Applying the distance-based approach to model flow uncertainty in structurally complex reservoirs: initial investigation

Aaditya Satija and Jef Caers
Stanford University

Abstract
Uncertainty in structural models is complex to quantify. Usually, only a limited number of structural models can be used in inverse problem workflows. This leads to an incomplete communication of the uncertainty represented by the available data. For facies models however, there are experimental design methods that can be used to approximate the underlying uncertainty well. Distance-kernel methods use distances between realisations in experimental design workflows to assess uncertainty in flow response. In this paper, distance kernel methods are being applied to structural models. However, the nature of uncertainty in the parameters of structural models needs for these parameters to be characterised. Parameters of the structural models are characterised using their sensitivity to the difference in approximation of flow response using streamline simulation. Moreover, the scenarios of well placement and fault geometry where the parameters are significant are presented.
1. **Introduction**

The structural model of a hydrocarbon reservoir is a common model for geological faults and horizons. In addition to forming the basis for volumetric calculations, well planning and reservoir grids, it forms the geometrical framework for the 3D grid providing the boundaries for the facies models that describe rock properties. Uncertainty in the structural model can have far-reaching consequences on the decisions the model is used for.

The effects of structural model parameters on the flow response are indirect, non-linear and complex. Moreover, the spatial extent of structural artefacts like faults and horizons is narrow. As a result, studying these effects is a more complex problem. However, a complete treatment of the uncertainty in the structural model is often neglected in most reservoir studies due to methodological, computational or algorithmic limitations of current methodologies.

The uncertainty in the parameters of facies models has been quantified using stochastic and deterministic parameters that in turn can be modeled using ranking (Saad 1996) and experimental design (Manceau 2001) methods respectively. Distance Kernel Methods (Scheidt 2007) use experimental design techniques to define a realization-based model of uncertainty which is in turn parameterized by a distance metric. As a result, estimates of quantiles of the response accurately with few flow simulations are obtained.

We are applying the Distance Kernel Method to structural models. We weren't able to apply the entire workflow to the structural models. To do this, we needed a characterisation of parameters of structural models and well-positions. We studied how these parameters relate to streamline simulation flow response of those models. We have studied multiple scenarios of well placement to identify the scenarios where this approach can be applied reliably.

2. **Methodology**

2.1. **Distance Kernel Methods**

The methodology aims to efficiently select a subset from a potentially large set of realizations such that the response of the subset exhibits the same statistical properties as the entire set of realisations. Since the transfer function for the realizations can be very
complex computationally, the methodology seeks to minimise the number of evaluations of the transfer function of the realisation. For a binary categorical variable, the methodology is illustrated in Figure 1.

A distance between two realizations is defined as a scalar that determines how dissimilar two realisations are. Starting with $n$ realizations, an $n \times n$ distance matrix is calculated by taking the distance between two of the $n$ realizations at a time. Usually, realizations can be added to the model set for a more complete representation of prior uncertainty. Multidimensional Scaling (Borg 1997) is used to map all the $n$ realisations to $n$ points in a Euclidean space $R$. Each point in this space represents a model. Its distance from another point corresponds to the dissimilarity distance between the realizations. However, in most cases, the points in $R$ are not linearly mapped. Kernel methods (Schöelkopf 2002) are used to transform the Euclidean space $R$ into a feature space $F$ (Scheidt 2007) such that the points in $F$ behave more linearly to facilitate the use of linear pattern detection and model

---

**Figure 1** Workflow for generic distance-kernel method (a) distance between two models (b) distance matrix (c) MDS in Euclidean Space (d) MDS in Feature Space (e) pre-image (f) Estimates of uncertainty characteristics (Scheidt 2007)
selection tools such as *Clustering algorithms* that divide the feature space $F$ into subsets of points that are more similar to others in the same subset than points outside the subset. Kernel k-means algorithm (Dhillon 2004) tries to assign points in $F$ to subsets by minimizing the expected distance between the points in a cluster and its centre.

Distance Kernel Methods (Figure 1) use the centers of the clusters as the selected typical points to use for uncertainty qualification. Once these points have been selected, they form a basis for a response surface analysis akin to the traditional experimental design for linear relationships. From this (proxy) response surface analysis, a good estimate of the characteristics of the uncertainty in the model parameters is ascertained.

### 2.2. Parameters in Structural Models

Structural models are parameterised as sets of surfaces that in turn have very narrow uncertainties associated with them. These parameters in turn are related to each other in complex manners. The uncertainty of these complex parameters expresses in the realizations built from these parameters and further, in responses predicted by 3DSL from those realizations. The transfer function between structural model parameters and flow responses is high dimensional and complex. As a result, Distance-kernel methods can’t directly be used to identify representative models from distances based on parameters. So, we need a way to characterise parameters and identify those that are more important to in turn, characterise the uncertainty.

