Improved Geological Modeling and Dynamic Data Integration Using
the Probability Perturbation Approach

by

Alvaro Enrique Barrera, B.S.

Report
Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science in Engineering

The University of Texas at Austin
May 2006
Improved Geological Modeling and Dynamic Data Integration Using
the Probability Perturbation Approach

APPROVED BY
SUPERVISING COMMITTEE:

__________________________________
Sanjay Srinivasan

__________________________________
Steven L. Bryant

__________________________________
Larry W. Lake
Abstract

Improved Geological Modeling and Dynamic Data Integration Using
the Probability Perturbation Approach

by

Alvaro Enrique Barrera, M.S.E.
The University of Texas at Austin, 2006

SUPERVISOR: Sanjay Srinivasan

The lack of geological information to appropriately reproduce the flow paths in typical reservoir modeling scenarios is one of the most significant sources of uncertainty and, consequently, it has been repeatedly addressed in multiple geological modeling approaches. However, current modeling techniques that combine geological and dynamic information are computationally extensive and frequently lead to geologically inconsistent models.

Practitioners and particularly geologists have found the traditional geostatistics, based on the variogram and kriging, suitable to describe geological heterogeneity within a single facies, but too limiting for describing more organized
geological features such as channels, fractures, and facies distributions among others. These geological features usually have the largest impact on the flow response, calling for a different approach using multiple-point statistics instead of variogram models to capture the required conditional information to generate more accurate models that exhibit more structural organization.

Honoring the geological model is an important objective during the generation of static geological models; however, it is commonly forgotten during the integration of dynamic information. Reconciliation between the model predictions and the field records represents a rather significant challenge, considering the highly non-linear relationship between the model parameters and the production response. Reproduction of the geological heterogeneity during the calibration of the model response with the production history is the motivation behind the selection of a probabilistic approach for dynamic data integration. This work focuses on a probabilistic approach to integrate dynamic data that ensures consistency between reservoir models developed from one stage to the next. The algorithm relies on efficient parameterization of the dynamic data integration problem and permits rapid assessment of the updated reservoir model at each stage.

This report is presented as part of the requirements to obtain a Master of Science degree in petroleum engineering under the Fast Track Option. It summarizes the proposal and preliminary results related with the Ph.D. dissertation that is currently on going.
**Table of Contents**

List of Tables ........................................................................................................................................ vii

List of Figures ....................................................................................................................................... viii

1 INTRODUCTION .................................................................................................................................. 1

2 LITERATURE REVIEW ......................................................................................................................... 4

3 IMPROVED GEOLOGICAL MODELING USING MULTIPLE-POINT STATISTICS ........................................ 8

   3.1 Multiple-Point Approach Overview ................................................................................................. 8

   3.2 Geological Feature Identification and Simulation .............................................................................. 12

      3.2.1 Scanning Process .................................................................................................................... 12

      3.2.2 Simulation Process .................................................................................................................. 14

   3.3 Template Selection Algorithm ......................................................................................................... 17

      3.3.1 Preliminary Studies of Template Selection ............................................................................... 20

4 DYNAMIC DATA INTEGRATION ......................................................................................................... 24

   4.1 Geostatistics Overview ..................................................................................................................... 26

   4.2 Sequential Indicator Simulation ....................................................................................................... 29

   4.3 Gradual Deformation of Geological Models using Dynamic Data ................................................. 33

      4.3.1 Perturbation of Local Conditional Distributions ...................................................................... 33

      4.3.2 Conditioning of Local Distributions to Dynamic Information .............................................. 42
4.3.3 Merging the Information from Dynamic and Static Sources ................. 46

4.4 Gradual Updating Procedure for History Matching ........................................ 48

4.5 Application ........................................................................................................ 52

5 CONCLUSIONS AND RECOMMENDATIONS ........................................... 58

Bibliography ........................................................................................................... 62

Vita ............................................................................................................................ 67
List of Tables

Table 1. Description of a simulation case for the evaluation the history matching algorithm.................................................................................................................. 53
List of Figures

Figure 1  Schematic for the proposed multiple point stochastic simulation approach 10

Figure 2. Simulation results for a 3D application of the multiple point stochastic
algorithm ................................................................................................................................. 11

Figure 3. Examples of optimal templates obtained with the Template Algorithm..... 19

Figure 4. Training image exhibiting North-South ellipsoidal lens. ......................... 21

Figure 5. Impact of template geometry on captured statistics .............................. 22

Figure 6. Impact of template size on captured pattern statistics ..... 23

Figure 7. Spatial interpolation obtained by: a) Kriging; b) sequential Gaussian
simulation and c) sequential indicator simulation ............................................. 28

Figure 8. Example of gradual deformation of geological models by probability
perturbation method ............................................................................................. 36

Figure 9: Effect of deformation parameter on local distribution under 3rd perturbation
scheme. ...................................................................................................................... 39

Figure 10. Example of gradual deformation of geological models by probability
perturbation scheme four .................................................................................. 41

Figure 11: Effect of the deformation parameter on the flow response of a geological
model. ...................................................................................................................... 43
Figure 12. Effect of the rD parameter in the water saturation distribution in a reservoir after 4000 days and 5200 days

Figure 13: Illustration of the Dekker-Brent inverse parabolic interpolation process.

Figure 14: Field pressure history match for the study case.

Figure 15. Field production history match for the study case.

Figure 16. Gradual deformation of the geological model (third layer) in the study case.

Figure 17. Third layer of the reference geological model in the study case.

Figure 18. Convergence of the objective function in the study case.
1 INTRODUCTION

The lack of geological information (subsurface measurements) to accurately or at least appropriately reproduce the flow paths in typical reservoir modeling scenarios is one of the most significant sources of uncertainty and consequently, it has been repeatedly addressed in multiple geological modeling approaches. However, regardless of the approach, techniques for the reproduction of geological structures in reservoir models continue to be limited and uncoupled with dynamic information. Current modeling techniques that combine geological and dynamic information are computationally extensive and frequently lead to geologically inconsistent models, where any updating to integrate recently acquired dynamic information require major adjustments in the model, or in worst cases the construction of new ones.

Stochastic simulation supported by the principles of geostatistics, has become a widespread tool for reservoir modeling and uncertainty assessment, by generating multiple equiprobable reservoir models preserving an apparent geological structure. Traditional geostatistics are based on two basic concepts, the variogram model as representation of the spatial heterogeneity or continuity (more statistical than geologic), and kriging for spatial interpolation. Nowadays, practitioners and particularly geologists have found the variogram suitable to describe geological
heterogeneity within a single facies, but too limiting for describing and reproducing more organized and sharp geological features such as channels, fractures, and facies distributions among others. These reservoir architecture related features usually have the largest impact on the flow response. The traditional variogram-based geostatistics fail to appropriately retain the required conditional information to reproduce these features; generating amorphous realizations that exhibit maximum entropy instead of systematically organized structures and patterns as expected from prior geological knowledge. These geological features call for a different approach utilizing multiple-point (MP) statistics instead of variogram models (two point statistics) to capture the required conditional information to generate fields that exhibit lower entropy, more structural organization and preserve reservoir heterogeneity.

Normally, reservoir models are initially built based on static information proceeding from different sources, such as seismic, stratigraphy and logs. Then, these models are tuned in with dynamic information from tests and production records, in order to validate them. Reproduction of the historical production data is usually the most important goal during the validation process of a reservoir model. Reconciliation between the model prediction and the field records however represents a rather significant challenge, considering the highly non-linear relationship between the model parameters and the production response. Consequently, the process requires an iterative approach that translates into a higher computational effort. In spite of intense research activity in this area and significant achievements during the
last decade, the number of relevant realistic field applications of assisted history matching is still rather limited and given the complexity of the problem, there are still several issues to be addressed.

The main objective of this report can be described as: i) To assess the potential of the multiple-point geostatistics approach for reproducing geological features in reservoir models. Several components of the MP simulation approach: the training image, the spatial template, the scanning technique and the simulation algorithm, are critically evaluated; ii) Develop a conditional MP simulation method implemented in multi-\textit{cpu} computational environment; iii) History matching of reservoir models using the probability perturbation method and demonstration of the method in some reservoir example.

