Integrating New Data in Reservoir Forecasting Without Building New Models

Sebastien Strebelle, Sarah Vitel, and Michael J. Pyrcz

Abstract In most hydrocarbon reservoir development projects, geological models are fully rebuilt on a regular basis to integrate new data, in particular observations from new wells, for up-to-date forecasts. Not only this common practice is very time consuming as rebuilding models can take weeks or even months, but it also leads to major, hard-to-justify, fluctuations in reservoir volume or flow performance forecasts, especially when the modeling staff changes or a new modeling technology, workflow, or software is adopted. Rationalizing the geological model updating process is required to provide stable and reliable forecasting and make timely, well-informed, reservoir management decisions. This paper presents an innovative methodology to quickly reassess model forecasts, such as reservoir oil-in-place or oil recovery, without rebuilding any geological models provided that the new data observations are reasonably consistent with the current models. The proposed methodology uses a Bayesian framework whereby the multivariate probability joint distribution of new data predictions and forecast variables needs to be modeled. Assuming that this joint distribution is multi-Gaussian, the first step consists in computing proxies, e.g., response surfaces using experimental design, to estimate from the set of current geological models the distribution (mean and variance) of new data predictions and forecast variables as a function of the input modeling parameters (e.g., property variograms or training images, trends, histograms). Because the model stochasticity (i.e., spatial uncertainty away from wells) typically entails significant uncertainty in the prediction of new local data observations, computing the previous proxies requires generating multiple stochastic realizations for each combination of input modeling parameters. Then, using those proxies and Monte Carlo simulation, the full multivariate probability joint distribution of new data predictions and forecast variables is estimated. Plugging the actual new data values into that joint distribution finally provides new updated

S. Strebelle (✉) • M.J. Pyrcz
Strategic Research Unit, Chevron Energy Technology Company, 1500 Louisiana St., Houston, USA
e-mail: stsb@chevron.com; mpyrcz@chevron.com

S. Vitel
Earth Sciences Department, Chevron Energy Technology Company, 1500 Louisiana St., Houston, TX, USA
e-mail: sarah.vitel@chevron.com

probabilistic distributions of the forecast variables. This new methodology is illustrated on a synthetic case study. In addition to quickly reassess reservoir volume and flow performance predictions, this new approach can be used to select new data observation types and impact maps to assess potential well locations that would optimally reduce forecasting uncertainties.

1   Introduction

Most reservoir modeling projects involve updating current models with new information (e.g., logs from new wells, new seismic processing, or early production data). Models are typically rebuilt from scratch, with very little quantitative effort to check the consistency of the new data with the current models and to estimate the impact of those same new data on the project forecasts (e.g., oil-in-place or ultimate recovery). Yet, if the new data are consistent with the current models, i.e., if the new data values could have been predicted by some of the current models, there may be no need to rebuild models; a relationship between new data measurements predicted from current models and corresponding forecasts could be developed and used to update current forecasts with the actual new data measurements. This would not only save a considerable amount of time, allowing rapid reservoir management decisions in response to new information, but it would also reduce the risk of irrational fluctuations of the model forecast uncertainty range due to successive subjective reinterpretations of the data, arbitrary changes in modeling decisions, and/or introduction of new modeling technologies.

The proposed approach identified as "direct forecast updating" is quite similar to "direct forecasting," a new reservoir modeling methodology that aims at making reservoir forecasts by integrating data without performing any complex conditioning or inversion (Scheidt et al. 2015a; Satija and Caers 2015); the main difference is that prior models, which are fully unconstrained in direct forecasting, are replaced with reservoir models constrained by previously collected data. One particular focus of this paper is the direct updating of reservoir forecasts using new well data or, to be more specific, using statistical measures computed from those new well data, for example, well net-to-gross or well hydrocarbon pore column. When building models to make global forecasts, modelers only assess and model global uncertainties, such as reservoir facies proportions or porosity and permeability distributions. However, to be able to build a relationship between new well data predictions and global forecasts, local variability at the new well locations also needs to be captured in the current models. That local variability is derived from both geostatistical simulation stochasticity (seed number) and local input modeling parameter uncertainties, for example, local petrophysical property trends. In this paper a new methodology is proposed to account for such local variability when updating current forecasts directly with new well data.

