Integration of Seismic Information in Reservoir Models: Global Stochastic Inversion

Hugo Caetano

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/52308

1. Introduction

The stochastic models for reservoir properties characterization are a known important tool for reserve management, as well as reservoir quality and thickness that play a key role in deciding optimal well locations in any producing fields. Today’s production reservoirs are each day more complex, and the majority of them, are of difficult access (off-shore), which in technical and cost terms, represents a lack of information.

In petroleum applications, stochastic modeling of internal properties (porosity and permeability), lithofacies and sand bodies of reservoirs, normally use core and log data which in the area provides detailed reservoir parameters, spatially it is limited to a few subsurface locations, scarce and expensive but it is reliable information.

The models created, with the lack of information, are models with great level of uncertainty. It is in this category of models that it is possible to find the stochastic simulation – Sequential Indicator Simulation to the morphological characterization of lithoclasses in [1], the Sequential Gaussian Simulation in [2] and recently, the Direct Sequential Simulation in [3].

The integration of different types of information in a unique and coherent stochastic model has been one of the most important, and still current, challenges of the geostatistical practice of modeling physical phenomena of natural resources and in order to make decisions regarding the development of well locations the geoscientists need to use all available data.

The recent trend of the scientific community regarding development and research for reservoir characterization is creating models which integrate other kind of information (secondary or auxiliary) normally available – the seismic information.
The seismic data which can cover the entire reservoir space has a high uncertainty given the quality and the vertical coarse resolution of seismic. This varies from 25 by 25 meters in horizontal and 1 to 4 milliseconds, in 3-D seismic acquisitions. This data sample is much coarser that the data measured in wells, which vary from some centimeters to a few feet. It is important information never the less, but in almost all applications the seismic data cannot have a direct link to the wells properties (lithology, porosity and permeability), and are difficult to use directly in the models one wishes to create.

The reservoir models based only in seismic information (3-D or 4-D), are normally limited to the structural information. This relationship derives from the major horizons and faults systems, interpreted in the coarse seismic, and it does not take in account the available well information, related to the internal characteristics of the reservoir (porosity, permeability and water saturation). On the other side, the characterization of reservoir models based only in the information of wells, like the recent geostatistical stochastic models, can have a great improvement by the integration of seismic information, which normally is available in the initial phases of prospecting and production.

The integration of these two types of information, with different special coverage and with totally different uncertainty levels is a challenge that even today dazzles the scientific community linked to the earth science modeling.

2. Objective

The main objective of this work is the development and implementation of a stochastic model algorithm for seismic inversion to improve reservoir characterization.

The methods of integration of seismic data can be roughly divided in two approaches. The methods that rely on a statistical relationship between seismic data and internal properties, or lithofacies, to characterize local distributions of these properties in any location of the reservoir by using, for example, co-simulations as [4,5], and others different approaches are posed as an inverse problem framework where the solution, the known amplitudes of seismic, are physically related with the unknown acoustic impedances (or porosity) by mean of a convolution model. Among them there is the so called geostatistical inversion in [6].

The major disadvantages and drawbacks of the direct models, are that the correlation found in the wells locations between the seismic and the internal properties (porosity and permeability), are normally low and sometimes spurious, condemning all the models from there to a great uncertainty.

From the recent deterministic inversion models, the great drawback, is caused by the lack of production of any measure of uncertainty and from not being robust (a great dependency from the seismic quality) and having little liability in almost all of the more complex reservoirs.

In 1993 Bortolli, launched the embryo of what is considered a liable alternative to the existing inverse models, the stochastic inversion. In this method the sequential gaussian
simulation is used to transform, using an interactive process, each of the N verticals columns of the seismic cube.

Since then, the geostatistical seismic inversion has been a commonly used technique to incorporate seismic information in stochastic fine grids models.

Essentially, geostatistical inversion methods as in [7-9], perform a sequential approach in two steps:

i. Acoustic impedance values are simulated in each trace (a column of a 3D grid) based on well data and spatial continuity pattern as revealed by the variograms;

ii. The acoustic impedance values are transformed, by a convolution with a known estimated wavelet, into amplitudes giving rise to a synthetic seismogram that can be compared with the real seismic.

A “best” simulated trace is retained, based on the match of an objective function (function of the similitude between real seismic trace and seismogram), and another trace is visited to be simulated and transformed. The sequential process continues until all traces of acoustic impedances are simulated. In each step as long as the “best” transformed trace is accepted, the traces of simulated acoustic impedances are incorporated as “real” data for the next sequential simulation step. This can lead to artificially good matches in local areas where the bad quality of seismic prevails.

The base idea of this research work is precisely to incorporate stochastic simulation and co-simulation methodologies to conceive and implement a model of global seismic inversion and creating uncertainty linked to areas with different seismic quality.

The use of geostatistics for the creation and transformation of images (acoustic impedances) and the genetic algorithms for the modification and generation of better images, allow the convergence of the inverse process.

The methodology is proposed based on a global perturbation, instead of trace-by-trace, to reach the objective function of the match between synthetic seismogram and real seismic. Using the sequential simulation and co-simulation approaches it creates several realizations of the entire 3D cube of acoustic impedances that are simulated in a first step, instead of individual traces or cells.

After the convolution, local areas of best fit of the different images are selected and “merged” into a secondary image of a direct co-simulation in the next iteration.

The iterative and convergent process continues until a given match with an objective function is reached. Spatial dispersion and patterns of acoustic impedances imposed are reproduced at the final acoustic impedance cube.

