Integrating Model Uncertainty in Probabilistic Decline-Curve Analysis for Unconventional-Oil-Production Forecasting

Aojie Hong and **Reidar B. Bratvold,** National IOR Centre of Norway and University of Stavanger, and **Larry W. Lake** and **Leopoldo M. Ruiz Maraggi,** University of Texas at Austin

Summary

Decline-curve analysis (DCA) for unconventional plays requires a model that can capture the characteristics of different flow regimes. Thus, various models have been proposed. Traditionally, in probabilistic DCA, an analyst chooses a single model that is believed to best fit the data. However, several models might fit the data almost equally well, and the one that best fits the data might not best represent the flow characteristics. Therefore, uncertainty remains regarding which is the "best" model. This work aims to integrate model uncertainty in probabilistic DCA for unconventional plays.

Instead of identifying a single "best" model, we propose to regard any model as potentially good, with goodness characterized by a probability. The probability of a model being good is interpreted as a measure of the relative truthfulness of this model compared with the other models. This probability is subsequently used to weight the model forecast. Bayes' law is used to assess the model probabilities for given data. Multiple samples of the model-parameter values are obtained using maximum likelihood estimation (MLE) with Monte Carlo simulation. Thus, the unique probabilistic forecasts of each individual model are aggregated into a single probabilistic forecast, which incorporates model uncertainty along with the intrinsic uncertainty (i.e., the measurement errors) in the given data.

We demonstrate and conclude that using the proposed approach can mitigate over/underestimates resulting from using a single decline-curve model for forecasting. The proposed approach performs well in propagating model uncertainty to uncertainty in production forecasting; that is, we determine a forecast that represents uncertainty given multiple possible models conditioned to the data. The field data show that no one model is the most probable to be good for all wells.

The novelties of this work are that probability is used to describe the goodness of a model; a Bayesian approach is used to integrate the model uncertainty in probabilistic DCA; the approach is applied to actual field data to identify the most-probable model given the data; and we demonstrate the value of using this approach to consider multiple models in probabilistic DCA for unconventional plays.

Introduction

Although numerical techniques for forecasting hydrocarbon production have developed rapidly over the past decades, DCA remains an industry-accepted method and is used extensively in the oil and gas industry. Decline-curve models are very computationally attractive because only production data, which can be easily acquired, are required for determining a few parameter values through history matching. The history-matched model is further used for forecasting the hydrocarbon production and reserves.

The Arps (1945) model[1] has been used extensively for both conventional and unconventional plays. However, the Arps model might not be ideal for unconventional plays because unconventional wells are usually completed with hydraulic fractures such that several flow regimes (formation linear flow, apparent boundary-dominated flow by fracture interference, linear flow in unstimulated matrix, and true boundary-dominated flow) might appear during the life of a well (Joshi and Lee 2013). Therefore, several alternative decline-curve models have been proposed to capture the characteristics of different flow regimes. Examples are the power-law exponential model (Ilk et al. 2008), the stretched exponential model (SEM) (Valko 2009), the Duong (2011) model, the logistic growth model (LGM) (Clark et al. 2011), and the combined capacitance-resistance model (CRM) proposed by Pan (2016). The Pan (2016) model is subsequently referred to in this paper as the Pan CRM.

Given these models, a question that arises naturally but has not been discussed widely is: Which is the best model? This question is subtler than it appears because the meaning of "best" is not well-defined. In many applications, the model that can best fit the data in a least-squares sense is regarded as the best model. This implies a common but wrong assumption: The model that best fits the data is the one that best represents the actual flow behaviors. This assumption ignores two facts: the best-fit model might not be the model that best describes the actual flow behaviors, and there might be several models that can fit the data almost equally well.

In deterministic DCA, a single model with its best-fit-parameter setting is used for forecasting. Deterministic analysis alone does not quantify the uncertainty in a forecast, and thus it often leads to a "precisely wrong" result. A "vaguely right" forecast in a probabilistic/uncertain sense is more useful and essential for decision making (Read 1920; Bratvold and Begg 2008).

Although probabilistic DCA has not been used extensively in the industry, it is gathering more focus. Jochen and Spivey (1996) and Cheng et al. (2010) used a bootstrap method with the Arps model to quantify forecast uncertainty for conventional plays. Gong et al. (2014) used a Markov-chain Monte Carlo method with the Arps model to quantify forecast uncertainty for unconventional plays. These authors considered measurement errors in DCA.[2] However, they assumed that a model—the Arps model—for forecasting flow behaviors is already given. Gonzalez et al. (2012) used a Markov-chain Monte Carlo method with different individual models to investigate

[1]Commonly, the Arps model is categorized into three types depending on the value of the decline exponent b: exponential ($b=0$), hyperbolic ($0 < b < 1$), and harmonic ($b=1$). In some applications, b is also allowed to be greater than unity. However, $b \geq 1$ yields unrealistic forecasts because the cumulative production is unbounded (Lee and Sidle 2010). To avoid unbounded cumulative production, we use "the Arps model" to refer to the case with $0 \leq b < 1$ in our application.

[2]In DCA, the data fluctuations caused by changes in operating conditions are treated as being caused by measurement errors (Jochen and Spivey 1996).

This paper (SPE 194503) was revised for publication from paper URTEC 2900625, presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Houston, 23–25 July 2018. Original manuscript received for review 6 August 2018. Revised manuscript received for review 30 October 2018. Paper peer approved 6 November 2018.

their performances in quantifying forecast uncertainty for production from unconventional plays; their objective was to identify a single "best" model for probabilistic DCA.

Because a decline-curve model is used to forecast flow behaviors, its goodness should be defined as how well it represents the actual flow behaviors. How well a model fits data does not tell directly how well it represents the actual flow behaviors, but how probably it can well represent the actual flow behaviors. That is, data fitting can only assess the probability of a model being good. Therefore, instead of identifying a single best-fit model for unconventional production, we propose to regard any model as potentially good and describe its goodness in a probabilistic sense. If a model has larger probability, we say that this model is more probable to be a good model.[3] If several models have close probabilities, we say that uncertainty in the models is large because it is difficult to tell which model is more probable to be a good model. In this manner, the model uncertainty[4] can be easily integrated in probabilistic DCA.

An analogy to our proposal is the aggregation of forecasts from diverse experts. Such a forecast is a probability distribution over an uncertain quantity provided as a subjective assessment or modeled by an expert. Thus, a decline-curve model can be analogized to an "expert." The topic of aggregating forecasts has been furthered significantly over the past decades in other fields (e.g., risk analysis and operational research). More than 40 years ago, Bunn (1975) reasoned: "The principal motivation for combining forecasts has been to avoid the a priori choice of which forecasting method to use by attempting to aggregate together all the information which each forecasting model embodies. In selecting the 'best' model, the forecaster is often discarding useful independent evidence in those models which are rejected. Hence, the methodology of combining forecasts is founded upon the axiom of maximal information usage."

Clemen and Winkler (1999) provided a comprehensive review on the aggregation approaches, ranging from simple linear aggregation (Stone 1961) to a complex copula approach (Jouini and Clemen 1996) that uses a copula function to include the dependences among experts' forecasts. Some of their conclusions, which provide support to our work, are "it is worthwhile to consult multiple experts and combine their probabilities"; "in general, simple combination approaches perform quite well"; "heterogeneity among experts is highly desirable," and it is suggested to use three to five experts; and rather than finding a single, all-purpose aggregation approach, an approach should be designed depending on the details of each individual situation.

Bunn (1975) quantified the weights in a linear aggregation approach using the probability of a model outperforming the other models. Although his interpretation of probability is different from ours, he also used Bayes' law to update probabilities using historical data, which is similar to our method. Dillon et al. (2002) presented a Monte Carlo-based approach to aggregate forecasts. They assigned each individual forecast the same weight. We use the same Monte Carlo-based approach but assign each individual forecast a different weight depending on its probability.

In the remainder of this paper, we first briefly review four decline-curve models. We then introduce an approach to assess model probabilities that can be used for probabilistic DCA. An example with synthetic data is used to illustrate the value of considering model uncertainty, followed by an application of our proposed approach, which combines multiple models in probabilistic DCA, with real data from two unconventional fields. Finally, we present a discussion and conclusions.

Decline-Curve Models

One of the most popular models for DCA is the Arps model. It has been used for more than 60 years. However, as the production from unconventional resources has increased over the past decade or more, it has been realized that the Arps model might not be ideal for forecasting unconventional production. Thus, several alternative models have been proposed. In this work, we focus on four models: the Arps model, the SEM, the LGM, and the Pan CRM.

Arps Model. The Arps model is an empirical model formulated as

$$q_t = \begin{cases} q_0(1+bD_i t)^{-1/b}, & 0 < b \le 1 \\ q_0 e^{-D_i t}, & b = 0, \end{cases} \quad (1)$$

where q_t is the rate at time t, q_0 is the initial rate at $t = 0$, D_i is the initial decline rate, and b is the decline exponent. With the assumption of boundary-dominated flow, the upper bound of b is unity. For the transient flow regime of unconventional production, b is often greater than unity (Valko and Lee 2010) and, as a result, the cumulative production is unbounded. To avoid unbounded cumulative production, we restrict the range of b to be $0 \le b < 1$ (the cumulative production will also be unbounded when $b = 1$) in our application.

SEM. On the basis of the analysis of Barnett Shale wells, Valko (2009) presented the SEM to formulate an empirical time–rate relation,

$$q_t = q_0 e^{-\left(\frac{t}{\tau}\right)^n}, \quad (2)$$

where τ is the characteristic time parameter and n is the exponent parameter. Valko and Lee (2010) interpreted the SEM the following way: The actual production decline is an integrated effect of multiple contributing volumes, each in its individual exponential decay, with a specific distribution of characteristic time constants. The distribution can be determined by τ and n—τ is the median of the characteristic time constants, and n describes the fatness of the tail of the distribution.

LGM. LGM is an empirical model originally developed for modeling population growth (Verhulst 1845; Pearl and Reed 1920). It describes the limit of biological population growth. Hubbert (1956) adopted this model to model production for entire fields or producing regions. Clark et al. (2011) applied it to model production in a single unconventional well. The specific formulation of the LGM used is

[3]In the remainder of this paper, "probability of model" or "model probability" is referred to as the probability of a model being relatively good to describe flow behaviors, compared with other models.

[4]We use "a model" to refer to a mathematical formulation for describing the physics of a phenomenon. For a given model, its parameter values are uncertain. This is referred to as "uncertainty in model parameters." On the other hand, we can be uncertain regarding which mathematical formulation (i.e., which model) should be used to describe the physics of a phenomenon. This is referred to as "model uncertainty" or "uncertainty in models." When only a single model is considered, many authors have used "model uncertainty" to refer to uncertainty in the model parameters, but it has a different meaning in this paper.

$$q_t = \frac{aK\eta t^{\eta-1}}{(a+t^{\eta})^2}, \quad \ldots\ldots (3)$$

where a is a constant, K is the carrying capacity, and η is the hyperbolic exponent parameter. The carrying capacity K is important for DCA because it has a physical meaning: the total amount of hydrocarbon that can be recovered (i.e., ultimate oil recovery) under the primary recovery mechanism. By fitting this model to oil-production data, the ultimate oil recovery can be estimated directly rather than by extrapolation. The hyperbolic exponent parameter η controls the decline behavior; the larger the parameter value, the slower the decline. The constant a and the hyperbolic exponent parameter η together determine the time when one-half of the carrying capacity K has been produced (i.e., the cumulative production at $t = \sqrt[\eta]{a}$ is $K/2$).

