Linking Geostatistics with Basin and Petroleum System Modeling: Assessment of Spatial Uncertainties and Comparison with Traditional Uncertainty Studies

Bin Jia and Tapan Mukerji

Department of Energy Resources Engineering
Stanford University

Abstract

Basin and Petroleum System Modeling process covers a large spatial and temporal interval. Many of the input parameters are highly uncertain. While probabilistic approaches based on Monte Carlo simulations have been used to address this uncertainty, the impact of spatial uncertainty on basin modeling remains unexplored. Facies is one of the key modeling inputs since the rock properties are wrapped into facies definition. Many techniques had been developed for facies modeling in reservoir characterization regime. These methods can be applied directly in basin modeling. In particular Multi-point Geostatistical Method had been proven effective in facies modeling given sound training images. Another important spatial parameter is the structure model. Present day structure model is the initial point for reconstruction the deposition history. In this work we first show the traditional uncertainty analysis in basin modeling. Then the impact of facies distribution and structure uncertainty from time-to-depth conversion are studied. It is concluded that facies distribution has great impact on the oil accumulation and different geological interpretation gives quite different results. Structure uncertainty from depth conversion has less impact in this case because the target area is considered quite homogeneous and also the global structure feature is maintained quite well.
1. Introduction

Basin and Petroleum System Modeling (BPSM) is a key technology in hydrocarbon exploration that reconstructs deposition and erosion history and forward simulates thermal history and the associated generation, migration and accumulation of petroleum (Peters, 2009).

BPSM involves solving coupled nonlinear partial differential equations with moving boundaries. The equations govern deformation and fluid flow in porous media, coupled with chemical reactions and energy transportation. The coupled system has to be solved numerically on discretized time and spatial grids with the integration of geological, geophysical, and geochemical input. PetroMod uses the finite element method to solve these equations. The workflow and key input parameters are summarized in Figure 1 (Peters, 2009).

Figure 1: Basin and Petroleum System Modeling Workflow (Peters, 2009)
The modeling process can cover very spatial and temporal interval. Many of the
input parameters are highly uncertain and yield very different simulation results. Thus
understanding the impact of input parameters is critical for exploratory decision
making.

The interest of uncertainty analysis in BPSM increases as the computer power
makes it possible to assess multiple models in a reasonable time. While much work
has been done on the uncertainty (e.g. Zwach and Carruthers 1998; Corradi et al.,
2003; Wendebourg, 2003), the focus is mainly on traditional Monte Carlo techniques
which randomly draws values from statistical distributions of the input parameter and
compare the difference in the result for each drawn input parameter. This gives an
estimate of parameter uncertainty but important spatial correlations are not taken into
account. The outputs from these parameter Monte Carlo simulations cannot be used
to assess the joint spatial uncertainty of the results. In earth sciences, one seldom has
sufficient data to accurately reveal the entire underlying subsurface conditions.
Typically in basin modeling one has to estimate the input parameters for the entire
area with only a few data point. Spatial modeling techniques have to be used to make
the best geological interpretation and understand the associated uncertainties.

Facies map is a key input for BPSM process because all the important
dependent and used in model building. Multi-point geostatistical (MPS) algorithm is
the state of art method which generates multiple geological models that honor the
general structure and the well data at the same time. It is more suitable for
facies modeling than traditional variogram-based methods. We will examine the
impact of facies distribution by generating multiple facies map realizations using
MPS method.

Present-day structure model is the starting point for compaction analysis.
Structure models are usually built based on picking and interpretation of seismic data
with constraints at the wells. Well data is considered exact, while each step of the
seismic processing chain (acquisition, preprocessing, stacking, migration,
interpretation, and time-to-depth conversion) has inherent uncertainty that must be
evaluated and integrated into the final result. It is also pointed out that the time-to-
depth conversion uncertainty often represents 50% or more of the total uncertainty
(Thore et al., 2002). In this paper we study the impact of structure uncertainty using
Bayesian Kriging for seismic time-to-depth conversion and COHIBA software
package is used to generate multiple structure realizations.

