Evaluation of the Discrete Latin Hypercube with Geostatistical Realizations Sampling for History Matching under Uncertainties for the Norne Benchmark Case

Carla J. Ferreira, Alessandra Davolio, Denis J. Schiozer, UNICAMP

Abstract

This work describes a methodology that evaluates the Discrete Latin Hypercube with Geostatistical Realizations (DLHG) sample size for complex models in the history matching under uncertainties process with application to the Norne Benchmark Case. The sample size affects the time demanded and results accuracy in a history matching process because a small sample size can yield inaccurate risk quantification and a high sample size can demand excessive time to reach good results. Both factors should be evaluated in order to improve the project’s efficiency and to obtain reliable results. Such evaluation gains greater importance in complex reservoir models because the number of tests to determine the reservoir scenarios that match dynamic data can be high due to the level of complexity. The methodology presented in this work is divided in three steps. First, we evaluate the ability of DLHG to produce output cumulative distribution functions (CDF) that replicate a more exhaustive sampling technique (Monte Carlo) using the Kolmogorov–Smirnov test. The output is the misfit between observed and simulated production rates; then, we compare the influence and correlation matrices obtained with DLHG and Monte Carlo samples. The influence matrix shows the impact of the uncertainty variation on the outputs and the correlation matrix measures the strength of the dependence between the uncertainty attributes and outputs. Finally, we perform the stability test. The methodology was applied to the Norne benchmark case; a field located in the Norwegian Sea. The main characteristics of the methodology are: (1) it uses a statistical technique to compare the output CDFs from the reference and DLHG samples and (2) it evaluates the ability of the DLHG sample to identify the reservoir attributes that affect the history match results. We evaluated DLHG sample sizes of 20, 50, 100 and 200, and considered a MC sample size of 5,000 to the Norne benchmark case. The DLHG CDFs for the 100 sample size was able to accurately replicate the corresponding MC CDFs, however it did not replicated the behavior of the influence and correlation matrices. The DLHG sample size of 200 was able to reproduce the CDFs outputs, the influence and correlation matrices and it was considered stable. The study showed that even if the sample size is able to represent the CDFs outputs from a reference solution, the influence and correlation matrices should be evaluated. The methodology presented can be incorporated into usual history match routines.
Introduction

The history matching under uncertainties process is an important step of the Closed Loop Reservoir Management and Development method described by Schiozer et. al (2015). The methodology is a tool for decision analysis related to the development and management of petroleum fields; it is based on 12 steps. The third, fourth and fifth steps perform the model calibration, scenarios generation and reduction of scenarios with dynamic data.

The description and characterization of uncertainty introduced into a reservoir model is a key component of the model calibration and reduction of scenarios with dynamic data steps. The most common methods to combine the reservoir uncertainties to generate reservoir scenarios are the Derivative Tree technique (DT), Monte Carlo (MC), and Latin Hypercube (LH).

The definition of the number of reservoir scenarios is crucial because the incorrect sample size yields to inaccurate risk quantification or demanding excessive time to reach good results. Moreover, the sample size affects the time demanded and results accuracy in a history matching under uncertainty process. Both factors should be evaluated in order to improve the project’s efficiency and obtain reliable results. Such evaluation gains greater importance in complex reservoir models because the number of tests to determine the reservoir scenarios that match dynamic data can be high due to the level of complexity.

Sampling techniques are used in different stages within the history matching process. They are mainly used (1) to evaluate the influence of attributes in the objective function and (2) to implement a probabilistic process of reducing uncertainties. The methodology described in this work refers to the first use, in which the sample size is determined according to the impact on the observed and simulated mismatch.

Schiozer et. al (2016) discussed the sampling techniques limitations and proposed a new method for risk quantification called Discrete Latin Hypercube with Geostatistical Realizations (DLHG). The authors comment that the DT technique is simple to use but can produce many combinations when considering several uncertainties, whereas the MC technique requires many simulations, therefore is usually combined with a proxy to model the reservoir, making it more complex. The LH method would be the most efficient because it generally requires fewer simulations, thus it could potentially be used directly with simulation runs.

The DLHG method, proposed by Schiozer et. al (2016), presents the efficiency of the LH sampling technique and the geological consistency of the geostatistical methods, which generate realizations for the reservoir properties related to the sedimentary model. The method integrates all types of uncertainty, such as rock, fluid and operational uncertainties. It discretizes attributes with continuous variation and combine with discrete attributes and realizations that are generated by the geostatistical algorithm (Schiozer et. al, 2016).