This report is presented as part of the requirements to obtain a Master of Science degree in petroleum engineering under the Fast Track Option. It summarizes the proposal and preliminary results related with the Ph.D. dissertation that is currently on going.
2 LITERATURE REVIEW

Geostatistics and its two original basis, the variogram model and the kriging methodology, were initiated by the work of Daniel Krige (1951) and built up by Georges Matheron (1962-1963, 1965), with the purpose of providing locally accurate grade estimates of mining blocks; however, its application has extended from the mining industry to many other related disciplines, including the oil industry.

Stochastic simulation was introduced by Matheron (1973) and Journel (1974) to correct for the smoothing effects and other artifacts of kriging, allowing the reproduction of spatial variance predicted by the variogram model. Different algorithms were developed including sequential simulation (Journel, 1983, Isaaks, 1990; Srivastava, 1992; Goovaerts, 1997; Chiles and Delfiner, 1999), which has become the workhorse for many current geostatistical applications. Stochastic simulation provides the capability to generate multiple equiprobable realizations, giving birth to the idea of assessing spatial uncertainty (Journel and Huijbregts, 1978).

Stochastic simulation stepped away from the variogram and kriging based core for the first time with the Boolean object-based algorithms, introduced by
Stoyan, Kendal and Mecke (1987), and Haldorsen and Damsleth (1990), in an attempt to reproduce geological features, like channels and fractures, by fitting parametric shapes. Initially Srivastava (1992), and later Caers (1998) and Strebelle (2000), proposed the idea of borrowing conditional probabilities directly from a training image, allowing the use of higher order or multiple point statistics to reproduce geological structures and patterns. Advantages of this approach over the Boolean object-based method include the pixel-based non-iterative process, and ease of integration of data of multiple types.

Training images and multiple point statistics methods remained largely untried until developments in computational capacity and multiple point scanning techniques took place. The search tree, a training image scanning technique proposed by Strebelle (2000), allowed wide application of the multiple point statistics approach. In the search tree approach, the scanned statistics are recorded in a dynamic data structure that renders it convenient to store nested information such as number of outcomes of data events (joint probabilities). Classification methods proposed by Arpat (2003) and Breiman et al. (1984); and neural net based models, allow the reproduction of “similar” data events by modeling and interpolating between found data events.

Currently, the basic components that characterize the application of multiple point geostatistics remain the major focus of continuous research and development
efforts, with the objective of defining a more consolidated and practical methodology. These components include, among others, the generation or acquisition of numerical spatial representations to be used as training images; the optimal definition of scanning templates to capture the proper conditional information from the training image; and development of computational schemes that render the application more practical and convenient. This research focus on the optimal selection of spatial templates, investigation of the dependence of simulated models on training images, and the definition of a feasible computational environment for the application of multiple point statistics on high resolution 3-Dimensional reservoir models.

The final purpose of every reservoir model is to provide reliable predictions of reservoir performance. Even though the accuracy of a prediction model can not be easily evaluated or measured before the actual event occurs, it seems reasonable that models with a more accurate description of the driving mechanisms will produce more realistic estimates. This is the case of reservoir models that honor the geological heterogeneity as a driving mechanism for flow. The motivation behind the attempt to reproduce more realistic geological features is to ensure consistency and accuracy in the flow response of the reservoir model.

Multiple approaches have been considered in the task of conditioning reservoir models to dynamic flow data. Reservoir models are characterized by a very large number of parameters. The optimization algorithms follow an iterative process
to update the reservoir model applying a perturbation scheme to calibrate parameters with the flow response.

Different approaches have been considered in the attempt to integrate dynamic data in reservoir models, including gradient based methods (Yeh, W., 1986; Anterion, F. et al 1989), Pilot point methods (La Venue el al. 1992; de Marsily et al. 1995), sequential self-calibration methods (Gomez-Hernandez, J.J, Capilla, J.E. et al 1997), Markov chain Monte Carlo methods (More H, Tjelmeland, H,. 1996), Gradual deformation methods (Roggero, F., Hu, L.Y., 1998) and Geostatistical Methods which suggest that the permeability fields could be conditioned to both hard data as well as dynamic data in a probabilistic sense by using a Bayesian formulation for the integration.

Literature addressing the issue of dynamic data integration utilizing multiple point statistics is very limited. Srinivasan and Caers (2000) proposed a neural network-based procedure for modeling multiple point averages of permeability in the neighborhood of the well, by filtering the information related to the permeability field in flow response data.
3 IMPROVED GEOLOGICAL MODELING USING MULTIPLE-POINT STATISTICS

3.1 MULTIPLE-POINT APPROACH OVERVIEW

Sequential Indicator Simulation is a classic stochastic simulation approach that is pixel-based and uses kriging/cokriging to obtain estimates of the necessary conditional distributions and generate various analogous reservoir realizations of the target reservoir. However, reproduction of complex 3D patterns is not possible using traditional two-point statistics (variogram) based approaches. Therefore, a new approach based on multiple point statistics has surfaced in the literature.

This approach requires a training image, a numerical representation of the spatial law or explicit non-conditional conceptual description of the geological structures and patterns to be reproduced in the model. Training images require some prior information of the reservoir depositional environment. Training images can be obtained from outcrops and photographs. Although it seems easier to define a variogram model than develop a training model depicting the critical reservoir heterogeneity; it is important to realize that variogram models are not any less
subjective, and moreover, they are more limiting and less intuitive in describing geological heterogeneity. Patterns to be reproduced can be explicitly depicted in the training model as compared to the hidden higher-order statistics implicit within traditional variogram-based geostatistical models.

A new multiple point stochastic simulation approach has been proposed and implemented in this research, based on geological feature identification and reproduction using a unique growth-based simulation algorithm that speeds up the simulation. The algorithm was developed in the java language and runs in multi-threaded computational environment. Features are grown starting from conditional data locations based on multiple point statistics inferred using optimized spatial templates. A fast and robust algorithm to derive the optimal configuration of spatial templates has also been implemented.

This approach is summarized in the following procedure (Figure 1): First an optimal spatial template is defined by the template optimization algorithm based on the training image. Then, the multiple point stochastic simulation algorithm uses the scanning template to capture the proper conditional information (multiple-point statistics) from a training image or numerical representation of the spatial law, applying a scanning technique to store the conditional information during the scanning process and retrieve it during the growth-based simulation. The selection of
the optimal spatial template is critical during the capture (scanning) and reproduction (simulation) of connectivity patterns representing geological features.

Since computational efficiency is a significant consideration in implementation of multiple point statistics-based modeling approaches, a multithreaded computational environment has been implemented to render the process efficient for field scale reservoir characterization.

![Figure 1](image.png)

Figure 1  Schematic for the proposed multiple point stochastic simulation approach

The Figure 2 shows an example of the results obtained with the algorithm for a particular 3D application. In this Figure a sample slice of the training image is shown on the left and the slices one through four of the simulation image on the right.
The size of the simulated image is 100x130x10. The training image consists of a 3D volume made of fluvial channels trending in the North-South direction. Conditioning data along vertical wells are assumed.

![Training Image](image)

![Simulation Images](image)

Figure 2. Simulation results for a 3D application of the multiple point stochastic algorithm

It is also to be noted that for this particular case, the best simulation images resulted with 2D templates. This is because the channels change in orientation from one layer of the training image to the next. This causes the 3D templates to introduce noise into the simulation. Despite the 2D templates used, the slices exhibit correlation from one slice to the next due to the profusion of conditioning data in the vertical direction (due to the presence of vertical wells).
3.2 GEOLICAL FEATURE IDENTIFICATION AND SIMULATION

3.2.1 SCANNING PROCESS

The purpose of the multiple point geostatistics approach is borrowing conditional proportions (joint probabilities) from a numerical representation of the spatial law (training image) describing the random function. Training images require some prior information. When this prior information is uncertain, alternative training images can be used to narrow the range of uncertainty and a Bayesian approach can be adopted to incorporate that uncertainty in resultant geological models (Liu et al., 2004). The training image requires the effort of geologists to generate one or more numerical representations of spatial heterogeneity, reflecting prior information about complex shapes and geometrical patterns.