The rest of this paper is organized as the follows. Section 2 the traditional
uncertainty study on parameters including TOC (Total Organic Carbon), HI
(Hydrogen Index) and Heat flow is reviewed using PetroMod risking functionality.
Section 3 studies the impact of facies uncertainty and section 4 studies structure
uncertainty from seismic time-to-depth conversion. The results are compared with the
traditional uncertainty assessment A sampling method is proposed to reduce the
number of models that are required in the ensemble-based workflow to evaluate the resulting uncertainty. The discussion ends with conclusions and future work.

2. Traditional uncertainty practices

In this section we show the results of a traditional parameter uncertainty analysis on important parameters including TOC, HI and Heat flow using a Monte-Carlo approach. PetroMod has the built-in functionality for such kind of uncertainty analysis.

The generation and maturation of hydrocarbon components, molecular biomarkers and coal macerals can be quantified by chemical kinetics including TOC, and HI (Hantschel and Kauerauf, 2009). The TOC (total organic carbon) is the ratio of the mass of all carbon atoms in the organic particles to the total mass of the rock matrix. HI (hydrogen index) is the ratio of the generative mass of hydrocarbon to the mass of organic carbon. The HI multiplied with the TOC and the rock mass is equal to the total generative mass of HCs in the rock (Hantschel and Kauerauf, 2009).

Another important parameter usually risked is the basal heat flow at the base of the sediment. Magnitude, orientation and distribution of the heat inflow at the base of the sediments are determined by mechanical and thermal processes of the crust and mantle (Allen and Allen, 2005). We will see that the heat flow not only affect the amount of HC accumulation but also the compositions.

2.1. The input model

A 3D synthetic layer cake model is used for our analysis (Figure 2). The model consists of five layers, and from bottom to top are: Under burden layer, Organic lean shale layer, Source rock layer, Reservoir rock layer and Over burden layer. The model has 120 grid cells in x direction and 30 in y direction. Each grid cell represents a region of 1km x 1km area. The total region covers an area of 120km x 30km. The total depth is more than 4500 meters.

Figure 2: 3D display of the layer cake model
Figure 3 shows the same model in 2D sections. On the left is the vertical cross-section of the model and the thin dark layer in the middle is our source rock layer. The top right figure shows facies map of reservoir layer in plan view. It consists two facies: sandstone (yellow) and organic lean shale (dark blue). The bottom right is the depth map of the reservoir layer in plan view. The reservoir top is deeper on the left side and becomes shallower towards the right side.

The deposition setting is summarized in Figure 4. Each layer is deposited in about 10 million years except for the source rock layer which is formed in 1 Ma. The reservoir layer has a channel fluvial depositional setting and is made of two types of facies. Erosion is not modeled in this example.

<table>
<thead>
<tr>
<th>Name</th>
<th>Color</th>
<th>Deposition</th>
<th>Erosion</th>
<th>Max. Time Step</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ob</td>
<td></td>
<td>Age from 10.00</td>
<td>Age to 0.00</td>
<td>Duration 10.00</td>
</tr>
<tr>
<td>Res</td>
<td></td>
<td>Age from 20.00</td>
<td>Age to 0.00</td>
<td>Fac_Rev Fac_Sel</td>
</tr>
<tr>
<td>Src_up</td>
<td></td>
<td>Age from 21.00</td>
<td>Age to 0.00</td>
<td>Fac_Src</td>
</tr>
<tr>
<td>Src_lean</td>
<td></td>
<td>Age from 30.00</td>
<td>Age to 21.00</td>
<td>Fac_Sed</td>
</tr>
<tr>
<td>Base</td>
<td></td>
<td>Age from 40.00</td>
<td>Age to 30.00</td>
<td>Fac_OB</td>
</tr>
</tbody>
</table>

The facies definition is summarized in Figure 5. Source rock layer has lithology of typical shale which has 5% TOC, 500 HI value. The kinetics of Pepper&Corvi(1995)_TII(B) is used which is oil prone. The reservoir layer has both typical sandstone organic lean shale. The sandstone acts as the trap for HC and shale acts as the barrier for fluid flow.
The Boundary conditions are shown in Figure 6. The default value are used for which heat flow is kept constant at 60mW/m², paleo water depth is kept constant at 0, and surface temperature is 20 degree Celsius.