Pan CRM. Pan (2016) proposed a model to capture the productivity-index behavior over both linear transient and boundary-dominated flow,

$$J = \frac{\beta}{\sqrt{t}} + J_{\infty}, \quad \ldots\ldots (4)$$

where J is the productivity, β is the linear transient flow parameter, and J_{∞} is the constant productivity index that a well will eventually reach at boundary-dominated flow. β is related to the permeability in the analytical solution of linear flow into hydraulically fractured wells presented by Wattenbarger et al. (1998). By combining Eq. 4 and a tank material-balance equation for a slightly compressible fluid, Pan (2016) derived the analytical solution of rate over time,

$$q_t = \Delta P\left(\frac{\beta}{\sqrt{t}} + J_{\infty}\right)e^{-(2\beta\sqrt{t}+J_{\infty}t)/(c_t V_p)}, \quad \ldots\ldots (5)$$

where ΔP is the difference between the initial reservoir pressure and the assumed constant flowing bottomhole pressure, c_t is the total compressibility, and V_p is the drainage pore volume. The Pan CRM gives an unrealistically large rate for small t, as q_t approaches infinity when t approaches zero, because the flow rate is not defined at $t = 0$. To deal with this issue, we use this model for production after 10 days (i.e., $t \geq 10$). The Pan CRM is the only one of the four models that is derived analytically and possesses all parameters related with physical quantities of a reservoir system, although it is derived for slightly compressible single-phase-fluid flow. Ruiz Maraggi (2018) provides a detailed description of the relationships between the model parameters and the physical quantities.

Determination of Model Parameters

In the context discussed in this paper, the interest is future oil production. Future oil production is predicted using production models (e.g., decline-curve models), which describe flow behaviors with uncertain model parameters; once the values for these parameters are established, the future oil production is calculated by forward modeling. For a given model, the parameters are determined through history matching, with the goal to minimize a predefined loss (or objective) function by adjusting the model parameters.

Deterministic Approach. In a deterministic approach, a model and the parameter setting that best fit the given data are found and used for forecasting. The best-fit model gives a single estimate of our interest (e.g., future production). A widely used approach is the least-squares estimation (LSE) that aims to minimize the difference between model forecast and measured data. Thus, the loss function of LSE is defined as

$$L_{LSE}(\boldsymbol{x}) = \sum_{k=1}^{T}[q_k(\boldsymbol{x}) - \hat{q}_k]^2, \quad \ldots\ldots (6)$$

where $L_{LSE}(\boldsymbol{x})$ is the loss function of the LSE, which is a function of the vector of model parameters $\boldsymbol{x}$; $q_k(\boldsymbol{x})$ is the model-forecast rate at timestep k; $\hat{q}_k$ is the measured rate (i.e., data) at timestep k; and T is the total number of data points.

Another approach is the MLE, which aims to maximize the likelihood function (i.e., the conditional probability of observing the data given a model). Assuming that the data measurements are independent, the likelihood function is

$$P(\hat{q}_1, \hat{q}_2, \ldots, \hat{q}_T|\boldsymbol{x}) = \prod_{k=1}^{T} P(\hat{q}_k|\boldsymbol{x}), \quad \ldots\ldots (7)$$

where P denotes probability and $P(\hat{q}_k|\boldsymbol{x})$ is the conditional probability of observing $\hat{q}_k$ given the model with its parameter setting $\boldsymbol{x}$. If the measurement of $\hat{q}_k$ has a Gaussian random error with zero mean and standard deviation (SD) σ_k, we have

$$P(\hat{q}_k|x) = \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-[q_k(x)-\hat{q}_k]^2/(2\sigma_k^2)}. \quad \ldots\ldots (8)$$

Thus,

$$P(\hat{q}_1, \hat{q}_2, \ldots, \hat{q}_T|\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^T \prod_{k=1}^{T}\sigma_k^2}} e^{-\frac{1}{2}\sum_{k=1}^{T}\frac{[q_k(x)-\hat{q}_k]^2}{\sigma_k^2}}. \quad \ldots\ldots (9)$$

Maximizing the likelihood function (Eq. 9) is equivalent to minimizing

$$L_{MLE}(\boldsymbol{x}) = \sum_{k=1}^{T}\frac{[q_k(x)-\hat{q}_k]^2}{\sigma_k^2}. \quad \ldots\ldots (10)$$

If $\sigma_1 = \sigma_2 = \cdots = \sigma_T$, Eq. 10 can be reduced to Eq. 6. Therefore, LSE is a special case of MLE. The advantage of using Eq. 10 instead of Eq. 6 is that σ_k acts as a weighting factor so that a more-accurate data point (i.e., with small σ_k) will have more weight than a less-accurate data point (i.e., with large σ_k).

A practical issue when using MLE is that σ_k, which is seldom known a priori, must be assessed. To estimate σ_k, a moving-window approach is used. For example, to assess σ_k, we first assign a value to the half-window width w, then take a subset of data $\{\hat{q}_{k-w}, \ldots, \hat{q}_{k-1}, \hat{q}_k, \hat{q}_{k+1}, \ldots, \hat{q}_{k+w}\}$ and calculate their residuals, and finally set σ_k to be equal to the sample SD of the residuals. In particular, we use the "rlowess" method in MATLAB (MathWorks 2017a) to construct a nonparametric fitting curve. The "rlowess" method regresses the subset of data in a moving window using weighted linear least squares and a first-degree polynomial model. The residuals between the data and the fitting curve are calculated. The SDs of the residuals calculated using the moving window are used as the SDs of the measurement errors. The width of the moving window is optimized using cross validation (Hastie et al. 2001; MathWorks 2017b). The steps of performing cross validation are the following:

1. The original data set is partitioned into κ ($\kappa = 10$ in our application) randomly chosen subsets.
2. One subset is selected, and the remaining subsets are used to construct the fitting curve.
3. The selected subset is used to validate the goodness of the fitting curve by calculating the error between the fitting curve and the selected subset.
4. This process is repeated for each subset.
5. The average cross-validation error is used as an indicator for the goodness of fitting.

The cross validation is performed for different values of the half-window width; the half-window width that leads to the lowest average cross-validation error is the optimal one. To avoid a too-small half-window width, we set the minimal half-window width to be 3. In this way, a subset of data with larger fluctuations will have larger σ_k, and consequently will be assigned less weight.

Probabilistic Approach. Instead of estimating a single value, a probabilistic approach assesses a distribution of interest with the consideration of uncertainties in measurements, inverse modeling[5], and model type. The latter will be detailed in the next section, Integrating Model Uncertainty in Probabilistic DCA.

Jochen and Spivey (1996) illustrated the use of the bootstrap method combined with DCA to provide a probabilistic reserves estimation. The bootstrap method assumes a noninformative (i.e., uniform) prior, so it does not require a priori knowledge on the distribution of model parameters. Because of unknown measurement errors, the bootstrap method includes uncertainties in measurements by using the Monte Carlo method to resample data sets from the original data set with replacement. For each resampled data set, the model parameters are determined using LSE. As the bootstrap repeats for numerous resampled data sets, the posterior distributions[6] of model parameters can be obtained.

Tavassoli et al. (2004) presented the issue of nonuniqueness in inverse modeling (the inherent uncertainty in history matching) and showed that different combinations of model parameters can give nearly equally good matching quality but different forecasts. Sayarpour et al. (2011) started with different sets of initial guesses of model parameters to history match data to generate numerous history-matched solutions of model parameters. To allow for the uncertainty in inverse modeling, we start with different sets of initial guesses of model parameters when history matching different sampled data sets.

Integrating Model Uncertainty in Probabilistic DCA

The approaches presented in the preceding section assume that a single model, which predicts reservoir performance, has been proposed. None of the reviewed publications discussed a systematic approach of determining which decline-curve model is the best for unconventional production. Often, it is difficult to select such a best model because satisfactory history-matching results can be produced by several different models. Instead of determining a single best model, we propose including the model uncertainty in the probabilistic approach by using a Bayesian approach to assess the probability of each candidate model being good to describe flow behaviors. The idea of considering multiple models and using Bayesian approaches to calculate their probabilities for probabilistic analysis is not abnormal because applications of this can be found in the oil and gas literature (Park et al. 2013; Aanonsen et al. 2019), although they are few.

This section illustrates our proposed approach to integrating model uncertainty in probabilistic DCA. We call this approach the multiple-model probabilistic MLE approach (MM-P) because it considers multiple models and applies probabilistic DCA with MLE for data matching. More specifically, in the following, we use single-model probabilistic MLE (SM-P) to refer to the case where probabilistic DCA with MLE is applied for a single model, and single-model deterministic MLE (SM-D) to refer to the case where deterministic DCA with MLE is applied for a single model.

The procedure of performing MM-P is as follows. A Gaussian random error with zero mean and SD σ_k is used to quantify uncertainty in the measurement of rate $\hat{q}_k$. The SD σ_k is estimated using the moving-window approach discussed previously. As in the bootstrap method, numerous data sets are sampled; however, because the SD of measurement error at each data point has been assessed, we can sample a data value from a Gaussian distribution with mean $\hat{q}_k$ (the originally measured data value) and the estimated SD σ_k for each data point. After matching many data points, all the resulting parameter settings of a given model tend to produce similar forecasts and lose the ability to reflect measurement uncertainty in forecasting. To address this issue, we sort the sampled values at each data point and pick the values in the same rank to form a sampled data set (i.e., all the largest sampled values at each data point constitute a sampled data set, all the second-largest sampled values at each data point constitute another sampled data set, and so forth). **Table 1** shows an example of the sampled data sets without and with sorting. For simplicity, the example contains only four sampled data sets for three data points. The first sampled data set is 1,766, 1,080, and 1,002 B/D over time. After sorting, it becomes 1,766, 1,426, and 1,002 B/D over time, which consists of the largest sampled values at each data point. The sorting will not introduce any bias because the means and SDs stay unchanged before and after sorting. Because these sampled data sets are generated using Monte Carlo simulation, each of them has a probability (or weight) of 1/4. **Fig. 1** illustrates the effect of sorting the sampled data values on the matching result. Without sorting, the P10 to P90[7] interval is very narrow after matching the data with Pan CRM. With sorting, the P10 to P90 interval is wider and better captures the fluctuations in the data.

[5]We use "uncertainty in inverse modeling" interchangeably with "nonuniqueness in inverse modeling" to describe the fact that multiple combinations of model-parameter values might give equally good matches to data.

[6]In the context of history matching, the "posterior distribution" is the distribution of the interested uncertain quantity given additional information.

[7]In this paper, a percentile is defined as the value below which a given percentage of estimates fall. For example, P10 is the value below which 10% of the estimates fall.