2 Methodology

To explain the methodology proposed in this paper, the following simple case study is considered: the original oil-in-place (OOIP) forecasts of a reservoir need to be updated after a new well was drilled and an average net-to-gross value NTG = NTG_m was estimated from the logs at the new well location. Statistically speaking, we want to compute $P\{\text{OOIP}|\text{NTG} = \text{NTG}_m\}$, which can be rewritten using Bayes' formulation as:

$$P\{\text{OOIP}|\text{NTG} = \text{NTG}_m\} = \frac{P\{\text{OOIP and NTG} = \text{NTG}_m\}}{P\{\text{NTG} = \text{NTG}_m\}} \qquad (1)$$

Let $\boldsymbol{\theta}$ be the set of input modeling parameters representing the major global geological uncertainties identified in the reservoir, for example, the global reservoir rock volume or the reservoir porosity distribution.

The numerator and denominator of Eq. 1 can be rewritten using integrals over the whole input modeling parameter uncertainty space:

$$P\{\text{OOIP}|\text{NTG} = \text{NTG}_m\} = \frac{\displaystyle\int P\{\text{OOIP and NTG} = \text{NTG}_m \big| \boldsymbol{\theta}\}d\boldsymbol{\theta}}{\displaystyle\int P\{\text{NTG} = \text{NTG}_m \big| \boldsymbol{\theta}\}d\boldsymbol{\theta}} \qquad (2)$$

The input modeling parameter uncertainty space can be sampled by drawing n equiprobable combinations $\boldsymbol{\theta}_i(i = 1 \ldots n)$ of global input modeling parameters:

$$P\{\text{OOIP}|\text{NTG} = \text{NTG}_m\} = \frac{\frac{1}{n}\sum P\{\text{OOIP and NTG} = \text{NTG}_m \big| \boldsymbol{\theta}_i\}}{\frac{1}{n}\sum P\{\text{NTG} = \text{NTG}_m \big| \boldsymbol{\theta}_i\}} \qquad (3)$$

For each combination $\boldsymbol{\theta}_i$ of global input modeling parameters, multiple realizations can be generated to capture local uncertainties. Very often, generating multiple stochastic realizations by changing the random seed numbers of the geostatistical simulations is sufficient to capture local variability. However, in more complex cases, additional local variability such as local property trends may need to be accounted for.

One solution to compute $P\{\text{OOIP and NTG} = \text{NTG}_m| \boldsymbol{\theta}_i\}$ for any combination $\boldsymbol{\theta}_i$ of input parameters is to use a multi-Gaussian model, in which case only the means and standard deviations of the NTG and OOIP, as well as the correlation coefficient between NTG and OOIP, need to be estimated as a function of $\boldsymbol{\theta}_i$. The multi-Gaussian assumption can be tested using various methods (Mecklin and Mundfrom 2005) by generating a large number of realizations for some representative combinations $\boldsymbol{\theta}_i$ of input parameters. Provided that the multi-Gaussian assumption is not rejected, the NTG and OOIP means and standard deviations, as

well as the correlation coefficient between NTG and OOIP, can be modeled using a design of experiments built for the global input modeling parameters space $\boldsymbol{\theta}$:

1. For each of the N experimental design runs, which correspond to a specific combination $\boldsymbol{\theta_i}(i = 1 \ldots N)$ of global input modeling parameters, generate L stochastic realizations.
2. For each of the L realizations, compute the OOIP and NTG value at the new well location.
3. Calculate the means and standard deviations of the L values of NTG and OOIP, as well as the correlation coefficient between the NTG and OOIP values.
4. Using the N experimental design runs, model a response surface for the NTG and OOIP means and standard deviations, as well as the correlation coefficient between NTG and OOIP.

Using the previous response surfaces, under the multi-Gaussian assumption, $P\{\text{OOIP and NTG} = \text{NTG}_m | \boldsymbol{\theta_i}\}$ and $P\{\text{NTG} = \text{NTG}_m | \boldsymbol{\theta_i}\}$ can be computed for a very large number n of combinations $\boldsymbol{\theta_i}$ randomly drawn from Monte Carlo simulation; this provides a new updated OOIP probability distribution according to Eq. 3.