In order to capture the relevant details of reservoir heterogeneity, the size of the training data as well as the size and complexity of the spatial template can be large. Consequently the process of scanning and storing multiple point conditional probabilities can have a rather high computational requirement. Taking advantage of recent advances in computational technology using multiple cpu processors, the task
of scanning training images is performed using multiple processing threads. The training image is segmented into multiple domains and the task of scanning each sub-domain using the specified spatial template(s) is assigned to a dedicated CPU thread. At the end of the scanning process, the statistics computed by each thread is compiled into a single statistical summary for the entire training image.

The objective of the scanning process is to obtain the joint probabilities:

$$\text{Prob}\{Z(\mathbf{u}_i) = z_k \forall i = 1, \ldots, N; k = 1, \ldots, K\}$$

(1)

$N$ is the number of nodes in the template, $\mathbf{u}_i$ is the position of the $i^{th}$ node, $z_k$ is the $k^{th}$ threshold of the random variable $Z(\mathbf{u})$. The scanning process is described by the following steps:

1. Select a cell location in the training image and superimpose the central node of the template at that location.
2. Capture the cell values corresponding to the template nodes.
3. Identify the pattern and store it. The pattern is stored in a long integer data type, $I$, by converting the $N$ individual template node values, $v_i$, to integers in the range [1-9] and looping over all the template nodes with the formula: $I_i = I_{i-1} + v_i \times 10^{(i-1)}$, for $i = 1, 2, \ldots, N$; and $I_0 = 0$.
4. Repeat steps 1 through 3 until the center of the template has been located in all training image cells excluding the marginal offset.
As the training image is being scanned, each obtained pattern integer representing a data event is put into an array that is large enough to hold all possible occurrences for that training image. The array is sorted and scanned for number of occurrences of each data event. The occurrences are stored in an alternate array at the same index as the first occurrence of the data event. Both arrays are then compressed by eliminating the repetitions and zeros.

### 3.2.2 SIMULATION PROCESS

During the simulation phase, the spatial template is placed at a location \( \mathbf{u}_o \). Some nodes \( \mathbf{u}_j, j = 1, \ldots, N' \) where \( N' \subset N \) is the subset of template nodes that already have simulated values. The requirement is of the probability of a simulation event \( \mathcal{A}(\mathbf{u}_o) \) conditional to the pattern \( \mathcal{A}(\mathbf{u}_j, j \in N') \) in the surrounding nodes of the template. The joint probability expression in Equation (1) can be used to calculate the requisite conditional probabilities:

\[
\text{Prob}\{\mathcal{A}(\mathbf{u}_o)\} = \frac{\text{Prob}\{\mathcal{A}(\mathbf{u}_o) \cup \mathcal{A}(\mathbf{u}_j)\}}{\text{Prob}\{\mathcal{A}(\mathbf{u}_j)\}}
\]

In terms of the notation in Equation (1), the numerator in Equation (2) is simply the joint probability: \( \text{Prob}\{\mathcal{A}(\mathbf{u}_i) = z_k \forall i \notin N'; \mathcal{A}(\mathbf{u}_j) = z_{k'} \forall j \in N'; k, k' = 1, \ldots, K \} \).
The denominator is the joint probability in the numerator summed over all occurrences of patterns $A(u_o)$:

$$\text{Prob}\{A(u_j)\} = \sum_{A(u_o)} \text{Prob}\{A(u_o) \cup A(u_j)\}$$

Knowing the conditional probability given by Equation (2), a simulation pattern can be obtained by drawing from the conditional probability distribution.

The simulation event $A(u_o)$ in traditional multiple point statistics implementations consists of a single point event i.e. the outcome at the central node of the template given the multiple point event in the surrounding nodes. In the implementation presented here, the simulation event is allowed to be a multiple point event. Thus both the conditioning and simulation events are multiple point events. This renders the simulation process to be fast and more important, the simulated patterns exhibit better continuity.

Following scanning, the process of stochastic simulation is commenced where the inferred multiple point histogram is used in conjunction with reservoir specific data distributions to obtain realizations with realistic spatial heterogeneity. In the traditional implementations of multiple point geostatistics, the simulation nodes are visited along a random path and the probability of the central node given the
configuration of pattern in the surrounding nodes is obtained from the search tree that contains a catalog of the scanned patterns. At the beginning of the simulation when only a few nodes have assigned values, considerable cpu time may be spent searching for unusual or infrequent patterns. In order to alleviate this problem, patterns or structures are grown from data locations in the method implemented in this study. This approach improves the continuity of the simulated patterns while reducing the artifacts in the simulated image.

Dividing the simulation domain into regions based on the density of conditional data, the conditioning data locations are visited along a random path. A node surrounding the data location is marked as “simulatable” based on whether the spatial template centered at that location contains at least a single conditional data. Corresponding to a randomly picked conditioning data location, a “simulatable” node in the vicinity of that data location is also randomly picked. The multiple point simulation event $A(u_o)$ is picked from the conditional probability distribution (Equation (2)) at that location. After all the conditioning data locations are visited, the list of “simulatable” nodes is updated and the next node is selected from this updated list. The simulation is continued until all the simulation nodes have been assigned a value. During the simulation process:

1. Select a cell as center cell of the spatial template
2. Detect the pattern $t(n)$ on the nodes of the spatial template by matching only non-negative values against the ones previously recorded during the scanning phase. Only when all these positive values and their positions are completely matched, then the record is chosen. Retrieve the subset of scanned patterns that exhibit the pattern $t(n)$.

3. Take the probability of each pattern in the subset and construct a step CDF with probability on y-axis and pattern index number on x-axis.

4. Randomly sample from that CDF.

5. Get the corresponding pattern by its index and fills the cells under the template that are marked as empty.

The conventional stochastic simulation algorithms can be rather slow and cumbersome since they only simulate one point at a time. The scanning process defined above allows the simulation algorithm employed here to simultaneously simulate all cells under the template. All simulation nodes except the conditioning data locations are assigned a negative one (-1) value initially.

### 3.3 TEMPLATE SELECTION ALGORITHM

The template plays the important role of an interface between the training image and the final conditional distributions used to reproduce the geological
features. A proper template should identify the most relevant geological features to be reproduced. The size and geometry of a suitable scanning template define the search neighborhood and the detection of multiple point data events describing the spatial law. No statistics beyond the size of the template will be captured from the training image.

A fast and robust algorithm to derive optimal spatial templates was developed, and is based on a semi-automated procedure where, given a number of nodes, an algorithm searches the training image to define the optimal locations around a central node that exhibit a higher correlation considering the spatial continuity/heterogeneity in different directions, thus identifying the optimal template configuration to better capture structures and patterns from the training image, i.e. geological features. Figure 3 shows some examples of optimal templates obtained with the Template Optimization Algorithm for training images exhibiting different patterns.

Consider a grid \( u \in (\Theta \subset \Omega) \) where \( \Theta \) is the size of the grid and \( \Omega \) is the size of the simulation domain. A stationary attribute \( Z(u) \) is assumed over the grid. Suppose a spatial template of size \( N \) is desired. The objective is to identify a template \( t(u_i), i = 1, \ldots, N \) within the grid \( \Theta \) such that the selected template optimally represents the dominant pattern of reservoir heterogeneity. The size and the scale of the template \( N \) are user-specified and are dependent on the available CPU, complexity.
of the geological image, etc. Designating the central node in the grid $\Theta$ as $\mathbf{u}_o$, the covariance $C(\mathbf{u}_j, \mathbf{u}_o)$ between pairs of nodes $\mathbf{u}_j \in \Theta : \mathbf{u}_j \neq \mathbf{u}_o$ and $\mathbf{u}_o$ is calculated on the basis of the particular training image. Under stationarity, the required covariance is calculated by translating a two-point template $\mathbf{h}_{j0} = |\mathbf{u}_j - \mathbf{u}_o|$ over the training image. The $\Theta - 1$ covariance values $C(\mathbf{u}_j, \mathbf{u}_o), j = 1, \ldots, \Theta - 1$ are ranked and the top $N$ values and the corresponding locations $\mathbf{u}_i, i = 1, \ldots, N$ define the optimal spatial template $\mathbf{t}(\mathbf{u})$. Spatial templates at multiple scales can be obtained by choosing the grid $\Theta$ of different resolution (multiple grids), accounting for geological patterns at different scales.

