calculated and plotted on the y-axis. This process was repeated for the P_{50} and the P_{90} estimates.

A methodology is well calibrated if, for all the estimates assigned the same probability, the proportion correct is equal to the probability assigned (Fig. 2.1**)**. For high probabilities (above P_{50}), underconfidence occurs when the probability assigned is lower than the proportion correct and overconfidence occurs when probability assigned is greater than the proportion correct. On the other hand, for low probabilities (below P_{50}), underconfidence occurs when the probability assigned is higher than the proportion correct and overconfidence occurs when probability assigned is lower than the proportion correct for high probabilities. In other words, being overconfident means having narrower ranges for our production forecasts.

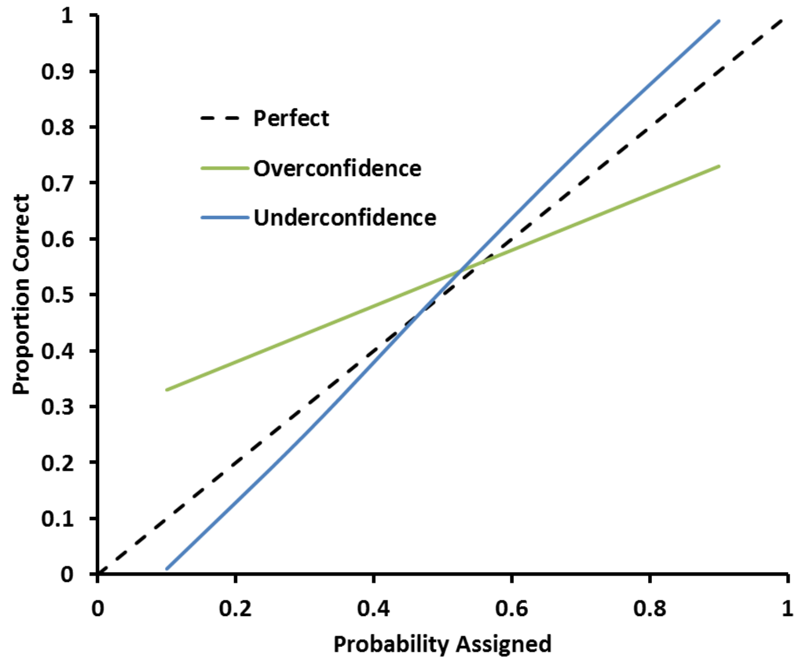


Fig. 2.1—Probability calibration plot. Well calibrated results should lie on the unit-slope line.

Two different options to match the DCA to the known production data: Cartesian and logarithmic regressions were used. The Cartesian regression is the most common regression used. This regression has the disadvantage that it honors higher production data point more than lower production data points. For gas wells that decline extremely fast in the first months, using Cartesian regression could yield to matches that do not honor the late production behavior. The logarithmic regression rate was introduced to account for gas wells that decline extremely fast within the first months. While using the log scale all the data available is honored.

2.2 Validation of MBM and Evaluation of MCMC on Conventional Wells

A validation studied was conducted for the MBM method for the same data set as Cheng et al. (2010). It was recommended to achieve stable results to use 120 bootstrap realizations (Cheng et al., 2010). An evaluation studied was conducted on the same data set for the MCMC method. Gong et al. (2011) used log-normal proposal distributions for D_i and q_i with given standard deviations of 0.4 and 0.2 for the proposal distribution. Gong et al. (2011) used a normal proposal distribution for b with a given standard distribution of 0.2 for the proposal distribution, and a length for the Markov chain of 2000. The bounds for Arps' equations were taken from Gong et al. (2011), which were in accordance to Cheng et al. (2010). The prior distributions for decline curve parameters were taken from Gong et al. (2011) as independent, uninformative uniform distributions.

The validation results (**Table 2.1**) are consistent with Cheng et al. (2010) results. As the creation of the synthetic data sets is random, the probability to obtain the exact same results twice is really small. The wells from which the predictions differ are wells that demonstrate the most spurious behavior and longest production. By having the longest production the amount of synthetic data sets that can be selected increases considerably and yields to a larger amount of possible predictions. However, the results are stable within an acceptable range as demonstrated by the low error value obtained. The MCMC method was equally well calibrated as the MBM. It was concluded that the MCMC and the MBM are both equally well calibrated for the given sample of mature conventional wells. The MBM and the MCMC are well calibrated for our sample of conventional wells. In, accordance to high coverage rate, a good calibration was achieved (**Fig. 2.2**). On the other hand, the JSM clearly shows overconfidence.

Table 2.1—Comparison of PDCA Methods on a Sample of Conventional Wells Using Arps Equations

	Cheng et al. (2010)	MBM	MCMC
No of Wells	100	100	100
Coverage rate	83%	84%	81%
Relative error: Average((P_{50}-CPEOH)/CPEOH)	15.44%	12.41%	-2.25%
Absolute Error: Average(|P_{50}-CPEOH|/CPEOH)	29.17%	28.57%	28.22%
Average(C.I./P_{50})	0.9566	1.0135	1.0676
True CPEOH, MSTBOE	4,557.24	4,295.22	4,204,459
Sum of P_{50} values, MSTBOE	4,114.54	4,114.54	4,114.54
Percentage Error in CPEOH	10.76%	4.38%	2.16%

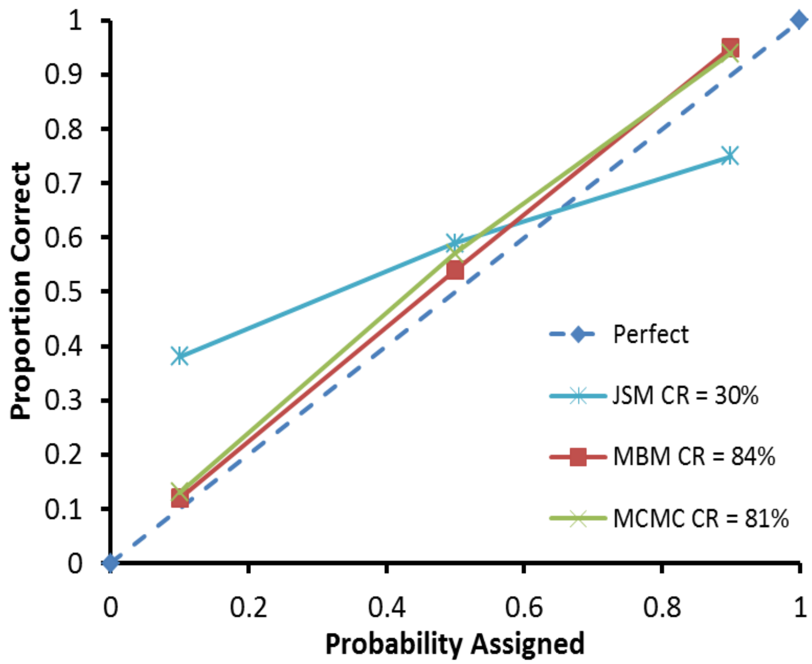


Fig. 2.2— MBM and MCMC PDCA methods provide the best calibration for the tested sample of conventional wells. Accordingly, the MBM and the MCMC provide the best coverage rate.

3. PERFORMANCE OF MULTIPLE PROBABILISTIC DECLINE CURVE ANALYSIS METHODS ON SHALE GAS RESERVOIRS

3.1 Selection of the Sample Data

The performance and benefits of using the JSM, MBM, MCMC PDCA methods in a sample of Barnett shale gas wells was evaluated. Three DCA models widely used in the industry: Arps, modified Arps, Arps with a minimum decline and four DCA models presents in the literature specially developed for shale gas resources: Power Law (Ilk et al., 2008), Stretched Exponential (Valko and Lee, 2010), Rate-Decline analysis for fractured-dominated shale gas reservoirs (Duong, 2011), and logistic growth Model (Clark et al., 2011) were tested. The objective of the analysis was to evaluate how the probabilistic methods compare to a deterministic approach, how well calibrated were the 80% confidence intervals results over time, their width and the overall calibration of the PDCA methods.

---

*Part of this chapter (portions of pages 27, 33-36, 38-46) is reprinted with permission from Gonzalez, R.A., Gong, X., and McVay, D.A. 2012. Probabilistic Decline Curve Analysis Reliably Quantifies Uncertainty in Shale Gas Reserves Regardless of Stage of Depletion. Paper presented at the SPE Eastern Regional Meeting, Lexington, Kentucky, USA. Society of Petroleum Engineers SPE-161300-MS. DOI: 10.2118/161300-ms.

Public production data from 197 horizontal wells from the Barnett shale was collected (**Fig. 3.1**). The Barnett shale was chosen because it is one of the longest-producing plays where multi-stage hydraulic fractures have been performed on horizontal wells. The chosen wells were active gas wells from Denton, Tarrant and Wise counties in Texas. Production data was edited for wells that had undergone an obvious stimulation/recompletion treatment. In wells in which the stimulation process appeared to be performed near the beginning of the known history, the starting date was shifted to model the dominant decline trend (**Fig. 3.2**). With these exclusions, the wells had useable production data ranging from 59 to 119 months, with an average of 80 months.

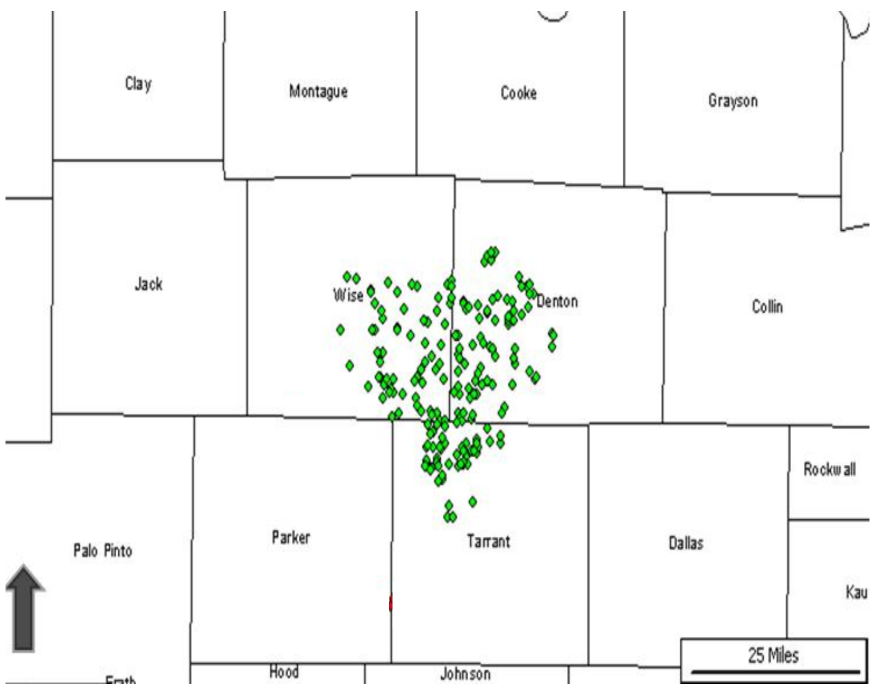


Fig. 3.1—Barnett shale gas wells location.