As the iterative process is based on global simulations and co-simulations of impedances, there is no local imposing of artificial good fit, i.e. areas of bad seismic tend to remain with bad match coefficients, as it does not happens in most trace-by-trace approaches.
At each iterative step one knows how close is one given generated image from the objective, by the global and local correlation coefficients between the transformed traces and the real seismic traces. These correlation coefficients of different simulated images are used as the affinity criterion to create the next generation of images until it converges to a given predefined threshold.

In a last step, porosity images can be derived from the seismic impedances obtained by seismic inversion and the uncertainty derived from the seismic quality is assessed based on the quality of match between synthetic seismogram and real seismic.

For the case of characterization of the reservoir in terms of facies distribution several methods for the integration of the seismic data in facies models have been proposed, several of which rely on the construction of a facies probability cube by calibration of the seismic data with wells. If only post-stack seismic data is considered, is typically inverts the seismic amplitudes into a 3D acoustic impedance cube, and then converts this impedance into a 3D facies probability using a calibration method of choice.

This facies probability can serve as input of several well-known geostatistical algorithms to create a facies realization, such as the use the cube as locally varying mean on indicator kriging or the use of the tau model in [10], in a multi-point simulation to integrate the facies probability cube with spatial continuity information provided by a training image.

While this provides satisfactory results in most cases, the resulting facies realization does not necessarily match the original seismic amplitude from which the acoustic impedance was inverted. Indeed, if one would forward simulate, for example a 1D convolution on a single facies realizations, then this procedure does not guarantee that the forward simulated seismic matches the field amplitudes.

Another objective of this work is to present a geostatistical methodology, based on multipoint technique that generates facies realization compatible with the field seismic amplitude data, and therefore this new procedure has two main advantages, matching field seismic amplitude data in a physical sense, not merely in a probabilistic sense and using multi-point statistics, not just the two-point statistics (variogram).

3. Seismic inversion

The seismic data is the best source of spatially extensive measurement over the reservoir, but as explained previously, is very coarse vertically. From the wells we can use the acoustic impedance (AI) which is the result of multiplying the lateral interval velocity by the layer density, that later can be transformed into seismic amplitude, this is possible because the difference between the different geological materials is linked to the response of the seismic signal.

The advantages of working with AI instead of the recorded seismic data are that is layer property rather than an interface property and hence more like geology and therefore it has
a more physical meaning and during the inversion process all the well data is tied to the seismic data giving better understanding of the quality of both datasets.

So, the integration of the amplitude seismic data with the acoustic impedance data from wells requires some kind of transformation. There are two methods that transform one in to the other, an inverse method and a direct or forward method.

Both of them use wavelets for the transformation. These can be summarized as one-dimensional pulse and the link between seismic data and geology and they are a kind of impulse of energy that is created when a marine air gun or land dynamite source is released during the acquisition of seismic surveys.

The first method, “inverse process”, transforms the amplitudes from seismic to acoustic impedance by removing the wavelet from the amplitudes and obtaining a model of the acoustic impedance. The problem of simultaneously inverting reservoir engineering and seismic data to estimate porosity and permeability involves complex processes such as fluid flow through porous media, and acoustic wave propagation and cannot be solved by linear inversion methods as [11]. On the other hand, the process of sending a wavelet into the earth and measure the reflection is called forward process. Here we need to know what it is in the subsurface in terms of geological models divided in layers and each characterized by his acoustic impedance. Forward modeling combines the sequence of differences in the acoustic impedances with a seismic pulse to obtain a synthetic amplitude trace.

In the forward model if we compare those sequences of synthetic amplitude variations with the real ones, we can identify and quantify the differences and try to minimize them.

### 3.1. Synthetic seismogram

It is known that the propagation of waves or sounds that pass through earth can be explained by an elastic wave equation. One simple, but powerful and commonly used approach for computing the seismic response of a certain earth model is the so-called convolutional model as in [12-14], which can be derived from an acoustic approximation of an elastic equation.

This convolutional model (equation 1) creates a synthetic seismic trace $S_y(t)$, that is the result of convolving a seismic wavelet $w(t)$ with a time series of reflectivity coefficients $r(t)$, with the addition of a noise component $n(t)$, as follows:

$$S_y(t) = w(t) * r(t) + n(t)$$

(1)

An even simpler assumption is to consider the noise component to be zero (equation 2), in which case the synthetic seismic trace is simply the convolution of a seismic wavelet with the earth’s reflectivity:

$$S_y(t) = w(t) * r(t)$$

(2)
The reflection coefficient is one of the fundamental physical concepts in the seismic method. Basically, each reflection coefficient may be thought of as the response of the seismic wavelet to an acoustic impedance change within the earth and represents the percentage of the energy that is emitted compared to the one that is reflected.

Mathematically, converting from acoustic impedance to reflectivity involves dividing the differences in the acoustic impedances by the sum of them. This gives the reflection coefficient at the boundary between the two layers.

Each layer can have different rock acoustic impedance, which is a function of two porosity-dependent rock properties: bulk density and velocity (equation 3). For instance, the P-wave acoustic impedance is given by:

\[ AI(t) = \text{density}(t) \times \text{P-Velocity}(t) \]  

So the reflectivity coefficient (equation 4) can be computed from:

\[ r_i = \frac{AI(t+1) - AI(t)}{AI(t+1) + AI(t)} \]  

where \( AI \) represents the acoustic impedance and the subscripts \( t \) and \( t+1 \) refer to two subsequent layers in the stratigraphic column.