Training Image

Template

Figure 3. Examples of optimal templates obtained with the Template Algorithm
3.3.1 PRELIMINARY STUDIES OF TEMPLATE SELECTION

Results of a Preliminary Sensibility Study on the Impact of template selection in multiple-point statistics captured during the scanning process, demonstrate the importance of an appropriate template selection for capturing more information about the multiple point statistics behind a particular geological feature. Hence, by defining a template that identifies the relevant size and geometry of the geological structures, the amount of conditional information retained from the training image during the scanning process is increased; improving the reproduction of the desired features and also optimizing the use of computational recourses by constraining the template configuration to the required specifications.

3.3.1.1 Impact of Template Geometry

In this study, a training image exhibiting ellipsoidal lens with North-South orientation (Figure 4) was generated and scanned with six templates of the same size but different geometry. All the templates were produced by the template selection algorithm, considering different numerical representations of lens with the same size but different orientations.
The multiple point statistics inference process was performed using a variety of spatial templates. It can be noticed in Figure 5, enhanced performance is obtained corresponding to Template 1, which has a North-South orientation. This template captured more relevant data events (with frequency higher than 10) from the training image (approximately 11% more than the second best Template) and consequently, the total number of occurrences retained as multiple point conditional information is increased. The difference in configuration between Templates 2 and 3 is the location of a single node, however this single node causes the number of relevant data events to drop in approximately 11%. Hence, by defining a template with a geometry that identifies the geological structures, the amount of conditional information retained during the scanning process is increased, improving the reproduction of the desired features.
3.3.1.2 Impact of Template Size.

In the second part of this study, a training image exhibiting ellipsoidal lens with Northeast-Southwest orientation was generated and scanned with nine templates of different size. All the templates were produced by the template selection algorithm considering different number of nodes in the template (from 6 to 14 nodes).

Figure 6 is useful to analyze the combined effect of the template and the training image size on the amount of conditional information retained during the
scanning process. On Figure 6 it can be notice that the number of relevant data events captured from the training image increases initially and then decreases, as the size of the template (number of nodes) increases. On the other side, the total number of occurrences always decreases with an increasing template size. The initial increment in the number of relevant data events is due to the increasing number of possible combinations caused by the increasing number of nodes in the template. This increment continues until the size of the training image starts limiting the number of possible outcomes considering the size of the template, i.e. the training image is not big enough to contain all the possible combinations for a given template size.

![Impact of Template Size on Captured Pattern Statistics](image)

**Figure 6.** Impact of template size on captured pattern statistics
4 DYNAMIC DATA INTEGRATION

Reservoir models are commonly generated in two steps. First, a geological model is obtained by the integration of static geological information from different sources such as seismic, well logging, core analysis, and sequence stratigraphy, among others. Then, during the process known as history matching or dynamic data integration, the model is adjusted or modified in order to match the flow response of the model with the production history.

Reproduction of historical production data is usually the most important goal during the model validation process. Reconciliation between the model prediction and the field production records however represents a rather significant challenge, considering the highly non-linear relationship between the model parameters (such as permeability, porosity etc.) and the production response. Consequently, the process requires an iterative approach that translates into higher computational effort. In spite of intense research activity in this area and significant achievements during the last decade, the number of relevant realistic field applications of assisted history matching is still rather limited and given the complexity of the problem, there are still several issues to be addressed.
Multiple approaches have been considered for conditioning reservoir models to dynamic flow data. Trial and error optimization algorithms follow an iterative process to update the reservoir model applying a perturbation scheme for converging to a set of model parameters that yield a match to the observed flow response. Some of the most popular methods include gradient based methods (Yeh, 1986; Anterion, F. et al., 1989), pilot point methods (La Venue et al., 1992; de Marsily et al., 1995), sequential self-calibration methods (Gomez-Hernandez et al., 1997), Markov chain Monte Carlo methods (Omre and Tjelmeland, 1996) and gradual deformation methods (Roggero and Hu, 1998). There are other geostatistical methods that presuppose that the permeability fields could be conditioned to both hard data as well as dynamic data in a probabilistic sense by using a Bayesian formulation for the integration.

The subtlety of permeability variations in a reservoir influence the flow of fluids in the reservoir thereby influencing the flow response recorded at the wells. This led us to one of the motivations behind the selection of a probabilistic approach for dynamic data integrations that integrates information from the geological model of heterogeneity. Progress on the validation and testing of the probabilistic approach for incorporating production data in reservoir models, is discussed below.
4.1 GEOSTATISTICS OVERVIEW

Reservoir models are generally constructed considering subsurface geological data obtained from different sources (such as seismic, well logging, well tests, sequence stratigraphy, etc), and a geological model of heterogeneity. The prior geological model for heterogeneity is commonly represented in the form of a variogram model that is inferred from the same conditional subsurface information. These two components are combined within a simulation/interpolation framework to generate geological models conditioned to static information.

Geostatistics as a geological modeling technique and its two original basis: i) the variogram model and ii) the kriging interpolation methodology, were initiated by the work of Daniel Krige (1951) and later developed by Georges Matheron (1962-1963, 1965), with the purpose of providing locally accurate grade estimates of mining blocks; however, its application has extended from the mining industry to many other related disciplines, including the oil industry.

The simple kriging (SK) estimator $k^*_s$ at each location $u_j$ of the target geological model $k(u)$ (such permeability or porosity field) is the best linear unbiased estimator and can be written as:
\[ k_{sk}^* - m(u_j) = \sum_{i=1}^{N} \lambda_i(u_j)[k(u_i) - m(u_i)] \] (4)

where, \( m(u_j) = E\{k(u_j)\}, j = 1, \ldots, J \), are the known stationary means of the random function \( k(u_j) \) at the locations \( u_j \); \( J \) is the size of the model; \( k(u_i), i = 1, \ldots, N \) are the conditional data; and \( \lambda_i(u_j) \) are the kriging weights for each conditional data at each location for the estimation at location \( u_j \). The weights are calculated from the following system of equations:

\[ \sum_{k=1}^{N} \lambda_k(u_j) C(h_k) = C(h_{ij}), \forall i = 1, \ldots, N \] (5)

where \( C(h_{ij}) \) or \( C(u_i - u_j) \) is the covariance at the lag \( h_k = u_i - u_j \), considering stationarity; and is related to the variogram by: \( \gamma(h_j) = C(0) - C(h_j) \) and \( C(0) = Var\{k(u_i)\} \). The corresponding minimum estimation (error) variance \( \sigma_{sk}^2 \) is:

\[ \sigma_{sk}^2(u_j) = C(0) - \sum_{i=1}^{N} \lambda_i(u_j)C(h_{ij}) \] (6)

Stochastic simulation was introduced by Matheron (1973) and Journel (1974) to correct for the smoothing effects and other artifacts of kriging (See Figure 7a.), and to enable the reproduction of the spatial variance predicted by the variogram model. Different algorithms have been developed including sequential simulation (Journel,
1983, Isaaks, 1990; Srivastava, 1992; Goovaerts, 1997; Chiles and Delfiner, 1999), which has become the workhorse for many current geostatistical applications.

The stochastic simulation approach is based on the calculation of probability distributions at individual locations, considering the conditional information and the spatial heterogeneity model. There are different methods for the construction of the local probability distributions, including the Gaussian approach where the kriging estimation and the estimation variance are used as the mean and the variance of the local normal conditional distribution. In another approach, the use of indicator transforms allows modeling multivariate distributions without relying on Gaussian assumption, to generate models that exhibit more connected and well-defined geological bodies. Figure 7 compares the results of the original kriging interpolation technique (a) with the Gaussian (b) and the Indicator Sequential Simulations (c), considering the same conditional information.

![Figure 7](image.png)

Figure 7. Spatial interpolation obtained by: a) Kriging; b) sequential Gaussian simulation and c) sequential indicator simulation.
Stochastic simulation also provides the capability to generate multiple equiprobable realizations, giving birth to the idea of assessing spatial uncertainty (Journel and Huijbregts, 1978) on reservoir models.