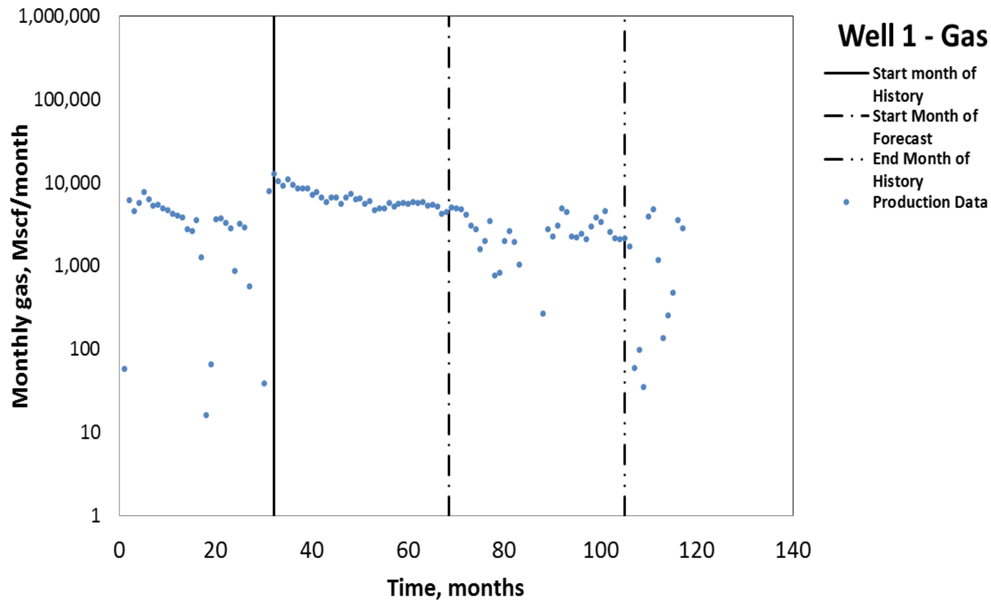


Fig. 3.2—Hindcast example of edited Gas well No. 1

3.2 Introduction to the Multimodel

The Multimodel was a simple and fast method developed during this research to assess the difference in the matches and forecasts provided by each DCA model. The method relies on the different matches that different DCA models provide. In the multimodel, for each well, the best fit using the following six DCA models: Arps, modified Arps, Power Law, SEPD, Duong and logistic growth model was calculated.

After the matches were obtained, the CPEOH was ranked for each of the DCA. The highest hindcast cumulative production obtained at the end of hindcast was considered to be the P_{10}. The lowest hindcast cumulative production obtained at the end of hindcast was considered to be the $P_{90.}$ The P_{50} was assumed to be the 4^{th} highest hindcast

CPEOH. The choice to use the 4th highest hindcast CPEOH was based on the performance of the probabilistic method. Hence, the selection was based on the procedure that brought the P_{50} closer to the actual CPEOH for the field.

3.3 Evaluation of PDCA Methods in Shale Gas Fields

A hindcast study for the MultiModel, JSM, MBM and MCMC models using Arps' equations on a single gas well was performed. The main objective was to determine the benefits of using PDCA methods over deterministic methods. And also to evaluate which PDCA methods work better on shale gas wells. The parameters to apply each PDCA method are the same ones as in Chapter II.

Arps' equations bracket the CPEOH of a single gas well (No. 47) when used with the JSM, the MBM and the MCMC method (**Table 3.1)**. The MultiModel did not bracket the CPEOH and yielded the narrowest C.I. (**Fig. 3.3a**). The JSM method brackets the CPEOH and gave the second narrowest C.I. results (**Fig. 3.3b**). The MBM also brackets the CPEOH and it yielded the widest C.I. (**Fig. 3.3c**). The MCMC PDCA method yielded the results that are closer to the actual CPEOH and provided the second wider C.I. (**Fig. 3.3d**). The MCMC, for this particular well, provided the closest estimate to the actual CPEOH, even better than the deterministic outcome.

Table 3.1—Comparison of PDCA Methods on Gas Well 47 Using Arps Equations				
	MultiModel	**JSM**	**MBM**	**MCMC**
Actual CPEOH, Mcf	182,825	182,825	182,825	182,825
Arps' Equations Best Fit, Mcf	121,124	121,124	121,124	121,124
CPEOH P90, Mcf	107,500	49,838	8,733	93,381
CPEOH P50, Mcf	117,992	102,528	102,630	167,876
CPEOH P10, Mcf	142,689	191,114	190,802	263,664
C Mean, Mcf		114,019	103,124	175,133

The reliability of a PDCA method cannot be based on a single well hindcast. C.I.s are probabilistic results and their reliability cannot be based on a single outcome. Capen (1976) stated that the proper method to evaluate the reliability of a probabilistic method is by evaluating multiple predictions, under the same conditions, over time. Hence, a hindcast study using Arps' equations as the DCA model was performed on the sample of wells from the Barnett shale to evaluate whether the PDCA methods are well calibrated and if they provide reliable quantification of uncertainty.

The MBM and the MCMC are the best calibrated PDCA methods (**Table 3.2**). The MBM and the MCMC have the closest coverage rate to the expected 80%, with 77% and 79% respectively. The MultiModel and the JSM provided poor coverage rates of around 30%. The MCMC yielded the smallest average absolute error between the estimated P_{50} and the actual CPEOH. The MultiModel yielded the smallest C.I in average, while the MBM yielded the widest C.I. in average. The MBM and the MCMC provided results that are closer to the sample of well's CPEOH, even better than the ones from a deterministic outcome.

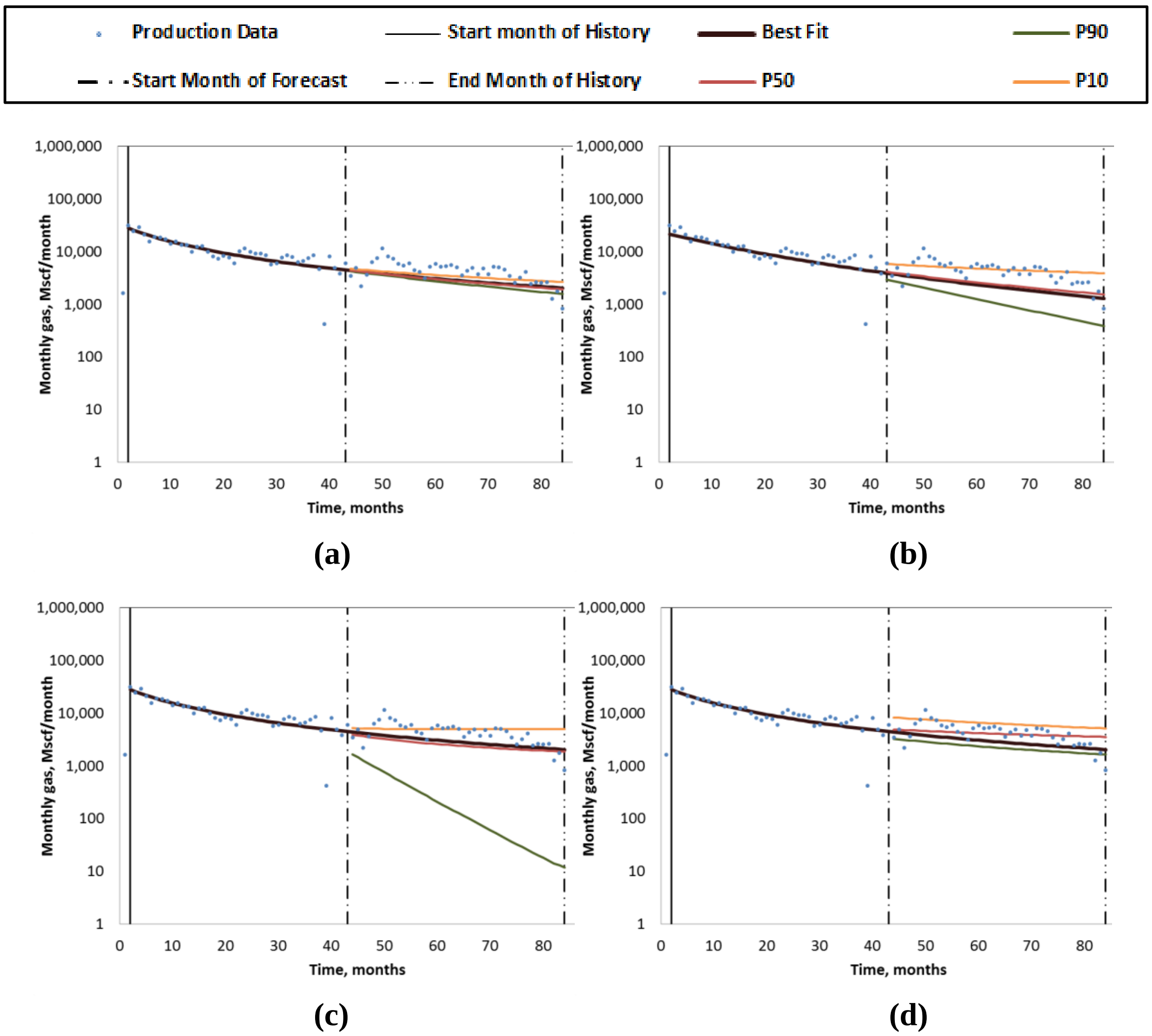


Fig. 3.3—PDCA on gas well 47, a) MultiModel, Arp's equations and b) JSM, c) MBM, d) MCMC. PDCA methods bracket actual production data in a single gas well hindcast.