Sampled Data Set Without Sorting		Data Point No.		
		1	2	3
Sampled Data Set No.	1	1,766	1,080	1,002
	2	1,763	1,426	903
	3	1,664	1,271	826
	4	1,346	1,092	991
	Mean	1,635	1,217	931
	SD	198	164	83

Sampled Data Set With Sorting		Data Point No.		
		1	2	3
Sampled Data Set No.	1	1,766	1,426	1,002
	2	1,763	1,271	991
	3	1,664	1,092	903
	4	1,346	1,080	826
	Mean	1,635	1,217	931
	SD	198	164	83

Table 1—Example of sampled data set without and with sorting.

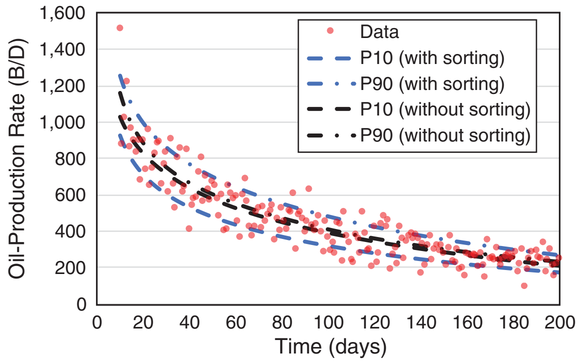

Fig. 1—Illustration of the effect of sorting sampled data values.

MLE is used to determine the parameters of each model considered for a given sampled data set. Bayes' law is then used to calculate the posterior probability of each model with its history-matched-parameter setting given the sampled data set,

$$P\left(\boldsymbol{x}_{ij}^{MLE}, m_i | \boldsymbol{d}_j, \Omega'\right) = \frac{P\left(\boldsymbol{d}_j | \boldsymbol{x}_{ij}^{MLE}, m_i\right) P\left(\boldsymbol{x}_{ij}^{MLE}, m_i | \Omega\right)}{\sum_{i'} P\left(\boldsymbol{d}_j | \boldsymbol{x}_{i'j}^{MLE}, m_{i'}\right) P\left(\boldsymbol{x}_{i'j}^{MLE}, m_{i'} | \Omega\right)}, \quad (11)$$

where $\boldsymbol{x}_{ij}^{MLE}$ is the parameters of the MLE solution given model m_i and sampled data set $\boldsymbol{d}_j$, Ω denotes a priori knowledge, and Ω' denotes a posteriori knowledge. Using a noninformative prior distribution [i.e., the probabilities of different models and different parameter settings are equal: $P(\boldsymbol{x}_{m1}^{MLE}, m_1|\Omega) = P(\boldsymbol{x}_{m2}^{MLE}, m_2|\Omega) = \cdots$], Eq. 11 is reduced to

$$P\left(\boldsymbol{x}_{ij}^{MLE}, m_i | \boldsymbol{d}_j, \Omega'\right) = \frac{P\left(\boldsymbol{d}_j | \boldsymbol{x}_{ij}^{MLE}, m_i\right)}{\sum_{i'} P\left(\boldsymbol{d}_j | \boldsymbol{x}_{i'j}^{MLE}, m_{i'}\right)}. \quad (12)$$

Inserting Eqs. 9 and 10 into Eq. 12, we obtain

$$P\left(x_{ij}^{MLE}, m_i | \boldsymbol{d}_j, \Omega'\right) = \frac{e^{-\frac{1}{2} L_{MLE}\left(x_{ij}^{MLE}\right)}}{\sum_{i'} e^{-\frac{1}{2} L_{MLE}\left(x_{i'j}^{MLE}\right)}} = 1 \Big/ \sum_{i'} e^{-\frac{1}{2}\left[L_{MLE}\left(x_{i'j}^{MLE}\right) - L_{MLE}\left(x_{ij}^{MLE}\right)\right]}. \quad (13)$$

The probability calculated using Eq. 13 is the probability of the estimate of interests forecasted by model m_i with its parameters $\boldsymbol{x}_{ij}^{MLE}$ given a sampled data set $\boldsymbol{d}_j$. Because $\boldsymbol{d}_j$ is an independent Monte Carlo sample, $P(\boldsymbol{d}_j) = 1/N$ and

$$P\left(\boldsymbol{x}_{ij}^{MLE}, m_i, \boldsymbol{d}_j | \Omega'\right) = P\left(\boldsymbol{x}_{ij}^{MLE}, m_i | \boldsymbol{d}_j, \Omega'\right) P(\boldsymbol{d}_j) = 1 \Big/ \left\{ N \sum_{i'} e^{-\frac{1}{2}\left[L_{MLE}\left(x_{i'j}^{MLE}\right) - L_{MLE}\left(x_{ij}^{MLE}\right)\right]} \right\}, \quad (14)$$

where N is the total number of Monte Carlo samples ($N = 1{,}000$ in the following applications). The probability calculated using Eq. 14 is used to weight the estimate forecasted by model m_i with its parameters $\boldsymbol{x}_{ij}^{MLE}$. By repeating this process over the models and the sampled data sets, we obtain a set of weighted samples of the estimate as shown in **Table 2,** where M is the total number of models considered, f_{ij} is the estimate forecasted using m_i and $\boldsymbol{x}_{ij}^{MLE}$, and P_{ij} denotes $P\left(\boldsymbol{x}_{ij}^{MLE}, m_i, \boldsymbol{d}_j | \Omega'\right)$. This set of weighted samples represents the posterior distribution of the estimate with the integration of model uncertainty. The statistics (including mean, mode, and percentiles) of the posterior distribution can be further used for probabilistic analysis of future oil production and reserves, which is more useful and essential for decision support than a deterministic analysis. For example, the mean of the estimates is $\bar{f} = \sum_{i=1}^{M} \sum_{j=1}^{N} P_{ij} f_{ij}$. As will be shown later, the statistics of MM-P and SM-P results can be very different, and using MM-P can quantify the uncertainty in the production forecasts better than SM-P.

Estimate forecasted using $\boldsymbol{m}_i$ and $\boldsymbol{x}_{ij}^{MLE}$	f_{11}	…	f_{M1}	f_{12}	…	f_{M2}	…	f_{1N}	…	f_{MN}
Weight assigned to the estimate	P_{11}	…	P_{M1}	P_{12}	…	P_{M2}	…	P_{1N}	…	P_{MN}

Table 2—Set of weighted samples of estimate.

The marginal posterior probability of a model over all the sampled data sets ($\boldsymbol{d}_j$ for $j = 1, 2, \ldots, N$) is calculated as

$$P(m_i|\Omega') = \sum_{j=1}^{N} P\left(\boldsymbol{x}_{ij}^{MLE}, m_i, \boldsymbol{d}_j|\Omega'\right). \quad (15)$$

$P(m_i|\Omega')$ is referred to as model probability, which is a measure of the relative truthfulness of model i in comparison with the other models for a given data set. The candidate models can be ranked according to their model probabilities.

Fig. 2 summarizes the work flow of using the proposed MM-P approach. The result provided by the proposed approach is referred to as the "combined" result in the remainder of this paper because it combines the results of multiple models.

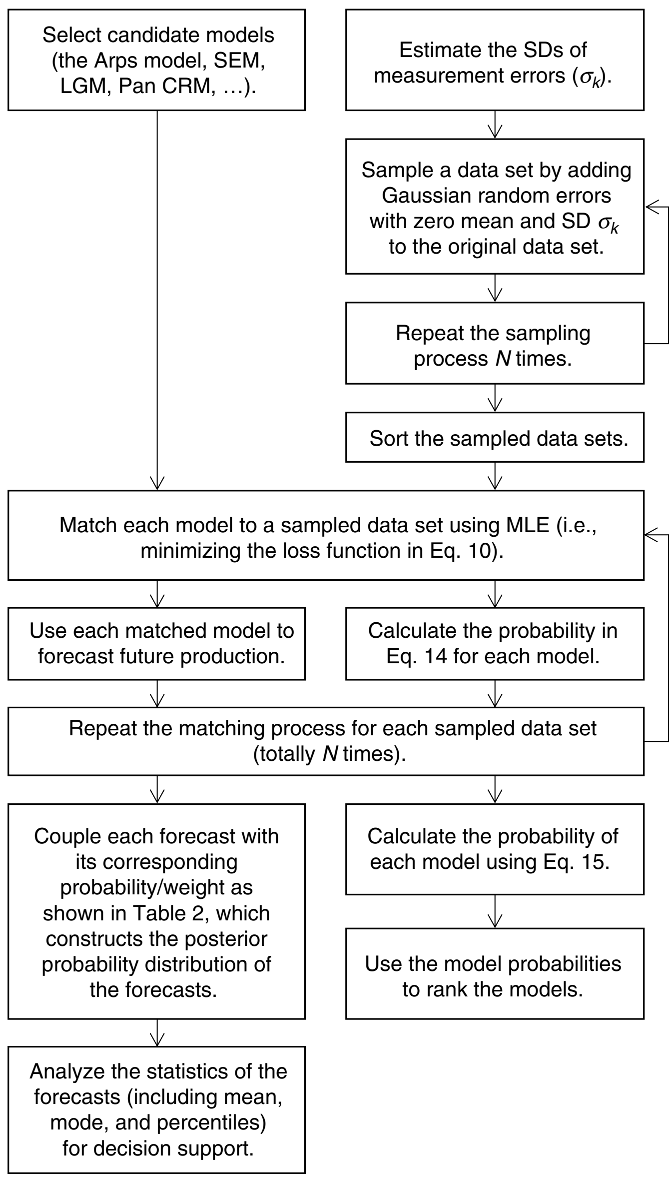

Fig. 2—Work flow of the proposed MM-P approach.

Illustrative Example With Synthetic Data

In this section, we use a synthetic data set to illustrate the effects of using the deterministic and probabilistic approaches without considering the model uncertainty (i.e., SM-D and SM-P, respectively) on cumulative-oil-production estimation, and to highlight the value of using the MM-P approach proposed in this work. We consider the Arps model, SEM, LGM, and Pan CRM for this example, but more models can easily be included.

The "true" decline is generated using the Pan CRM. Random errors, drawn from Gaussian distributions with zero mean and SD of 20% of the "true" rate, are added to the "true" decline to form the synthetic data set. **Fig. 3** illustrates the synthetic data set and the "true" decline of the oil-production rate as well as the estimated SDs of measurement errors. Our interest is the cumulative oil production from the time of the last data point to Day 10,950 (Year 30). The "true" cumulative oil production given by the Pan CRM is 48.1 million bbl (Mbbl). This value is used as a reference for subsequent estimates.

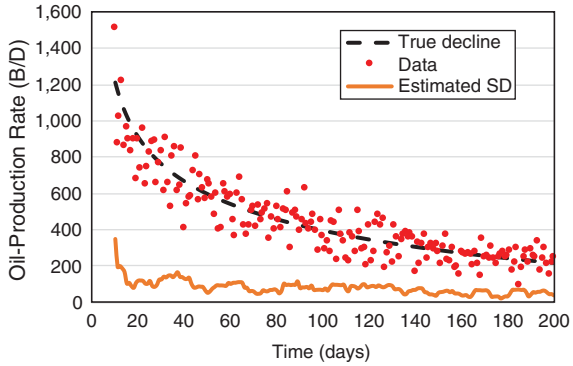

Fig. 3—Synthetic data set to Day 200. The "true" decline is used to generate the "data."