4.2 SEQUENTIAL INDICATOR SIMULATION

Spatial distributions can be modeled following a non-parametric approach, where the local probability distributions $F(u; z_i)$ can be calculated from the available conditional information, by defining a set of thresholds $z_i, i = 1,..., NT$ to discretize the range of variability of the spatial variable, and subsequently performing indicator kriging using the indicator transformed variables.

The indicator transform of a random variable is simply a binary transform: the value one is assigned if the value at a location is less than the threshold and zero if not. The expected value of an indicator random variable is therefore equivalent to the probability of that particular threshold. Hence, the probability distribution can be calculated by sequentially calculating the expected value of the indicator random variable corresponding to different thresholds. A multi-valued indicator variable can be defined as:

$$I(u, z_i) = \begin{cases} i & \text{if } k(u) \leq z_i \quad \forall i = 1,..., NT \\ NT + 1 & \text{if } k(u) > z_i \end{cases}$$

(7)
Hence the indicator transform discretizes a continuous variable (such as permeability) into classes or categories. The expected value of the indicator corresponding to a particular category is:

\[ E \{ I(u, z_j) \} = \text{Prob} \{ k(u) \leq z_j \} = F_k(z_j) \]  

(8)

The indicator coded data is used to infer the experimental variogram at each threshold, allowing the usage of different heterogeneity models (variograms) for different thresholds. Then, at a particular location, the conditional expectation of the indicator random function for each threshold is determined by applying indicator kriging with the available indicator coded conditional information.

\[ I^* \left( u_\alpha ; z_i \right | (n)) = F^*_k \left( z_i \right | (n)) = \sum_{i=1}^{n} \lambda_i(u) i(u_\alpha ; z_i) \]  

(9)

where \( i(u_\alpha ; z_i) \) is the indicator coded data at location \( u_\alpha \), with \( n \) conditional data; and the weights \( \lambda_i(u) \) are obtained by solving a kriging system that utilizes indicator covariances:

\[ \sum_{\beta=1}^{n} \lambda_\beta(u, z_i) C_i(h_{\alpha\beta}; z_i) = C_i(h_{\alpha\alpha}; z_i), \forall \alpha = 1, ..., n \]  

(10)

The probabilities (conditional expectations) for the local conditional distributions are evaluated at a limited set of thresholds. Therefore, interpolation and extrapolation methods are required to obtain a continuous conditional cumulative
distribution function. Interpolation between the thresholds and tail extrapolations can be obtained by applying different approaches such linear or hyperbolic interpolation/extrapolation or using tabulated values.

Following the sequential simulation approach, a realization of the target reservoir model is generated by sequentially sampling from the local conditional distributions following a random path, where the previously sampled values become conditional information for the construction of subsequent local conditional distributions. The process is repeated until all the uninformed locations in the model are populated. Monte Carlo or other sampling technique can be applied on the local conditional distributions. Multiple realizations of the target reservoir can be obtained by altering the random path and/or changing the random draw from the local conditional distributions.

The application of the Indicator sequential simulation approach can be summarized by the following:

1. Select appropriate thresholds consistent with the spatial phenomena.
2. Indicator code the data corresponding to different thresholds.
3. Infer indicator variogram/covariance model(s) for different thresholds.
4. Define a random path to visit all uninformed locations. On each subsequent location of the random path apply the following sub-procedure:

4.1. Calculate the conditional expectation of the indicator random function for all the thresholds, applying indicator kriging with the available indicator coded conditional information.

4.2. Correct for order relations (non-monotonicity of the distributions) on evaluated probabilities (conditional expectations).

4.3. Randomly sample a value from the local conditional distribution. In the sampling process, use interpolation/extrapolation methods to model a continuous ccdf from the discrete probabilities evaluated at the thresholds.

4.4. Include the sampled value in the list of conditional information for subsequent estimations.

5. A single realization of the target reservoir is obtained after all the uninformed locations have been visited following the random path. To generate multiple realization repeat step 4 with different random paths.

An indicator sequential simulator has been implemented on C++ language, and validated with other algorithms available on public domain. This algorithm is the base code for the probability updating method utilized in the research project. In subsequent sections, additions including modules for gradual deformation of
geological models, interface with flow simulators, and optimization schemes are also developed.

### 4.3 GRADUAL DEFORMATION OF GEOLOGICAL MODELS USING DYNAMIC DATA

Honoring the geological model is an important objective during the generation of static geological models; however, it is commonly forgotten during the integration of dynamic information. During the final stages of reservoir modeling, the history matching process, the perturbations or modification in the model should be performed to ensure a match to the flow history, while preserving the geological model of heterogeneity. In this research, that goal is accomplished by applying a probabilistic approach for gradual deformation of geological models. The gradual deformation is obtained by systematically perturbing the local conditional distributions with a deformation parameter, $r_d$ that is calibrated using the available dynamic information.

#### 4.3.1 PERTURBATION OF LOCAL CONDITIONAL DISTRIBUTIONS

The gradual deformation starts with a particular realization of the target reservoir, such that all locations in the reservoir, cells or nodes have attribute values.
In the case of continuous variables, the initial realization is transformed into an indicator random field such that the value at a particular location falls within an indicator category referred to as the initial class, \( z_I \). The initial realization is obtained by following a particular random path through the reservoir model, sequentially generating the local conditional distribution at each location conditioned to the original data and previously simulated values and randomly sampling values from the constructed distributions.

During the gradual deformation of the geological model, the local conditional distributions are perturbed using a deformation parameter \( r_d \). The optimum value of \( r_d \) is calibrated in such a fashion as to minimize the deviation from the observed production history. In addition, the random path along which the visited as well as the random draws from local conditional distributions are modified in order to search for a global optimum for \( r_d \). Different perturbation schemes for establishing the optimal \( r_d \) have been evaluated in order to define the methodology that better fits the objectives of the research.

In the first perturbation scheme, the deformation parameter, \( r_d \), reduces the probabilities of all indicator categories in the local conditional distribution, except that of the initial class, \( z_I \), which is proportionally increased. In this case the deformation parameter multiplies the probabilities of the off-class indicator categories.
and the probability of the initial class always increases (or remains the same for \( r_d = 1 \)). This perturbation can be written in terms of conditional probabilities as:

\[
F_i^\prime (z_i | (n)) = \begin{cases} 
 r_D \cdot F_k (z_i | (n)) & i \neq I \\
 1 - \sum_{j=1}^{NT+1} r_D \cdot F_k (z_j | (n)) & i = I 
\end{cases}
\]  

(11)

where \( F_i^\prime (z_i | (n)) \) is the perturbed local conditional probability. A deformation parameter of value zero will generate a distribution with probability 1 for the initial class, \( z_i \), ensuring the reproduction of the initial realization. A deformation parameter of value one will recover the local distribution conditioned to geological information and hence an independent realization from the original conditional distribution \( F_k (z_i | (n)) \) is sampled. Since the deformation parameter only increases the probability of the initial class, the maximum deformation of the model is rather small, slowing the process of gradual deformation. This drawback encouraged the evaluation of other perturbation schemes.

Figure 8 shows an example of gradual deformation of geological models by probability perturbation method under the first scheme. A single parameter, \( r_d \) determines the magnitude of the perturbation in this model.
Figure 8. Example of gradual deformation of geological models by probability perturbation method

In the second perturbation scheme, the deformation parameter reduces the probability of the initial class in the local distributions, while the probabilities of the other indicator categories increase proportionally. In this case the probability of the initial class is always reduced (or remains the same for \( r_d = 1 \)). This perturbation scheme can be written as:

\[
F'_k(z_i \mid n) = \begin{cases} 
  r_d \cdot F_k(z_i \mid n) & i = I \\
  \frac{F_k(z_i \mid n)(1 - r_d \cdot F_k(z_I \mid n))}{\sum_{j=1 \atop j \neq I}^{N_T} F_k(z_j \mid n)} & i \neq I
\end{cases}
\]

(12)

The perturbed conditional probability for \( i \neq I \) is written such that the perturbed probability for a particular class is scaled according to its initial value. A
deformation parameter value of zero will generate a distribution with probability 0 for the initial class, producing a notable deformation of the geological model. A deformation parameter of value one will recover the local distribution conditioned to geological information. Now, the deformation parameter only decreases the probability of the initial class, ensuring a large deformation of the geological model, speeding the process of gradual deformation. However, this perturbation scheme does not allow the reproduction of the initial realization and consequently the deformation process is no longer systematic and controlled.