Table 3.2—Comparison of PDCA Methods on a Sample of Barnett Shale Gas Wells Using Arps Equations					
	MULTI	**JSM**	**MBM**	**MCMC**	**Deterministic**
No of Wells	199	199	199	199	199
Coverage rate	30%	31%	77%	79%	N/A
Relative error: Average((P_{50}-CPEOH)/ CPEOH)	-0.91%	-5.07%	0.76%	1.00%	-1.23%
Absolute Error: Average(|P_{50}- CPEOH |/CPEOH)	25.76%	26.96%	29.82%	20.93%	25.49%
Average(C.I./P_{50})	0.3202	1.0329	1.0402	0.8031	N/A
True CPEOH, MSCF	77,528,345	77,528,345	77,477,694	77,528,345	77,528,345
Sum of P_{50} values, MSCF	77,515,554	75,933,759	77,809,417	78,003,468	78,151,911
Percentage Error in CPEOH	0.02%	2.06%	0.43%	0.61%	0.80%

In accordance to low coverage rates, poor calibration was obtained for the JSM and Multimodel (Fig. 3.4). The MultiModel and the JSM clearly show underconfidence. The MBM and the MCMC are well calibrated for our sample of wells.

Based on the results from **Table 3.2** it was decided to discard the MultiModel and the JSM as candidates for further study. Hence, the application of multiple DCA models was only evaluated when coupled to the MBM and the MCMC as PDCA methods.

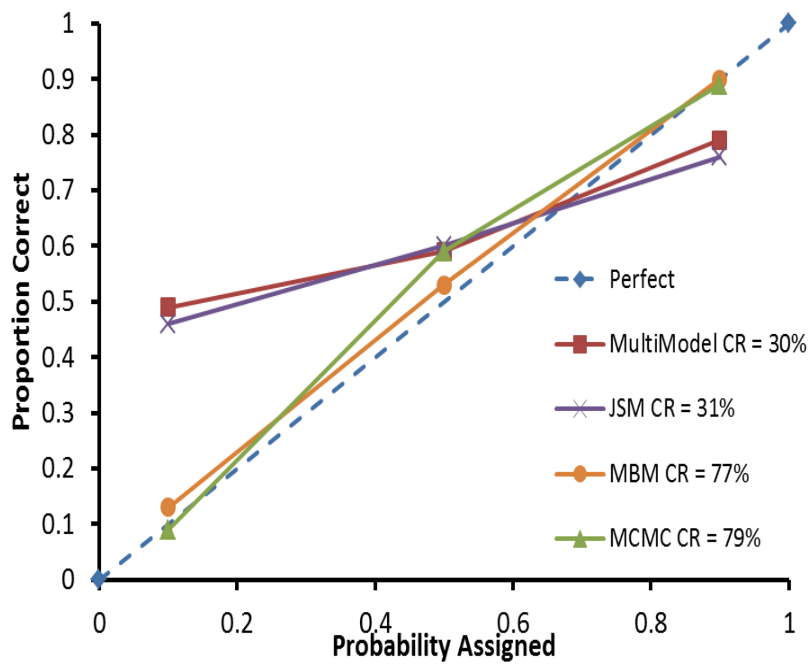


Fig. 3.4—MBM and MCMC PDCA methods provide the best calibration for the sample of Barnett shale gas wells. In accordance, the MBM and the MCMC provide the best coverage rate.

3.4 Evaluation of Multiple DCA Models with MBM and MCMC as PDCA Methods

A 50% of the known history hindcast study was performed for all the reviewed DCA models with MBM and the MCMC methods. The bounds for the DCA parameters (**Table 3.3**) were defined wide enough to accommodate all realistic scenarios provided by the synthetic data sets for MBM and the matches for MCMC. The MCMC requires that a prior distribution be specified for each of the decline curve parameters used in each of the DCA models. The prior distributions for decline curve parameters were defined as independent, uninformative uniform distributions. The bounds for these uniform prior distributions (**Table 3.3**) were also used in the likelihood function; no values outside these bounds were considered. The bounds for all the variables are wide enough so any reasonable combination of decline curve parameters can be considered.

In the MCMC method the proposal distribution, given the previous step candidates of DCA parameters is $\delta_{i,\mathrm{proposal}} \sim N(\delta_{i,t-1}, \sigma_i)$. The only unknowns are the standard deviations for each of the decline curve parameters. Hence, the MCMC requires that a proposal distribution be specified for each of the decline curve parameters used in each of the DCA models. The standards deviations were chosen to obtain the best calibrated results in the MCMC method (**Table 3.4**).

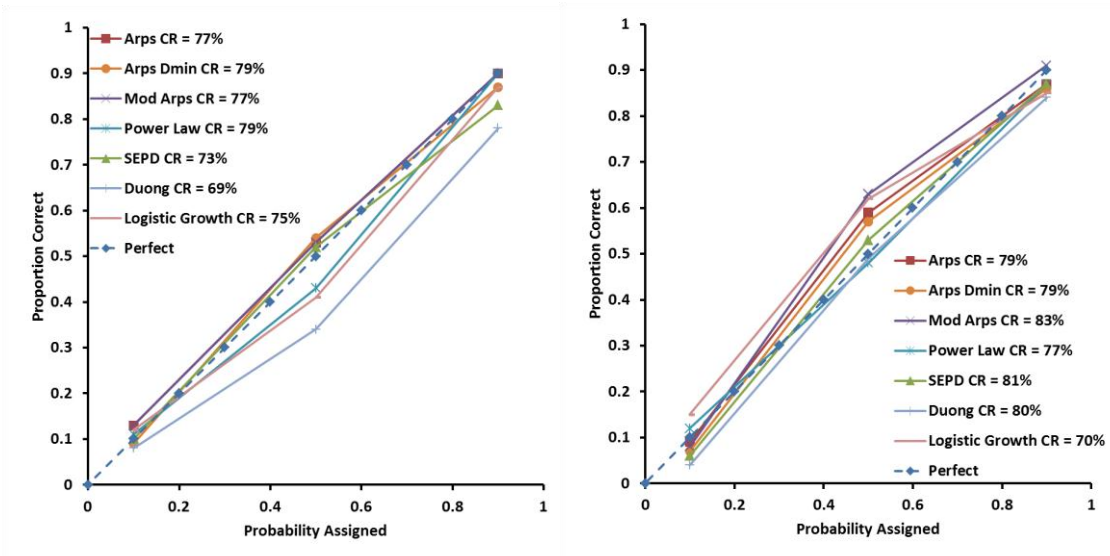
The MCMC method provided slightly better graphical results than the MBM method. The MCMC was well calibrated for all the DCA models at the 50% hindcast (**Fig. 3.5**). Overall the MBM was well calibrated for methods based on Arps' equations. This was expected as it was developed, calibrated and tested using Arps' equations (Cheng et al., 2010).

For the MBM the best coverage rates, were provided when used with Arp's equations. In the other hand, the MCMC provided fairly good coverage rates despite the DCA model used. One of the main advantages of the MCMC over the MBM is that is faster. Because of this, it was decided to discard the MBM method from further study and focus on the calibration of the DCA models with the MCMC method over time.

Table 3.3—Bounds for Decline Curve Model Equations

DCA Parameter (δ)	Upper bound	Lower bound
Arps' Equations		
q_i, Mcf/D	1,000,000	0.01
D_i, 1/year	50	0.1
b	2	0
Arps' with 5% Min Decline		
q_i, Mcf/D	1,000,000	0.01
D_i, 1/year	50	0.1
b	2	0
Modified Arps' Equations		
q_i, Mcf/D	1,000,000	0.01
D_i, 1/year	50	0.1
b	2	0
T_0, months	10,000	3
Power Law Model		
q_i, Mcf/D	1,000,000	0.01
D_{hat}	10	0.001
D_∞	1	1 Exp-09
n	2	0.001
Stretched Exponential Model		
q_i, Mcf/D	1,000,000	0.01
η	5	0.01
τ	10	0.15
Duong's Model		
q_i, Mcf/D	1,000,000	0.01
a	5	0.5
m	2	0.5
Logistic Growth Model		
K, Mcf	100,000,000	1,000
a_L, months	1	1000
n_L	0.01	1

Table 3.4— Proposal Distribution for Each DCA Model	
DCA Parameter (δ)	σ
Arps' Equations	
Ln (q_i)	0.2
Ln (D_i)	0.4
b	0.2
Arps' with 5% Min Decline	
Ln (q_i)	0.2
Ln (D_i)	0.4
b	0.2
Modified Arps' Equations	
Ln (q_i)	0.2
Ln (D_i)	0.4
b	0.2
T_0	1
Power Law Model	
Ln (q_i)	0.2
Ln (D_{hat})	0.4
Ln (D_∞)	0.2
n	0.4
Stretched Exponential Model	
Ln (q_i)	0.2
Ln(η)	0.4
τ	0.2
Duong's Model	
Ln (q_i)	0.2
a	0.2
m	0.2
Logistic Growth Model	
Ln(K)	0.4
Ln(a_L)	0.2
n_L	0.3



a) **b)**

Fig. 3.5—Various DCA models with PDCA a) MBM and b) MCMC. MCMC provided a slightly better calibration for the sample of Barnett shale gas wells. The MCMC coverage rates are slightly better than the ones from the MBM.

The results have been presented for hindcasts predictions from 59-119 months. The use of Arps' equations with a 5% minimum decline is almost equivalent to the use of Arps without any constraints (Fig. 3.5). Therefore, for the reminder of the study the behavior of Arps with a 5% minimum decline was discarded from further study as it is equivalent to Arps without any constraints within the hindcast prediction window.

3.5 Evaluation of MCMC Using Multiple DCA Models and Stages of Depletion on Shale Gas Fields

The analysis was performed using several hindcasts by increasing the amount of months analyzed. Six hindcasts were performed by increasing the amount of production data available to match from 6 to 36 months using 6-month steps. The rest of the months not used to hindcast were assumed to be the actual future production. Moreover, following Gong et al. (2011) a "white noise was added to the production data to model the inherent error". The white noise was added to the standard deviation from the residuals.

In addition to the calibration plot, the C.I.s evolution over time was calculated. To calculate the P_{90}, P_{50}, and P_{10} values for the set of wells, the individual-well P_{90}, P_{50} and P_{10} estimates were aggregated assuming the wells' individual estimates are perfectly correlated.

The hindcast cumulative production P_{90} to P_{10} range for the set of wells narrows as the amount of production data matched increases (**Fig. 3.6a** to **3.6f**). In other words, uncertainty in the CPEOH period decreases over time, as expected. In general, the results are biased if 18 months or less of production are matched. The bias for the CPEOH is the difference between the mean value of CPEOH and the actual CPEOH. As more production data are matched, the results become less biased. For most of the models (modified Arps, Power Law, SEPD, Duong's Model, and logistic growth) but

Arps' the P_{50} estimate for the set of wells is more accurate than the deterministic estimate when limited production data are available to be matched (less than 18 months).