The wavelet is usually extracted from the seismic survey through deconvolution (the methodology of extraction the wavelet from the seismic signal). During the deconvolution, the wavelet extraction does not take in account the low frequency of the earth model, this represents the compaction of the porous media in depth, which in some seismic inversion algorithm and in the inverse modeling, are ignored or only take in account adjacent geology units, not the full vertical analysis.

In this work, the problem is resolved by the algorithm of generation of the acoustic impedance model, consider the trend that are represented in the wells log data, by creating initially a model only conditioned by the wells data (AI).

### 3.2. Correlation coefficient comparison

Using the convolution algorithm described previously, one can forward an entire model (or cube if one considers a model as a cartesian grid) of AI created only by using the wells and its spatial continuity to a synthetic seismic amplitude cube that can be compared with the real amplitude seismic cube. This comparison will give the mismatch between those two cubes of amplitudes.

The mismatch between the synthetic and the field amplitudes is calculated as a simple collocated correlation coefficient as in [15-17]. Alternatively, this mismatch could be calculated in the form of a least-square difference between both amplitude cubes.
In a first step, each of the N synthetic seismic amplitude cubes are divided into layers (figure 1). The choice of number of layers is a tuning parameter of the algorithm.

The amount of mismatch between each pair of columns of the synthetic and the real amplitude is calculated and stored in a new cube (figure 1), named the mismatch cube. Essentially, this new cube provides information on which locations in the reservoir are fitting the seismic and which need improvement.

![Figure 1. Calculus of the mismatch cube.](image)

A mismatch cube is calculated for each of the simulated models using their corresponding synthetic seismic models.

### 3.3. Co-generation of acoustic images

The generation of an acoustic impedance model is the main objective of this work, but until an optimized one is created, it is submitted to a convergence genetic modification.

These images are generated through stochastic simulation, i.e. generation of different models with the same probabilistic and spatial distribution. This means that if the parameters that are used to the creation of the models are too tight, the produced images are almost identical, if the parameters allow more freedom, these models are totally different.

Using this idea, the proposed methodology uses two approaches for image generation, two different programs are used. The first one is the DSS (Direct Sequential Simulation) for the generation of the first unconditioned images and then for the convergence is used the Co-DSS (Direct Sequential Co-Simulation) that uses a soft or secondary data to simulation conditioning as in [3,18,19].

The second algorithm is the Snesim (Single Normal Equation SIMulator), a multi point simulation algorithm described in [20,21]. This version is used to create the first unconditional facies models. A second version with modification in [2], uses the influence of secondary data to create the models that are conditioned to more information. Both these
programs use secondary data for the co-simulations, and this secondary or soft data is the one that makes the algorithm converge.

For the use of Direct Sequential Co-Simulation for global transformation of images as in [17], let us consider that one wishes to obtain a transformed image $Z_t(x)$, based on a set of $N_i$ images $Z_1(x)$, $Z_2(x)$,...$Z_{N_i}(x)$, with the same spatial dispersion statistics (e.g. correlogram or variogram and global histogram): $C_i(h)$, $\gamma_i(h)$, $F_{z_i}(z)$.

Direct co-simulation of $Z_t(x)$, having $Z_1(x)$, $Z_2(x)$,...$Z_{N_i}(x)$ as auxiliary variables, can be applied as in [3]. The collocated cokriging estimator (equation 5) of $Z_t(x)$ becomes:

$$Z_t(x_0) - m_t(x_0) = \sum_{\alpha} \lambda_{\alpha} \left( x_0 \right) \left[ Z_t(x_\alpha) - m_t(x_\alpha) \right] + \sum_{i=1}^{N_i} \hat{\lambda}_i \left( x_0 \right) \left[ Z_i(x_0) - m_i(x_0) \right]$$

(5)

Since the models $\gamma_i(h)$, $i=1$, $N_i$ and $\gamma(h)$ are the same, the application of Markov approximation is, in this case, quite appropriated, i.e., the co-regionalization models are totally defined with the correlation coefficients $\rho_{t,i}(0)$ between $Z_t(x)$ and $Z_i(x)$.

The affinity of the transformed image $Z_t(x)$ with the multiple images $Z_i(x)$ are determined by the correlation coefficients $\rho_{t,i}(0)$. Hence, one can select the images with which characteristics we wish to “preserve” in the transformed image $Z_t(x)$. So, a local screening effect approximation can be done, by assuming that to estimate $Z_t(x_0)$ (equation 5) the collocated value $Z_i(x_0)$ of a specific image $Z_i(x)$, with the highest correlation coefficient $\rho_{t,i}(0)$, screens out the influence of the effect of remaining collocated values $Z_j(x_0)$, $j \neq i$. Hence, equation 6 can be written with just one auxiliary variable: the “best” at location $x_0$:

$$Z_t(x_0) - m_t(x_0) = \sum_{\alpha} \lambda_{\alpha} \left( x_0 \right) \left[ Z_t(x_\alpha) - m_t(x_\alpha) \right] + \hat{\lambda}_i \left( x_0 \right) \left[ Z_i(x_0) - m_i(x_0) \right]$$

(6)

With the local screening effect, $N_i$ images $Z_i(x)$ give rise to just one auxiliary variable. The $N_i$ images are merged in on a single image based on the local correlation coefficient criterion.

Basically, the correlation or least mismatch of each of the previously simulated images at each $x_0$ is converted to a single image where the best of each $x_0$ is selected, this way.