In the third perturbation scheme, the two previous schemes were combined to ensure a more controlled, but at the same time fast gradual deformation of the geological model. In this case the probability of the initial class has a range of variation from 1 to 0 (for $r_d = 0$ to $r_d = 1$), ensuring the reproduction of the initial realization for $r_d = 0.0$, the recovery of the distribution conditioned to geological information for $r_d = 0.5$, and rejecting the initial class, generating a new realization and corresponding large deformation for $r_d = 1.0$. In this approach the range of variation of the deformation parameter, $r_d$, is divided in two intervals, values below and above 0.5. For values of $r_d$ below 0.5 the first perturbation scheme is applied using a transformed value of $r_d$, $r'_d = 2 \times r_d$. For $r_d$ values above 0.5, the second perturbation scheme is applied with the transformed value of $r_d$, $r'_d = (2 - 2 \times r_d)$. The perturbation scheme can be written as:
\[ r_D' = \begin{cases} 
2r_D & r_D \leq 0.5 \\
2 - 2r_D & r_D > 0.5 
\end{cases} \tag{13} \]

\[
F_k'(z_i | (n)) = \begin{cases} 
\frac{r_D' \cdot F_k(z_i | (n))}{\sum_{j=1}^{N+1} F_k'(z_j | (n))} & i \neq I; r_D \leq 0.5 \\
1 - \sum_{j=1}^{N+1} r_D \cdot F_k(z_j | (n)) & i = I; r_D \leq 0.5 \\
\frac{F_k(z_i | (n)) \left(1 - r_D' \cdot F_k(z_i | (n))\right)}{\sum_{j=1}^{N+1} F_k(z_j | (n))} & i \neq I; r_D > 0.5 
\end{cases} \tag{14} 
\]

Consequently, the probability of the initial class is increased from the original value for \( r_D \) values below 0.5; and decreased for \( r_D \) values above 0.5, ensuring a wider but controlled range of variation in the perturbation of the local distribution. Figure 9 shows the effect of the deformation parameter \( r_D \) on the local conditional distribution under the perturbation scheme 3. For \( r_D = 0.5 \), the local distribution conditioned to geological information is recovered.

The above three perturbation schemes focus on gradually transitioning an initial realization to a new realization or proposal. Multiple equiprobable proposals can be generated using sequential indicator simulation, simply by modifying the random path and/or the set of random drawings for the local conditional distributions. However, in the three previous schemes, starting from an initial realization, the transition to only one new realization is evaluated at a time. This implies that only the
range of variability between these two limiting realizations is searched for an optimal value of $r_d$. This limits the range of variation during the gradual deformation of the geological model. A fourth perturbation scheme is proposed to circumvent this drawback.

![EFFECT OF CALIBRATING PARAMETER RD ON LOCAL DISTRIBUTION](image)

Figure 9: Effect of deformation parameter on local distribution under 3rd perturbation scheme.

In the fourth perturbation scheme the transition to two different new realizations are evaluated at the same time. In this case the perturbation introduced by the deformation parameter in the local distributions is controlled in order to generate a gradual transition between the initial realization and two different equiprobable
realizations obtained with different random paths and different sampling draws. In this scheme the initial realization is reproduced for \( r_d = 0.5 \) (probability of the initial class is 1), and the distributions conditioned to geological information corresponding to the two proposals are recovered for \( r_d \) values of 0.0 and 1.0. Consequently, in this approach the range of variation of the deformation parameter, \( r_d \), is divided in two intervals, values below and above 0.5. The transitions between the initial realization and the first and second proposals are obtained with values of \( r_d \) below and above 0.5 respectively. These transitions follow the first perturbation scheme using a transformed value of \( r_d \), \( r_d' = 1 - 2 r_d \), for \( r_d \leq 0.5 \); and \( r_d' = 2 r_d - 1 \), for \( r_d > 0.5 \). The perturbation scheme can be written as:

\[
r_d' = \begin{cases} 
1 - 2 r_d & r_d \leq 0.5 \\
2 r_d - 1 & r_d > 0.5 
\end{cases}
\]  

\[ (15) \]

\[
F'_k (z_i | (n)) = \begin{cases} 
1 - \sum_{j=1}^{N+1} r_d' \cdot F_k^a (z_j | (n)) & i \neq I; \quad r_d \leq 0.5 \\
\sum_{j=1}^{N+1} r_d' \cdot F_k^b (z_j | (n)) & i = I; \quad r_d \leq 0.5 \\
1 - \sum_{j=1}^{N+1} r_d' \cdot F_k^b (z_j | (n)) & i \neq I; \quad r_d > 0.5 \\
\sum_{j=1}^{N+1} r_d' \cdot F_k^a (z_j | (n)) & i = I; \quad r_d > 0.5 
\end{cases}
\]  

\[ (16) \]

Where \( F_k^a \) and \( F_k^b \) are the probabilities corresponding to the two proposals \( a \) and \( b \).
The fourth scheme compared to the first three, evaluates twice the number of proposals, increasing the rate of convergence and reducing the probability of getting trapped on local minima. The following Figure shows an example of gradual deformation of geological models by probability perturbation under scheme four. A single parameter, rd determines the transition between an initial realization and two proposals. The range of variation of the geological model is enlarged under this scheme.

Figure 10. Example of gradual deformation of geological models by probability perturbation scheme four.
4.3.2 CONDITIONING OF LOCAL DISTRIBUTIONS TO DYNAMIC INFORMATION

Now that the methodology for the perturbation of local conditional distributions in geological models has been defined, the next step is to calibrate the deformation parameter on the basis of the observed production history of the reservoir. A primary goal is to estimate the local probability distributions of the geological event A, (permeability, porosity, etc) conditioned to the dynamic information, C, i.e. P(A|C). This requires the implementation of an optimization scheme to calibrate the deformation parameter, and the development of interfaces between the geological modeling algorithm and a flow simulator. Figures 11 and 12 show the impact of the deformation parameter on the flow response during an application of the probability perturbation method for the gradual deformation of geological models.

The deformation parameter $r_D$ is calibrated using the Dekker-Brent iterative optimization procedure where the objective is to improve the fit between the flow response of the model (from the simulator) and the production history. The Dekker-Brent algorithm is an inverse parabolic interpolation method that has the advantage of being a non-gradient based approach that only requires the calculation of the objective function corresponding to different values of the deformation parameter.
The algorithm yields an optimal value of the deformation parameter, $r_d^*$ (abscissa), corresponding to a minimum value of the objective function, $\Delta O = f(r_d^*)$ (ordinate). Three abscissa values are required, $a$, $b$ and $c$ with the corresponding values of the objective function $f(a)$, $f(b)$ and $f(c)$; and $b$ chosen such that $a < b < c$ and $f(a) > f(b) < f(c)$. The estimated location of the abscissa ($r_d^*$) with the apparent minimum ordinate (objective function) is calculated by fitting a parabola through these three points.

$$r_d^* = \frac{(b+c)f(a) + (a+c)f(b) + (a+b)f(c)}{2\left(\frac{f(a)}{(a-b)(a-c)} + \frac{f(b)}{(b-a)(b-c)} + \frac{f(c)}{(c-a)(c-b)}\right)}$$

(17)

Figure 11: Effect of the deformation parameter on the flow response of a geological model.
Figure 12. Effect of the rD parameter in the water saturation distribution in a reservoir after 4000 days and 5200 days

The next figure illustrates the process of convergence of an objective function for a single-parameter problem using the Dekker-Brent inverse parabolic interpolation. An inverse parabola (red) is fitted through the initial values of rd a, b and c and the correspondent objective functions for these values, f(a), f(b) and f(c), resulting in the first apparent optimal value of rd = x1. Then, the real objective function correspondent to x1, f(x1) is calculated and based on the result, the three points for the next parabolic interpolation (green) are selected (points a, b and x1).
This iterative procedure continues until a minimum in the objective function is reached.

