Uncertainty Quantification for History-Matching of Non-Stationary Models Using Geostatistical Algorithms

Maria Helena Caeiro¹, Vasily Demyanov², Mike Christie³ and Amilcar Soares⁴

Abstract The quantification of the uncertainty associated to the characterization of a petroleum reservoir is a challenge in reservoir modelling practices. Describe the fluid flow course of meandering channels is a prime issue in the history matching of non-stationary processes. A novel methodology to integrate image transforming method based on direct sequential simulation with local anisotropy correction (DSS-LA) and sampling optimisation techniques in the history matching process is presented in this paper. DSS-LA is a variogram based algorithm which has the advantage of using the original variable without requiring any processing, very relevant for the simulation of continuous variables. It tackles the problem of non-stationarity and connectivity of the model by introducing spatial trends, which represent local anisotropy variations (direction of maximum continuity and anisotropy ratio). The parameters of the local anisotropy model are optimised to match the production history data, which leads to the uncertainty quantification through the Bayesian inference. The methodology was applied to a synthetic petroleum reservoir with fluvial deltaic structures (Stanford VI). As a result, the stochastic models of reservoir’s properties (porosity) and the respective dynamic responses are obtained. Multiple history matched models quantify uncertainty related to the trend model. They have good fitting to the production data and provide more precise uncertainty evaluation for the predictions. This application aims to demonstrate the feasibility of linking DSS-LA with an optimisation algorithm to integrate dynamic data and quantify the uncertainty associated to the characterization of a non-stationary reservoir in the history matching process. The results reveal that this has been successfully achieved and the proposed approach demonstrates efficiency and robustness for the generation of multiple history matching models.

¹ CERENA, Instituto Superior Técnico, Av. Rovisco Pais, 1049-001 Lisboa, Portugal, helena.caeiro@ist.utl.pt
² PET, Heriot Watt University, Riccarton Campus, Edinburgh, Scotland, Vasily.Demyanov@pet.hw.ac.uk
³ PET, Heriot Watt University, Riccarton Campus, Edinburgh, Scotland, Mike.Christie@pet.hw.ac.uk
⁴ CERENA, Instituto Superior Técnico, Av. Rovisco Pais, 1049-001 Lisboa, Portugal, asoares@ist.utl.pt

Introduction

History matching and uncertainty quantification are two important steps in reservoir modelling procedures, which must be integrated to predict production performance of a field. History matching is an inverse problem where the model is calibrated to reproduce the historical observations of the field. The parameters of the model of the reservoir are perturbed and the dynamic responses are compared with the historical known productions in the wells. This is a non-unique inverse problem, where one can achieve different property models that have a good match with historical production data.

The challenge gets a higher dimension when one intends to calibrate non-stationary models. This means that the structures to be simulated are not regular and have local change of direction and high variability in local scale. Sand channels features, such is the current application, is one example. The challenge focuses in the simultaneous simulation of both morphology and distribution of the petrophysical properties. An accurate representation of the geometry of the reservoir is essential for the prediction of its performance.

There are many methods presented in the literature to do stochastic characterization of sand channel reservoirs. One of them are the object-based methods (Boolean models) [11, 12, 16, 24]. They work by locating in the space whole objects or set of pixels at a time, having a defined geometry, such is a curvilinear channel [8]. Objects are added until the target number or volume proportion is matched. Through a trial and error process controlled by an objective function the objects sizes are modified, moved around and added or removed. The advantage of these algorithms is the crisp reproduction of the target geometries. Its drawback is the difficulty to honour a combination of prior proportion curves and dense well facies indicator data [15]. Nowadays, the multi-point statistic is a very popular strategy to characterize curvilinear features [2]. It allows capturing structures from “training images” by borrowing the patterns and then anchoring them to subsurface well log, seismic and production data. Even not a traditional variogram based method, it does not escape from the same principals as traditional variogram-based geostatistics [3]. It is still a stochastic method and is very capable models in the reproduction of the connectivity. Nevertheless, there is always associated a huge uncertainty associated to the training image, which intend to reproduce the geological patterns. Also neural networks applications have been applied to describe non-stationary structures as an alternative to geostatistical tools [7]. It captures dependencies from available input data and produces realistic patterns. The drawback of this procedure is related with the need of a previous morphological model to constrain the simulation of the property.