In practice, the algorithm chooses in each $x_0$ of the simulated acoustic impedance model the best genes and the correlation that those genes have. Based on figure 2, the process is explained as follows:

The process starts by analyzing each of the mismatch cubes that were created previously, and in each position of the cubes ($x_0$) it compares which has the higher value of correlation or lower mismatch. In this case the example used is the correlation, so, the points 1a and 1b are compared, and since it is the 1a that has the higher value of correlation, that value is copied (1c) to the Best Correlation Cube.
At the same time, since it was from the first simulation that has better coefficient of correlation (CC) with the real seismic, another cube is created with the correspondent acoustic impedance (AI) value (1d). Next, it starts comparing the next set of values from the CC cubes, and the process is the same. With a comparison between the sets 2a and 2b, and since in this case it is the second cube that has the higher value of cc, it is this values that is copied to the Best CC cube (2c). The same is done from the AI cube of this simulation, copying the acoustic impedance values to the Best AI (2d). This process continues until all the N simulated models have been analyzed and to all the values of the cubes. Through this way two new cubes are created, one with the best genes from each acoustic impedance model (Best AI) generated previously and another cube with the confidence factor of each part (Best CC).

![Figure 2. Process of creating “Best Correlation Cube” and “Best Acoustic Impedance cube”](image)

Since the process is an iterative one, in the very first steps of the iterative process, the secondary image (Best AI) do not have the spatial continuity pattern of the primary (simulated AI), as it results from a composition of different parts of a set of simulated images. As the process continues, the secondary image tends to have the same spatial pattern of generated images by co-simulation, because the correlation values are becoming higher and the freedom of the co-simulation is diminishing. Finally these two new data sets are used as soft data for the next iteration.

4. Algorithm description using direct sequential simulation

The proposed stochastic inversion algorithm is settled with the following key ideas in mind: generation of stochastic images, transformation of the images in synthetic seismograms, chose and keep the best “genes” from each of the images and then use them through the exercise of the genetics algorithm formalism and the stochastic co-simulation to create a new generation of images and the convergence of the process. It can be summarized in the following steps;
i. Generate a set of initial 3D images of acoustic impedances by using direct sequential simulation. Instead of individual traces of cells;

ii. Create the synthetic seismogram of amplitudes, by convolving the reflectivity, derived from acoustic impedances, with a known wavelet;

iii. Evaluate the match of the synthetic seismograms, of entire 3D image, and the real seismic by computing, for example local correlation coefficients;

iv. Ranking the “best” images based on the match (e.g. the average value or a percentile of correlation coefficients for the entire image). From them, one selects the best parts (the columns or the horizons with the best correlation coefficient) of each image. Compose one auxiliary image with the selected “best” parts, for the next simulation step;

v. Generate a new set of images, by direct sequential co-simulation, using the best locals correlation as the local co-regionalization model and return to step ii) starting an iterative process that will end when the match between the synthetic seismogram and the real seismic is considered satisfactory or until a given threshold of the objective function is reached;

vi. The last step of the process is the transformation of the optimized cube of acoustic impedance in internal reservoir characteristics.

At each iterative step one knows how closer is one given generated image from the objective, by the global and local correlation coefficients between the synthetic seismogram and the real seismic. These correlation coefficients of different simulated images are used as the affinity criterion to create the next generation of images until it converges to a given predefined threshold. A simplified diagram is show in figure 3.

Figure 3. Diagram of the proposed algorithm
4.1. Case study

A case study of a Middle East reservoir will be presented but only a small part of the full reservoir is studied due to data confidentiality and the coordinates presented are modified.

The field described in [22,23] is a carbonate reservoir with a deposition geology that holds in some zones a strong internal geometry with clinoforms. These are sedimentary deposits that have a sigmoidal or S shape. They can range in size from meter, like sand dunes, to kilometers and can grow horizontally in response to sediment supply and physical limits on sediment accumulation.

This type of geological phenomenon increases the complexity of the internal reservoir, making the geophysics interpretation of seismic a very difficult job, mainly when the resolution of acquisition is very coarse. It also causes a great impact in oil production in wells and reservoir characterization and modeling.

From the calibration of the acoustic impedance data of the wells with the seismic, a wavelet was extracted and there were 19 wells in the study area but only two were used, because only these two have the velocity log, and without the velocity log the acoustic impedance data cannot be calculated. On those wells the acoustic impedance was determined, then calibrated with the seismic and finally upscaled to fit the seismic scale of 4 ms.

The convolution for well 11 is show in figure 4.

![Figure 4. Calculation of the AI, convolution and match with seismic of well 11.](image-url)
As one can notice, there is a very good match with the synthetic amplitude (red line on right side) calculated with the acoustic impedance (white line in the middle) from the wells and the real seismic amplitude extracted from the seismic cube (cyan line in the right side).

4.2. Results

The first sets of 32 images of acoustic impedances are generated with the direct sequential simulation conditioned to well data (AI) and the chosen variogram model. In the first iteration, the generated acoustic models (figure 5) are not constrained to any soft data, so
only the wells are reproduced. This causes a high variability between the synthetic models (figure 6) and wide range standard deviation when calculated using all 32 simulations.

As it can be noticed, the two models have a totally different spatial distribution, although the histograms and variograms are the same.

That different spatial distribution is visible in the correlation cubes between the synthetic models and the real seismic (figure 7).

Figure 7. Correlation cubes between the synthetic seismic model and the real seismic, of simulation #1 (left) and #15 (right).

Figure 8. Average (left) and Standard Deviation (right) of acoustic impedance of all 32 simulations of this iteration.
To confirm this variability, the average cube and standard deviation (figure 8) of all 32 simulations were determined and the influence of the wells in the average of the unconditioned simulations is highly visible principally around well 1, as the variogram is reproducing the well log spatial variability.