![Graph illustrating the Dekker-Brent inverse parabolic interpolation process.](image)

**Figure 13:** Illustration of the Dekker-Brent inverse parabolic interpolation process.

In this case, the objective function to be minimized is a measurement of the deviation between the simulated production response and the production history. Different production variables can be included in the objective function, including field and well pressures, single phase rates, two phase ratios, and basically any other variable for which a historical record is available. The mismatches of individual variables are normalized in order to level their influence on the objective function. However, in some cases it might be useful to assign higher weights to some production variables in order to emphasize the relevance of their reproduction in the
target model. The proposed objective function for $N$ production variables over $T$ time steps is:

$$\text{Obj Fun} = \sum_{i=1}^{N} \sum_{t=1}^{T} \frac{(\text{Sim}_{i,t} - \text{Hist}_{i,t})^2}{\text{Var}(\text{Hist}_i)}$$

(18)

Where $\text{Sim}_{i,t}$ represents the simulated value of the production variable $i$ at time $t$, and $\text{Hist}_{i,t}$ represents the corresponding historical value for the same variable at the same time. The square of the difference between the simulated and the historical values at a particular time is normalized by the variance of the historical values over time, in order to control the influence of each variable on the objective function. Other objective functions can be easily implemented applying different norms and normalization methods.

4.3.3 MERGING THE INFORMATION FROM DYNAMIC AND STATIC SOURCES

At this point, the methodology to estimate the local distribution conditioned to dynamic information, $P(A|C)$, has been presented, i.e. perturbing the local conditional distributions with a parameter calibrated with the production history data. In order to ensure consistency with the geological model during all stages of the history matching procedure, the conditional distribution $P(A|C)$ has to be merged with the distribution inferred from static information, $P(A|B)$. Realizations sampled from the
resultant merged distribution \( P(A|B,C) \) will honor both the static information (well data and geologic interpretation) as well as the historic production data.

The Permanence of Ratio Hypothesis (Journel, 2002) is the methodology applied to combine the individual distributions conditioned to dynamic and static information. The following distance or information measures \( a, b, c \) and \( x \) are defined:

\[
\begin{align*}
    a &= 1 - \frac{P(A)}{P(A)} \\
    b &= 1 - \frac{P(A|B)}{P(A|B)} \\
    c &= 1 - \frac{P(A|C)}{P(A|C)} \\
    x &= 1 - \frac{P(A|B,C)}{P(A|B,C)}
\end{align*}
\]  

The quantity \( a \), for example, denotes the relative distance to the event \( A \) occurring given \( P(A) \). If \( P(A) \) is one, the relative distance \( a \) is zero. The relative distance \( a \) is infinity, if \( P(A) \) is zero. The measures \( b, c \) and \( x \) can be interpreted similarly.

According to the permanence of ratios hypothesis, the relative updating of a simulation event \( (A) \) due to a dynamic event \( (C) \) remains the same irrespective of the presence of the static event \( (B) \). This can be written in terms of \( a, b, c \) and \( x \) as:

\[
\frac{x}{b} = \frac{c}{a}
\]  

Consequently, the joint probability of the simulation event given the dynamic and static information can be calculated from the elemental probabilities – the prior
probability for A: P(A), the conditional probability of A given the dynamic information: P(A|C), and the conditional probability of A given the static information B: P(A|B).

\[ P(A | B, C) = \frac{a}{a + bc} \]  

(21)

In this approach, the marginal probabilities of the static data P(B), and that of the dynamic data P(C), which are difficult to estimate in practice, are not required.

### 4.4 GRADUAL UPDATING PROCEDURE FOR HISTORY MATCHING

The calibration of the dynamic parameter \( r_D \) and the subsequent merging of conditional probability distributions requires the implementation of an interface between the geological modeling algorithm and the flow simulator. In this project, the geological modeling algorithm (Sequential Indicator Simulation) is the main program that has been expanded to include all the tasks in the probabilistic approach for dynamic data integration. The flow simulator (@Eclipse) is also executed within the program. The task of combining the conditional probability distributions \( P(A|B) \) and \( P(A|C) \) into a joint-conditional distribution \( P(A|B,C) \) is also implemented within the main program. The geological model used in the flow simulator to calibrate the deformation parameter \( r_D \) is sampled from this jointly conditioned distribution, \( P(A|B,C) \). The complete probability updating procedure therefore consists of:
Performing indicator kriging and obtaining the conditional distributions $P(A|B)$ at each unsampled location in the reservoir. Sample a realization from the $P(A|B)$ by sequential simulation.

Corresponding to that realization of the permeability field and making an initial guess for $r_D$ obtain the corresponding $P(A|C)$.

Merge $P(A|C)$ with $P(A|B)$ to obtain $P(A|B,C)$ and sample a realization from $P(A|B,C)$.

Perform flow simulation and obtain objective function. Revise estimate of $r_D$ using Dekker-Brent approach and repeat until objective function is minimized.

The reproduction of historical production data is a complex non-linear inverse problem. This implies that the probability updating cannot be accomplished in a single perturbation loop starting from an initial realization of the permeability field; calibrating an optimal $r_D$ value and obtaining an updated probability distribution reflecting the dynamic characteristics of the reservoir. Instead, a multi-loop iterative process is required to update the geological model using the dynamic data.

A Markov-Chain is a stochastic updating procedure where the parameter state at any step of the procedure is assumed to be dependent only on the state immediately...
prior to that step. Thus the proposed realization at any stage of the process depends only on the preceding realization in the sequence, and the convergence towards the desired target depends on carefully specifying the transition from one realization to the next one, i.e. the methodology for the new proposed realization. In this case, the parameter $r_D$ controls the transition of the permeability value at a location from one category to the next.

In the implemented Markov chain approach, at every updating step, from iteration step $l$ to step $l+1$ of an outer loop, the probability distributions conditioned to dynamic and static information, $P(A|B,C)^l$, is obtained by applying the permanence of ratio hypothesis to combine distributions conditioned to static and dynamic information. The distribution $P(A|B)^l$ is obtained from geological data and heterogeneity model and the distribution conditioned to dynamic information, $P(A|C)^l$ is estimated knowing the indicator category at each location from the realization sampled from $P(A|B,C)^\ell$, the prior distribution $P(A)$ and the deformation parameter, $r_D$, calibrated using the Dekker-Brent iterative optimization procedure. At the end of each inner Dekker-Brent loop the converged distribution $P(A|B,C)^l$ is used to sample the initial realization for the next outer loop, $l+1$, until global match to the historic data is attained. The calibration of the deformation parameter $r_D$ to honor the available dynamic information thus represents the internal
optimization scheme. The converged model and realization at the end of the inner
loop is used as the starting realization for the next sequence of inner Dekker-Brent
optimization runs to determine the conditional distribution $P(A|C)$ and that constitutes
the outer optimization loop.

Even though the two-loop Markov chain procedure ensures global
convergence, the introduction of multiple sets of inner optimization schemes requires
multiple evaluations of the flow response. However, the dynamic-calibrated gradual
deformation methodology renders the history match process faster and more
controlled, increasing the consistency between the initial and the proposed
realizations at every step, and improving the rate of convergence of the objective
function.

The implemented interface between the geological modeling algorithm and
the flow simulator is summarized in the following steps:

1. The geological modeling algorithm generates a file with a realization of
the target permeability model in the appropriate format. This file will be
used by the flow simulator as an include file.
2. The modeling algorithm invokes the flow simulator. The flow simulator
(ECLIPSE®) is run, generating an output file with the flow response.
3. The simulated flow response is read from the output file and the objective function is evaluated.

4. The Dekker-Brent optimization is implemented in order to generate a new realization of the geological model.

4.5 APPLICATION

An approach that uses a probability perturbation method for gradual deformation of geological models conditioned to dynamic information has been presented. This approach, compared to other perturbation methods, offers the important advantages of preserving the prior geological heterogeneity model and simplifying the history match process to a single (or few) parameter(s) optimization problem.