The proposed approach has several advantages:

- First, the multi-Gaussian assumption, combined with the use of response surfaces to estimate the parameters of the multi-Gaussian model for any combination $\boldsymbol{\theta_i}$ of input parameters, allows fully determining the bivariate distribution P {OOIP and NTG}; there is no need to use any arbitrary interpolation technique such as the traditional Kernel smoothing (Park et al. 2013; Scheidt et al. 2015b).
- Then, the exact NTG_m value can be directly plugged into the multi-Gaussian function $P\{\text{NTG} = \text{NTG}_m | \boldsymbol{\theta_i}\}$ for any combination $\boldsymbol{\theta_i}$ of input parameters; there is no need to determine a quite arbitrary bandwidth around the new data measurements (Scheidt et al. 2015a).
- $P\{\text{NTG} = \text{NTG}_m | \boldsymbol{\theta_i}\}$ provides the probability that the NTG value measured at the new well location will be observed for the specific combination $\boldsymbol{\theta_i}$ of global input modeling parameters. Typically, one would expect the correlation between NTG and OOIP over multiple stochastic realizations to be quite low. If this is indeed the case, i.e., if NTG and OOIP are conditionally independent, which can be tested, Eq. 3 can be rewritten as:

$$P\{\text{OOIP}|\text{NTG} = \text{NTG}_m\} = \frac{\frac{1}{n}\sum P\{\text{OOIP}|\theta_i\}P\{\text{NTG} = \text{NTG}_m|\theta_i\}}{\frac{1}{n}\sum P\{\text{NTG} = \text{NTG}_m|\theta_i\}} \quad (4)$$

In that new Eq. 4, the updated OOIP forecasts can be interpreted as the linear combination of the OOIP forecasts corresponding to each possible combination $\boldsymbol{\theta_i}$ of input parameters weighted by the probabilities that the NTG value at the new well location be observed in the stochastic model realizations generated for $\boldsymbol{\theta_i}$.

Extending the previous methodology to multiple new measurements, e.g., the NTG values from several new wells, is straightforward, provided that the multi-Gaussian assumption holds.

3   Illustrative Case Study

As an illustrative example, the following synthetic data set mimicking an actual Chevron reservoir is used with a reference model corresponding to a tidal dominated reservoir with 25 % sandbars. The synthetic example was generated using the multiple-point statistical simulation program snesim (Strebelle 2002). There is no horizontal facies proportion trend, but a significant vertical trend; see the three horizontal sections of the model displayed in Fig. 1.

The tidal sandbar porosity distribution is approximatively normal, with a 19 % mean, and the permeability distribution is lognormal with a 50 md mean. Porosity was simulated using SGS, while permeability was simulated from porosity using SGS with collocated cokriging and a 0.9 correlation coefficient. The background shale porosity and permeability are assumed to be close to 0. The reference OOIP is about 80 M bbl.

Table 1 provides the list of global input modeling parameters and their corresponding uncertainties.

The reference model is consistent with the uncertainty ranges defined for the different global input modeling parameters. A D-optimal design of experiments (Atkinson et al. 2007) was used to generate 99 models, and a quadratic response surface was computed to estimate the initial OOIP probabilistic distribution displayed in Fig. 2. The initial P50 value (93 M bbl.) significantly overestimates the OOIP from the reference model, while the uncertainty range is relatively broad with a P10 value of 61 M bbl. and a P90 value of 128 M bbl.

Three wells were used to condition all the initial models. The objective of this case study is to update the initial OOIP forecasts using the NTG values measured at two alternative new well locations. The first location was randomly selected and has a relatively low NTG of 4 %, whereas the second location is very close to an existing well and has a NTG value of 20 % similar to the NTG of that well (see Fig. 3).

The methodology presented in the previous section was applied by generating 10 stochastic realizations for each of the 99 runs of the D-optimal experimental design. For each alternative new well location, the NTG and OOIP means and standard deviations, as well as the correlation coefficient between NTG and OOIP, were computed for the 99 experimental design runs, and quadratic response surfaces were computed as a function of the input modeling parameters.