As perceived in the average model, the small scale variability has disappeared, however this cube can be considered the Low Frequency Model, because it reproduces the main trends that are represented in the wells logs. Also noticeable the practically constant value of standard deviation, representing the variability of unconditional stochastic simulations, except the small decline of values in the middle of the figure, which is the influence of the well 1.

Still, these first 32 simulations were used to build the “Best Correlation Cube” and “Best Acoustic Impedance cube” that will be used as soft data for the next iteration (figure 9).

![Figure 9. “Best Acoustic Impedance cube” (left) and “Best Correlation Cube” (right) derived from first iteration.](image)

It is visible the delineation of the structural layers in the Best Acoustic Impedance model (BAI) and in the Best Correlation Cube (BCC) a selection of correlation coefficients high values.

This assembled acoustic impedance model, has lost all the spatial distribution that the original acoustic impedance models had (figure 5), but this is a simple intermediate result and not the final one since these will be used as secondary or soft data for the next iteration that will impose the variogram spatial distribution and wells global histogram.

The algorithm has made six iterations, one of them, the first, was unconditional to any secondary or soft data, only the last five had the Best Correlation Cube and Best Acoustic Impedance cube as secondary or soft data imposition.
Since the algorithm will always choose the best genes from each iteration, the patterns that are in the real seismic will start to become more visible in each iteration and the correlation will become higher and more continuous in all cube positions.

The convergence of the process is inevitable until a local maximum correlation is attained (figure 10). This maximum does not represent the best that the original seismic can produce, but rather a simulation that the entire process has created. If one has run the same process with the same data but with a different seed for the generation of the acoustic impedance model the result could be a different one, but not so totally different.

![Figure 10. Convergence progress of the algorithm.](image)

The 32 images of acoustic impedances of iteration 5 (considering that the first iteration is called 0, because it is an unconditional to soft data) are generated with the direct sequential co-simulation conditioned to well data (acoustic impedances), the chosen variogram model and the soft or secondary data (BAI and BCC) of the forth iteration.

One can clearly see the fast convergence from iteration zero to iteration 1 and afterwards the process starts to stabilize. The algorithm chooses the parts that have higher correlation values in the end of iteration 0 and after iteration 1, almost every simulation has its correlation values around 1, making the selection of each part of simulation, a very detailed event (sometimes in the third of forth decimal of the correlation value).

The obtained final results demonstrate that the conditional data is imposing a very strong effect in all 32 simulations. The variability of different models generated with different seeds is now almost none existing (figure 11 and 12) and they practically look the same.

As it can be noticed the two models have an almost equally spatial distribution, but some differences can be found, since they are two independent realizations.
Those small differences are easier to distinguish in the correlation cubes between the synthetic models and the real seismic (figure 13), since the correlation coefficient is very sensible to little variations in patterns.

In these examples the layer set sizes were big enough to show the differences.
To confirm this lack of variability (figure 14), the average cube is almost equal to a simulation, and validated by the lower values of standard deviation of all 32 simulations.

In figure 14 the same color scale for standard deviation, was used, for comparison with the standard deviation of the first iteration model (figure 8).

The values of standard deviation for the final iteration only vary from 0 to 6900, with an average of 1600 which is a reduction of almost 80% in variability, comparing with the variation between 0 and 11500 and average of 8000 of the first iteration.
The influence of the wells is no longer visible because all data of wells are integrated in the full model.

4.3. Remarks

In spite of the scarcity of the log data, the proposed method achieved extremely good results. In case of data abundance, if the data is not well calibrated it could compromise the quality of the convergence and the influence of the not calibrated wells in the final model would be noticeable.

In this case, both synthetic seismic data and acoustic impedance cube captured the main geologic features of these complex reservoirs, noticeable in the correlation coefficients between the seismic and the synthetic amplitudes. The quality of the seismic data takes a minor role since the method overcomes the situation of imposition of artificial correlations as it happens in the standard methods;

Since the co-simulation of the impedances uses a local coefficient correlation, it is possible to compute the local uncertainty associated to the seismic acoustic impedances;

The uncertainty of the seismic acoustic impedance could be used to access the uncertainty associated with the porosity model, as presented in [24].

Tests prove that the variation of the final correlation coefficient is about 2% with the modification of the initial seed, it means that others parameters such as the number of layers and the size of it, as other parameters can be optimized to produce better results. But these results are more difficult to reach when the complexity of the geology and the structural model became more elaborated.

To handle different geology scenarios such channels or specific shape reservoirs, an adapted approach is proposed in the next part of this work.

5. Multipoint statistics

The objective of this work is to build a reservoir model with multiple alternatives, thereby assessing uncertainty about the reservoir, integrating information from different sources obtained at different resolutions:

- Well-data which is sparse but of high resolution, in the scale of a foot;
- Seismic data which is exhaustive but of much lower resolution, in the scale of 10’s feet in the vertical direction;
- Conceptual geological models, which could quantify reservoir heterogeneity from the layer scale to the basin scale.

This last point has been possible using a variogram in algorithms such as DSS and Co-DSS, previously described, which allow integration of well and seismic data using a pixel-based approach. Those variogram based models are inadequate in integrating geological concepts since the variogram is too limited in capturing complex geological heterogeneity and is a
two-point statistics that poorly reflects the geologist’s prior conceptual vision of the reservoir architecture, e.g., sand channels as in [20].

Integration of geological information beyond two-point variogram reproduction becomes critical in order to quantify more accurately heterogeneity and assess realistically the uncertainty of model description.