Tests using the DLHG method in synthetic complex models showed that a number between 100 and 300 properly quantifies the reservoir risk (Schiozer et. al, 2016). However, the number of simulations should be evaluated in real reservoir models considering the simulation run time, the importance of the study and available work time.

Thus, this work describes a methodology that evaluates the Discrete Latin Hypercube with Geostatistical Realizations (DLHG) sample size for complex models in the history matching under uncertainties process. Moreover, it presents the results of the methodology application to the Norne Benchmark Case.

Objective and Premises

The objective of this paper is to evaluate the Discrete Latin Hypercube with Geostatistical Realizations (DLHG) sample size to the Norne Benchmark Case. It evaluates the ability of the DLHG technique to replicate the results of a more exhaustive random sampling technique (Monte Carlo) and its
stability. This work also verifies if different DLHG runs of the same sample size are able to produce CDFs that accurately replicate corresponding MC CDFs.

The following premises are considered:

- We evaluate the quality of the sample size by comparing the output CDFs using the two sample Kolmogorov–Smirnov test, the influence matrix and the correlation matrix;
- The output is the misfit between the observed and simulated production rates represented by the index Normalized Quadratic Distance with Sign (NQDS);
- We considered production data up to 2001, because this work was part of broad study that applied a history-matching under uncertainty methodology to the Norne Benchmark case; and the remaining part of the production history (2001-2006) was used to estimate the quality of the models compared with history data.

**Theoretical Background**

In this section, a basic explanation of the main concepts used to formulate the proposed methodology is presented. For a deeper comprehension, interested readers are referred to some key references on each topic.

**History Matching**

The history matching concepts are taken from Avansi et. al, (2016), that describes a probabilistic and multi-objective history matching methodology. The methodology uses the DLHG technique to combine the critical uncertainties of the reservoir, evaluates different objective functions (well rates and pressure), reviews the reservoir characterization and filters the models that match historical data.

The definitions necessary for better comprehension of the results are the following:

- **Attribute**: is the reservoir property such as porosity, absolute permeability, relative permeability and oil-water contact. An attribute may be characterized by an image generated by geostatistical, as well as by a multiplier, value, or a table;
- **Normalized Quadratic Distance with Sign (NQDS)**: index that measures the misfit between the simulated and observed production rates. The NQDS acceptable range is defined by the decision maker and depends on the complexity of the reservoir model under analysis;

**Influence Matrix**

The influence matrix is used within the history matching process to identify the impact of the reservoir uncertainties on the NQDS results. The index

\[ I_{prod}^l = \frac{\sum_{i=1}^{n} |NQDS_{prod i}^l|}{n^l} \]  

measures the mean of the NQDS for a specific production data (prod), such as oil rate, and attribute level (l). N is the total number of reservoir scenarios with the same attribute level.

The \( I_{prod} \) variation for different attribute levels shows which level improves the quality of the history match. Figure 1 shows a schematic influence matrix; the index shows that the reservoir attribute \( A_1 \) affects the NQDS results and level 2 increases the NQDS value.

![Figure 1 - Schematic influence matrix.](figure1.png)
Correlation Matrix

Maschio and Schiozer (2016) described the use of a correlation matrix in a probabilistic history matching methodology. The covariance of two random variables $x$ and $y$ is

$$cov(x, y) = \frac{1}{n-1} \sum_{k=1}^{n} (x_k - \bar{x})(y_k - \bar{y})$$

(2)

where $\bar{x}$ and $\bar{y}$ are the mean of $x$ and $y$, respectively.

Supposing three variables $x$, $y$ and $z$, the covariance matrix ($C$) is

$$C = \begin{bmatrix}
cov(x, x) & cov(x, y) & cov(x, z) \\
cov(y, x) & cov(y, y) & cov(y, z) \\
cov(z, x) & cov(z, y) & cov(z, z)
\end{bmatrix}$$

(3)

where $cov(x,x)$ is the variance of $x$ and $cov(x,y) = cov(y,x)$ is the covariance between $x$ and $y$, computed according to Eq. 2, similarly for the other pairs $(x,z)$ and $(y,z)$.

The correlation coefficient (cc) is a normalization of the covariance. Thus, the correlation matrix is

$$R(i, j) = \frac{C(i, j)}{\sqrt{C(i, i) \times C(j, j)}}$$

(4)

where $i$ and $j$ denote the row and column position, respectively, of each entry in the matrix $C$.