The result for the base case is shown below, for which the oil accumulation is about 1378 MMbbls. The oil was generated from the source rock layer and migrates up to the reservoir layer. The oil then continues move toward the high elevation area and stops until reaches a barrier. The organic lean shale in reservoir layer acts as the flow barrier and the oil was accumulated at different places.
2.2. TOC

We first performed the risking analysis on TOC values. An extensive range of TOC value from 1% to 10% is studied. We see the oil accumulation increases as the TOC value increases. The total accumulation varies from 500 to 2300 MMbbls. Typical values of TOC range from 3% to 7%, and correspondingly the oil accumulation varies from 1100 to 1700 MMbbls. Figure 8 shows the simulation results and also the estimated oil accumulation distributions.

![Figure 8: TOC risking results. The estimated distribution gives P10 of 973, P50 of 1395 and P90 of 1992 MMbbls.](image)

2.3. HI

A similar study is performed for hydrogen index. For the typical HI value ranging from 300 to 700 mgHC/gTOC, the oil accumulation is between 1100 and 1700 MMbbls. We see the result is similar to the result of TOC risking. It makes sense because as mentioned before it is the product of TOC and HI that decide the oil accumulation.

2.4. Heat flow

Heat flow is another uncertain parameter for which people commonly do risk analysis in basin modeling. We tested a range of heat flow values around the base value of 60 mW/m². Obviously at very low heat flow values there is no oil accumulation because oil is not generated. As the value increases to above 40 mW/m² the oil accumulation increases dramatically and reaches the peak at value 50 mW/m².
Then the oil accumulation starts to decrease because of secondary cracking and oil is cracked into gas. Correspondingly we see at the right side that gas starts to accumulate when heat flow is above 60 mW/m$^2$.

Figure 9: HI risking result. The estimated distribution gives P10 of 981, P50 of 1368 and P90 of 1905 MMbbls

Figure 10: Heat flow results. Oil accumulation starts at 43 mW/m$^2$ and gas generation activates at heat flow above 60 mw/m$^2$
3. Facies uncertainty

We now turn to the spatial uncertainty and first examine the impact of spatially heterogeneous facies distribution. Assume that the reservoir consists of mainly two facies, sand and shale in a channel depositional environment. From estimated channel width, wavelength and amplitude, a training image is created representing the best conceptual geological understanding of the spatial distribution. In addition a few well logs are available as hard constraints. With all these available knowledge one can build multiple facies maps for the reservoir layer using multiple-point geostatistical algorithms. We used snesim (Strebelle, 2002) to generate facies realizations from the training image, conditioned to the well data. Figure 11 shows 4 possible realizations out of the 50 realizations from snesim. All these facies maps have the same sand ratio of 50%.

Figure 11: Four facies map selected from the 50 realizations. Yellow color represents sandstone and dark blue is organic lean shale

All these models were input into PetroMod for simulation, and we get 50 different oil accumulations using the same parameters and boundary conditions as in the base case. For difference facies distribution, we get P10 of 558, P50 of 965 and P90 of 1848 MMbbls.

Figure 12: Oil accumulations for 50 realizations
To compare the impact of uncertainty in spatial distribution of facies to other modeling inputs, the mean and standard deviation are calculated and plotted shown in Figure 13. We can see that the facies distribution has similar level impact as TOC and HI. Thus it should also be considered as an important parameter for risking analysis.

![Figure 13: Mean and standard deviation for different parameters](image)

Another important factor is that besides the amount of oil accumulation, the spatial pattern of oil accumulation is also different. Four simulation results are shown below in which the green blobs represent the accumulated oil. We can see that for different facies distribution, the places that oil accumulates are very different. This obviously could be important parameters for exploration.

![Figure 14: Different oil accumulation patterns](image)
3.1. Homogeneous reservoir layer

Basin and Petroleum System Modeling is typically applied at the very early stage of exploration when the geological knowledge is limited. It is common that people will use homogeneous lithology map for the entire layer. We studied this scenario by assigning sandstone for the entire reservoir layer. And the accumulation result is shown in Figure 15. The result is surprisingly low with only 310 MMbbls of oil accumulations though it is common to expect the oil accumulation should increase with more space for trapping. However what could happen is that all the oil migration towards the high elevation area and when the trap structure is full, the remaining oil will be lost.