One of the narrowest C.I.s are produced with Arps' equations (**Fig. 3.6a** to **Fig**. 3.6**b)**. However, the results for the Arps models tend to be more biased if one year or less of production is matched. For the two models based on Arps' equations the P_{50} underestimated the actual production for the set of wells; i.e., the models were pessimistic.

The C.I. for the Power Law model is the second widest (**Fig. 3.6c**). However, the results are less biased than Arps' results at 6 months. The Power Law model P_{50} underestimates true production for the set of wells if one year of production is matched, while it overestimates true production for the rest of the hindcasts. This is the only method that crosses; all the other methods either consistently underestimate or overestimate. The results for 6 months using the SEPD model were not plotted because the model yielded unrealistic results for this limited amount of data. Nevertheless, the SEPD model provided one of the least biased results overall (**Fig. 3.6d**). The SEPD model P_{50} slightly overestimated actual production for the set of wells; i.e., the results were optimistic. Duong's model yields a wider C.I. (**Fig. 3.6e**), and it required the most production data to produce relatively unbiased results. The Duong model P_{50} significantly overestimated actual production for the set of wells; i.e., the results were optimistic. The Duong model P_{50} might overestimates production for the set of wells because it assumes long-term linear flow for all the wells analyzed, even though some wells may deviate from that trend. The C.I. for the logistic growth model is one of the narrower from all the models (**Fig. 3.6f**). Furthermore, the logistic growth model provided one of the least biased results overall. The logistic growth model P_{50} slightly overestimated actual production for the set of wells; i.e., the results were optimistic.

In general, the hindcasts were reasonably well calibrated (i.e., close to the unit-slope line), particularly the P_{10} and P_{90} values (**Fig. 3.7a** to **Fig.** 3.7**f**). Most of the deviations from the unit-slope line occur in the Arps' P_{50} values at early times. Calibration improves when more production data is matched in the hindcasts. Except for the Duong model, the curves lie at or above the unit-slope line, indicating the P_{10}, P_{50} and P_{90} estimates are too low. Overall, the coverage rate ranges from 64% to 85% for the 80% C.I.s, with most hindcasts in the 77-82% range.

With the Arps and modified Arps models, more production data is required to be well calibrated compared to the other models (Fig. 3.7a to Fig. 3.7f). This is likely related to having smaller C.I.s, as seen in Fig. 3.6a to 3.6c. The P_{10}, P_{50} and P_{90} estimates all lie on or above the unit-slope line. Thus, the hindcast production estimates should be greater to be well calibrated. Coverage rates are near 80% for the Arps' models. This means that approximately 80% of the P_{10}-P_{90} ranges bracket the actual production.

The hindcasts appear to be better calibrated for the Power Law model, although coverage rates are below 70% for 6 and 12 months of production data matched and are no greater than 74% for more production data (**Fig. 3.7c**). This indicates that coverage rate alone is not sufficient to measure the quality of the probabilistic hindcast. With the SEPD model, 12 months of production data were required for the model to be stable (**Fig. 3.7d**). The coverage rate ranges from 71% to 78% for matches of 12 to 26 months of production data. The SEPD exhibits the best calibration overall, and is well calibrated for all production data periods matched. The Duong model has the best calibration and coverage rate with 6 months of production data matched (**Fig. 3.7e**), and calibration and coverage rate decrease with more production data matched. Compared to the other models, the Duong model predictions are consistently optimistic; i.e., the estimates should be smaller to be well calibrated. The hindcasts appear to be better calibrated for the logistic growth model, although coverage rates are below 70% for 6 to 18 months of production data matched and are no greater than 71% for more production data (**Fig. 3.7f)**, similar to the behavior exhibit by the Power Law model.

The different models have different assumptions, which affect their performance in shale gas reservoirs. In general, when limited production data are available (6–12 months), models that were not developed for shale gas resources, such as Arps, tend to underestimate future production. This is because with little data, the steep early declines are fit best with either exponential declines or hyperbolic declines with relatively low *b* values, which do not capture the subsequent flattening of the decline curve.

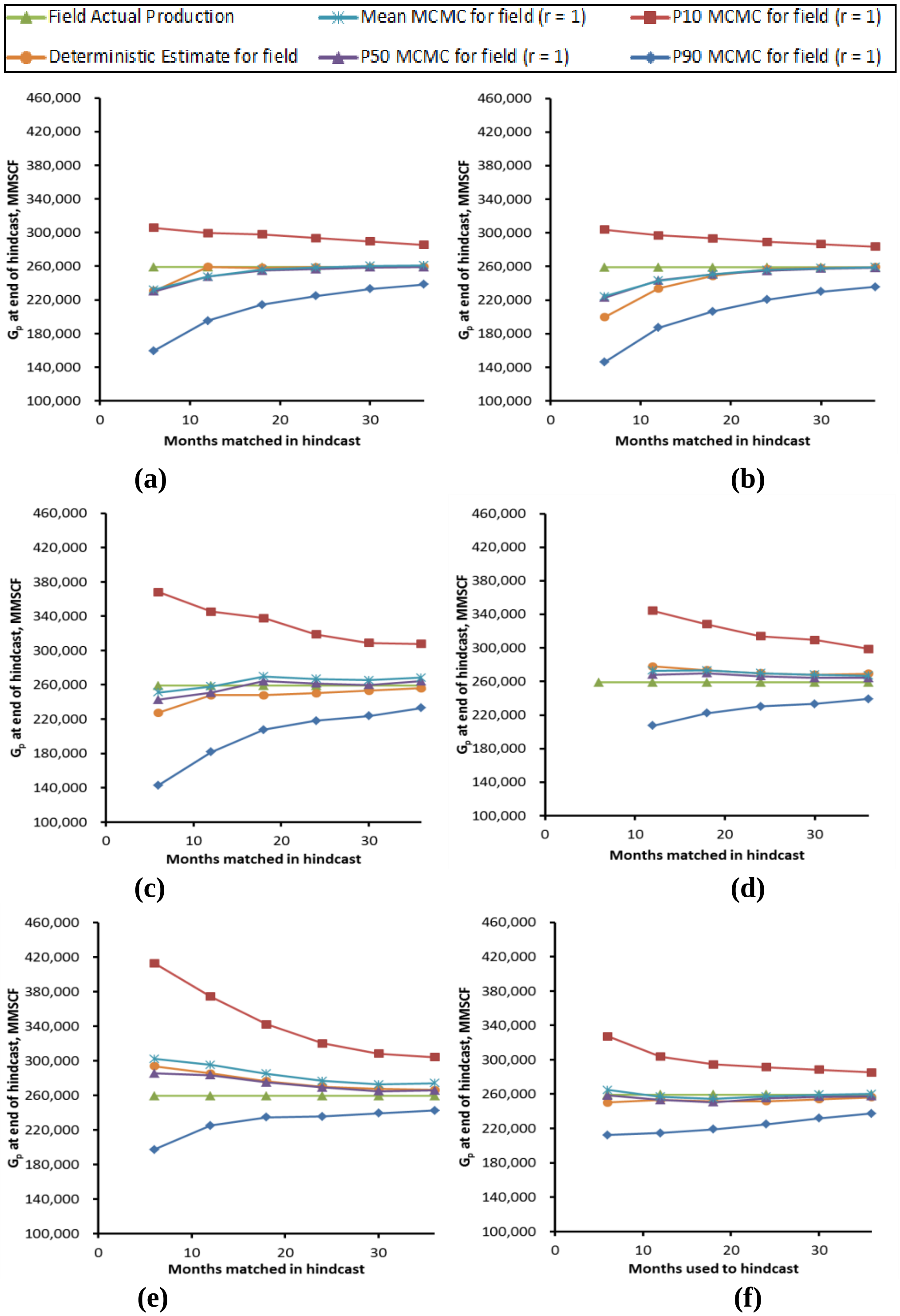


Fig. 3.6—MCMC and a) Arps' equations, b) Modified Arps' equations, c) Power Law, d) Stretched Exponential, e) Duong's model, d) logistic growth model. In general, 80% C.I decrease in size as the amount of production analyzed increases. Also, the results are biased if less than 18 months of production data are available to match in the hindcast.

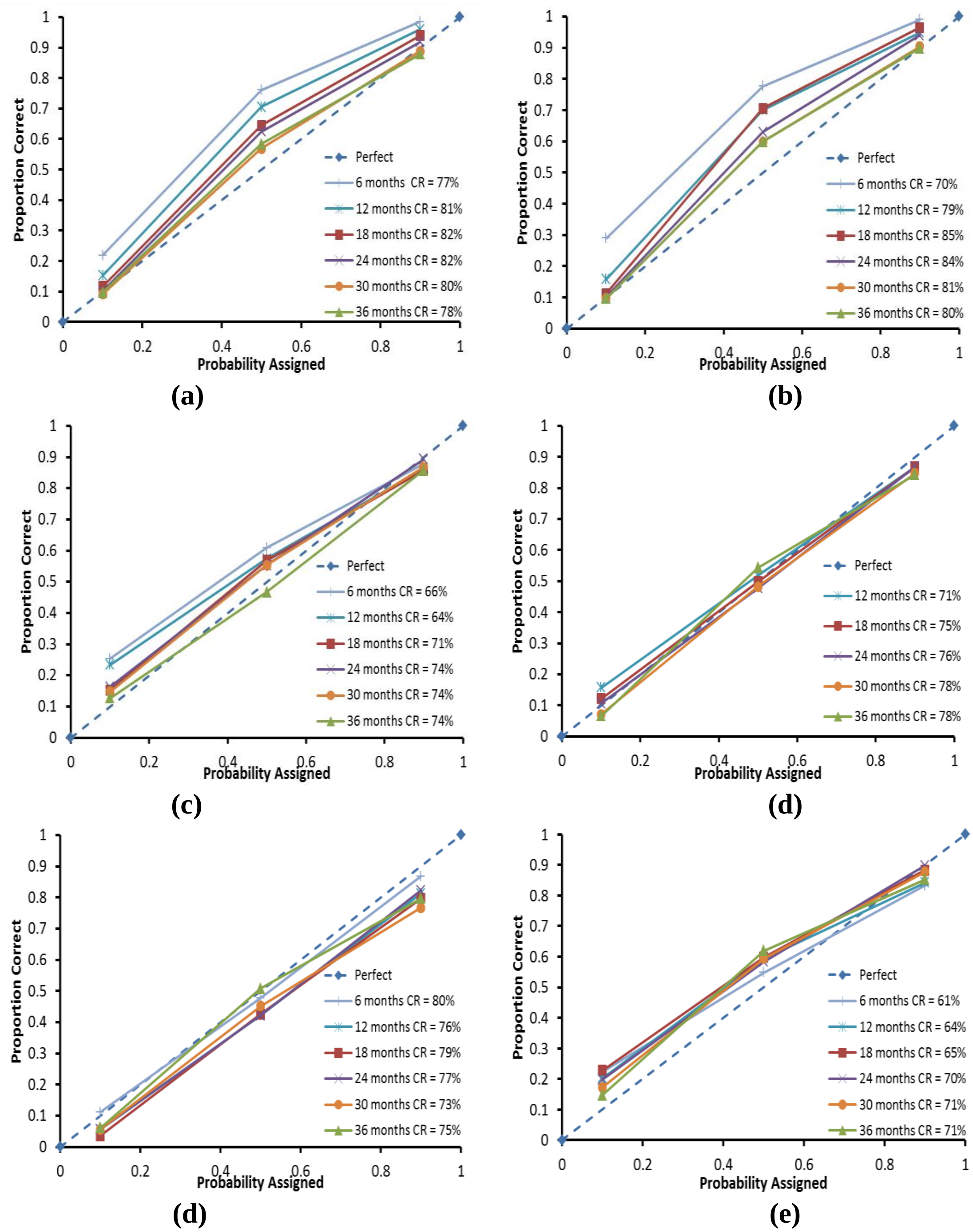


Fig. 3.7—MCMC and a) Arps' equations, b) Modified Arps' equations, c) Power Law, d) Stretched Exponential, e) Duong's model, d) logistic growth model. In general, calibration is enhanced over time using MCMC as probabilistic method.