The parameters of faults in structural models can be classified as:

- *Number of Faults*
- *Fault Orientation*: strike and dip angles
- *Fault Placement*
- *Fault Hierarchy* (Cherpeau 2009)

To use the Distance Kernel Methods to characterise the parameters for structural models, the parameter distances need to be measured. A summary of the distances used is given in Figure 1. As is evident, the parameters are very different in range and type. An intuitive approach to use distance-based methods would be to use combinations of these distances. However, the aforementioned differences in range and type of the underlying parameters pose serious challenges in normalising these distances. For example, the fault hierarchy is not even an ordered parameter.
### Table 1 Description of Parameter Distances

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Type</th>
<th>Upper Bound</th>
<th>Lower Bound</th>
<th>Distance Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strike Angle</td>
<td>Continuous</td>
<td>180</td>
<td>0</td>
<td>Norm of difference of vectors of strike angles of faults in hierarchy</td>
</tr>
<tr>
<td>Fault Hierarchy</td>
<td>Categorical</td>
<td>n/a</td>
<td>n/a</td>
<td>Tree Distance (Satija 2010)</td>
</tr>
<tr>
<td>Dip Angle</td>
<td>Continuous</td>
<td>90</td>
<td>0</td>
<td>Norm of difference of vectors of dip angles of faults in hierarchy</td>
</tr>
<tr>
<td>Number of faults</td>
<td>Discrete</td>
<td>n/a</td>
<td>1</td>
<td>Difference in number of faults</td>
</tr>
</tbody>
</table>

Since fluid flow measurements like water saturation at the producer well or the total oil production are strongly dependent on the placement of faults and wells, distances that use well positions and fault placements also need to be defined in addition to the parameters. The details of these distances are in Table 2.

### Table 2 Description of Placement Distances where $f_i$ is the vector of fault origin points in model $i$, $w_{inj}$ is the injector well position, and $w_{pro}$ is the producer well position

<table>
<thead>
<tr>
<th>Distance</th>
<th>Description</th>
<th>Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placement Distance One</td>
<td>How far are faults in one model from faults in the other model?</td>
<td>$</td>
</tr>
<tr>
<td>Placement Distance Two</td>
<td>Difference in how far are faults from the injector in the two models</td>
<td>$</td>
</tr>
<tr>
<td>Placement Distance Three</td>
<td>Difference in how far are faults from the producer in the two models</td>
<td>$</td>
</tr>
<tr>
<td>Placement Distance Four</td>
<td>Difference in how far are faults from the a well in the two models</td>
<td>$</td>
</tr>
</tbody>
</table>
Once the distances are available, distance matrices are constructed. Using MDS Techniques (§2.1), these distance matrices are used to represent the underlying models in first, a metric space and then, a feature space. Models in the feature space are clustered into control levels. This can be separated into subspaces that in turn can represent control levels to use in Central Composite Design. For example, if it is intended to use a two degree experimental design technique to identify the quantiles to represent uncertainty, the feature space $F$ is divided into three clusters that in turn represent the three control levels to be used for corresponding Central Composite Design to study the effects of the underlying parameter and its interactions with other parameters on the flow response. This eliminates the problem of normalizing discrete and categorical parameters.

2.3. Assessing Distances
To assess the utility of the parameters of a structural model, their individual sensitivity to the flow response are approximated using 3DSL simulations is examined.

![Figure 2 Workflow for flow-simulating the structural models](image)

To do this, structural models are represented in SKUA a set of fault surfaces. Assuming a common volume of interest across multiple models to be the convex hull of the fault surfaces, the resultant volume is gridded for each fault model. This requires carefully
calculating fault throws and intersections with other faults for each fault in each model. For most fault geometries, it is computationally very complex to grid the volume.

Once the surfaces are represented as a gridded volume, the volume itself is modelled as a layer-cake model of homogeneous petrophysical properties to control for the stochastic nature of these properties in a real case. A transmissibility multiplier associated with the faults effectively represents the surfaces of a structural model as regions of heterogeneity in the respective flow simulation grid. Next, a set of well placements and completions is decided upon and the grids are simulated using 3DSL. The flow response can be any relevant measurement that can be used in a broader inverse problem workflow based upon the application e.g. total oil production, the water saturation at production well etc.