Soares [21] did one of the first attempts to model the morphology of geological curvilinear shapes using geostatistics. He proposed the use of local anisotropy directions to estimate folded geological data (morphological kriging). The importance of this result for petroleum applications is that it provided the
possibility to identify different structures for the numerical modelling of reservoirs. Later in 1997, Luis and Almeida [17] and Xu [25] applied this idea in order to condition sequential simulation procedures for the characterization of sand channels geometry in a fluvial reservoir. The pixel-based models do not have the pleasing geometry of the boolean models but reproduce the continuity statistics measured by the experimental data.

Considering the state of the art, this article presents a methodology to simulate continuous variables (delta front) which is pixel based and is relies on the proposed by Horta [13]. It presents an application of direct sequential simulation (DSS) [22] to the characterization of continuous variables with a spatial distribution conditioned to a meander structure, accounting the local anisotropy information (direction of maximum continuity and anisotropy ratio). Thus, the b-point statistics are reproduced plus the local anisotropy information.

One of the major topics of the studies about petroleum is the characterization of internal properties, in spite of the impossibility to check results directly with the reality. A main step of the implemented methodology is to create a model of local anisotropy to define the local changes of spatial continuity of the fluid flow paths. It is composed by the model of directions of maximum continuity $\theta$ (given by the azimuth) and the ratio of anisotropy $r$ (ratio between the major and the minor amplitude distances). In the presented case, the model of anisotropy was created using smooth trends, a deterministic approach that considers regions based on prior beliefs about the depositional process. The anisotropy model is composed by changing the parameters, which deterministically define the smooth poly-linear trend function: the coordinates of index point of the channel direction angle and the anisotropy ratio trend surfaces and the angle and ratio boundary values themselves. This anisotropy model defines the channels directions as variable in linear bands and linearly thinning along the flow direction. The model of local anisotropy was integrated in the stochastic simulation of internal properties of the reservoir to explore uncertainty associated to the geometry of the non-stationary delta geometry structure [1].

The uncertainty quantification is a problem with huge importance in reservoir engineering. The scheme of having just one good model of history matching is no longer a goal. Now, one intends to generate multiple models, which are “good enough” to honour the production history data and at the same time include the effects of uncertainty in the model components. The production of multiple reservoir models enables the quantification of the probability of the production forecast, which includes the most likely scenario, bounded by confidence intervals, and characterization the chance on the unknown true solution all within the uncertainty bounds [6]. The main difficulties on doing probabilistic predictions are the incorporation of all sources of uncertainty and the necessity of a high number of reservoir simulations [9].

In this paper is presented a genetic based and sequential method to quantify the morphological uncertainty associated to the case study field, thought the optimization of the history matching procedure. The convergence is done by a
sequential simulation technique, using the variogram based algorithm direct sequential simulation with local anisotropy correction (DSS-LA) [14].

This paper is organized in three main sections. The first one is theoretical and presents a summary of the applied procedures. It is divided into two sub sections which correspond to the explanation of the principal steps of the implemented methodology. In the second section we present the application, including the case study description and the main results and discussion topics. Finally, the conclusions are summarized in the third section.

**Uncertainty Quantification Framework for optimized History Matching**

One important point of this paper is the integration of an anisotropy model of continuous variables with an optimisation and convergent iterative process in a way to assess the fitting of the dynamic responses from the static models of porosity produced in by DSS-LA and the posterior uncertainty quantification. The general scheme of the implemented methodology presented in this article is shown in Figure 1.

![Flowchart of the process](image)

**Describing the fluid flow course using DSS-LA as image transforming method**

The problem of conditioning simulation to channel structures was raised in the context of characterization the morphology of the case study field. Since the fluid flow trend is defined, the anisotropy parameters can be estimated. The use of a stochastic simulation method is a reliable option to determine the spatial distribution of the internal properties of a reservoir (porosity and permeability) conditioned to the curvilinear structures of a channel. Simulation algorithms allow the spatial assessment of an attribute and provide information about the uncertainty involved on that evaluation. The spatial trends representing
local anisotropy variations which reproduce the meander structures where a variable is simulated were introduced in the DSS algorithm [14].