A solution was initiated by practitioners from the oil industry in Norway, where Boolean objects-based algorithms were introduced in the late 1980’s to simulate random geometry as in [25,26]. These parametric shapes, such as sinusoidal channel or ellipsoidal lenses, are placed in simulated volume. Through an iterative process this shape is changed, displaced or removed to fit the conditioning statistics and local data.

The simulated objects finally resemble the geologist drawings or photographs of present day depositions

Passed the enthusiasm, the limitation of objects-based simulation algorithms became obvious, this iterative, perturbation type, algorithm for data conditioning did not converge in the presence of dense hard data or could not account for diverse data types, such as seismic used as soft data.

Also the limitation of not simulating continuous variables, time and CPU demanding in large 3D cases, enroll on the drawbacks of the methodology.

Strebelle in [20], following the works of Srivastava in [27] and Caers in [28] has proposed an alternative approach of Multipoint statistics that combines the easy conditioning of pixel-based algorithms with the ability to reproduce “shapes” of object-based techniques, without relying on excessive CPU demand.

Multipoint statistics uses a training image instead of a variogram to account for geological information. The training image describes the geometrical facies patterns believed to represent the subsurface.

Training images does not need to carry any local information of the actual reservoir, they only reflect a prior geological/structural concept.

Basically, multipoint statistics consists of extracting patterns from the training image, and anchoring them to local data, i.e. well logs and seismic data. Several training images corresponding to alternative geological interpretations can be used to account for the uncertainty about the reservoir architecture.

Training images, identical to the variogram, can be a questionable subject when the model to be generated has few hard data or no pre-conceptual model has been studied by geologists or geomodelers.

Using this new tool a new objective is to illustrate the multipoint statistical methodology to seismic inversion.
5.1. Seismic inversion using multipoint statistics

For the case study, the Stanford VI synthetic reservoir dataset described in [29], was used. This synthetic reservoir (figure 15), consists of meandering channels, sand vs mud system, and each facies were populated with rock impedance, using sequential simulation.

To mimic a seismic inversion, this high resolution rock impedance model were smoothed by a low-pass filter to obtain a typical acoustic impedance response and then convoluted to an amplitude model that will be considered as reference seismic amplitude.

![Figure 15. Facies model (left), high resolution rock impedance (middle left), smothered rock impedance (middle right) and forwarded seismic synthetic amplitude (right)](image)

5.2. Probabilistic approach

For comparison of the methodology, is presented the application of a traditional probabilistic modeling on the Stanford VI acoustic impedance, i.e. calibrate acoustic impedance into a 3D facies probability cube, and then use it as soft data constraint for multiple-point geostatistical simulation.

Several techniques exist to calibrate one or more seismic attributes with well data, such as PCA, ANN, etc., one use a simple Bayes’ approach (equation 7) as documented in [30]. In a Bayesian method one uses the histogram of impedance for each facies denoted as:

\[ P(facies_f) = \frac{\text{histogram of impedance for facies } f}{\text{total number of samples}} \]

Using Bayes’ rule one can calculate the probability for each facies for given impedance values as (equation 8):

\[ P(AI \mid facies_f) = \frac{P(AI \mid facies_f) \cdot P(facies_f)}{P(AI)} \]

where \( P(facies) \) is the global proportion of facies and \( P(AI) \) is derived from the histogram of the impedance values. The final result is a cube of probabilities for each facies type based on the acoustic impedance cube. In figure 16, one can see the result of the application of this method to the case study data.

This probability data can be used by different geostatistical algorithms to create a facies model. In our case the simulated facies are obtained with the multi-point method snesim, conditioned to the probability cubes previously calculated, to the training image and to the rotation and affinity cubes.
Figure 16. Result of the Bayesian approach for the calculus of the probabilities cubes for case study and the simulated facies model.

To verify our hypothesis that this facies model does not match the field amplitude, the facies model is forwarded simulated to a synthetic amplitude dataset. We assumed the ideal situation where an exact forward model was available. Figure 17 confirm that the forward modeled amplitude does not match the field data. In fact, the co-located correlation coefficient is 0.20.

Figure 17. “Real” seismic amplitude (left) vs forward modeled amplitude derived from the simulated facies model using a traditional probabilistic approach (right).

5.3. Algorithm description using multipoint statistics

For the case of using multi-point statistics, the seismic inversion algorithm has gone through some changes, mainly caused by the simulation algorithm that does not creates acoustic impedance models directly, but facies models, that can be populated with acoustic impedance values later.

So the method is initialized by simulating N facies models only conditioned to the wells and training images using snesim, in this case, additional channel azimuth and affinity (as interpreted from seismic or geological understanding) constraints are enforced (figure 18).
Figure 18. Initial simulations conditioned only to the training image and rotation/affinity.

The facies models are then populated with acoustic impedances, converted in synthetic seismograms and chose and keep the best “genes” from each of the images (see figure 19). In the next step there is also a modification, since one cannot use directly the correlation values in the snesim simulator. So a modification of the probability perturbation method as in [31] is used to create a probability for each facies.