First order sensitivity analysis techniques are then used to test the sensitivity of the response to the control levels. Once all the parameter values are separated into control levels, Classical Sensitivity Analysis can be used on the coefficients calculated from Central Composite Design for multiple parameters to identify the more significant parameters or interactions thereof for the flow response from a Pareto plot. A higher value of the sensitivity factor represents a stronger and more 'linear' relationship between the parameter and the flow response. The level of the sensitivity allows for the choice of parameters, or a combination thereof to use in the Distance Kernel algorithm.

3. Problem Setting

To apply the proposed methodology to a geological scenario, a synthetic test-case was created. Using the FaultMod plugin developed at the gOcad group at the Nancy School of Geology, 50 structural models were generated. Each of these models had four faults. Other than the number of faults, parameters in a structural model being modeled by FaultMod can be summarized under three categories:

- **Fault Orientation** (Figure 3) is simulated from distributions of strike and dip angles. For this case, realistic dip angles were limited to a uniform distribution between 80° and 90°. The strike angles however, were allowed their full uniform range between 0° and 180°.
• *Fault Placement* is restricted by specifying a fault zone in FaultMod. This restricts seed-points of the fault surfaces in the horizontal space to a zone in the center of the model.

• *Fault Hierarchy* (Cherpeau 2009) can be extracted from a model on structural simulation. When there are four faults, there are three fault hierarchies possible (Figure 4).

![Figure 3 Fault Orientation in terms of Strike and Dip angles](image3)

![Figure 4 Three Network Hierarchies possible with three faults. Red Triangles are fault nodes and Blue rectangles are fault block nodes (see Cherpeau, 2011).](image4)

These 50 models were then grid as $50 \times 50 \times 25$ using SKUA assuming layer-cake models of petrophysical properties of 25% porosity and a permeability of 100 md to control for effects of difference in facies on flow simulation. A transmissibility multiplier of 10% for all faults was used to control for differences in the fluid flow characteristics of the fault surfaces. This allowed for the isolation of the effects of structural parameters on fluid flow. Lastly, we used wells on the major diagonal of the model. Having this well configuration
allowed to account for the effects of all the faults in the model on fluid flow between the injector and the producer. These models were simulated for 9000 days using 3DSL to obtain the volumes and rates of production during the simulation period.

The parameter distances were measured. A summary of the distances used is given in Table 3. The treatment of placement distances is described in §4.

Table 3 Description of Parameter Distances used in Problem Setting

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Type</th>
<th>Upper Bound</th>
<th>Lower Bound</th>
<th>Distance Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strike Angle</td>
<td>Continuous</td>
<td>180</td>
<td>0</td>
<td>Norm of difference of vectors of strike angles of faults in hierarchy</td>
</tr>
<tr>
<td>Dip Angle</td>
<td>Continuous</td>
<td>90</td>
<td>80</td>
<td>Norm of difference of vectors of dip angles of faults in hierarchy</td>
</tr>
<tr>
<td>Fault Hierarchy</td>
<td>Categorical</td>
<td>Three unique fault hierarchies</td>
<td>Tree Distance (Satija 2010)</td>
<td></td>
</tr>
</tbody>
</table>

Once the distances were available, distance matrices were constructed. Using MDS Techniques (§2.1), these distance matrices are used to represent the underlying models on first, a metric space and then, a feature space. Models on a feature space were clustered into control levels. First order sensitivity analysis techniques are then used to test the sensitivity of the response to the control levels. The level of the sensitivity allows for the choice of parameter distances, or a combination thereof to use to create the proxy response surface.

4. Effects of placement

Amongst the distances in Table 2, a lot of the information carried by the distances can be made redundant. To narrow the number of distances down, a scenario was chosen such that

- All the models had the same number of faults
- All the models had the same fault hierarchy
- All the models had a very narrow distribution for strike and dip angles. The narrow strike distribution primarily applies to the top fault in the hierarchy with the daughter faults having a wider strike distribution available.
- The wells were on opposite diagonal corners of the reservoir model to allow for maximum interaction of the fluid flow with the faults.

This scenario was akin to making the parameter distances (Table 3) very low values. Some models from this scenario are seen in Figure 5.