We first use SM-D to estimate the cumulative oil production. The resulting best-fit parameters and corresponding loss-function value of MLE for each model are listed in **Table 3.** The Arps model with $q_0 = 1,185.5$ B/D, $b = 0.88$, and $D_i = 0.019$ day^{-1} best fits the data (minimal loss function). Therefore, it is used for forecasting, and provides an estimated cumulative oil production of 174.2 Mbbl. This estimate is 3.6 times the "true" value. **Fig. 4** shows the best-fit curve of each model. From Fig. 4a, we can see that all the models fit the data quite well; this is also indicated by the fact that the loss-function values (L_{MLE}) of the models in Table 3 are very close to each other. However, the models' forecasts are very different, as shown in Fig. 4b, and result in very different values of estimated cumulative oil production, as listed in Table 3. It makes sense that the estimated cumulative oil production of the "correct" model (the Pan CRM) is the closest to the "truth." However, the Pan CRM does not give the smallest loss-function value. The reason for this will be discussed later.

Arps Model		SEM		Pan CRM		LGM	
q_0 (B/D)	1,185.5	q_0 (B/D)	1,657.0	ΔP (psi)	812.6	a (days)	113.5
b	0.88	τ (days)	47.5	J_∞ (B/D/psi)	0.72	K (Mbbl)	206.8
D_i (day^{-1})	0.019	n	0.48	$c_t V_p$ (bbl/psi)	179.8	η	0.87
				β (B/D$^{1/2}$/psi)	2.6		
L_{MLE}	213.3	L_{MLE}	213.6	L_{MLE}	215.0	L_{MLE}	213.5
Estimated cumulative production (Mbbl)	174.2	Estimated cumulative production (Mbbl)	71.4	Estimated cumulative production (Mbbl)	43.7	Estimated cumulative production (Mbbl)	102.8

Table 3—Best-fit parameters for the Arps model, SEM, Pan CRM, and LGM for the synthetic data set to Day 200 and estimated cumulative oil production to Day 10,950.

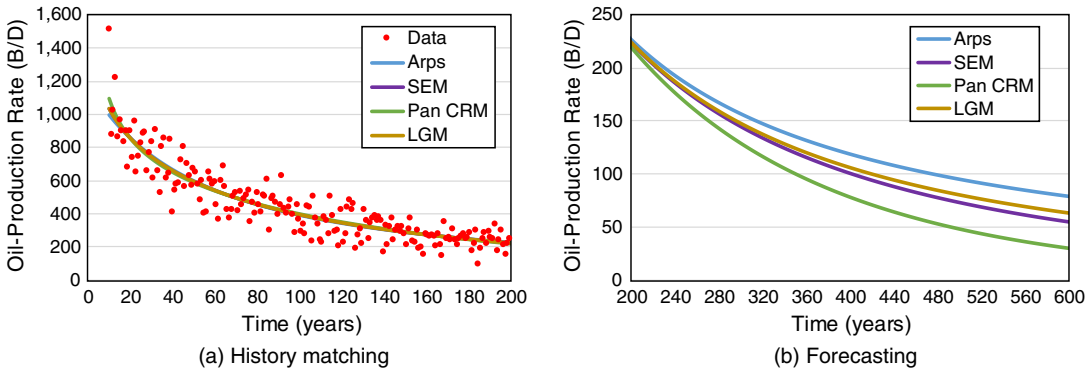

Fig. 4—Plots of the deterministic history matching and forecasting results for the Arps model, SEM, Pan CRM, and LGM.

To take the uncertainties in measurements and inverse modeling into consideration, we use SM-P with the Arps model. This produces a distribution of the cumulative oil production, the P10, P50, mean, and P90 of which are listed in **Table 4.** The forecast of the Arps model is biased, far from the "truth" (48.1 Mbbl), and the 80% confidence interval (from P10 to P90) does not contain the "truth."

We use MM-P to integrate the model uncertainty in this analysis. Indeed, the minimized loss-function values of the four models (Table 3) are very close to each other, which means no model is superior to the others. The marginal posterior probabilities calculated

using Eq. 15 are 33.9, 27.8, 18.2, and 20.1% for the Arps model, SEM, Pan CRM, and LGM, respectively. This result seems counterintuitive because the "correct" model (the Pan CRM) is the least-probable one. This is because the original data set shows no obvious linear transient flow behavior that the Arps model cannot capture. The linear transient behavior is masked by the noise in the original data set. Thus, the model uncertainty remains large even when the data set is given.

	P10	Mean	P50	P90
Cumulative oil production (Mbbl)	128.9	168.0	171.4	200.6

Table 4—Statistics of the cumulative oil production forecast using the Arps model given the synthetic data set to Day 200.

Fig. 5 shows the box plots (indicating P10, P25, P50, P75, P90, and mean) of cumulative oil production forecast using SM-P with only the Arps model, SEM, Pan CRM, or LGM, and using MM-P to combine the models' forecasts. Among the four models, the LGM has the largest uncertainty in inverse modeling because it gives the largest 80% confidence interval, whereas the Pan CRM has the smallest uncertainty in inverse modeling. Because the Pan CRM is the model used to generate the data, its estimate is the best. Although the 80% confidence interval of the combined forecast does not contain the "truth" (the P10 is slightly larger than the "truth"), using MM-P improves the forecast, compared with using SM-P with the best-fit model (the Arps model). Compared with the Arps model's forecast, the 80% confidence interval of the combined forecast is wider because of the inclusion of model uncertainty, and is closer to the "truth."

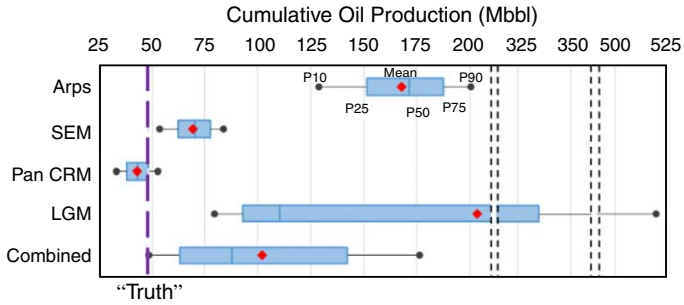

Fig. 5—Box plots of cumulative oil production forecast using a single model, compared with that of the combined forecast, given the synthetic data to Day 200[8].

We extend the data set to Day 400 and use MM-P to estimate the cumulative oil production from Day 400 to Day 10,950 (Year 30). The data set and estimated SDs of measurement errors are shown in **Fig. 6.** The resulting marginal posterior probabilities are 8.4, 30.9, 56.3, and 4.3% for the Arps model, SEM, Pan CRM, and LGM, respectively. Given the additional data from Day 200 to Day 400, the Pan CRM becomes the most probable, whereas the LGM is the least probable and the Arps model is the second-least probable. **Fig. 7** shows the matching qualities of each model. **Fig. 8** shows the box plots of cumulative oil production forecast using SM-P with only the Arps model, SEM, Pan CRM, or LGM, and using MM-P. Because the Pan CRM and the SEM now have much-larger probability than the Arps model and the LGM, the combined distribution using MM-P is skewed toward the distributions forecasted by only the SEM and the Pan CRM with a tail toward that of the Arps model and LGM. When considering the Arps model, SEM, or LGM individually, none of these models provides an 80% confidence interval containing the "truth." However, when all the four models are considered such that the "correct" model is included, the estimate is improved because the 80% confidence interval of combined distribution contains the "truth," which is very close to the P10.

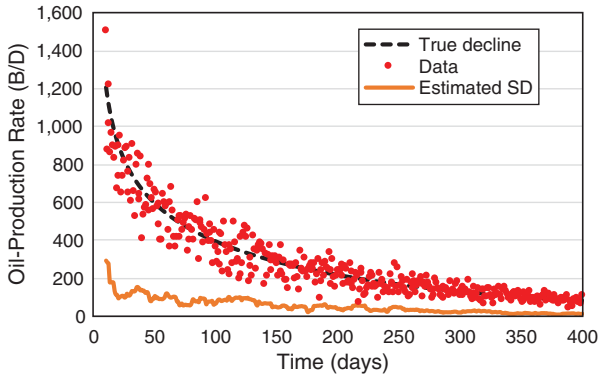

Fig. 6—Synthetic data set to Day 400.

This example indicates that including multiple models for forecasting reduces the risk of rejecting a good model by selecting a single best-fit model.

[8]The dashed black lines in Figs. 5, 8, 12, and 15 mark where the figures are cut for saving space.

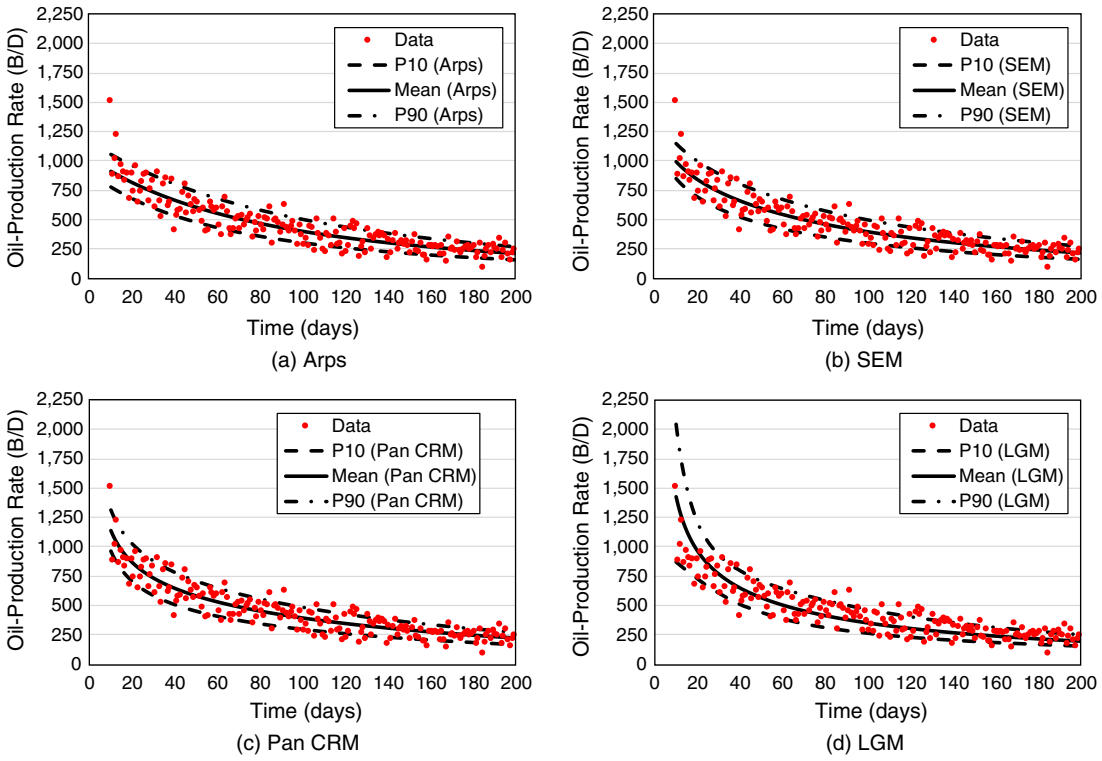

Fig. 7—Plots of the probabilistic history-matching results for the Arps model, SEM, Pan CRM, and LGM.