Sampling Techniques

The Monte Carlo (MC) technique is a traditional method of sampling random variables in simulation modelling (Rubinstein, 1981). It is useful to simulate systems with many coupled degrees of freedom, including modeling phenomena with significant uncertainty in inputs, such as calculating risk analysis (Hughes, 1995; Kwak and Ingall, 2007).

Amongst the strengths of MC are that it is easy to program and apply; amenable to analytical and numerical models; and produces unbiased estimates of the mean and variance of the output variables (Saltelli et al., 2006)

A major disadvantage of MC is that it is computationally intensive, requires much simulation runs to represent problems and for long running models the total simulation time may be in itself prohibitive. MC analysis may be computationally prohibitive because it is important to ensure that the MC simulations preserve the input probability distributions by exhaustively sampling from various points of the input distributions (Lye, 2008)

Conduct a large number of simulations is not practical in the history matching process of real complex reservoir models, because the time spent to determine the reservoir scenarios that match dynamic data would make the history match unfeasible.

The DLHG integrates the Latin Hypercube (LH) sampling with geostatistical realizations of spatial uncertainties. The LH is better than MC in that it is more efficient because it ensures that the ensemble of random numbers represents the real variability but with fewer points (Risso et al., 2011).

The LH has limitations dealing with discrete variables, especially the parameters represented by spatial distributions. The DLHG was designed to overcome such limitations by combining the efficiency of the LH sampling technique with the geological consistency of the geostatistical methods.

The main characteristics of the DLHG method are: efficiency combining different types of variables, simple application in real cases, considers geostatistical realizations and few samples to represent uncertainty (this number depends on the precision required for each application). Moreover, the number of simulation is not a strong function of the number of attributes and levels; therefore the method has potential to be used for complex reservoirs with many uncertain attributes (Schiozer et al., 2016).
Kolmogorov–Smirnov Test

The so-called Kolmogorov–Smirnov goodness-of-fit test, referred to as the K-S test, is based on a statistic measurement of the deviation of the observed cumulative histogram from the hypothesized cumulative distribution function (one-sample K–S test), or from a second sample (two-sample K–S test), for more detail see Soong (2004).

The Kolmogorov–Smirnov test is a hypothesis test procedure for determining if two samples of data are from the same distribution. The test is non-parametric and entirely agnostic to what this distribution actually is. The KS-test tries to determine if two datasets differ significantly. It has the advantage of making no assumption about the distribution of data.

Suppose that the first sample has size m with an observed cumulative distribution function of F(x) and that the second sample has size n with an observed cumulative distribution function of G(x). Define:

$$D_{m,n} = \max |F(x) - G(x)|$$

The null hypothesis is H0: both samples come from a population with the same distribution. As for the Kolmogorov-Smirnov test for normality, we reject the null hypothesis (at significance level α) if $D_{m,n} > D_{m,n,α}$ where $D_{m,n,α}$ is the critical value. Level of significance is the probability of rejecting the null hypothesis in a statistical test when it is true (Frank & Massey, 1951).

The critical value ($D_{critical}$) for a level of significance equal to 10% is

$$D_{m,n,0.1} = 1.22 \sqrt{\frac{n+m}{n \cdot m}}$$

where \(n\) and \(m\) are first and second sample sizes, respectively.

The two-sample test checks whether the two data samples come from the same distribution. This does not specify what that common distribution is (e.g. whether it’s normal or not normal). For example, at a 0.10 level of significance, the critical value D for \(n = 20\) and \(m = 5000\) is 0.2733; this means that in 10 per cent of samples of size 20, the maximum absolute deviation between the \(n\) sample cumulative distribution and \(m\) sample cumulative distribution will be 0.2733.

However, the $D_{critical}$ does not provide a precise indication of the correspondence between the DLHG and MC samples. Thus, we defined a second index called $D_{limit}$ to assist in the definition of the acceptable sample size. The $D_{limit}$ is the maximum acceptable distance for a specific point in the CDF curve between the DLHG and MC samples. It is defined by the user according to the acceptance level needed to the process, and in the case studied we considered equal 10%.

The results are considered acceptable if the D value obtained for a DLHG sample is lower than both critical ($D_{critical}$) and limit ($D_{limit}$) values. That is, the DLHG and MC samples come from the same distribution and there is a correspondence between the DLHG and MC samples.