![Figure 5 Selected Structural Models from Scenario (strike restricted from 40 to 140)](image)

Now, the individual parameter ranges were separated into many control levels. Since clustering is not deterministic, using 7 and 15 control levels on placement parameters as explained in Table 2 led to one-way sensitivities of streamline simulation flow response to the placement parameter as seen in Figure 6 and Figure 8. Not surprisingly, a multiway analysis of these sensitivities leads to the conclusion that the response is more sensitive to Placement One than Placement Four. A higher sensitivity is a measure of a more ‘linear’ relationship between the streamline simulation flow responses and the underlying structural model parameters. To capture the characteristics of the Placement distances, Placement One and Placement Four were chosen for further examining interactions with structural parameters. Placement One represents the positions of faults relative to each other and Placement four represents the positions of faults relative to wells.
Figure 6 One way sensitivity of oil production to Placement One with 9 control levels

Figure 7 One way sensitivity of oil production to Placement Four with 9 control levels: lower sensitivity.
Figure 8 One way sensitivity of oil production to Placement Four with 15 control levels: lower sensitivity

Figure 9 Multiway sensitivity analyses of placement parameters and their interactions on the total production response of the models
5. Examining effects of structure and placement

Once placement parameters suitable for calculating distances have been selected, the effect of placement in conjunction with that of structural parameters such as fault orientation and hierarchy can be studied.

5.1. Fault Strike

Another dataset of structural models was simulated with a wider distribution of strike angles than in §4 to allow for larger distance measures for strike distance. Using exactly the same method taking into account the fault orientation and placement distances for MDS and clustering, control levels for fault orientation and placement were developed. The sensitivity of total oil production to fault strike is represented in Figure 10. A comparison of the sensitivity to fault strike and fault placement using the parameters selected in §4 is seen in Figure 11.

Interestingly, the interaction of the placement of faults with the fault strike is more sensitive than the placement interactions on their own. Also with a wider strike distribution, the faults can act as both barriers and conduits between the injection and production wells. As a result the Placement Four parameter that depends upon the placement of the fault relative to the wells becomes more important in this case.

Figure 10 One way sensitivity of oil production to fault strike
5.2. Fault Hierarchy
A dataset of structural model was simulated to have all three fault networks in addition to the conditions in §Fault Strike5.1 to allow for tree distance to be measured. Since there are only three fault networks possible (Figure 4), the analysis is done with three control levels. A comparison of the sensitivity values is in Figure 12.

The interaction of the fault hierarchy with the strike is just about as sensitive as the effect of the fault hierarchy on its own. In general, the response of models with the first network in Figure 4 with only one fault at every level tends to be lower oil production in all. This can be due to earlier water cut with the injected fluid getting finding conduits in the faults one after the other.
6. Case for exception

However, there are situations in which the suggested scheme to characterise parameters is not as effective. To model one of such situations, a scenario of models with four faults that is the same as the case in §5 other than the following:

- The dip angles of the faults in the models, instead of being distributed in the narrow range between 80 and 90 degrees, are distributed uniformly between 0 and 90 degrees.
- The wells instead of being placed at the diagonal corners of the reservoirs, are in the interior
- The response considered instead of being the total production of the reservoir is the saturation of water in the production well.
When the same multiway sensitivity analyses as §4 are performed after clustering the parameter space into control levels, the sensitivities are found to be really low in the order of magnitude of $10^{-15}$ as seen in Figure 15. This is likely to have happened since the unrealistically low dip angles caused the formation of almost horizontal faults as in Figure 13 that provided ready conduits between the injector and the producer wells. So, it can be asserted that there are cases where there is little relationship between any of the parameters of a structural model to the streamline simulation flow response. The flow responses are due to a “stroke of luck”, not because of any significant parameter effect. In this case, the scenario is a poor candidate for use of Distance Kernel Methods.
Figure 15 Multiway sensitivity analyses of placement and parameters on the streamline simulation flow response of the models. When the wells are too close to each other in the presence of horizontal faults, there is very little sensitivity.

7. Conclusions

The methods proposed by (Scheidt 2007) for characterizing the uncertainty in facies models can be used to characterize the uncertainty in structural models when the distances used for the finding the representative models to build a response surface have a good correlation with the difference in flow responses of the underlying models.

In structural models, the streamline simulation flow response is sensitive to well placement, fault orientation and fault placement parameters. In this paper, we identify fault orientation to be more of an influence on the streamline simulation flow response than placement parameters. There are scenarios identified based upon well placement and fault geometry (hierarchy and orientation) to use in Distance Kernel Methods. However, there are also scenarios identified where these parameters do not form the basis for a good distance to use in Distance Kernel Methods, particularly when wells are close and too many random effects impact flow.

More work is needed to identify a generic rule as to classify the appropriateness of scenarios for Distance Kernel Methods for uncertainty characterization based on objective measurable criteria such as well placement and fault geometry. In an appropriate scenario, the sensitive parameters can be used to select ‘representative’ realisations on which we can
run a commercial flow simulator like Eclipse and obtain the statistical quantiles for a good representation of the uncertainty as described in (Scheidt 2007).

**Bibliography**