The opposite occurs for the Duong model. The assumption of long-term linear tends to overestimate future production, at least for the well set examined in this study. I believe that both model types can be improved with use the use of different prior distributions (Chapter IV)

Overall, there seems to be a correspondence between size of the confidence intervals and quality of calibration. The two Arps models tend to have smaller confidence intervals (Fig. 3.6a and 3.6b) and, correspondingly, do not appear to be as well calibrated (Fig. 3.7a and 3.7b), in spite of the coverage rates near 80%.

The Power Law and Duong models have larger confidence intervals (Fig. 3.6c and 3.6e) and, correspondingly, appear to be better calibrated (Fig. 3.7c and 3.7e), despite the lower coverage rates. The SEPD model appears to be somewhat of an exception to this; it has small confidence intervals (Fig. 3.6d) but is also very well calibrated (Fig. 3.7d). The primary disadvantage of the SEPD model is that it did not work well with only 6 months of production data to match. A special case is the logistic growth model that exhibits smaller confidence intervals (Fig. 3.6f) and is well calibrated (Fig. 3.7f). Nevertheless, the coverage rate is the lowest one of only 70%.

It does not appear to be a correspondence between visual quality of the calibration and the coverage rate. Models that appear to be better calibrated visually (Power Law, SEPD, Duong and logistic growth model) tend to have coverage rates that deviate more from 80% than models that appear to be poorer calibrated visually (the Arps models).

Despite the comparisons between models and qualifications offered, all of the DCA models are reasonably well calibrated when used in conjunction with the MCMC probabilistic methodology, which is remarkable given the small amount of production data that is being analyzed in these cases. In the next section, I investigate if the poorer calibration at early production stages (6-12 months of production) for some of the models can be improved even further.

4. INTEGRATION OF VOLUMETRIC DATA INTO THE MARKOV CHAIN MONTE CARLO USING THE LOGISTIC GROWTH DECLINE CURVE MODEL

4.1 Evaluation of MCMC-Logistic Growth Models at Different Stages of Depletion with Informative Prior Distribution

In Chapter III, the evolution of the calibration of different combinations of MCMC with DCA models was studied. It was concluded that the SEPD and the logistic growth models were the best calibrated DCA models overall. Nevertheless, the SEPD presented some limitations when using only 6 months of data. On the other hand, the logistic growth model was well calibrated at 6 months and it allows for the integration of volumetric information because of its decline parameters. The *K* parameter in the logistic growth model is known as the carrying capacity, and is defined similarly as the EUR of a well (Clark et al., 2011).

Because of the flexibility of the logistic model, it can provide some non-unique matches if the carrying capacity cannot be obtained from other types of assessments, such as volumetric (Clark et al., 2011). Clark et al. (2011) validated their model on a 1000-well sample base. From this sample base, 400 wells were not matched because they yielded unreasonable DCA parameters. The authors did not describe what kind of values for DCA parameters were considered unreasonable. However, the distributions of parameters provided by Clark et al. (2011) give us an idea of the range of values they consider reasonable.

The MCMC requires that a prior distribution be specified for each of the decline curve parameters used. For the logistic growth model, the prior distribution for the carrying capacity, *K*, was taken from Clark et al. (2011) (**Fig. 4.1**). Clark et al. (2011) performed DCA using the logistic growth model on a sample of 600 Barnett shale gas wells to obtain a distribution for the carrying capacity, *K* (**Table 4.1**). Despite using Clark et al. (2011) boundaries as constraints for the prior *K* distribution, the boundaries for the likelihood function were defined as before in **Table 3.3**. The boundaries for the likelihood function are wider than the boundaries for the prior distribution. The boundaries for the prior distribution are based on prior knowledge, while the boundaries for the likelihood were set wider so any reasonable combination of decline curve parameters that match the data can be considered. The prior distributions for the other two DCA parameters, a_L and n_L, were kept as uniform distributions, with boundaries defined in **Table 3.3**.

Gas well No. 47 was used to evaluate the use of a lognormal distribution on a single gas well (**Fig. 4.2**). The use of a lognormal prior distribution increased the P_{50} for the CPEOH when using only six months of data (**Table 4.2**). By using a lognormal prior distribution the P_{50} for the CPEOH estimate was shifted further away from the actual CPEOH than when using a uniform prior distribution. When only 6 months of production are matched, the prior distribution has a higher weight in the posterior distribution than the likelihood function. For this well, the estimated CPEOH when using a lognormal prior distribution was higher than the one given by the likelihood function (the uniform distribution does not affect the results of the posterior distribution and are only based on the likelihood function). Hence, a poor choice of a prior distribution could result in poor calibrated results at 6 months. When more data become available, the likelihood function should correct the estimates for CPEOH.

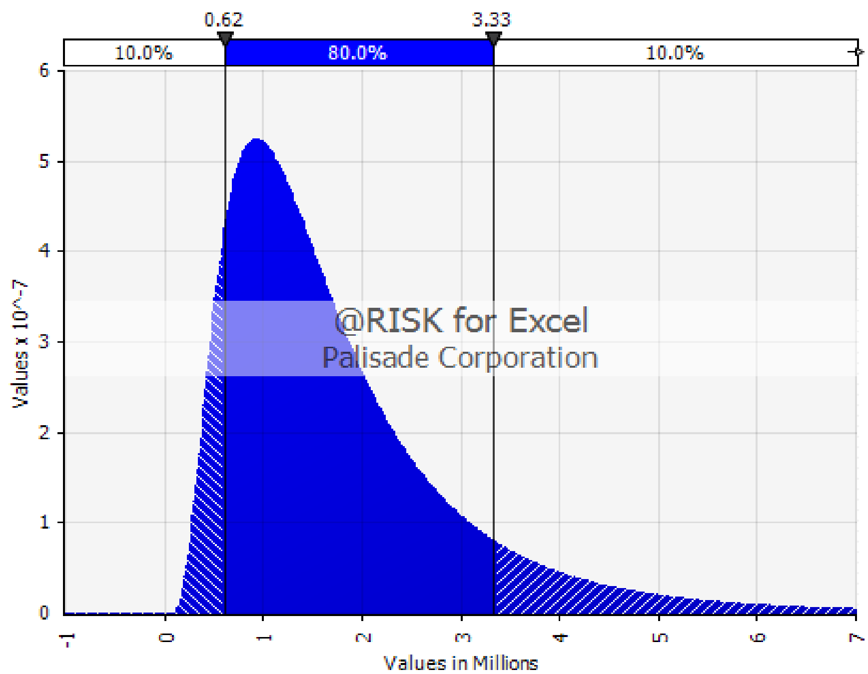
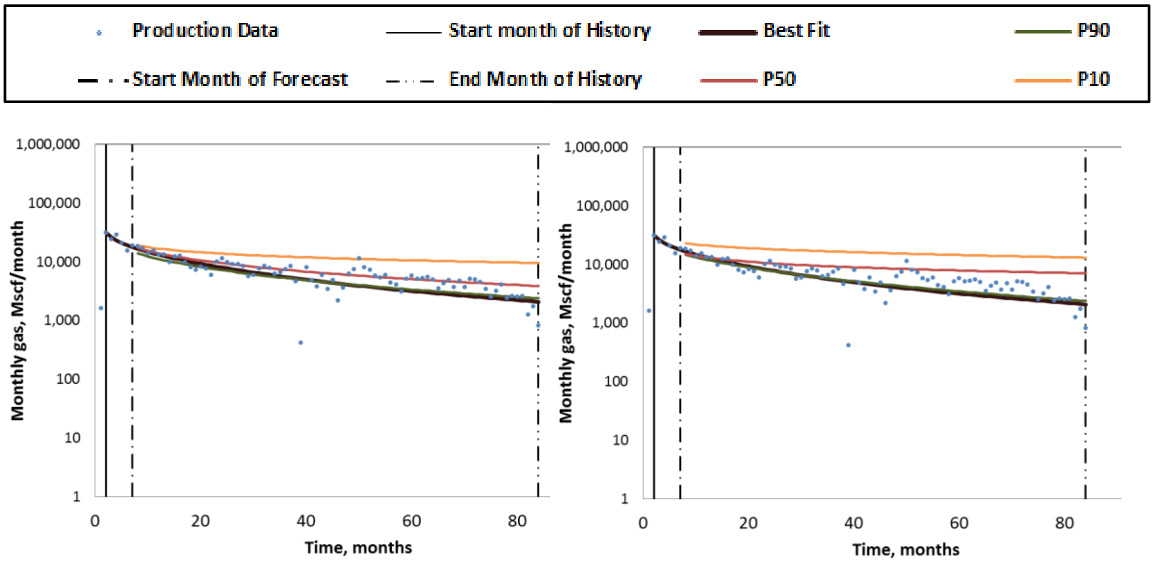


Fig. 4.1—Informative prior distribution for *K*, from Barnett Shale analogous wells.

Table 4.1—Properties of the Logistic Prior Distribution

Logistic Growth Model			
DCA Parameter	μ_L	σ_L	Type
K, Mcf	1,782,392	1,309,691	Logarithmic



(a) **(b)**

Fig. 4.2—PDCA MCMC-logistic growth on gas well 47, a) uniform prior and b) lognormal prior. Probabilistic methods bracket actual production data in a single gas well hindcast.