Methodology

The methodology flowchart is presented in Figure 2 and described as follows:

1. Generate reservoir scenarios using the DLHG technique considering the sample size to be evaluated and generate reservoir scenarios using the MC technique;
2. Simulate the reservoir scenarios from both techniques;
3. Calculate the misfit between history and simulated data (NQDS) for the production data available, such as oil ($Q_o$), water ($Q_w$) and gas ($Q_g$) rates, and bottom-hole pressure (BHP) for all production wells;
4. Calculate the CDF for each NQDS ($Q_o$, $Q_w$, $Q_g$ and BHP) and compare all CDFs from DLHG and MC samples using the Kolmogorov-Smirnov test;
5. Generate the influence and correlation matrices;
6. Compare the influence and correlation matrices derived from the DHLG and MC samples;
7. The DLHG sample size is considered acceptable if: (1) the D value is lower than $D_{\text{limit}}$ and $D_{\text{critical}}$ for all outputs and (2) the uncertain attributes that affects the NQDS results, identified through the influence and correlation matrices, are the same from those using the MC sample;
8. The sample size is discarded if the results are not acceptable;
9. Generation of new DLHG runs using the same sample size under analysis and verify if the new DLHG runs are also considered acceptable;
10. Verification if all the results are considered acceptable for all runs using the same sample size;
11. Acceptance of the sample size if the results are stable, and use in the history-matching process as a consequence.

Application

The analysis of DLHG sample size for history matching purposes was conducted to the Norne benchmark case. The Norne field is an oil and gas field located in the Norwegian Sea in 380 water depth. It was discovered in December 1991 and production started in November 1997 (Verlo and Hetland, 2008 and Ferreira et al., 2017).

The field consists of five formations: Garn, Not, Ile, Tofte and Tilje. The Tilje, Tofte, Ile and Garn formations are dominated by fine-grained arenites, whereas the Not formation behaves as a cap rock, that leads to little or no communication between Garn formation and the rest of the reservoir. Gas is mainly found in the Garn formation and oil in the Ile and Tofte formation. The Tilje formation is mostly saturated with water (Verlo and Hetland, 2008 and Ferreira et al., 2017).

Norne field consists of two separate oil compartments; Norne Main Structure (Norne C, D and E segments, discovered in 1991), which contains 97% of the oil in place, and the North-East Segment (Norne G-segment). Figure 3 shows the Norne field segments.
An Eclipse reservoir simulation model was released as part of the benchmark study by the Center for Integrated Operations in the Petroleum Industry (IO Center) at NTNU. The dimensions of the model are 46 x 112 x 22, with 44431 active blocks. It has five equilibrium regions with different depths of initial fluid contacts.

The current study used the simulation model converted from Eclipse to CMG-Imex (Mazo, 2014) and modified by Mazo (2015). Ferreira et. al (2017) presents the description of the simulation model. The main modifications carried out were:

- Fluid treatment: Black-Oil modeling;
- Well control based on production history data: Control based on oil rate;
- Model initialization: Equilibrium calculation;
- Petrophysical properties of the model: Stochastic sets.
- Transmissibility of vertical barriers: changes performed as a result of 4D seismic data was discarded.

The reservoir comprises fourteen production wells, four water/gas injection wells and three water injection wells. The outputs evaluated were the misfit between production and simulated data (oil, gas and water rates and BHP) from production wells, totaling fifty-six outputs. The misfit is calculated using the NQDS index and history data until from 1997 to 2001 was used in this study.

Results
The results are divided in four topics. First, we describe the reservoir characterization and attributes that were considered uncertain; then we present the K-S results and compare the influence and correlation matrices. Finally, we present the stability test results.

Characterization under Uncertainties
The uncertain attributes of the reservoir model were divided into two groups: discrete attributes and geostatistical attributes.

The discrete uncertain attributes were:
- Rock compressibility (cp);
- Water-oil contact (WOC) from Regions 1 to 5;
- Gas-oil contact (GOC) from Regions 1 to 5;
- Horizontal permeability multiplier (kh);
- Relative permeability set (kr);
• Vertical permeability multiplier (kz) from layers 1, 8, 11, 12, 15, 18 and 20;
• Fault transmissibility (t) from eight faults;
• Volume multiplier for Segment G.

The geostatistical attributes were: (1) permeability map; (2) porosity map and (3) net to gross map. Two hundred geological realizations of each map type (1, 2 and 3) were generated considering well data until 2001. We assumed the probability distribution function (PDF) of the continuous attributes as uniform.

**Kolmogorov–Smirnov Test Analysis**

The maximum absolute difference between the DLHG and MC CDF samples (D) was determined to fifty-six outputs and compared to the critical value (D_{critical}) and limit value (D_{limit}). The level of significance considered was equal 10%.