Then, 10,000 combinations θ_i of input modeling parameters were generated from Monte Carlo simulation. Figure 4 shows the histograms of predicted NTG values for each new well location. The actual NTG value observed at the first new well location corresponds to the 9th percentile of the prediction distribution, while
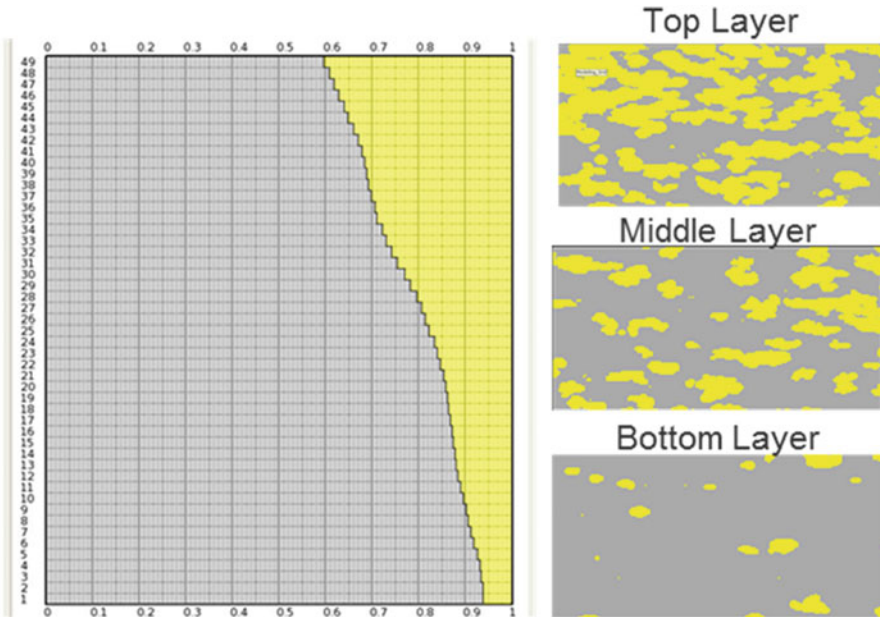
Fig. 1 Facies proportion curve and three horizontal sections of the reference model

Table 1 Global input modeling parameters, uncertainty ranges, and reference model

Input modeling parameters	P10	P50	P90	Reference
Reservoir sand proportion	20 %	30 %	40 %	25 %
Sand geobody shapes	Ellipses	Variogram-based	Channels	Ellipses
Vertical trend	None	Medium	High	Medium
Horizontal trend	None	Medium	High	None
Porosity average	15 %	17 %	19 %	19 %
Permeability average	10 md	50 md	250 md	50 md
Poro/perm correlation	0.5	0.7	0.9	0.9

the actual NTG value observed at the second new well location corresponds to the 64th percentile. Therefore, in both cases, the new NTG measurements can be considered as consistent with the existing models, and the previously described forecast updating process can be applied.

Figure 5 provides a bubble graph displaying the bivariate distribution P {OOIP and NTG} resulting from Eq. 3 for the first new well location.

Table 2 provides the P10, P50, and P90 OOIP values for the two alternative new well locations.