Preliminary evaluations of the algorithm for dynamic data integration using the proposed probabilistic approach have been pursued with the synthetic case described in the Table 1. Figures 14-18 present the results of the history match for the study case. Results after forty flow simulation runs, distributed on 5 outer iterations of 8 inner Dekker-Brent iterations, are presented. Figures 14 and 15 show the field pressure and production history match obtained with the probabilistic dynamic data integration algorithm for the simulation case. Figure 16 show the gradual deformation
of the third layer of the geological model through the history matching process.

Figure 17 shows the third layer of the reference geological model used to generate the production history. Figure 18 shows the convergence of the objective function in the study case. During the convergence, the relatively small increments observed in the objective function help to avoid falling into a local minimum.

<table>
<thead>
<tr>
<th>SIMULATION PROPERTY/DESCRIPTION</th>
<th>VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation Model</td>
<td>Black Oil</td>
</tr>
<tr>
<td>Solution</td>
<td>Implicit</td>
</tr>
<tr>
<td>Simulation Period, years</td>
<td>2</td>
</tr>
<tr>
<td>Grid (Cartesian)</td>
<td>50x50x5</td>
</tr>
<tr>
<td>Active Grid blocks</td>
<td>12500</td>
</tr>
<tr>
<td>Grid block dimensions, ft³</td>
<td>80x80x4</td>
</tr>
<tr>
<td>Porosity</td>
<td>0.22</td>
</tr>
<tr>
<td>Kx = Ky (Mean – Std Dev), md</td>
<td>200 - 250</td>
</tr>
<tr>
<td>Kz/Kx</td>
<td>0.15</td>
</tr>
<tr>
<td>Saturation Pressure, psi</td>
<td>5064</td>
</tr>
<tr>
<td>Water-Oil Contact, ft</td>
<td>9000</td>
</tr>
<tr>
<td>Gas-Oil Contact, ft</td>
<td>4000</td>
</tr>
<tr>
<td>Reference Depth, ft</td>
<td>7300</td>
</tr>
<tr>
<td>Initial Pressure @ 7300 ft, psi</td>
<td>6000</td>
</tr>
<tr>
<td>Residual Water Saturation</td>
<td>0.18</td>
</tr>
<tr>
<td>Residual Oil Saturation</td>
<td>0.24</td>
</tr>
<tr>
<td>Oil Relative Permeability Endpoint</td>
<td>0.7</td>
</tr>
<tr>
<td>Water Relative Permeability Endpoint</td>
<td>0.5</td>
</tr>
<tr>
<td>Oil Gravity (API)</td>
<td>35</td>
</tr>
<tr>
<td>Water Injectors</td>
<td>1</td>
</tr>
<tr>
<td>Injection – Control Rate, Stb/day</td>
<td>5000</td>
</tr>
<tr>
<td>Injection – BHP upper Limit, psi</td>
<td>8000</td>
</tr>
<tr>
<td>Oil Producers</td>
<td>2</td>
</tr>
<tr>
<td>Production – Control BHP Lower limit, psi</td>
<td>2000</td>
</tr>
<tr>
<td>Production – Minimum rate, Stb/day</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 1. Description of a simulation case for the evaluation the history matching algorithm.
Figure 14: Field pressure history match for the study case.

Figure 15. Field production history match for the study case
Figure 16. Gradual deformation of the geological model (third layer) in the study case.

Figure 17. Third layer of the reference geological model in the study case.
Figure 18. Convergence of the objective function in the study case.

The implemented computer code for history matching can be described in the following steps:

1. An indicator sequential simulator is used to generate an initial stochastic realization of the target reservoir model and calculate the local probability distributions conditioned to static information.

2. A Markov chain iterative updating process is started with the initial realization. The Markov chain forms the outer loop of the procedure and every outer step or outer iteration includes the following sub procedures:

2.1. Generate new random paths and sets of random draws to sample from the local conditional distributions. The random paths and the sampling draws
are fixed during each outer iteration, but changes from one outer iteration to the next.

2.2. Evaluate the Objective function at different values of the deformation parameter, $r_d$, screening the whole range of variability [0, 1]. Usually 5 different values are enough to get started. For each value of $r_d$, a different permeability model is obtained by merging with $P(A|B)$ and that model is run in the flow simulator; in order to obtain the objective function.

2.3. Pick the value of the deformation parameter with the minimum objective function and start the calibration process of $r_d$ with the dynamic data using the Dekker-Brent iterative algorithm. This calibration process is called the inner loop, and the number of inner steps or inner iterations can be fixed or controlled by a tolerance in the change of the objective function in consecutive steps.

2.4. Use the best model (with the minimum objective function) to update the stochastic realization. When the best model is obtained with a deformation parameter of 0.5, no updating is required (the realization remains invariant). This is the final step of the inner loop.

3. Repeat step 2 (outer loop) until a tolerance in the objective function (history match) has been reached or for a fix number of outer iterations.

4. Print out final permeability realization with the corresponding flow response.
5 CONCLUSIONS AND RECOMMENDATIONS

A 3D multiple point stochastic algorithm to model connectivity patterns representing geological features such fractures, channels and facies distributions, has been implemented on an efficient multi-threaded computational environment. This algorithm uses a scanning template to capture the proper conditional information (multiple-point statistics) from a training image or numerical representation of the spatial law, applying a scanning technique to store the conditional information during the scanning process and retrieve it during the simulation.

A unique stochastic simulation approach based on growth of objects within the simulation domain is presented. The object growth is controlled by the multiple point statistics inferred on training images. In order to render the simulation computationally efficient, the process is implemented on a multi-threaded environment. The scanning as well as the simulation processes both take advantage of multiple cpus.

Compared to traditional multiple point statistics applications, where the simulation event consists of a single point event, the implemented algorithm allows the simulation event to be a multiple point event. Thus both the conditioning and
simulation events are multiple point events. This renders the simulation process to be fast and more important, the simulated patterns exhibit better continuity.

The selection of an optimal spatial template for retrieving the multiple point statistics is an important aspect of the proposed simulation algorithm. A fast and robust approach to derive optimal spatial templates is presented. The results show that the number of template nodes and their geometry influence the robustness of the retrieved statistics.

An algorithm that uses a probability perturbation method for gradual deformation of geological models in a history matching context has been implemented. This approach, compared to other perturbation methods, offers the important advantages of preserving the prior geological heterogeneity model and simplifying the history match to a single (or few) parameter(s) optimization problem. The algorithm couples a modified Indicator Sequential Simulator with a flow simulator in a Markov-Chain approach where the reservoir model is gradually distorted in an iterative process by updating the conditional distributions of permeability using a deformation parameter that is calibrated using the production information. This dynamic parameter, which is calibrated using the Dekker-Brent Iterative optimization procedure, determines the magnitude of the perturbation on the local distribution and is used to attune the dynamic information pertaining to the reservoir heterogeneity in a probabilistic manner. Local distributions conditioned to
static and dynamic information are iteratively updated and sampled until a global match to the historic data is attained.

The dynamic-calibrated gradual deformation method renders the history match process faster and more controlled, increasing the consistency between the initial and the proposed realizations at every step, and improving the rate of convergence during the minimization of the objective function.

Different schemes are recommended to improve this history match optimization framework:

- Better selection of the initial stochastic realization by picking the best model from a set obtained with different random paths and random sampling draws
- Implementation of a perturbation scheme where the perturbation of the local conditional distributions is controlled so as to transition between an initial realization and three different proposals. This perturbation scheme is expected to further improve the convergence rate during the history matching procedure.
- The proposed method for assisted history matching should be evaluated on multiple 3D cases of different geological complexity in order to identify the capacity and limitations of the proposed approach. 3D cases with layer-to-layer variations in the angle of permeability anisotropy and histograms will be
set up and the production data from the reference models will be used to perturb permeability models that initially exhibit heterogeneity patterns that are radically different from the reference. The results are expected to provide interesting insights into the heterogeneity related information contained in dynamic data. Guidelines for implementing the procedure in complex reservoir environments can be developed and tested.
Bibliography