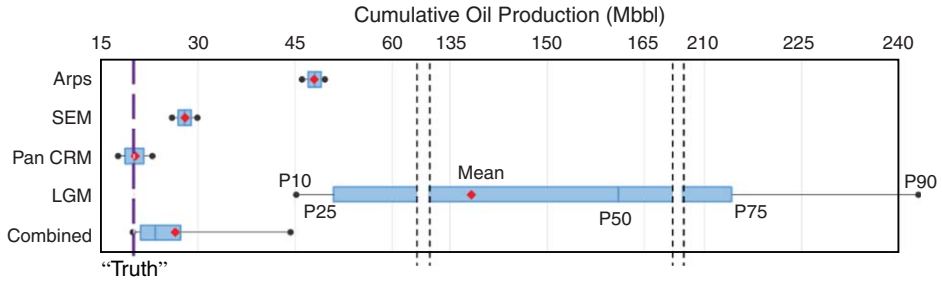

Fig. 8—Box plots of cumulative oil production forecast using solely one model, compared with that of the combined forecast, given the synthetic data to Day 400.

Application to Bakken Field Data

The proposed approach, MM-P, is applied to the daily oil-production-rate data from the Bakken field. The Bakken wells we study in this work are completed in a hydraulically fractured shale reservoir. All the wells for this study are completed in the same formation. First, we conduct a hindcast test, where the first part of a data set is used for history matching and the second part is used for comparison with the forecast. We then use the proposed approach to forecast the cumulative oil production for the selected wells.

Hindcast Test. The oil-production-rate data are from Bakken Well UT-ID 220 (red dots in **Fig. 9**). There are 660 data points in total, covering the period from Day 10 to Day 695. The rates are measured daily. We use the data from Day 10 to Day 100 for history matching, and we then use the resulting models for forecasting. The forecasts are then compared with the measured data from Day 100 forward. The Arps model, SEM, Pan CRM, and LGM are considered. The resulting model probabilities are 6.1, 27.2, 36.2, and 30.5% for the Arps model, SEM, Pan CRM, and LGM, respectively. Although the Pan CRM is the most-probable model, the probabilities of the Arps model, SEM, and LGM are significant, and thus these models should also be included in the forecast using MM-P analysis. The combined forecast from Day 100 to Day 695 is shown in Fig. 9. Most of the data fall within the 80% confidence interval.

We use additional data points to investigate the effect of more data on the forecast. Given data from Day 10 to Day 200, the probability of the Pan CRM increases from 36.2% (given the data from Day 10 to Day 100) to 70.6%, the probability of the LGM drops to 22.5%, the probability of the SEM drops to 6.9%, and the probability of the Arps model drops to 0%. The forecast of the oil-production rate from Day 200 to Day 695 is shown in **Fig. 10.** The forecast mean matches the data better and the uncertainty is reduced, compared with that shown in Fig. 9. The 80% confidence interval is very narrow between Days 200 and 350 because the models are "forced" to match the data points between Days 180 and 200. The fluctuation of these data points is very small, and thus they are assigned much-greater weights for history matching than the other data points. Overweighting the data points between Days 180 and 200 leads to the failure of the 80% confidence interval to cover most of the data points between Days 200 and 350 (most of these data points are above the P90 curve).

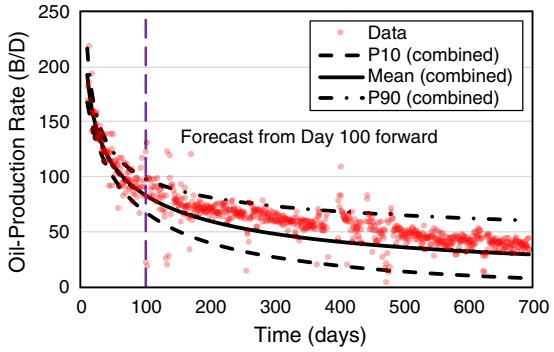

Fig. 9—Forecast from Day 100 to Day 695 compared with the corresponding data from Bakken Well UT-ID 220.

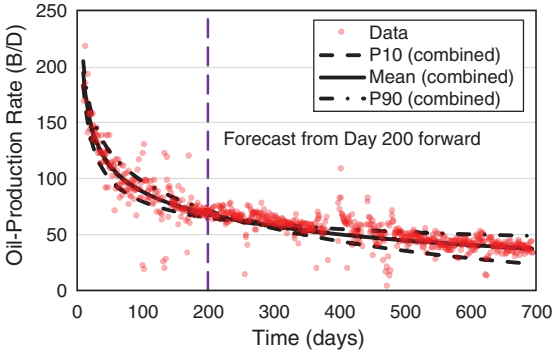

Fig. 10—Forecast from Day 200 to Day 695 compared with the corresponding data from Bakken Well UT-ID 220.

Model Probabilities and Cumulative-Oil-Production Estimates for Bakken Wells. We select 28 Bakken wells with relatively less-noisy data. The model probabilities and cumulative oil production are assessed for these wells. The wells have different operation times, with short and long production histories. The production-history length of each well is listed in **Table 5.**

Well UT-ID	222	208	401	195	69	194	413	187	216	804	78	223	80	209
Length (days)	207	218	233	260	290	309	335	337	369	375	387	392	484	497

Well UT-ID	232	265	385	227	220	391	386	228	67	4	198	197	81	82
Length (days)	515	527	545	672	695	764	885	993	1,132	1,222	1,306	1,376	1,531	1,642

Table 5—Production-history lengths of selected Bakken wells.

This analysis considers the Arps model, SEM, LGM, and Pan CRM. The marginal posterior probabilities of the decline-curve models are shown in **Fig. 11.** Among the 28 selected wells, the Pan CRM is the most probable for 12 wells, the SEM for nine wells, the Arps model for four wells, and the LGM for three wells. This confirms that the Arps model might not be suitable for unconventional plays and indicates that the Pan CRM, as the only physics-based model, is more probable to well-describe the unconventional-flow behaviors than the empirical models. Nevertheless, no one model is the most probable for all the wells.

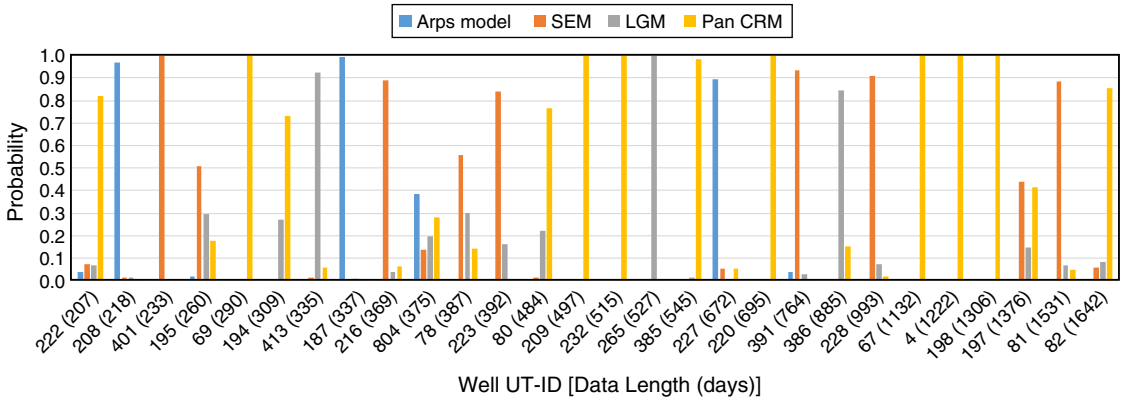

Fig. 11—Marginal posterior probabilities of the decline-curve models for the 28 selected Bakken wells.

For most of the wells, one of the models is dominating with a probability greater than 90%. For some wells, the model uncertainty remains large, even for some wells with long production histories (such as Well UT-ID 197). Thus, using solely one model for further analysis might underestimate the uncertainty, even when there is a large amount of data.

We see two wells (Wells UT-ID 208 and 187) for which the Arps model dominates (i.e., model probability greater than 90%). The reason might be that the data are so noisy that this masks the oil-decline features, or that the fracture flow is indeed not dominating in those wells.

Considering a well's life cycle of 30 years, we are interested in how much oil is left in each well (i.e., the cumulative oil production from the last day with data to Day 10,950). The mean value of the cumulative oil production for each of the selected Bakken wells is shown in **Fig. 12.** Fig. 12 includes the mean values estimated using only one of the models as well as the mean value using MM-P to combine the individual models' estimates. The box plots of the cumulative oil production estimated by MM-P are in **Fig. 13.** In the box plots, a percentile value on the *y*-axis is its normalized deviation from the mean [i.e., normalized percentile = (percentile–mean)/mean].

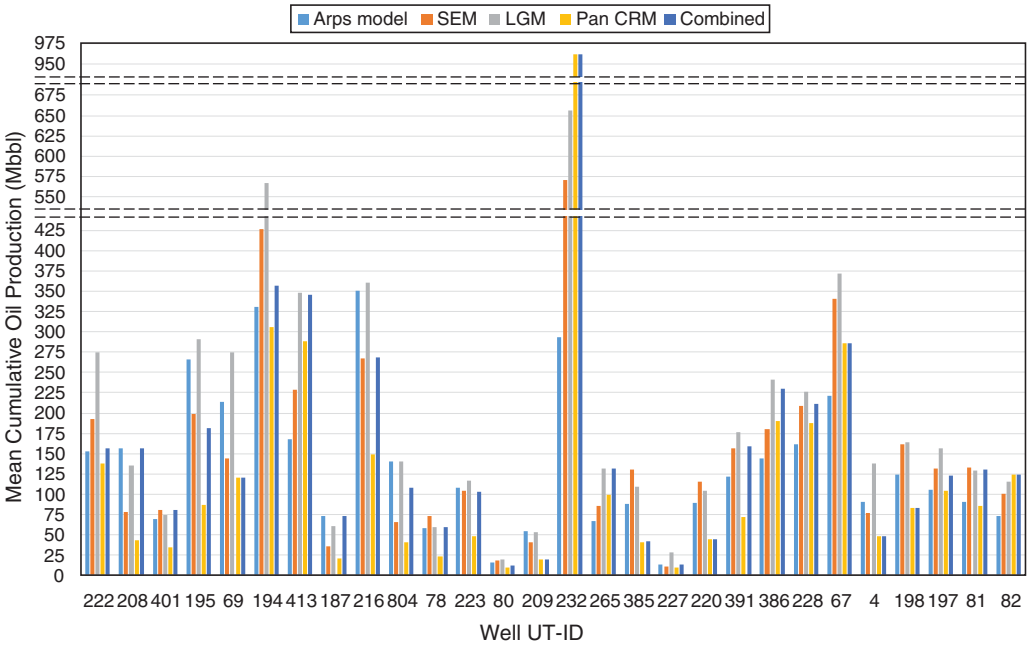

Fig. 12—Mean values of estimated cumulative oil production for the 28 selected Bakken wells.