For the first well, the P50 value of the updated forecasts is closer to the reference OOIP value (80 M bbl.), while the uncertainty forecast range significantly decreased: the new P10-P90 difference is 46 M bbl. vs. 67 M bbl. initially. In contrast, as expected, because the second well is very close to an existing well, its
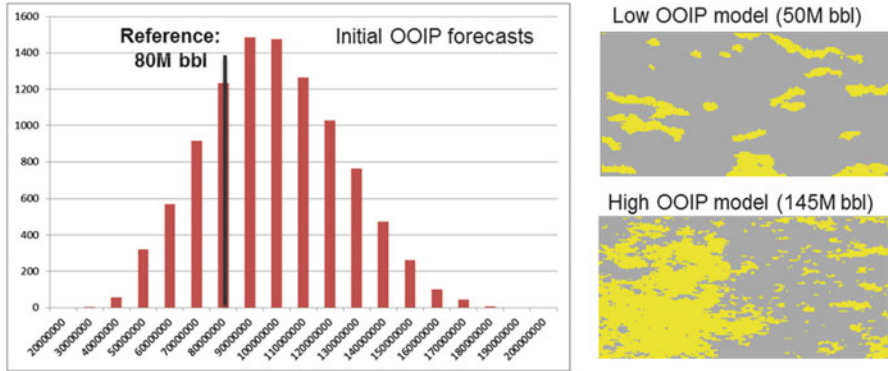
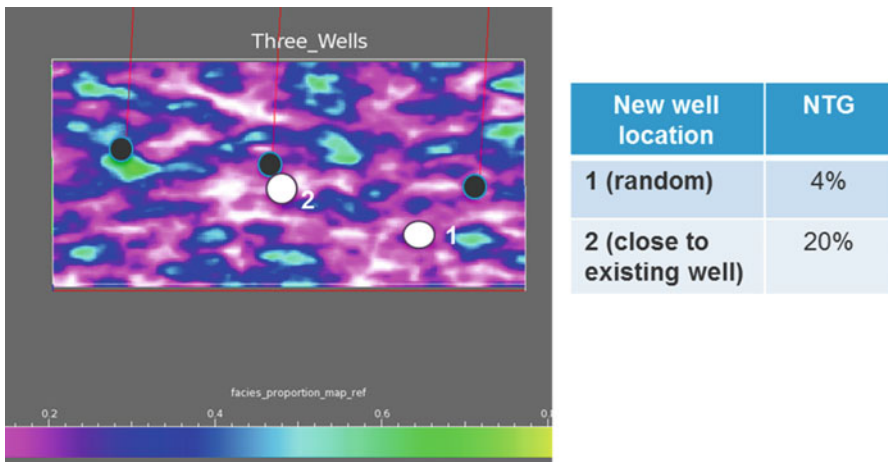**Fig. 2** Initial forecasts and horizontal sections of two models generated using the D-optimal experimental design

Fig. 3 NTG map from reference model, with locations of the three existing wells (*black dots*) and two alternative new wells (*white dots*)
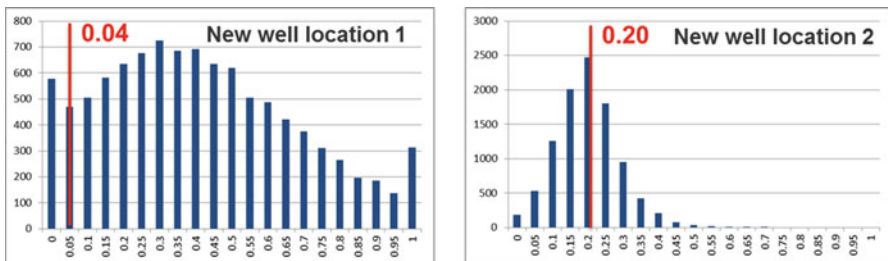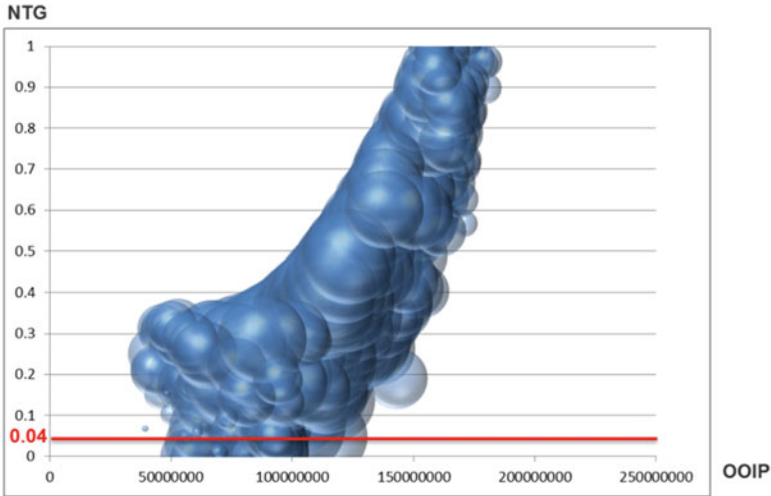