Figure 4 to Figure 7 show the maximum absolute difference between the DLHG with 20, 50, 100 and 200 samples and MC with 5,000 samples, respectively. Table 1 shows the number of outputs that present the maximum difference D higher than D_{critical} and D_{limit}.

**Table 1 - Number of outputs considered not acceptable.**

<table>
<thead>
<tr>
<th>Limit Values</th>
<th>DLHG 20</th>
<th>DLHG 50</th>
<th>DLHG 100</th>
<th>DLHG 200</th>
</tr>
</thead>
<tbody>
<tr>
<td>D_{critical}</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D_{limit}</td>
<td>49</td>
<td>30</td>
<td>6</td>
<td>0</td>
</tr>
</tbody>
</table>

The samples sizes equal 20 and 50 were not acceptable considering both critical and limit values. The CDFs derived from the sample size equal 100 came from the same distribution that the MC sample, however 6 outputs were not considered acceptable because of the limit value.

Figure 8 shows an example of an output from the DLHG with 20 samples that was considered not acceptable by both criteria (Figure 8a) and rejected only by D_{limit} (Figure 8b). Figure 8b shows the need of the D_{limit}, because we did not consider that the CDF from the DLHG with 20 samples represents the MC sample in a visual way. Figure 9 shows two outputs considered not acceptable by the D_{limit} for the DLHG with 100, 50 and 20 samples.
Figure 4 – D value for all output considered: DLHG 20 sample.

Figure 5 – D value for all output considered: DLHG 50 sample.
Figure 6 – D value for all output considered – DLHG 100 sample.
Figure 7 – D value – DLHG 200 sample.

Figure 8 – Comparison of CDFs: (a) DLHG 20 rejected by $D_{\text{critical}}$ and $D_{\text{limit}}$  (b) DLHG 20 rejected only by $D_{\text{limit}}$.

Figure 9 – Comparison of CDFs: DLHG 100 rejected by $D_{\text{limit}}$. 
The variability of the CDF from $n$ runs of the same sample size decreases as the sample size increases. Figure 10 and Figure 11 show a comparison of CDFs of well E-3H oil rate NQDS and well E-4AH gas rate NQDS, respectively, from three different runs of the same sample size for the DLHG under study. There was a low variability on the CDF results from the runs of the DLHG with 200 samples.

Therefore, only the results from the DLHG with 200 samples were considered acceptable by the Kolmogorov–Smirnov test and presented low results variability. The comparison of the influence and correlation matrices for all samples is presented in the next section.

Figure 10 – CDFs of well E-3H oil rate from different samples of the same size: (a) DLHG 20 (b) DLHG 50 (c) DLHG 100 (d) DLHG 200.
Influence and Correlation Matrices Analysis

The influence and correlation matrices are used in the history matching process to identify the impact of the reservoir uncertainties on the NQDS results. The NQDS results were affected by the WOC and GOC levels from region 5 in accordance to the influence and correlation matrices from the MC sample. Figure 12 shows the correlation coefficients between gas rate NQDS and GOC_{Region 5} and water rate NQDS and WOC_{Region 5} for MC and DHLG samples.

The DHLG with 20 samples did not show the correlation between WOC_{Region 5} and NQDS_{Qw} for wells D-4H and E-3H., whereas the DHLG with 50 samples did not show the correlation for well B-3H. The correlation between GOC_{Region 5} and NQDS_{Qg} for DHLG with 20 samples was significantly higher than for the MC sample; this could lead to inaccurate impact of the attribute in the NQDS results. Both, DHLG with 100 and 200 samples were able to identify correlation between the outputs and attributes that were identified by the MC sample.

We also compared the influence matrix index variation in relation to the attribute levels from the DHLG and MC samples. Figure 13 shows the I_{Qw} variation over the WOC_{Region 5} levels for wells D-4H and B-3H. The results corroborated with the correlation coefficients presented in Figure 12a.
Figure 14 shows that there was a slight difference between the DLHG with 100 samples behavior and MC influence matrix results, even though the DLHG with 100 samples presented the same behavior as MC sample for the correlation matrix (Figure 12).

The DLHG with 200 samples replicated the same behavior from the MC sample. Thus, we considered that only the DLHG with 200 samples was able to replicate the same behavior from the MC sample. The next step was to perform the stability test. The DLHG with 20, 50 and 100 samples were discarded from the process.

Figure 12 – Correlation coefficient for all samples: (a) WOC Region 5 versus Qw (b) GOC Region 5 versus Qg.