![Figure 15: Oil accumulation is only 310 MMbbls for a homogeneous reservoir layer](image)

3.2. Using a different training image

Different geologist gives different interpretations sometimes. Thus the training image for our MPS simulation could be different. For example it could be a delta environment rather than channel or more typically the channel wave length or thickness could be estimated differently. We tested another geological scenario which is more heterogeneous. Selected realizations of the facies distributions are shown below.

![Figure 16: Facies distribution maps using a more heterogeneous training image](image)
Again 50 simulations were performed with all other settings exactly the same as before. The result is compared against the results above. We can see that the oil accumulation distribution is quite different from the previous models. There is much more oil accumulated with the mean value of 3082 MMbbls. The variance is also higher with a standard deviation of 1006.

Figure 17: Results for a different training image is quite different from the example above. Both mean and variance are higher for the more heterogeneous scenario.

4. Structural uncertainty

Structural models for a basin are usually constructed based on depths determined from two-way seismic travels times supplemented by depths from well logs. The well data are accurate to within a few meters, but are usually available at only a few locations. Seismic travel times are usually available on a spatially extensive grid which allows an almost continuous but inexact description of the lateral depth trends. Many geostatistical methods have been developed to combine the exact well measurements and seismic travel time data to make the best prediction of the structure model and quantify the associated uncertainties. Abrahamsen (1996) compared different methods and concluded that Bayesian Kriging is one of the more suitable approaches for depth prediction because all data are included and all intercorrelations between subsurfaces and interval velocity fields are considered.

Traditionally the time-to-depth conversion is performed for each layer independently. Abrahamsen (1993) proposed an approach to integrate all subsurfaces and the estimated velocity field in one consistent model using Bayesian Kriging as described in Omre and Halvorsen (1989). The approach is briefly summarized here.
The travel time $t_i(x)$ is considered as an average of an area, so the ‘true’ travel time to the reflector $l$ is modeled with a residual as

$$T_i(x) = t_i(x) + R_i(x).$$

The interval velocity is modeled with a lateral trend and residual as

$$V_i(x) = \sum_{p=1}^{p_i} A_{l_p} g_{l_p}(x) + R_i(x),$$

where $A_{l_p}$ are the prior coefficient parameters, $g_{l_p}$ are the known regression functions which are typically interval velocities from stacking velocities or functions of interpreted travel times. The model can be formulated as

$$Z(x) = \sum_{l=1}^{L} \sum_{p=1}^{p_l} A_{l_p} g_{l_p}(x)\Delta t_l(x) + \sum_{l=1}^{L} R_i(x)\Delta t_l(x) + R_i(x).$$

where $Z(x)$ is the depth to subsurfaces. The Bayesian Kriging predictor and the corresponding variance are

$$Z^*(x) = f(x) \cdot \mu_0 + k_{\zeta}(x)K_{\zeta}^{-1}(Z - F\mu_0)$$

$$\sigma^*(x) = k_{\zeta}(x) - k_{\zeta}(x)K_{\zeta}^{-1}k_{\zeta}^T(x),$$

where

$$f(x) = \sum_{l=1}^{L} g_l(x)\Delta t_l(x).$$

$\mu_0$ is the prior mean of coefficient parameters $A$, $k_{\zeta}(x)$ is the prior variance of $Z(x)$, $k_{\zeta}(x)$ is the prior covariance between $Z(x)$ and the data vector $Z$, and $K_{\zeta}$ is the covariance matrix.

The key advantage of Bayesian Kriging is that unlike Kriging with Trend, it is stable for any number of coefficients and data, including cases without well observations (Abrahamsen 1996). This is important for basin structural modeling because it is usually done at an early stage of exploration when very few well data are available.

4.1. Depth maps

The structure of reservoir rock is an important factor for oil accumulations. We are interested to assess the impact on uncertainty in the oil accumulation due to
uncertainty in reservoir layer interpretation. A synthetic time map is generated for the reservoir layer shown below.