This facies probabilities cube is then used as soft constraint to generate the next set of N cubes. This is done using Journel’s tau-model as in [10], to integrate probabilistic information from various sources. The process converges as the previous algorithm and it can be summarized in the following steps:

i. Generate a set of initial 3D images of facies by using snesim.
ii. Populate the facies models with acoustic impedances.
iii. Create the synthetic seismogram of amplitudes, by convolving the reflectivity, derived from acoustic impedances, with a known wavelet.
iv. Evaluate the match of the synthetic seismograms, of entire 3D image, and the real seismic by computing, for example local correlation coefficients.
v. Ranking the “best” images based on the match (e.g. the average value or a percentile of correlation coefficients for the entire image). From it, one selects the best parts (the columns or the horizons with the best correlation coefficient) of each image. Compose one auxiliary image with the selected “best” parts, for the next simulation step.
vi. Transform the cube with the best parts and the corresponded best values to a probability cube to each facies to be simulated.
vii. Generate a new set of images using snesim adaptation to integrate secondary information, with the probability cubes used as the local co-regionalization model and return to step ii) starting an iterative process that will end when the match between the synthetic seismogram and the real seismic is considered satisfactory or until a given threshold of the objective function is reached.

viii. The last step of the process can be the transformation of the optimized facies cube in acoustic impedance, porosity or any internal reservoir characteristics.

As noted in previously algorithm, in step vi), in the very first steps of the iterative process, the probability data can not have the spatial continuity pattern of the primary simulated facies, as it results from a composition of different parts of a set of simulated images. As the process continues, that soft secondary image (probability cubes) tend to have the same spatial pattern of generated images by co-simulation, i.e. the imposed training image altered by the affinity and azimuth.

5.4. Co-generation of acoustic images with multipoint statistics

The major change made for the adaptation to MPS, is that the acoustic impedance values are estimated or populated in the simulated facies model. This can cause some reservations but if one considered that the main objective is comparing the reflection coefficients between the real and the simulated seismic amplitude, those reflection coefficients are more accentuated in the change of terrain type, and those can be considered a channel or crevasse. Inside the channels the rock type does not vary too much and that is visible in the seismic profiles by the presence of low amplitude seismic. So this difference in the algorithm can be considered a valid one.

The second difference is that to co-simulate a facies model, one can not use acoustic impedance values (a continuous variable) as soft data, so instead of choosing from the simulated acoustics model, the algorithm build the Best Facies cube from the simulated facies models (figure 19).

Figure 19. Process of creating “Best Mismatch cube” and “Best Facies cube”.
As the methodology for acoustic impedance models, the process starts by analyzing each of the mismatch cubes that were created previously, and in each position of the cubes (x) it compares which has the higher value of correlation.

The point set 1a and point set 1b are compared, from those is the 1a that has a higher value of correlation so, it is copied (1c) to the Best Mismatch cube.

At the same time, since the first simulation showed better coefficient of correlation (CC) with the real seismic, another cube is created with the correspondent facies values (1d).

The process follow to the next set of values from the CC cubes, and it is the same, comparison between the sets 2a and 2b, that in this case it is the second cube that has the higher value of cc (or lower mismatch), it is these values that are copied to the Best Mismatch cube (2c). The same is done from the facies cube of this simulation, copy of the facies values to the Best Facies (2d). This process continues until all the N simulated models have been analyzed and to all the values of the cubes.

Through this way two new cubes are created, one with the best genes from each facies model (Best Facies) generated previously and another cube with the confidence factor of each part (Best Mismatch). As noticed above, the “least mismatch cube” can be seen as a summary of the least mismatch of all N realizations. The “best facies cube” can be seen as facies model combined from all N facies models that best matches the seismic data. However the “best facies cube” does not have the same geological concept as the training image and may have various artifacts, since it is constructed by copying several sets from independently generated facies models.

The third difference is that, these two new cubes can not be used for the co-simulation of the new generation of facies models, but the “best facies cube” can be used indirectly to improve the existing N facies models.

In order to do this, one uses a modification of the probability perturbation method as in [24]. Caers’ method was developed to solve non-linear inverse problems under a prior model constraint. It allows the conditioning of stochastic simulations to any type of non-linear data. The principle of this method relies on perturbing the probability models used to generate a chain of realizations that converge to match any type of data.

In this methodology the unknown pre-posterior probability – Prob(A)\(|\ C\) – is modeled using a single parameter model in the following equation 9:

\begin{equation}
\begin{align*}
P(A_j | C) &= P(I(u_j) = 1|C) = (1 - r_c) \times i^{(0)}(u_{j}) + r_c \times P(A_j), j = 1, ..., N
\end{align*}
\end{equation}

where A is unknown data, B is well data and previously simulated facies indicators and C will be the “best facies cube”, \( r_c \) is not dependent on \( u_j \) and is between [0,1], and \( \{u^{(0)}_{j} \}, j = 1, ..., N \} \) is an initial realization conditioned to the B-data only.

According to the methodology proposed in this work, the equation 9 is adapted with \( r_c \) now representing the mismatch between the field and synthetic seismic as summarized with the
least-mismatch cube (correlation coefficient), and $I_{t_k}^{(0)}$ corresponding to the presence of the facies with the least mismatch. This adaptation leads to the following equation 10:

$$P(A_j | C) = \rho_c(u) \times I(u_j) + (1 - \rho_c(u)) \times P(A_j), \ j = 1, \ldots, N$$

where $\rho_c$ is the correlation coefficient, between [0,1], extracted from the “least mismatch cube”, $I(u_j)$ is the sand indicator extracted from the “best facies cube” with $j=1,\ldots,N$ facies, and $P(A_j)$ is the global proportion of the considered facies. This expression generates a “facies-probability cube” which is a mixture of global proportion and the “best facies cube”.

This facies probabilities cube is then used as soft constraint to generate the next set of N cubes. This is done using Journel’s tau-model to integrate probabilistic information from various sources (see equation 11 and figure 20);

$$\frac{x}{b} = \frac{c}{a}$$

where: $x = \frac{1 - P(A|B,C)}{P(A|B,C)}$, $b = \frac{1 - P(A|B)}{P(A|B)}$, $c = \frac{1 - P(A|C)}{P(A|C)}$ and $a = \frac{1 - P(A)}{P(A)}$

$a$ is the information of the global facies proportion, $b$ is the influence of the training image, and $c$ is the conditioning of the soft probability cube.