Figure 13 – Influence coefficient for all samples: (a) well D-4H (b) well B-3H.

Figure 14 – Influence coefficient for all samples: (a) well B-3H (b) well E-1H.
Stability Analysis

The DLHG with 200 samples was considered acceptable from the Kolmogorov–Smirnov test perspective and it replicated the same influence and correlation matrices behavior from the MC sample. The last step of the methodology was to verify the stability of the DLHG with 200 samples.

Two new DLHG runs with 200 samples were generated and the Kolmogorov-Smirnov test was performed. The D values from the three DLHG runs with 200 samples were compared to evaluate the stability of the sample size.

Figure 15 to Figure 18 show the D value from the three DLHG runs for NQDS\(_{Qo}\), NQDS\(_{Qw}\), NQDS\(_{Qg}\) and NQDS\(_{BHP}\), respectively. Despite the variation on the three DLHG runs results, they were considered acceptable from the Kolmogorov–Smirnov test perspective because all values were lower than \(D_{\text{critical}}\) and \(D_{\text{limit}}\). Thus, the DLHG with 200 samples was considered stable.

![Figure 15 – DLHG 200 stability test: NQDS \(Q_o\).](image1)

![Figure 16 – DLHG 200 stability test: NQDS \(Q_w\).](image2)
Conclusions

We presented a method that evaluates the ability of the Discretized Latin Hypercube Integrated with Geostatistical realizations (DLHG) to replicate a reference solution (generated by an exhaustive Monte Carlo sampling) in the history matching process. The main characteristics of the method are: (1) it uses a statistical technique to compare the output CDF’s from the reference and DLHG samples and (2) it evaluates the ability of the DLHG sample to identify the reservoir attributes that affect the history match results.

The results showed that the DHLG technique can be used as a sampling method in the history matching process under uncertainties in a complex case such as the Norne benchmark case, because it was able to represent the variability of the misfit between production and simulated data.
The study also showed that a $D_{\text{limit}}$ should be used along with the $D_{\text{critical}}$ to verify if the DLHG sample size is able to represent the CDFs outputs from a reference solution. The influence and correlation matrices also should be evaluated, because even though a sample size represents the reference outputs CDFs, it must identify the reservoir attributes that affect the history match outputs.

The correlation and influence matrices presented similar conclusions about which attributes affects the NQDS results, however the correlation matrix is easier to implement and evaluate. Thus, the user can apply only the correlation matrix in the sample size evaluation to reduce the effort spent in the process.

We have compared four DLHG sample sizes with Monte Carlo sample, which can be considered a reference solution due to the amount of simulations. A DLHG with 200 samples was able to reproduce the CDFs outputs, it was able to reproduce the influence and correlation matrices behavior, and it was considered stable. Therefore, the DLHG with 200 samples can be used in the history-matching process to the case studied considering the described reservoir characterization.

It is important to highlight that if the number of uncertain attributes increases significantly during the history-matching process the sample size evaluation should be performed again. The results obtained were in accordance to what Schiozer et. al (2016) have proposed that a number between 100 and 300 properly quantifies the reservoir risk.

The sample size affects the time demanded and the accuracy of the history match results. Both factors should be evaluated before starting the history match in order to obtain reliable results. Such evaluation gains greater importance in real reservoir models due to the complexity of reservoir. Therefore the methodology here proposed can be a valuable tool to ensure: (1) a proper history matching procedure and (2) a more robust risk analysis.

### Acknowledgments

This work was carried out in association with the ongoing Project registered as "BG-07 Reduction of uncertainties through the incorporation of 4D seismic data in the modeling of the reservoir" (UNICAMP/BG Brazil/ANP) funded by BG Brazil under the ANP R&D levy as "Investment Commitment to Research and Development". The authors thank also UNISIM, DE-FEM-UNICAMP, CEPETRO for supporting this work and CMG and Schlumberger for software licenses. We also would like to thank Statoil (operator of the Norne field) and its license partners ENI and Petoro for the release of the Norne data. Further, the authors acknowledge the Center for Integrated Operations at NTNU for cooperation and coordination of the Norne Cases. The view expressed in this paper are the views of the authors and do not necessarily reflect the views of Statoil and the Norne license partners.

### References


Verlo, S., Hetland, M., "Development Of A Field Case with Real Production and 4D Data from the Norne Field as a Benchmark Case for Future Reservoir Simulation Model Testing". Master’s degree, NTNU, 2008.