DSS is an algorithm of stochastic simulation, from which one can obtain equiprobable images from models of the probability distribution of random variables. Although the petroleum reservoirs are not a result of random processes, they have attributes that make them behave as if they were. DSS does not require any processing of the original variable. This is very relevant for the simulation of the continuous variables and a great advantage over stochastic indicator simulation (SIS) and stochastic gaussian simulation (SGS). In classical DSS that the spatial correlation is evaluated across the Euclidean space and the algorithm does not account for specific features, because the algorithm performs the search of conditioning data according to the global variogram parameters (direction, range and ratio of anisotropy). It assumes that the global variogram parameters are representative for the entire study area following the intrinsic hypothesis. Therefore, in a non-stationary situation classic DSS is not able to reproduce the curvilinear shapes. Horta [14] in her thesis, provided a solution to the non-stationary case: introducing local spatial trends to reproduce the meander features of the structures, by the representation of the local anisotropy variations (direction of maximum continuity and anisotropy ratio) [1].

An elliptical search radius to select the conditioning data, which is defined by the global variogram parameters, is used to solve the simple kriging equations of DSS. With DSS-LA, the direction of maximum continuity (given by azimuth θ) and anisotropy ratio \((r=a_θ/a_Φ)\) which are the local anisotropy parameters are used to change the search radius from node to node, determining that the variogram model is non-stationary. With this change, the simple kriging estimate of the local mean \((Z_{sk}(x_u)^*)\) becomes function of \(θ(x)\) and \(r(x)\), but to estimate a local cumulative distribution function at given location \(x_i\) only the local angle of \(x_i\) is retained. The matrix of data-to-data co-variances and the vector of data-to-unknown covariances are calculated with corrected local covariances \(C_{θr}(h)\) by the local values of \(θ(x)\) and \(r(x)\) [1]. The problem of non-stationarity of the geostatistical model is tackled by the implementation of local anisotropy by imposing a trend in spatial correlation structure, allowing the description of complex continuous and connected structures. One of the issues with the local anisotropy model becomes its source and the uncertainty associated with it [1].

The optimisation is performed by perturbing the local anisotropy model parameters, which condition the petrophysical DSS-LA models of porosity. In sum, the methodology performed was, firstly, the creation of a model of anisotropy, which defines the main trends of anisotropy’s channels (\(θ,r\)); then the integration of DSS-LA (and the anisotropy model creation) in a global iterative algorithm that also contains the dynamic simulation (Eclipse Black Oil Simulator) and the optimisation algorithm. This parallel iterative process is looped until the lowest misfit (determined for a general number of created models) is achieved. The anisotropy model was modified having in account the tuned anisotropy parameters which allow the convergence to the best model [1].
Generating multiple models in the sampling space and posterior inference in a Bayesian framework

The optimization algorithm used to sample in the multiparameter space was the Neighbourhood Approximation Algorithm (NA) method. It was developed by Sambridge [27] for solving a seismic waveform inversion problem. It is capable of finding models of acceptable data fit in a multidimensional parameter space. Subbey [32] did the first successfully application in reservoir modelling. This derivative-free method aims finding an ensemble of acceptable models rather than seeking a single solution and uses a structure of population members to find good fitting regions in the search space. To describe and search the multidimensional space, NA uses geometrical features of Voronoi polygons in high dimensions. Also, the posterior probabilities of the model are estimated using the Gibbs sampler [1]. This allows an accurate quantification of uncertainty and variability of the parameters of the models using a Bayesian extension of the framework [5].

Quantifying the uncertainty using NA includes two phases:
- **Search phase**: generation of multiple models that match the history data (solutions of the inverse problem) by sampling the multiparameter space [27];
- **Appraisal phase**: posterior inference in which NA-Bayes (NAB) [28] computes the posterior probability on the basis of the misfits of the sampled models and the Voronoi approximation of the misfit surface [5].