**Figure 20.** Next generation of simulations conditioned to the training image, rotation/affinity and probability cubes.
This iterative algorithm is run until the global mismatch between the synthetic seismic of the facies model and the field seismic data reach an optimal minimum.

5.5. Results

To illustrate the method, 6 iterations were computed with N=30 simulations on the Stanford VI dataset. Note that one assumed the availability of perfect geological information and a perfect forward model.

The algorithm converges as shown by a systematic increase in the global correlation coefficient between model and amplitude data (figure 21). The facies model with the highest correlation out of 30 models in the last iteration has a correlation of 0.88.

![Figure 21. Convergence of the algorithm.](image)

![Figure 22. Comparison between the two methods](image)

(Proposed Approach)  
Facies Cube  
Forward Modeling  
Seismic Amplitude  
\( \rho = 0.88 \)

(Reference Data)  
Facies Cube  
Forward Modeling  
Field Seismic  
\( \rho = 0.20 \)

(Probabilistic Approach)  
Facies Cube  
Forward Modeling  
Seismic Amplitude
The final facies model (top left in figure 22) does not contain any artifacts, i.e. reproduces well the geological continuity of the training image.

To check how well the seismic amplitude data is reproduced, is forward modeling the best facies model to seismic synthetic (see figure 22). Clearly the method matches well the field seismic, particularly when compared with the probabilistic approach.

In figure 23 some slices of the results are presented, where it is possible to see some similarities between the seismic response of the proposed approach and the reference data seismic amplitude.

![Figure 23. Correlation slices](image)

In summary, as stated previously, the local maximum correlation or local minimum mismatch depends on the parameters used to calculate the mismatches. Some parameters sensitivity is further analyzed:

- Number of iterations

In the first iteration the main optimization is obtained, after this the process tends to stabilize (figure 24). Hence, without changing any of the other parameters, the number of iterations does not have a big influence on the optimization.

Figure 24 compares the values of correlation and average difference of the simulation model with the best result for different number of facies models N simulated. These values are presented only for the first and last iterations.

For the first iteration the difference is not very noticeable, but in the last iteration the correlation value is higher as the number of simulations increases.
• Number of facies simulations (N) per iterations

This parameter can have a major influence on the optimization: with numerous simulations the process has a wide variety of possibilities to choose from, and can build a more precise best-facies cube, but if the simulations are very alike, the choice of N will not matter.

Figure 24. Influence of different number of iterations (1 to 6) and simulations (20, 30 and 40).

• Criterion of convergence

The criterion of differences is more precise than the correlation, since the correlation gives us a value of the similarity between two sets of data, while the difference is more susceptible to different patterns or small variation in the patterns. But to calculate the probabilities the correlations are needed.

• Size of the columns

In the previous test the cubes were divided in 4 layers with 20 values each column. Since the correlation has some sensitivity to the number of data used, the results also change. For a largest column the convergence will tend to be slower, since it is more difficult to match an entire column than piece-wise matching. In general we recommend that the number of layers chosen depends on the vertical resolution of the seismic data, i.e. for lower resolution less layer could be retained.

6. Conclusions

The algorithm proposed uses two different approaches to create a stochastic model of a reservoir property that are conditioned to both well and seismic data, since this two data are totally different, this process was to be an iterative one, that used the maximum correlation or minimum mismatch as the objective function.
The first approach uses the direct sequential co-simulation as the method of “transforming” 3D images, in an iterative process. Hence it generates, at each iterative step, global acoustic impedances images with the same spatial pattern, without imposing artificially good match in areas of low confidence in seismic quality.

It means that in those areas the final images will reflect high uncertainty. This uncertainty of the seismic acoustic impedance is used to access the uncertainty associated with the porosity model, by co-simulate the porosity model and using the generated final acoustic impedance model and its correlation coefficient cube as soft or secondary data.

Another example is the capacity to improve the seismic resolution, if the acoustic impedance model is created below the original resolution of the seismic. This is possible since the data from wells, that is mainly used to create the model, has a much lower resolution than the seismic.

Time consuming is no more a drawback as it is implemented in a parallel computing platform in [32].

This algorithm has been implemented and benchmarked tested with several real reservoirs, e.g. [33].

For the case of the multipoint approach, it is a new application for the newest stochastic generation of images and is a complement for the first approach.

The modifications made from the first approach to the second, are not too complex, except the population of the acoustic impedance cube, which for this example, a technique of simple populate the channels was used with previous simulated acoustic impedance values.

To resolve this bypass, an algorithm by [34], simulates any variable within the channel, making this step a more realistic one.

The program still needs to be improved and the algorithm can be optimized for real case applications and the lack of training images that could implement it to all kind or examples is still in research.

Author details

Hugo Caetano
Partex Oil and Gas, Portugal

Acknowledgement

This work was created based on the research made at: Centro de Modelização de Reservatórios Petrolíferos (CMRP), of Instituto Superior Técnico (IST), Universidade Técnica de Lisboa (UTL) with sponsorship from Partex - Oil and Gas and Stanford Center for
Reservoir Forecast (SCRF), Stanford University, Califórnia, EUA, with financial support from Fulbright Fundation and Fundação Luso-Americana.

7. References


Information in the Stochastic Models”, SPE 111305, 2007 SPE/EAGE Reservoir Characterization and Simulation Conference held in Abu Dhabi