Table 4.2—Comparison of PDCA Methods on Gas Well 47 Using a Uniform and a Lognormal Prior Distribution

	<u>Uniform prior</u>	<u>Lognormal prior</u>
Actual CPEOH, Mcf	507,713	507,713
Arps' Equations Best Fit, Mcf	432,614	432,614
CPEOH P90, Mcf	406,667	431,565
CPEOH P50, Mcf	557,327	705,456
CPEOH P10, Mcf	932,132	1,246,104
CPEOH Mean, Mcf	634,028	770,034

Nevertheless, as mentioned before, a single-well estimate cannot provide a good measurement of the calibration and reliability of a PDCA method. The analysis was performed using several hindcasts with increasing amounts of production data. Six hindcasts, using an uninformative (uniform) prior distribution, were performed by increasing the months of production data analyzed from 6 to 36 months using 6-month steps. The previous process was repeated using different informative (lognormal) prior distributions. The rest of the months not used to hindcast were assumed to be the actual “future” production.

The performance of the logistic growth model was evaluated using different lognormal prior distributions for *K*. One of the prior distributions for *K* was taken from Clark et al. (2011) (Fig. 4.1). The rest of the prior distributions studied for *K* were proportional to the one by Clark et al. (2011). The prior that yielded the best calibrated visually results was 1.5 times higher than the one reported by Clark et al. (2011) (**Table 4. 3** and **Fig. 4.3**).

The use of a lognormal prior distribution for *K* increased the coverage rate from 60% to around 70% (**Table 4.4**). The use of a lognormal prior distribution can either increase our decrease the percentage error in CPEOH (**Table 4.4**). By using the *K* distribution from Clark et al. (2011) we increase the percentage error in CPEOH, because the estimate for CPEOH from the prior distribution was too low. The corrected prior was chosen because it yielded the best calibrated results; hence we obtained a lower percentage error in CPEOH (**Fig. 4.4**). Therefore, a poor choice of a prior distribution could result in poor calibrated results at 6 months. The use of a lognormal prior distribution (from Clark et al. (2011) or the corrected prior) enhances the calibration (**Fig. 4.4**).

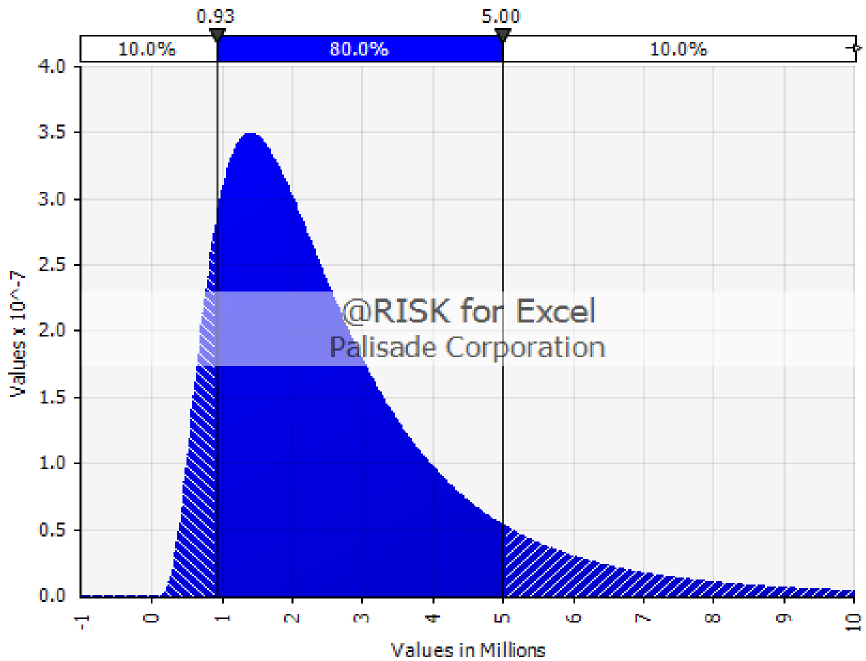


Fig. 4.3—Corrected informative prior distribution for *K*.

Table 4.3—Properties of the Corrected Logistic Prior Distribution

Logistic Growth Model			
DCA Parameter	μ_L	σ_L	Type
K, Mcf	2,673,588	1,964,537	Logarithmic

Table 4.4—Comparison of PDCA Methods With the Logistic Growth Model on a Sample of Barnett Shale Gas Wells Using Uniform and Volumetric Prior Distributions

	Uniform Prior	Lognormal Prior 1	Lognormal Prior 2
No of Wells	197	197	197
Coverage rate	61%	66%	68%
Relative error: Average((P_{50}-CPEOH)/CPEOH)	2.31%	0.812%	4.08%
Absolute Error: Average(|P_{50}-CPEOH|/CPEOH)	26.56%	25.328%	24.72%
Average(C.I./P_{50})	0.6510	0.6665	0.7026
True CPEOH MSCF	204,328,815	204,039,437	204,039,437
Sum of P_{50} values, MSCF	203,475,374	194,351,296	207,526,061
Percentage Error in CPEOH	0.41%	4.74%	1.71%

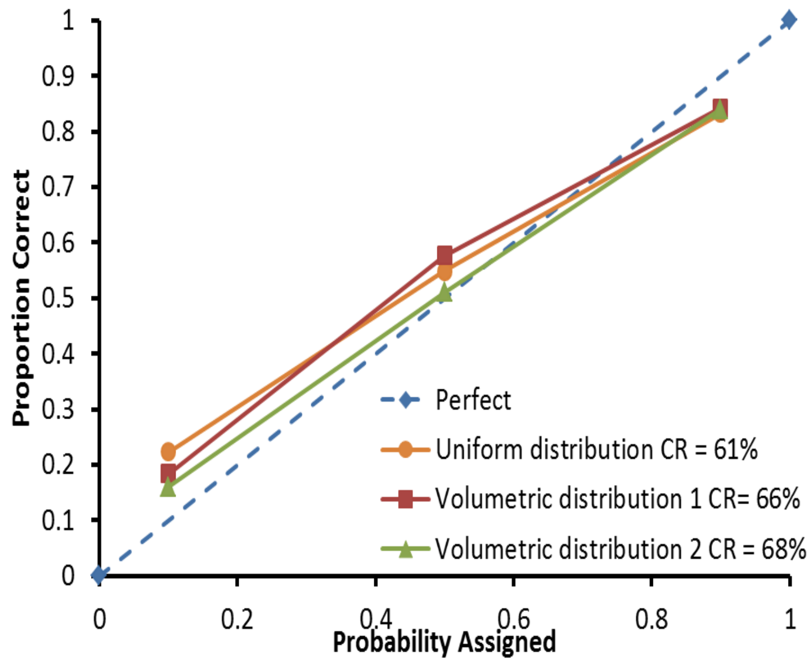


Fig. 4.4—PDCA MCMC-DCA logistic growth using a volumetric prior enhanced the calibration for the sample of Barnett shale gas wells.

The calibration for the sample of wells with the corrected prior still shows a clear overconfidence. This behavior can be corrected by incorporating a function that calibrates the prior distribution over time. In a true Bayesian approach, as we obtain more production data, our prior distribution should be updated from the posterior distribution obtained using the previous prior distribution. We can roughly approximate this trend, by modeling the prior as a power function that decreases over time (Eq. 9),

$$\pi_{new}(\theta) = \pi(K)^{g(t)} \quad (9)$$

The function g(t) was approximated using the best value of g (that which provided the best calibration for our sample of wells, i.e., achieved the best P_{90}, P_{50} and P_{10}) at each hindcast. It was found that the relation between g(t) and t can be approximated by a power law relation, Eq. 10, plotted in **Fig. 4. 6**.

$$g(t) = 4.556\, t^{-1.198} \quad (10)$$

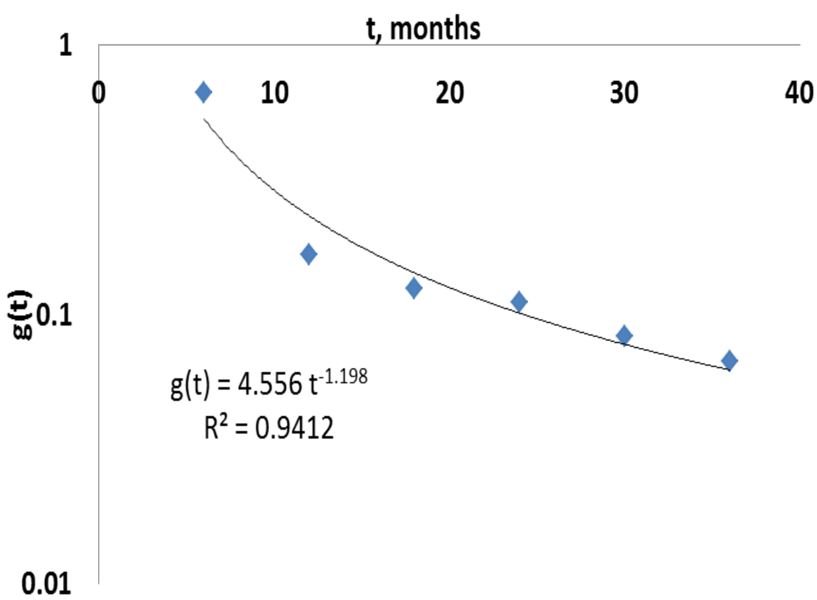


Fig. 4.5—g(t) decreases as the available amount of production, t, increases. The prior distribution has less impact as more production data become available to match.

The g(t) function for a lognormal prior distribution for *K* was chosen to reduce the bias from the mean of the CPEOH regardless of the amount of data available to match (**Fig. 4.6a and 4.6b**). The use of a lognormal prior distribution provides results that are better calibrated, regardless of the amount of data available to match, than when using a uniform prior distribution (**Fig 4.7a and 4.7b**). In addition, the use of a lognormal prior distribution enhances the CR than when using a uniform prior distribution (**Fig. 4.7b**).