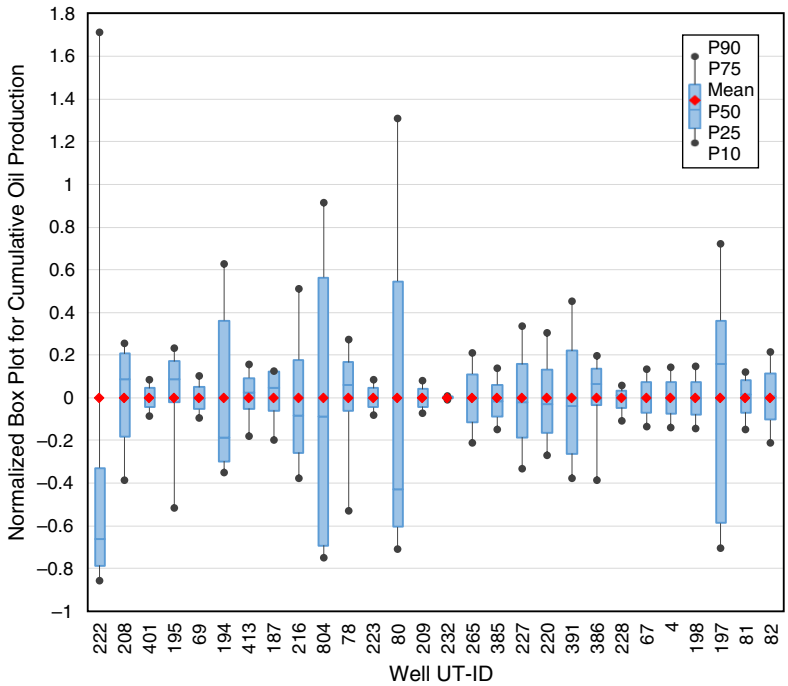

Fig. 13—Box plots of estimated cumulative oil production using MM-P for the 28 selected Bakken wells.

From Fig. 12, we see that the LGM provides the highest estimate of mean cumulative oil production for 18 wells, the SEM for five wells, and the Arps model for three wells. The Pan CRM provides the smallest estimate of mean cumulative oil production for 21 wells. Because the Pan CRM is the most-probable model for most of the wells, we will tend to overestimate the cumulative oil production if only the Arps model, SEM, or LGM is used. When MM-P is used, the mean values of 16 wells are somewhere between the highest and lowest mean values estimated using the models individually; six wells hit the highest mean values estimated using the models individually, and six wells hit the lowest mean values estimated using the models individually. Thus, using MM-P to include model uncertainty can adjust the estimate so that it is neither too optimistic nor too pessimistic.

In Fig. 13, the box plots of most of the wells are quite symmetric. Nearly half of the wells have an 80% confidence interval bounded between –0.2 and 0.2; that is, the P10 or P90 deviates less than 20% from the mean. Well UT-ID 222 has an extremely skewed distribution and wide 80% confidence interval because its data are very noisy. For nearly all the wells with wide 80% confidence interval (Wells UT-ID 222, 195, 194, 216, 804, 78, 80, 386, and 197), their model uncertainties are evident because no model dominates with a probability greater than 90%.

Application to Midland Field Data

We performed the same analysis for the Midland field. The Midland wells we study in this work are completed in a hydraulically fractured reservoir. Unlike the daily oil-production data from the Bakken field, the Midland field data are monthly. Thus, a Midland oil-production-rate data set is smoother than a Bakken data set. We selected 31 Midland wells with relatively less-noisy data. Each well has been operated over different time periods. **Table 6** lists the time period (production-history length) for each selected well.

Well Identification	28	24	47	9	22	11	14	48	25	15	29	32	42	27	39	45
Length (months)	62	68	71	72	75	76	76	76	77	78	78	78	78	80	81	82
Well Identification	43	49	26	13	31	33	34	40	41	44	17	18	19	30	16	
Length (months)	83	83	84	85	86	86	87	89	93	94	97	97	97	97	104	

Table 6—Production-history length of selected Midland wells.

Fig. 14 shows the marginal posterior probabilities, calculated using MM-P, for the Arps model, SEM, LGM, and Pan CRM. Among the 31 selected Midland wells, the Pan CRM is the most probable for 15 wells, the Arps model for five wells, the SEM for six wells, and the LGM for five wells. This indicates that, in general, the Pan CRM has a higher chance to be a good model for describing the Midland field data than the other three models. Nevertheless, no one model is the most probable for all wells. For some wells, the Pan CRM can be the least probable. Moreover, even if the Pan CRM is the most-probable model, other models might have probabilities close to that of the Pan CRM; for example, the marginal posterior probabilities of Well 14 are 34.7% for the Pan CRM, 34.0% for the LGM, and 24.9% for the SEM. Compared with the Bakken wells, the model uncertainty is relatively high for the Midland wells because a small number of the wells have a single model dominating with a probability greater than 90%. This is because the Bakken wells' daily production data contain more data points than the Midland wells' monthly production data. As the number of data points increases, the model uncertainty decreases.

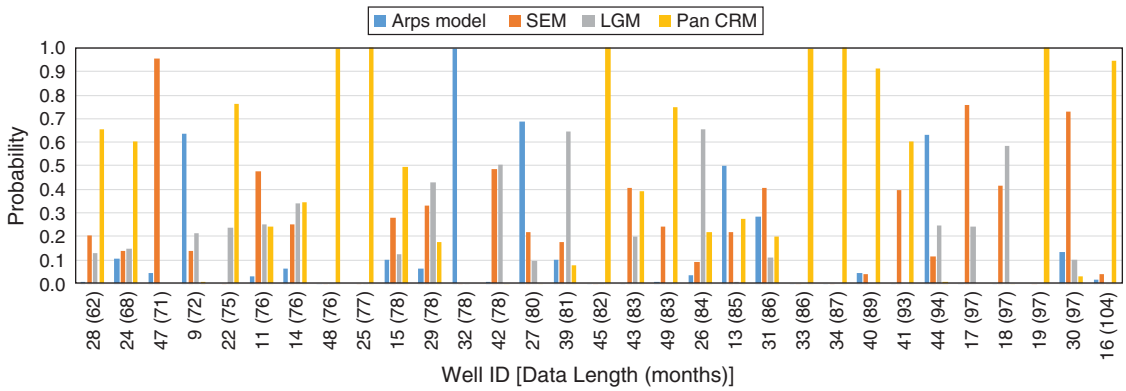

Fig. 14—Marginal posterior probabilities of the decline-curve models for the 31 selected Midland wells.

Fig. 15 illustrates the mean values of estimated cumulative oil production from the last data point to the end of the life cycle (Year 30) of a well using SM-P with only the Arps model, SEM, LGM, or Pan CRM, and MM-P combining the four models. When we evaluate these models individually, the LGM tends to give the highest estimate, whereas the Pan CRM tends to give the lowest estimate for most of the wells. Although the Arps model does not tend to give an estimate as high as the LGM, it tends to give an estimate higher than the SEM and Pan CRM. Using MM-P to weight each individual model produces a moderate estimate.

Fig. 16 illustrates the box plots of normalized estimated cumulative oil production using MM-P for the 31 selected Midland wells. The *y*-axis is scaled as in Fig. 13. For most of the wells, the 80% confidence interval is between –0.4 and 0.4; that is, the P10 or P90 deviates less than 40% from the mean. This can be used as a reference to distinguish between moderately uncertain (both P10 and P90 deviate less than 40% from the mean) and very uncertain (both P10 and P90 deviate more than 40% from the mean) production. The cumulative oil production of Well 13 is very uncertain because the data are quite noisy. Its distribution is highly skewed because after fitting the data, most realizations forecast that the oil rate will gradually decrease to 0 B/D before Year 30 (low cumulative oil production), whereas some realizations forecast that the oil production will continue to Year 30 (high cumulative oil production). This results in a long tail toward high values.

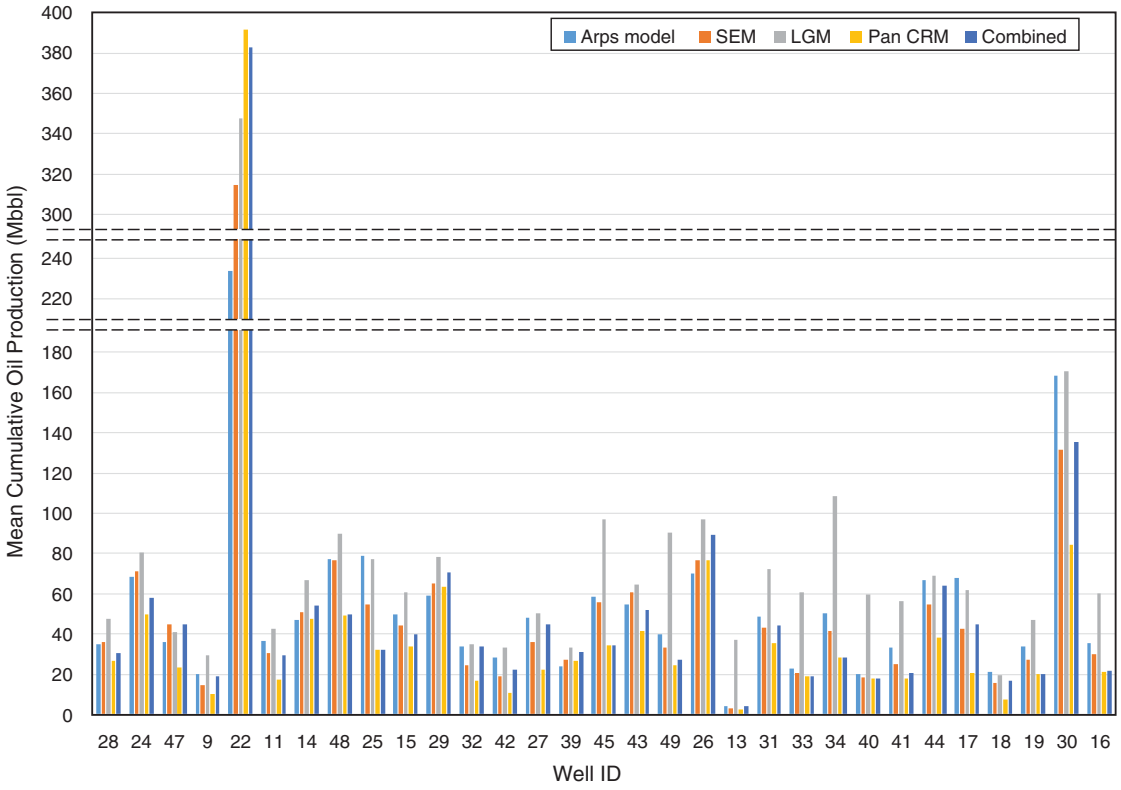

Fig. 15—Mean values of estimated cumulative oil production for the 31 selected Midland wells.