The algorithm is initialized with the randomly generation of \( n_s \) models in the search space. The algorithm generates \( n_s \) new models (as combinations of the model parameters) and calculates the objective function for each of them, followed by the ranking of all models according to their misfit score. This is done for each iteration. The latter describes the goodness of the model fit to the observed production data. Hence, the best \( n_r \) models which have lowest misfits are chosen and then new \( n_s \) models are generated by uniform random walk in the Voronoi cells of these best \( n_r \) cells. The algorithm behaviour is controlled by the ratio \( n_s/n_r \). Meaning that, \( n_s/n_r=1 \) aims to explore the space and find multiple regions of good fitting models. Increasing the ratio value, the algorithm tends to improve the matches obtained by finding multiple clusters of good models [25]. This process is repeated and stops just when the predetermined stopping criterion is met (usually the maximum number of iterations) [10].

The Bayes theorem is an important tool to help in the uncertainty quantification. If multiple NA-models are generated, the uncertainty can be quantified from its posterior distribution, namely using the Bayesian extension of NA (NAB) [20].

The updated current knowledge of a system on the basis of new information is provided by a systematic procedure using the Bayesian framework, a formal approach for statistical inference. The system is represented by a simulation model.
that includes all the information needed to solve a given problem. However not generally includes the parameters used to specify the numerical solution procedure itself. So, any information in \( m \) may be uncertain. Thus, \( M \) represents an ensemble of models in which the uncertainty is going to be determined. Meaning that is \( m \in M \) and a probability distribution on is defined \( p(m) \).

Using the Bayes formula, one can determine the updated estimate of probability for \( m \) – posterior distribution (Equation 1) [5].

\[
p(m|O) = \frac{p(O|m) p(m)}{\int \frac{p(O|m) p(m)}{p(O)}}
\]

Whereas:

- \( p(m|O) \) is the posterior probability (probability of the model \( m \), given the observed values \( O \) – data)
- \( p(O|m) \) is the likelihood term (probability of the data assuming that the model is true), representing the information that observed data provides
- \( p(m) \) is the prior probability (can be given as the sum of independent probabilities for model parameters or as more complex combination of model inputs)
- \( p(O) \) is a normalized constant (sometimes referred as evidence; if the term is small, it suggests that the model doesn’t fit the data well)

The misfit (\( M_f \)) is most of the times considered as the negative logarithm of the likelihood function (\( L \)) (Equation 2):

\[
M_f = - \log L.
\]

Assuming that the measurement errors are Gaussian, independent and identically distributed the misfit function can be computed as the conventional least-square formula (Equation 3). The misfit value represents how well a model fits the data.

\[
M = \sum_{i=1}^{N_w} \sum_{j=1}^{N_v} \sum_{k=1}^{N_t} \left( \frac{q_{ijk}^{obs} - q_{ijk}^{sim}}{2\sigma_{ij}^2} \right)^2 = \sum_{i=1}^{N_w} \sum_{j=1}^{N_v} \sum_{k=1}^{N_t} \frac{1}{2\sigma_{ij}^2} \left( q_{ijk}^{obs} - q_{ijk}^{sim} \right)^2.
\]
each well. The data standard deviation (σ) is 10% of the average of history data values for each variable of each well and the history period corresponds to 5 387 days [1]. This is Markov Chain Monte Carlo resampling technique, based on the multiple history matched models which uses the properties of Voronoi cells in high dimensions to accomplish multi history matched models.