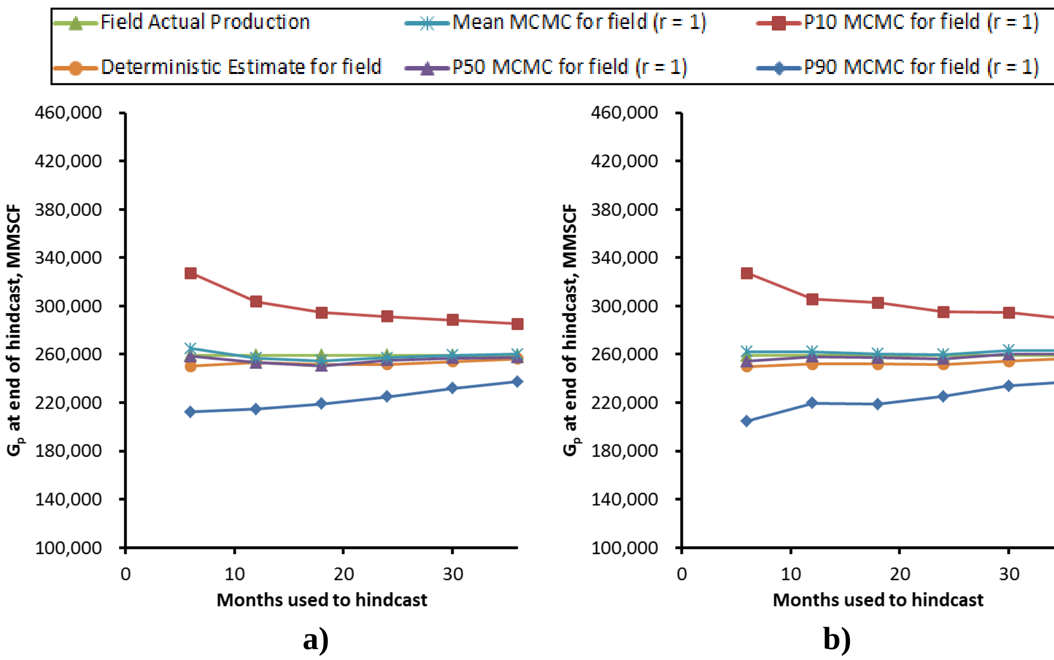


Fig. 4.6—MCMC and logistic growth model with a) uniform prior distribution, b) lognormal prior distribution. In general, 80% C.I decrease in size as the amount of production analyzed increases. The results are less biased if a lognormal prior is used.

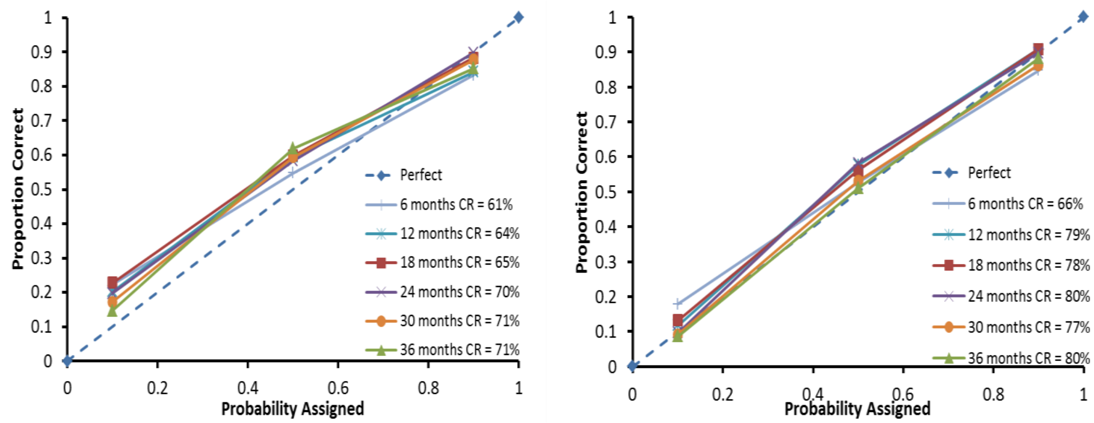


Fig. 4.7—MCMC and logistic growth model with a) uniform prior distribution, b) lognormal prior distribution. The calibration is enhanced over time using a lognormal prior distribution. Furthermore, the coverage rate is enhanced.

4.2 Application to Barnett Shale of MCMC-Logistic Growth Models at Different Stages of Depletion with Volumetric Prior Distribution

The MCMC-logistic growth model combination was tested and calibrated using a prior distribution for *K*, obtained from the DCA of 600 wells from the Barnett Shale. As in the previous section, during any hindcast the prior distribution was the same for all the wells in the sample.

A hindcast test was conducted in which *K* was now defined as the technically recoverable resources (TRR) for a Barnett shale gas well. The distribution for TRR for the Barnett shale was taken from Dong et al. (2012) (**Fig. 4.8** and **Table 4.5**). Dong et al. (2012) estimated the TRR for the Barnett shale based on analytical simulation. Dong et al. (2012) also calculated the distribution for recovery factors for the Barnett Shale.

Hence, for any given well with known volumetric assessment, a distribution of TRR for each well can be approximated based on the recovery factor distribution from Dong et al. (2012). The distribution for TRR from Dong et al. (2011) (Fig. 4.8) is similar to the corrected lognormal prior distribution tested in the previous section (Fig. 4.3)

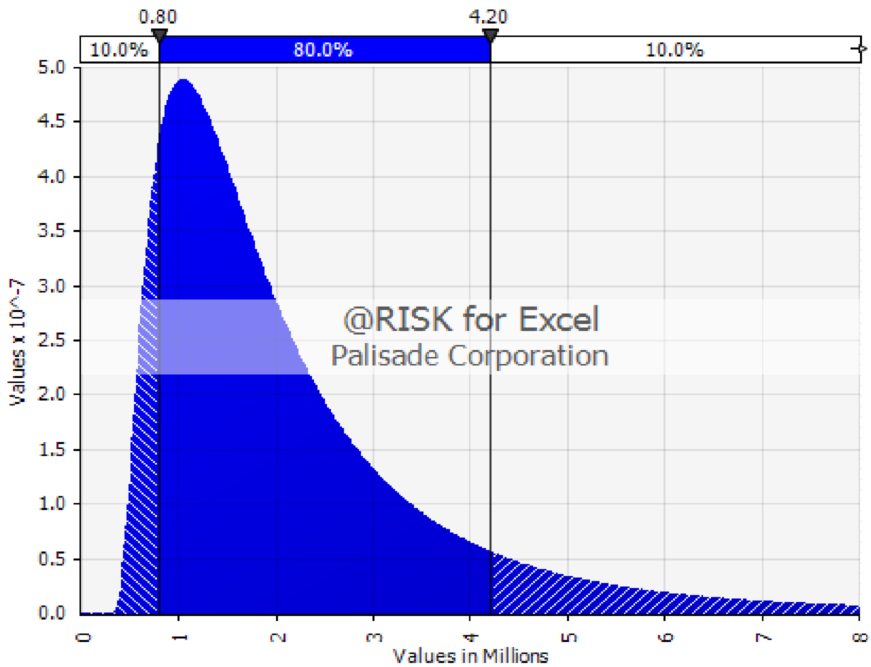


Fig. 4.8—Informative prior distribution for *K*, from TRR from the Barnett Shale.

Table 4.5—Properties of the Technically Recoverable Resources Distribution

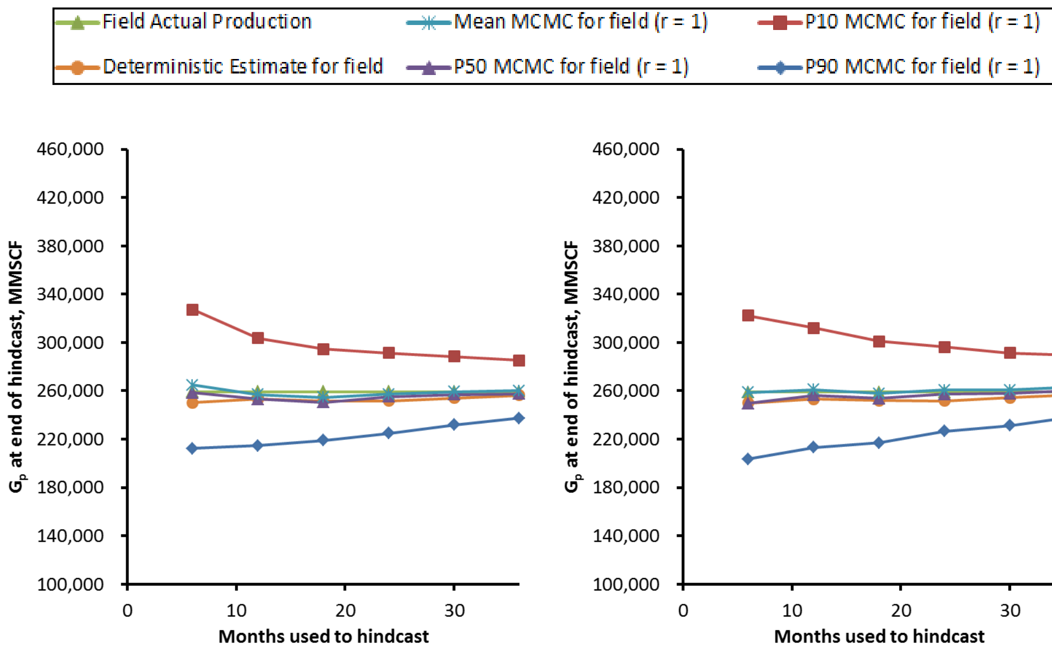
Logistic Growth Model			
DCA Parameter	μ_L	σ_L	Type
K, Mcf	2,226,008	1,820,854	Logarithmic

Six hindcasts were performed by increasing the amount of production data analyzed from 6 to 36 months using 6-month steps, once with a uniform prior and once with the

TRR lognormal prior distribution. The rest of the months not used to hindcast were assumed to be the actual “future” production.