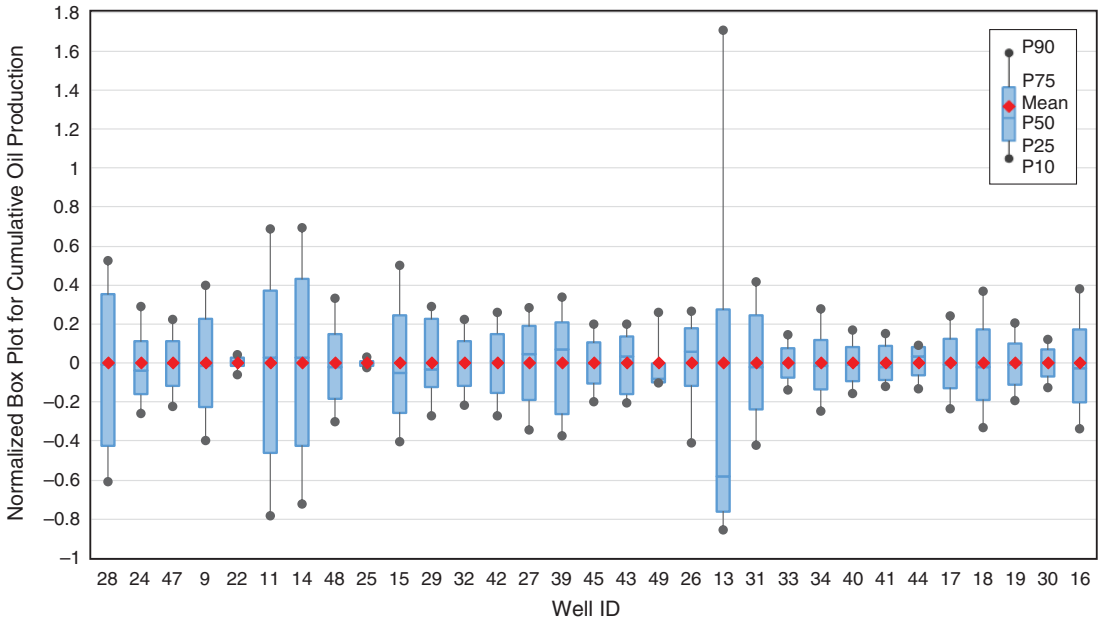

Fig. 16—Box plots of estimated cumulative oil production using MM-P for the 31 selected Midland wells.

Discussion

The ambiguous choice of a single model for DCA can be caused by cognitive biases[9,10]. Overconfidence is one of the most-common cognitive biases in the oil and gas industry (Welsh et al. 2005). When a single model is selected for DCA, one gives a 100% weight to this selected model (i.e., trust this model 100%). However, as stated by Box (1979), "all models are wrong but some are useful." Thus, selecting a single model indicates overconfidence (i.e., only the selected model is trusted rather than the others). When diverse models are considered, the effect of overconfidence is reduced, and the estimated production range will be closer to the actual production. The

[9]"Ambiguous" describes that a choice is made without sound reasoning and systematic analysis. An example of ambiguous choice is if one uses the Arps model for unconventional wells because it has been extensively used by others.

[10]A cognitive bias is the unconscious deviation from rationality in judgment (Haselton et al. 2005).

MM-P approach proposed provides a consistent and systematic framework for including multiple models in probability DCA and updating the model probabilities/weights using the measured data. Therefore, the risk of significant production over/underestimates is reduced.

Traditional measures of the goodness of fit, including (weighted) sum of squares and (weighted) R^2, can be used to rank different models, but they do not provide any useful interpretation for probabilistic analysis. Converting these measures into corresponding probability representations is more useful for probabilistic analysis. Moreover, probability has a more-intuitive interpretation to many people than sum of squares and R^2.

In our analysis, we used a noninformative prior distribution because we want to have a fair start for different models without a priori weighting one model more than the others. The noninformative prior distribution will be updated, as data become available, to the posterior distribution where the probability of a model is increased or reduced accordingly. In this case, the probability of a model is dependent only on how well it matches the data. One can definitely use an informative prior. For example, the results have provided the information that the Pan CRM has a greater probability than the Arps model for most of the wells, so we can include this information in new analyses by assigning a greater prior probability to the Pan CRM than the Arps model.

The Pan CRM is the most-probable model for most of the unconventional wells selected for this study. It tends to give the lowest estimate on cumulative oil production over long time periods among the four models studied in this work. This implies that it is very likely that the actual cumulative oil production will be lower than what is forecasted by the other models (including the widely used Arps model). This might be one contributor to the fact that most future-oil-production estimates are too optimistic (Nandurdikar and Wallace 2011).

However, no one model is the best in all circumstances; for some wells, the Arps model, SEM, or LGM might perform better than the others. Analyzing multiple models addresses the different dynamics observed in each well. Whether one model will be more probable than others depends on the flow behaviors of each well. The MM-P approach avoids the a priori choice of which model to use and, as a result of this choice, the rejection of a possible good model. Thus, we are open minded to other models and weight them depending on their data-fitting qualities. Any model holds the possibility of being the best when the measured data support it. In this sense, we assess the model uncertainty.

When the number of data points is small, the model uncertainty remains large because the data do not provide sufficient information to uniquely reflect the characteristics of different flow regimes of unconventional production. With increasing data volume, the characteristics of different flow regimes emerge, and any model that cannot well-capture the physics of different flow regimes will be gradually eliminated (i.e., its probability approaches zero).

The Pan CRM is the only physics-based model we considered in this work. It is designed to capture the major flow regime (transient- and semisteady-state-flow regimes) relevant for an unconventional well. Therefore, it is not surprising that the data from two unconventional fields show that the Pan CRM is the most-probable model to describe the flow behaviors, compared with the other three models, for most of the selected wells. For some wells, one of the other models evaluated is the most probable. The reasons for this might be that the measurement errors mask the characteristics of different flow regimes, different flow regimes show up simultaneously, minor flow regimes appear, fracture flow is not dominating, or unclear flow behaviors appear. In addition, all the parameters of the Pan CRM are physically defined, as opposed to the parameters of an empirical model, which have no physical meaning. Thus, the fitted-parameter values of the Pan CRM reveal important reservoir/well properties. For example, V_p in the Pan CRM reveals the effective drainage volume of a producer.

However, the number of fitting parameters is also important for the goodness of fit. The Pan CRM has four parameters, whereas the other three models have three parameters. Thus, the model comparison is dependent on a slightly different basis. It is hard to separate a better fit resulting from better formulating the flow behaviors from a better fit because of more fitting parameters. To take both goodness of fit and the number of fitting parameters into consideration, the Akaike (1974) information criterion can be used. A future study will consider incorporating the Akaike (1974) information criterion in the proposed approach.

Our investigation was limited to the Arps model, SEM, LGM, and Pan CRM. More models can be easily included in the analysis. Other models (e.g., a dual-porosity reservoir-simulation model) might provide better representations of the flow characteristics in a given situation compared with the Pan CRM. However, using a more-complex model with many more parameters (e.g., a grid-based reservoir-simulation model) can make probabilistic analysis too computationally intensive, or even prohibitive, and might not create much, if any, additional value [see the discussion in Bratvold and Begg (2009) and Hong et al. (2017)]. One of the advantages of DCA is its speed. Thus, models with a small number of parameters for DCA are preferable.

A limitation of the proposed approach is that the model probability tells nothing regarding the absolute truthfulness of a model, but only its relative truthfulness to other models. Therefore, for example, if all the candidate models fit the data poorly and overestimate future production, using the proposed approach will also result in an overestimated forecast. An ideal circumstance is that some candidate models will overestimate and some will underestimate. Then, the proposed approach will help neutralize the effects of over/underestimation and provide a better estimate. Nevertheless, we believe our proposed approach is helpful to account for multiple models in probabilistic DCA, when the selection of a model is not obvious, for improving production forecasts for unconventional reservoirs.

Conclusions

The MM-P approach proposed in this work, integrates model uncertainty in probabilistic DCA for unconventional-oil-production forecasting. As opposed to the traditional approach, where only a single best-fit model is used for further analysis, the proposed approach interprets the goodness of fit of a model using a probability representation that can be applied for uncertainty analysis and decision support. The uncertainty in the Arps model ($0 \leq b < 1$), SEM, LGM, and Pan CRM for analyzing unconventional plays, as well as the effect of model uncertainty on uncertain future-production estimates, was investigated.

An example illustrated that the best-fit model might not be the model that best represents actual flow behaviors. Using the proposed approach can reduce the risk of using a single best-fit model that does a very poor job of forecasting.

The proposed approach was applied to real oil-production data from the Bakken and Midland fields. The hindcast test showed that the model uncertainty was reduced, and the forecast was improved as more data points became available. The model probabilities were assessed, and the cumulative oil productions were estimated using the proposed approach for 28 selected Bakken wells and 31 selected Midland wells. The proposed approach performed well in propagating the model uncertainty to forecast uncertainty.

The main conclusions are the following:

1. It is confirmed that the Arps model ($0 \leq b < 1$) might not be ideal for unconventional plays.
2. No one model is the most probable for all wells, although the Pan CRM, as the only physics-based model, is more probable to well-describe the unconventional flow behaviors than the other three empirical models.
3. The model uncertainty can remain large even when the data length is long.

4. The Arps model and LGM tend to give a larger estimate of the expected cumulative oil production than the SEM and Pan CRM. Thus, using the Arps model or LGM can result in a too-optimistic estimate in cumulative oil production. By weighting the models, the proposed approach gives an estimate of the cumulative oil production that is neither too optimistic nor too pessimistic.

Nomenclature

a = constant parameter in the LGM
b = decline exponent in the Arps model
c_t = total compressibility
$\boldsymbol{d}$ = vector of data points
D_i = initial decline rate in the Arps model
f = model forecast
i and i' = index of model
j = index of sampled data set
J = productivity
k = timestep index
K = carrying capacity
L = loss-function value
m = model
M = total number of models
n = exponent parameter in the SEM
N = total number of Monte Carlo samples
P = probability
q = modeled oil-production rate
$\hat{q}$ = measured oil-production rate
Q = cumulative oil production
R^2 = coefficient of determination
t = time
T = total number of data points
V_p = drainage PV
$\boldsymbol{x}$ = vector of model parameters
β = linear transient flow parameter in the Pan CRM
J_∞ = constant productivity at boundary-dominated flow
ΔP = pressure difference
η = hyperbolic exponent in the LGM
σ = standard deviation
τ = characteristic time parameter in the SEM
Ω = a priori knowledge
Ω' = a posteriori knowledge

Acknowledgments

The authors acknowledge the Research Council of Norway and the industry partners ConocoPhillips Skandinavia A/S, Aker BP ASA, Eni Norge A/S, Maersk Oil Norway A/S, Statoil Petroleum A/S, Neptune Energy Norge A/S, Lundin Norway A/S, Halliburton A/S, Schlumberger Norge A/S, Wintershall Norge A/S, and DEA Norge A/S of The National IOR Centre of Norway for support. Author Aojie Hong would like to thank the staff and colleagues at the University of Texas at Austin for their support during his visit.

References

Aanonsen, S. I., Tveit, S., and Alerini, M. 2019. Using Bayesian Model Probability for Ranking Different Prior Scenarios in Reservoir History Matching. *SPE J.* **24** (4): 1490–1507. SPE-194505-PA. https://doi.org/10.2118/194505-PA.