![Synthetic time map](image)

**Figure 18: Synthetic time map**

Multiple depth maps are then generated from the time map using the COHIBA software package that runs Bayesian Kriging. Two well points are available as hard data located at (102, 8) and (11, 21). Figure 19 shows the predicted depth maps at the top part and the corresponding error at the bottom part. The depth prediction is exact at the well location and the error is small at the near wellbore region. It is clear that with only two well points the depth can be off as large as more than 300 meters.

![Multiple depth maps and error](image)

**Figure 19: Multiple realizations of depth map and the corresponding errors**

10 realizations of depth map are generated and input into PetroMod for simulation. The result is shown in Figure 20. We see that there are differences in the oil accumulation but the uncertainty is much smaller comparing to the previous examples. One reason could be that the depth variations are localized and the overall structure is still very similar. Thus the trapping capacity does not vary too much so that the oil accumulation has smaller difference.
5. Reducing the number of simulations

We have just shown that spatial uncertainty of facies distributions is an important factor in the overall uncertainty, and one could get the output distributions by doing simulation with multiple realizations. In the example above 50 facies map realizations were studied. But the problem is that BPSM simulation is quite time consuming. For the small (120 x 30) models we have, it still requires tens of hours to do the simulation for all the models. Since spatial uncertainty in basin and petroleum system modeling is still in its infancy, modeling with ensembles of facies distributions is not yet an automated functionality in any BPSM package. This means that one has to do a significant amount of manual work. The question is can we reduce the amount of models that are required for simulation and still get a good approximation of the uncertainty profile? One approach is to use distance and kernel methods (Scheidt and Caers, 2009). Using multi-dimensional scaling (MDS), with an appropriate distance metric, multiple models are mapped to a low-dimensional space. Kernel clustering method is then used in the low-dimensional space to select a subset of realizations that are representative for the uncertainty space. We used the Euclidean distance to do MDS using the 50 facies realizations. Figure 21 shows the plot of the 50 realizations in two dimensions, and the 7 clusters from kernel K-means clustering.
Is the subset of 7 realizations a good representative of the uncertainty space from 50 realizations? Figure 22 shows the oil accumulation for the 7 selected models (blue dots) on top of all the 50 realizations (red stars). The oil accumulation at different time is plotted for each model. We can see that the 7 selected realizations span most part of the original uncertainty space though some extreme values are missed.
Figure 23 compares the pdf and cdf of oil accumulation from the subset of selected realizations to the ones obtained from all 50 realizations. Again we see a good match showing that the 7 models can reasonably capture the distribution from the 50 models.

![Comparison of pdf and cdf](image)

Figure 23: Comparison of 7 models from kernel K-mean clustering to the 50 prior models shows a good match for estimated pdf and cdf

6. Conclusions and future work

Uncertainty is an intrinsic problem in Basin and Petroleum System Modeling because of the large spatial and temporal scale of the problem. The traditional methodology is to do Monte Carlo simulations on the input parameters. This is done in the study as the benchmark. The spatial uncertainties including facies and structure are not studied before. We applied the geostatistical methods to generate multiple realizations of facies and structure map. The corresponding uncertainty in oil accumulation and oil distribution is assessed. We showed that facies distribution has a great impact on the simulation results and different geological interpretation could lead to very different results. It is equally if not more important than the typical modeling inputs such as TOC or HI. Basin modeler should not only identify the lithology type but also the spatial correlation among them. Structure uncertainty on the other hand does not have as great impact in this example as we expected. One
reason is that we only considered the uncertainty in time-to-depth conversion assuming the seismic picks are perfect. One important future work is to investigate different time horizon interpretations. Another thing is that in this work we only studied 2D facies map due to the limitation of the software package. The impact of 3D facies map is definitely first thing to analyze when the software and technology becomes available.

7. Acknowledgement

The authors would like to thank Ken Peters, Les Magoon, Allegra Scheirer and Oliver Schenk for their help. Special thanks to Petter Abrahamsen for providing the COHIBA software package. We would also like to thank Schlumberger for the PetroMod license.

8. References