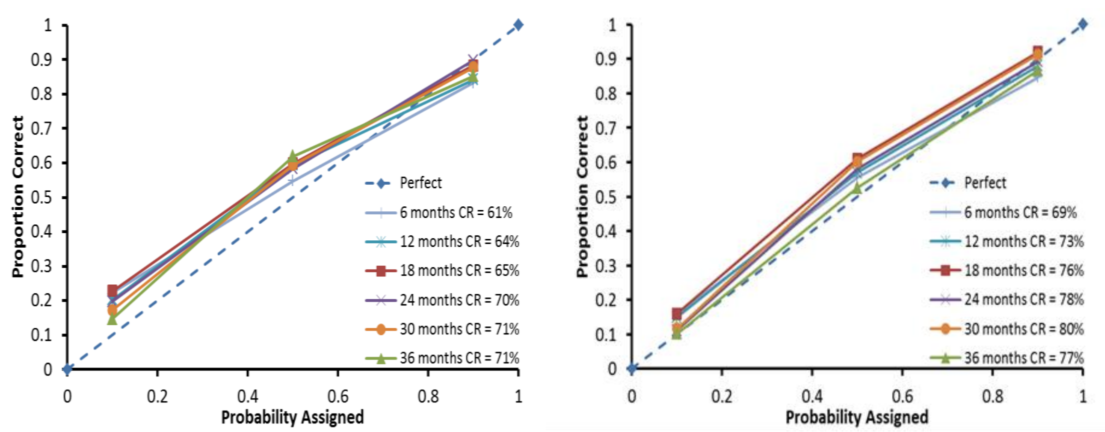
The use of a TRR prior distribution reduces the bias from the mean of the CPEOH regardless of the amount of data available to match (**Fig. 4.9a and Fig 4.9b**). We expected this since the TRR distribution reported by Dong et al. (2012) is similar to the corrected lognormal prior distribution used to calibrate the model. The use of a TRR distribution as the prior distribution for *K* provides results that are better calibrated regardless of the amount of data available to match than when using a uniform prior distribution (**Fig 4.10a and 4.10b**). In addition, the use of a TRR prior distribution enhances the CR than when using a uniform prior distribution (**Fig. 4.10b**).

The proposed combination of MCMC, logistic growth DCA model and the incorporation of volumetric data provides an integrated procedure to reliably quantify the uncertainty in production forecasts and reserves estimates in shale gas reservoirs. The reliability of the results has been demonstrated for predictions to 59-119 months, and not for long-term forecasts. Thus, it cannot be claimed that these methods will reliably quantify uncertainty in long-term forecasts and reserves estimates. However, I believe it is safe to say that long-term probabilistic forecasts and reserves estimates will be more reliable that they would be if they were not reliable in the short term.



a) b)

Fig. 4.9—MCMC and logistic growth model with a) uniform prior distribution, b) TRR prior distribution. In general, 80% C.I decrease in size as the amount of production analyzed increases. The results are less biased if a volumetric prior is used.



b) b)

Fig. 4.10—MCMC and logistic growth model with a) uniform prior distribution, b) TRR prior distribution. The calibration is enhanced over time using a volumetric prior distribution. Furthermore, the coverage rate is enhanced.

5. CONCLUSIONS

Based on hindcasts of 197 hydraulically fractured horizontal Barnett-shale gas wells with 59-119 months of production data available, I conclude the following:

- The uncertainty in cumulative production at the end of hindcast, CPEOH, decreases as the amount of monthly production data available for matching increases.
- The MBM and the MCMC probabilistic methods reliably quantify uncertainty in CPEOH, matching 50% of known production history, for all of the DCA models studied.
- The MCMC probabilistic method combined with the DCA models developed for shale gas resources reliably quantifies the uncertainty, regardless of the amount of monthly data available. Even with DCA models based on Arps' equations, the MCMC method is reasonably well calibrated.
- The DCA models developed for shale gas wells were better calibrated, but have wider confidence intervals for CPEOH, than models based on Arps' equations.
- For the modified Arps, Power Law, SEPD, Duong, and logistic growth models, the MCMC P_{50} estimate is more accurate than the deterministic estimate for the sample of wells when only 6-12 months of production data are available for matching.

- When a uniform prior distribution is used for the MCMC probabilistic method, the SEPD and logistic growth models are the best calibrated visually and have the best coverage rate overall.
- The use of a lognormal prior distribution for the carrying capacity, *K,* enhances the calibration of the logistic growth model and improves the overall coverage rate when used with the MCMC probabilistic method.
- The use of a volumetric TRR prior distribution for the carrying capacity, *K,* enhances the calibration of the logistic growth model, reduces the bias at early times and improves the overall coverage rate when used with the MCMC probabilistic method.

While uncertainty will always be present in any production forecast and reserves estimate, and will likely be quite large early in the producing life, reliable assessment of uncertainty enables better assessment of upside and downside potential, as well as better assessment of the expected value of reserve estimates. This can be particularly valuable early in the development of a play, because decisions regarding continued development are based to a large degree on production forecasts and reserves estimates for early wells in the play.

NOMENCLATURE

a	Duong intercept constant, 1/day
a_L	Time to the power n_l at which half of K has been produced, months
b	Arps decline exponent, dimensionless
CR	Coverage rate
C.I.	Confidence intervals
CPEOH	Cumulative production at the end of hindcast, Mcf
DCA	Decline curve analysis
D_i	Arps initial decline rate, 1/year
D_∞	Power Law decline at "infinite time" constant, 1/day
$\hat{D}$	Power Law decline constant, 1/day
EUR	Estimated ultimate recovery, Mcf
f	Likelihood function
K	Logistic growth Model carrying capacity, Mcf
JSM	Jochen and Spivey method
MBM	Modified Bootstrap method
MCMC	Markov Chain Monte Carlo
m	Duong slope
n	Power Law time exponent, dimensionless
n_L	Duong decline exponent, dimensionless
PDCA	Probabilistic decline curve analysis

P_{10}	Value at confidence level 10%
P_{50}	Value at confidence level 50%
P_{90}	Value at confidence level 90%
q_i	Initial gas rate, Mcf/D
t	Available monthly production data, months
T_0	Modified Arps time go into exponential decline, months
y	Monthly production, Mcf
δ	DCA parameter
θ	DCA parameter acting as random variable
θ_j	Parameter at step j in MCMC
$\theta_{proposal}$	Candidate drawn from proposal distribution
η	Exponent parameter or SEPD model, dimensionless
τ	Characteristic time parameter for SPED model, month
σ	Sample variance from best fit
σ_j	Sample variance at step j in MCMC
$\sigma_{proposal}$	Sample variance from proposal distribution
π	Prior or posterior probability

REFERENCES

Agarwal, A., Wei, Y., and Holditch, S. 2012. A Technical and Economic Study of Completion Techniques in Five Emerging Us Gas Shales: A Woodford Shale Example. *SPE Drilling & Completion* **27** (1): pp. 39-49. DOI: 10.2118/135396-pa

Aprilia, A.W., Li, Z., Mcvay, D.A. et al. 2006. Quantifying Uncertainty in Original-Gas-in-Place Estimates with Bayesian Integration of Volumetric and Material Balance Analyses. Paper presented at the SPE Gas Technology Symposium, Calgary, Alberta, Canada. SPE 100575. DOI: 10.2118/100575-ms.

Arps, J.J. 1945. Analysis of Decline Curves. *AIME* **160**: pp. 228-247.

Capen, E.C. 1976. The Difficulty of Assessing Uncertainty (Includes Associated Papers 6422 and 6423 and 6424 and 6425). *SPE Journal of Petroleum Technology* **28** (8): 843-850. DOI: 10.2118/5579-pa

Cheng, Y., Wang, Y., Mcvay, D. et al. 2010. Practical Application of a Probabilistic Approach to Estimate Reserves Using Production Decline Data. *SPE Economics & Management* **2** (1): pp. 19-31. DOI: 10.2118/95974-pa

Clark, A.J., Lake, L.W., and Patzek, T.W. 2011. Production Forecasting with Logistic Growth Models. Paper presented at the SPE Annual Technical Conference and Exhibition, Denver, Colorado, USA. SPE 144790-ms. DOI: 10.2118/144790-ms.

Dong, Z., Holditch, S.A., and Mcvay, D.A. 2012. Resource Evaluation for Shale Gas Reservoirs. Paper presented at the SPE Hydraulic Fracturing Technology Conference, The Woodlands, Texas, USA. SPE 152066-ms. DOI: 10.2118/152066-ms.

Duong, A.N. 2011. Rate-Decline Analysis for Fracture-Dominated Shale Reservoirs. *SPE Reservoir Evaluation & Engineering* **14** (3): pp. 377-387 DOI: 10.2118/137748-pa

Fetkovich, M.J., Fetkovich, E.J., and Fetkovich, M.D. 1996. Useful Concepts for Decline Curve Forecasting, Reserve Estimation, and Analysis. *SPE Reservoir Engineering* **11** (1): 13-22. DOI: 10.2118/28628-pa

Gong, X., Gonzalez, R., Mcvay, D. et al. 2011. Bayesian Probabilistic Decline Curve Analysis Quantifies Shale Gas Reserves Uncertainty. Paper presented at the Canadian Unconventional Resources Conference, Alberta, Canada. SPE 147588. DOI: 10.2118/147588-ms.

Ilk, D., Rushing, J.A., Perego, A.D. et al. 2008. Exponential Vs. Hyperbolic Decline in Tight Gas Sands—Understanding the Origin and Implications for Reserve Estimates Using Arps' Decline Curves. Paper presented at the SPE Annual Technical Conference and Exhibition, Denver, Colorado, USA. SPE 116731. DOI: 10.2118/116731-ms.

Jochen, V.A. and Spivey, J.P. 1996. Probabilistic Reserves Estimation Using Decline Curve Analysis with the Bootstrap Method. Paper presented at the SPE Annual

Technical Conference and Exhibition, Denver, Colorado, USA.SPE 36633-ms. DOI: 10.2118/36633-ms.

Lee, W.J. and Sidle, R. 2010. Gas-Reserves Estimation in Resource Plays. SPE Economics & Management **2** (2): pp. 86-91. DOI: 10.2118/130102-pa

Mattar, L. 2008. Production Analysis and Forecasting of Shale Gas Reservoirs: Case History-Based Approach. Paper presented at the SPE Shale Gas Production Conference, Fort Worth, Texas, USA. SPE 119897-ms. DOI: 10.2118/119897-ms

McVay, D.A. and Dossary, M. 2012. The Value of Assessing Uncertainty. Paper presented at the SPE Annual Technical Conference and Exhibition, San Antonio, Texas, USA. SPE 160189-ms. DOI: 10.2118/160189-ms.

Mengal, S.A. and Wattenbarger, R.A. 2011. Accounting for Adsorbed Gas in Shale Gas Reservoirs. Paper presented at the SPE Middle East Oil and Gas Show and Conference, Manama, Bahrain. SPE 141085-MS. DOI: 10.2118/141085-ms.

Ogele, C., Daoud, A., Mcvay, D.A. et al. 2006. Integration of Volumetric and Material Balance Analyses Using a Bayesian Framework to Estimate Ohip and Quantify Uncertainty. Paper presented at the SPE Europec/EAGE Annual Conference and Exhibition, Vienna, Austria. SPE 100257. DOI: 10.2118/100257-ms.

Valko, P.P. and Lee, W.J. 2010. A Better Way to Forecast Production from Unconventional Gas Wells. Paper presented at the SPE Annual Technical Conference and Exhibition, Florence, Italy. SPE 134231. DOI: 10.2118/134231-ms.