Application to a non-stationary case

Case Study

Stanford VI [4], a synthetic reservoir model with a fluvial channel system, is the case study in which the proposed methodology was applied. This reservoir is divided into three layers but for this specific study only one 2D slice was considered. The dimensions are 150 blocks in x axis and 200 blocks in y axis. The stochastic simulation is done using the DSS-LA algorithm and the prior ranges of anisotropy tuning parameters used in the NA parameterization are the angles of the direction and the values of the ratio of anisotropy. The direction of maximum continuity varies along the x axis of the trend model and the range of angles is defined according with the geological assumption of the channels, it means, the value represents the angle which the channels makes with the N-S direction (0° corresponds to N-S direction). The ratio of anisotropy varies along the y axis and it represents how thinner the channel is (higher values of ratio means thinner channels). Relating these ranges to the geological thinking, it means that the channels are turning sideways from the delta axis direction and are getting thinner in the direction of the depositional flow [1]. This means that the uncertainty analysis was performed based on the misfit in a 10-dimension parameter space. For this run, the application of NA framework were generated 1120 multiple history matched models, by initializing with 50 models (n1), chose the 7 models with lower misfit (n7) and produced 50 new models (n50) in the 7 Voronoi cells of the models. The number of iterations was 50. The hard conditioning data for DSS-LA of porosity and permeability is shown in Figure 2. They represent the variable distribution and enable the DSS-LA validation. In the porosity histogram two populations can be distinguished: for low porosity is mud and for high porosity the sand channels. Were considered 31 wells with porosity and permeability data to condition the stochastic simulation with DSS-LA but only 13 were used to assess the match of the dynamic responses of the static models. Figure 3 presents the location of the wells used to do the dynamic match.
Figure 2: Histogram and statistics of porosity of the wells used as hard data in DSS-LA

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>0.14%</td>
</tr>
<tr>
<td>variance</td>
<td>0.008%^2</td>
</tr>
</tbody>
</table>

Figure 3: Location of the wells

- wells used to do History Matching
- wells used in DSS-LA
Results and Discussion

The resultant models of porosity honour the static hard data (spatial variability) and the dynamic history data observed in the wells. Figure 4 presents the maximum likelihood model of porosity, its histogram and basic statistics. It is shown that the geometry of the structures (channels) is reproduced, meaning this that the method can provide morphological simulation of the models (spatial distribution). Also, when comparing Figure 4 with Figure 2, is evident that the histogram, mean and variance are honoured in the generated models, i.e., the method preserves the statistics of the simulated variable (porosity).

Figure 4: Porosity maximum likelihood model, its histogram and statistics
The lowest misfit value (M) of the models is 739.77 and the misfit convergence of the generated multiple history matched models is shown in Figure 5. The misfit value represents how the model fits the dynamic history data. The lowest the misfit, the better the fitting of the parameters used to define the trend anisotropy model.

The NAB framework was used to evaluate the uncertainty of the model parameters of the tuned anisotropy trend model. The evaluation of the uncertainty of the model parameters was done using the posterior probability distribution obtained for the resampled models. The NAB algorithm resampled 83 unique models from the NA-ensemble. The number of random samples generated by NA-walks is 500 000 and the total number of Voronoi cells sampled by NA-walk is 100. Figure 6 presents the histogram of the NAB generated samples. The most frequented model was visited 43.9% of the times (red) 95% of the samples generation of NAB is from the 8 models (they have 95% of probability of visit). NAB selects only models with low misfit, preserving the global character of the misfit distributions.

![Figure 5: Misfit convergence of the generated multiple history matched models](image)
Figure 6: Histogram of the NAB generated samples

To represent the inference, Figure 7 presents the coverage of 95% of the posterior probability distribution of the production curves of the field oil production ratio (FOPR). They were produced by the implemented geostatistical methodology and their dynamic responses are compared with the historical values and the true case model ones. The maximum likelihood model is also represented. The represented FOPR production curves have a similar reproduction and an acceptable fit for the history data production. The variability space is reduced and it shows that the run is convergent.

Figure 7: FOPR production curves (representing 95% of the posterior probability distribution)
Conclusions

The stochastic sequential algorithm DSS-LA was combined with the NA and NAB frameworks and applied to a non-stationary case, a delta front 2D slice of the synthetic reservoir Stanford VI. The generation of multiple history matched models using NA linked with DSS-LA demonstrated efficiency and feasibility. The proposed approach provided the assessment of the uncertainty of the geological trend model, very relevant for a non-stationary case.

The integration of DSS-LA in a stochastic perturbation framework of optimization in history matching revealed viability to generate models of porosity. The space of uncertainty is reduced when the method converges and the resultant static models of porosity have a good conformity between the morphology of the channels, the porosity distribution and the dynamic responses.

Acknowledgments

Sponsors of Heriot-Watt Uncertainty Quantification project, Phase III: BG group, BP, ConocoPhillips, JOGMEC
Sponsors of Instituto Superior Técnico: Joint agreement Partex Oil & Gas and Petrobras/CENPES

Bibliography