Fig. 4 Histograms of the NTG predictions for both new well locations. The *red line* corresponds to the actual observed NTG value

Fig. 5 Bivariate distribution $P\{OOIP$ and NTG$\}$ for the first new well location. Each bubble corresponds to a particular combination θ_i of global input modeling parameters; it is centered at the estimated NTG and OOIP mean values, and its size is proportional to the estimated NTG standard deviation (only 200 bubbles are displayed). The *red line* corresponds to the actual NTG value observed at that first new well location

**Table 2** Updated forecasts for the two alternative new well locations

New well location	P10	P50	P90
1 (random)	57	79	103
2 (close to existing well)	66	97	125
Initial forecasts	61	93	128

impact on OOIP forecasts is extremely limited; thus the P10, P50, and P90 values of the updated forecasts (66, 97, and 125 M bbl.) are very close to the initial forecasts (61, 93, and 128 M bbl.).

The same methodology based on Eq. 3 can be applied to the case where both wells 1 and 2 are drilled. This requires the computation of an additional response surface: the correlation between NTG values at wells 1 and 2 for any combination θ_i of global input parameters. Figure 6 provides the scatterplot of predicted NTG values at well location 1 versus predicted NTG values at well location 2 for 10,000 combinations θ_i of input modeling parameters. The actual NTG values (0.04 for the first well and 0.2 for the second well) are in the predicted ranges. Thus the combination of the two new NTG measurements can be considered as consistent with the existing models, which confirms that the previously described forecast updating process can be applied. Note that several methods exist to quantitatively check that consistency between new observed values and predictions, in particular the Mahalanobis distance (Mahalanobis 1936).

When both new well locations are used, the P10, P50, and P90 values of the updated forecasts are 65, 86, and 105 M bbl. The new P10-P90 difference is 40 M
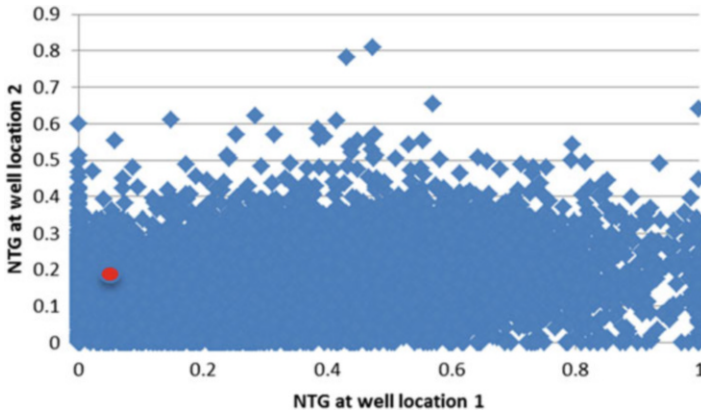
Fig. 6 Scatterplot of predicted NTG values at well location 1 versus predicted NTG values at well location 2 for 10,000 combinations θ_i of input modeling parameters. The *red dot* corresponds to the actual NTG values observed at the new well location (0.04 for well location 1 and 0.2 for well location 2)

bbl., which is, as expected but not guaranteed, smaller than the P10-P90 difference obtained for each well considered individually.

4 Discussion

The results obtained in the case study above show that reservoir forecasts can be directly updated in the presence of new information without rebuilding any models provided that the new information is consistent with the existing reservoir models. It should be noted that there is no guarantee for the P50 value of the updated forecasts to be closer to the true reservoir value or for the updated uncertainty range to systematically decrease; it all depends on the new data measured value. However, getting more accurate and precise forecasts is expected on average as the number of additional new wells increases.