Akaike, H. 1974. A New Look at the Statistical Model Identification. *IEEE Trans. Automat. Contr.* **19** (6): 716–723. https://doi.org/10.1109/TAC.1974.1100705.

Arps, J. J. 1945. Analysis of Decline Curves. *Trans. AIME* **160** (1): 228–247. SPE-945228-G. https://doi.org/10.2118/945228-G.

Box, G. E. P. 1979. Robustness in the Strategy of Scientific Model Building. In *Robustness in Statistics*, ed. R. L. Launer and G. N. Wilkinson, 201–236. New York City: Academic Press.

Bratvold, R. B. and Begg, S. H. 2008. I Would Rather be Vaguely Right Than Precisely Wrong: A New Approach to Decision Making in the Petroleum Exploration and Production Industry. *AAPG Bull.* **92** (10): 1373–1392. https://doi.org/10.1306/06040808070.

Bratvold, R. B. and Begg, S. 2009. Would You Know a Good Decision if You Saw One? *The Way Ahead* **5** (2): 21–23. SPE-0209-021-TWA. https://doi.org/10.2118/0209-021-TWA.

Bunn, D. W. 1975. A Bayesian Approach to the Linear Combination of Forecasts. *J. Oper. Res. Soc.* **26** (2): 325–329. https://doi.org/10.1057/jors.1975.67.

Cheng, Y., Wang, Y., McVay, D. et al. 2010. Practical Application of a Probabilistic Approach to Estimate Reserves Using Production Decline Data. *SPE Econ & Mgmt* **2** (1): 19–31. SPE-95974-PA. https://doi.org/10.2118/95974-PA.

Clark, A. J., Lake, L. W., and Patzek, T. W. 2011. Production Forecasting With Logistic Growth Models. Presented at the SPE Annual Technical Conference and Exhibition, Denver, 30 October–2 November. SPE-144790-MS. https://doi.org/10.2118/144790-MS.

Clemen, R. T. and Winkler, R. L. 1999. Combining Probability Distributions From Experts in Risk Analysis. *Risk Analysis* **19** (2): 187–203. https://doi.org/10.1111/j.1539-6924.1999.tb00399.x.

Dillon, R. L., John, R., and von Winterfeldt, D. 2002. Assessment of Cost Uncertainties for Large Technology Projects: A Methodology and an Application. *Interfaces* **32** (4): 52–66. https://doi.org/10.1287/inte.32.4.52.56.

Duong, A. N. 2011. Rate-Decline Analysis for Fracture-Dominated Shale Reservoirs. *SPE Res Eval & Eng* **14** (3): 377–387. SPE-137748-PA. https://doi.org/10.2118/137748-PA.

Gong, X., Gonzalez, R., McVay, D. A. et al. 2014. Bayesian Probabilistic Decline-Curve Analysis Reliably Quantifies Uncertainty in Shale-Well-Production Forecasts. *SPE J.* **19** (6): 1047–1057. SPE-147588-PA. https://doi.org/10.2118/147588-PA.

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. Presented at the SPE Eastern Regional Meeting, Lexington, Kentucky, 3–5 October. SPE-161300-MS. https://doi.org/10.2118/161300-MS.

Haselton, M. G., Nettle, D., and Murray, D. R. 2005. The Evolution of Cognitive Bias. In *The Handbook of Evolutionary Psychology*, ed. D. M. Buss, 724–746. Hoboken, New Jersey: John Wiley & Sons. https://doi.org/10.1002/9781119125563.evpsych241.

Hastie, T., Tibshirani, R., and Friedman, J. 2001. *The Elements of Statistical Learning*. New York City: Springer.

Hong, A. J., Bratvold, R. B., and Nævdal, G. 2017. Robust Production Optimization With Capacitance-Resistance Model as Precursor. *Computat. Geosci.* **21** (5–6): 1423–1442. https://doi.org/10.1007/s10596-017-9666-8.

Hubbert, M. K. 1956. Nuclear Energy and the Fossil Fuel. In *Drilling and Production Practice*, 1 January, New York. API-56-007.

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. Presented at the SPE Annual Technical Conference and Exhibition, Denver, 21–24 September. SPE-116731-MS. https://doi.org/10.2118/116731-MS.

Jochen, V. A. and Spivey, J. P. 1996. Probabilistic Reserves Estimation Using Decline Curve Analysis With the Bootstrap Method. Presented at the SPE Annual Technical Conference and Exhibition, Denver, 6–9 October. SPE-36633-MS. https://doi.org/10.2118/36633-MS.

Joshi, K. and Lee, W. J. 2013. Comparison of Various Deterministic Forecasting Techniques in Shale Gas Reservoirs. Presented at the SPE Hydraulic Fracturing Technology Conference, The Woodlands, Texas, 4–6 February. SPE-163870-MS. https://doi.org/10.2118/163870-MS.

Jouini, M. N. and Clemen, R. T. 1996. Copula Models for Aggregating Expert Opinions. *Oper. Res.* **44** (3): 444–457. https://doi.org/10.1287/opre.44.3.444.

Lee, W. J. and Sidle, R. 2010. Gas-Reserves Estimation in Resource Plays. *SPE Econ & Mgmt* **2** (2): 86–91. SPE-130102-PA. https://doi.org/10.2118/130102-PA.

MathWorks. 2017a. MATLAB Curve Fitting Toolbox User's Guide. https://se.mathworks.com/help/pdf_doc/curvefit/curvefit.pdf (accessed 8 March 2018).

MathWorks. 2017b. MATLAB Statistics and Machine Learning Toolbox User's Guide. https://se.mathworks.com/help/pdf_doc/stats/stats.pdf (accessed 8 March 2018).

Nandurdikar, N. S. and Wallace, L. 2011. Failure to Produce: An Investigation of Deficiencies in Production Attainment. Presented at the SPE Annual Technical Conference and Exhibition, Denver, 30 October–2 November. SPE-145437-MS. https://doi.org/10.2118/145437-MS.

Pan, Z. 2016. *Revised Productivity Index Equation to Improve Transient History Match for the Capacitance Resistance Model*. Master's thesis, University of Texas at Austin, Austin, Texas.

Park, H., Scheidt, C., Fenwick, D. et al. 2013. History Matching and Uncertainty Quantification of Facies Models With Multiple Geological Interpretations. *Computat. Geosci.* **17** (4): 609–621. https://doi.org/10.1007/s10596-013-9343-5.

Pearl, R. and Reed, L. J. 1920. On the Rate of Growth of the Population of the United States Since 1790 and Its Mathematical Representation. *Proc. Natl. Acad. Sci. USA* **6** (6): 275–288.

Read, C. 1920. *Logic: Deductive and Inductive*. London: Simkin.

Ruiz Maraggi, L. M. 2018. *Mechanistic and Probabilistic Rate-Time Analysis of Unconventional Reservoirs*. Master's thesis, University of Texas at Austin, Austin, Texas.

Sayarpour, M., Kabir, C. S., Sepehrnoori, K. et al. 2011. Probabilistic History Matching With the Capacitance-Resistance Model in Waterfloods: A Precursor to Numerical Modeling. *J. Pet. Sci. Eng.* **78** (1): 96–108. https://doi.org/10.1016/j.petrol.2011.05.005.

Stone, M. 1961. The Opinion Pool. *Ann. Math. Stat.* **32** (4): 1339–1342. https://doi.org/10.1214/aoms/1177704873.

Tavassoli, Z., Carter, J. N., and King, P. R. 2004. Errors in History Matching. *SPE J.* **9** (3): 352–361. SPE-86883-PA. https://doi.org/10.2118/86883-PA.

Valko, P. P. 2009. Assigning Value to Stimulation in the Barnett Shale: A Simultaneous Analysis of 7000 Plus Production Histories and Well Completion Records. Presented at the SPE Hydraulic Fracturing Technology Conference, The Woodlands, Texas, 19–21 January. SPE-119369-MS. https://doi.org/10.2118/119369-MS.

Valko, P. P. and Lee, W. J. 2010. A Better Way to Forecast Production From Unconventional Gas Wells. Presented at the SPE Annual Technical Conference and Exhibition, Florence, Italy, 19–22 September. SPE-134231-MS. https://doi.org/10.2118/134231-MS.

Verhulst, P. F. 1845. Mathematical Researches Into the Law of Population Growth Increase. *Nouveaux Mémoires de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles* **18:** 1–42.

Wattenbarger, R. A., El-Banbi, A. H., Villegas, M. E. et al. 1998. Production Analysis of Linear Flow Into Fractured Tight Gas Wells. Presented at the SPE Rocky Mountain Regional/Low-Permeability Reservoirs Symposium, Denver, 5–8 April. SPE-39931-MS. https://doi.org/10.2118/39931-MS.

Welsh, M. B., Bratvold, R. B., and Begg, S. H. 2005. Cognitive Biases in the Petroleum Industry: Impact and Remediation. Presented at the SPE Annual Technical Conference and Exhibition, Dallas, 9–12 October. SPE-96423-MS. https://doi.org/10.2118/96423-MS.

Aojie Hong holds a post-doctoral-degree position at the University of Stavanger. He was previously a PhD-degree candidate at the National IOR Centre of Norway. Hong's research interests include production optimization, history matching, and implementation of uncertainty analysis, decision analysis, data analytics, and machine learning in reservoir management. He holds a PhD degree from the University of Stavanger and a master's degree from the Norwegian University of Science and Technology (NTNU), both in petroleum engineering. Hong is a member of SPE.

Reidar B. Bratvold is a professor of decision and data analytics at the University of Stavanger and NTNU. Before academia, he spent 15 years in the industry in various technical and management roles. Bratvold's research interests include decision analysis, data analytics and assimilation, project valuation, portfolio analysis, real-option valuation, and behavioral challenges in decision making. He is a fellow of the Society of Decision Professionals and of the Norwegian Academy of Technological Sciences. Bratvold holds a PhD degree in petroleum engineering and a master's degree in mathematics, both from Stanford University; holds a master's degree in petroleum engineering from the University of Tulsa; and has business and management science education from INSEAD and Stanford University. He is the first author of the SPE book *Making Good Decisions*. Bratvold has three times served as an SPE Distinguished Lecturer. He was the recipient of the 2017 SPE Management and Information Award and was the executive editor for *SPE Economics & Management* from 2015 through 2017.

Larry W. Lake is a professor in the Hildebrand Department of Petroleum and Geosystems Engineering at the University of Texas at Austin, where he holds the Shahid and Sharon Ullah Chair. His research interests are in reservoir engineering and enhanced oil recovery. Lake has authored or coauthored more than 100 technical articles and four textbooks. He is a member of the US National Academy of Engineers. Lake holds a bachelor's degree from Arizona State University and a PhD degree from Rice University.

Leopoldo M. Ruiz Maraggi is a PhD-degree student in petroleum engineering at the University of Texas at Austin. Previously, he worked for 5 years as a process-design/safety engineer in the oil and gas sector in Argentina. Ruiz Maraggi's current research interests include rate/time and rate-transient analysis of unconventional reservoirs. He holds a bachelor's degree in chemical engineering from Universidad de Buenos Aires and a master's degree in petroleum engineering from the University of Texas at Austin.