In most cases, the previous methodology can be simplified by replacing some response surface with constant values or straightforward functions. For example, in the previous case study, it can be observed that, as expected, OOIP varies very little across multiple stochastic realizations for any particular combination θ_i of global input parameters. On average over the 99 experimental design runs, the coefficient of variation (ratio between standard deviation and mean) is only 0.004. This means that only the response surface for the OOIP mean need be modeled; the OOIP standard deviation could be directly estimated by multiplying the OOIP mean by 0.004. This simplification provides updated P10, P50, and P90 OOIP values very close (less than 0.5 % relative difference) to the updated forecasts obtained using a full response surface for the OOIP standard deviation. Ignoring completely the

OOIP standard deviation, i.e., setting it to a constant 0, still provides updated forecasts with less than 1 % relative difference compared to the initial methodology. In contrast, the coefficient of variation for the NTG is 0.803 for the first new well location and 0.243 for the second well location (close to an existing well) on average over the same 99 experimental design runs, which demonstrates the importance of capturing all the local variability at new well locations, a quite challenging exercise. Finally, the correlation coefficient between NTG and OOIP is systematically low for the 99 experimental design runs, 0.03 on average. Thus Eq. 4, which assumes the conditional independence of NTG and OOIP for any combination θ_i of global input modeling parameters, could have been used. This simplification would have again yielded very similar updated P10, P50, and P90 OOIP values (less than 0.1 % difference).

Expanding the proposed methodology to more than two new measurements and/or forecast variables is straightforward from a theoretical point of view, but will need to be tested in future work. In particular, when a large amount of new data is available, our methodology may require the identification of a limited number of physical metrics representing or summarizing the new data. For example, in the illustrative case study presented in this paper, the average NTG at the new well location was used instead of the whole facies log. If summary physical metrics are difficult to identify or compute, brute-force dimensionality reduction techniques such as nonlinear PCA (Scheidt et al. 2015a) could be applied.

Also, the proposed approach calls for the use of experimental design and the construction of response surfaces, which limits the application to continuous and ordinal input modeling parameters. However, other direct forecasting methodologies could reuse the main idea of this paper: explicitly account for local variability at new measurement locations by modeling new measurement predictions and global forecasts as the sum of an average value over multiple simulated realizations and a residual. The average value captures the impact of the global modeling uncertainty parameters, while the residual captures local variability, especially geostatistical simulation stochasticity.

5   Conclusions

A simple methodology using a Bayesian framework with a traditional multi-Gaussian assumption is presented in this paper to update forecasts after acquiring new well log data, which allows making rapid reservoir management decisions in response to new information. One main advantage of this methodology is that it fully accounts for local uncertainties, in particular model stochasticity, when estimating the impact of new local data on reservoir forecasts.

The proposed methodology was successfully applied to a simple synthetic case study, but it needs to be further tested on more complex synthetic data sets and actual reservoir modeling projects. Another next step is to use that methodology to select what new data should be collected and where it should be collected to

optimally decrease reservoir forecasts uncertainty similar to the impact map approach of Zagayevskiy and Deutsch (2013).

Bibliography

Atkinson A, Donev A, Tobias R (2007) Optimum experimental designs, with SAS. UOP Oxford, Oxford

Mahalanobis P (1936) On the generalised distance in statistics. Proc Natl Inst Sci India 2(1):49–55

Mecklin C, Mundfrom D (2005) A monte carlo comparison of the type I and type II error rates of tests of multivariate normality. J Stat Comput Simul 75(2):93–107

Park H, Scheidt C, Fenwick DH et al (2013) History matching and uncertainty quantification of facies models with multiple geological interpretations. Comput Geosci 17:609–621

Satija A, Caers J (2015) Direct forecasting of subsurface flow response from non-linear dynamic data by linear least-squares in canonical functional principal component space. Advances in Water Resources 77:69–81

Scheidt C, Renard P, Caers J (2015a) Prediction-focused subsurface modeling: investigating the need for accuracy in flow-based inverse modeling. Math Geosci 47(2):173–191

Scheidt C, Tahmasebi P, Pontiggia M et al (2015b) Updating joint uncertainty in trend and depositional scenario for reservoir exploration and early appraisal. Comput Geosci 19:805–820

Strebelle S (2002) Conditional simulation of complex geological structures using multiple-point statistics. Math Geol 34(1):1–21

Zagayevskiy Y, Deutsch CV (2013) Impact map for assessment of new delineation well locations. JCPT 2013:441–462