Using Metric Space Methods to Analyse Reservoir Uncertainty

Darryl Fenwick
Streamsim Technologies, Inc., Palo Alto, CA, USA
fenwick@streamsim.com

Rod Batycky
Streamsim Technologies, Inc., Calgary, Alberta, CA
batycky@streamsim.com

Abstract

In mathematics, a metric space is a set where a distance (called a metric) is defined between elements of the set. In this paper, we introduce the concept of a metric space in the framework of reservoir modelling and reservoir uncertainty. The distance between two models is a single measure that can be easily understood by the reservoir team, and can be tailored to the application of interest. We describe how placing an ensemble of reservoir models in metric space allows for novel methods for model visualization and analysis. Example applications are presented in the context of uncertainty quantification, sensitivity analysis, and history matching. Although established methods exist in these domains, placing the reservoir models in metric space allows for a complementary approach which has several advantages compared to traditional methods.

Introduction

Metric space methods have been employed for decades in various applications, for example in internet search engines, image classification, or protein classification [1]. However, in petroleum reservoir modelling, metric space methods are not widely known, and few applications of metric space methods have been presented. This is due most likely to the fact that, even now, a significant majority of reservoir modelling studies performed build a single reservoir model which is used for analysis and decision making.

Metric space methods are best employed to quickly analyse and interpret a group (ensemble) of reservoir models, and are thus attractive methods for uncertainty studies or sensitivity analysis when a large ensemble of reservoir models is needed. The creation of a metric space requires the definition of a dissimilarity distance, which measures the dissimilarity between two different models. The distance measure has two main requirements. First, since the distance must be calculated between each model pair, it should be rapid to calculate for large ensembles of models. Second, the distance measure must be tailored to the problem which is being addressed. No single distance measure is applicable to all situations.

In applications that involve flow in the subsurface, streamline-based flow simulation (SLS) has been shown to be an appropriate tool for measuring distances [3-6]. SLS follows the first requirement for a flow-based distance measure in that it is usually very CPU efficient compared to standard grid-based flow simulations. SLS can capture first order flow effects like well locations, fractional flow, flow barriers, reservoir heterogeneity, and production history.

We illustrate the metric space methods using an example application of a reservoir model shown in **Figure 1**. The model is composed of 43,000 active grid blocks, with over 100 producing wells, supported by over 20 water injectors, with 9000 days of production . 72 reservoir models have been created, varying the spatial correlation length of the reservoir properties (low,high), angle of correlation (45,90,135 degrees), Kv/Kh ratio (0.1, 0.01, 0.001), transmissibility between the upper and lower layers (0.001, 1), and the residual oil saturation (0.2, 0.3). Note that these parameters have been selected for illustrative purposes – there is no restriction on the number of parameters, nor on the type of parameters that can be used when employing metric space methods.

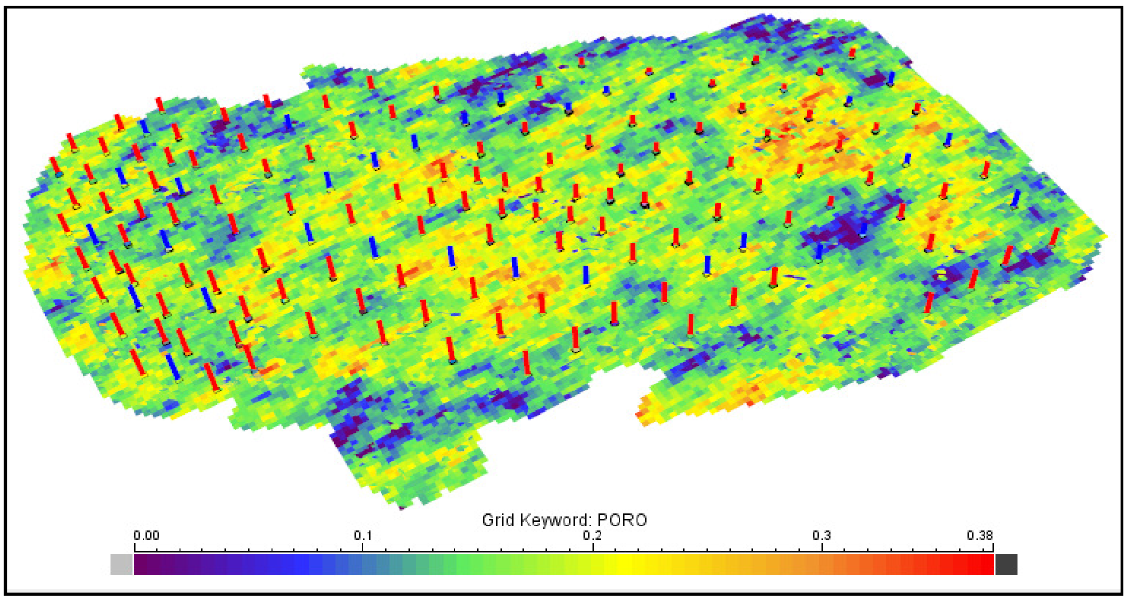


Figure 1: A realization of the porosity for the reservoir model, with the producers in red, and injectors in blue. The well configuration shows all open wells at 6000 days of production.

A typical distance measure δ_{ij} between model i and model j may be the sum of the absolute difference of the flow simulations of a particular reservoir response over the N_{ts} time steps of the simulation, such as the oil field production rate q_o.

$$\delta_{ij} = \sum_{ts=1}^{N_{ts}} \left| (q_{o,ts}^{i} - q_{o,ts}^{j}) \right|$$

This measure requires a flow simulation for all 72 models (determining 2556 unique δ_{ij} values), which may not be feasible for large models using traditional flow simulation software. However, in this case, streamline simulation is efficient, and flow responses are calculated in a little over an hour. Another example is the absolute difference of oil production for each well over the time steps of the simulation, which in this case requires a double sum, one over the time steps and another over the number of wells (N_w).

$$\delta_{ij} = \sum_{ts=1}^{N_{ts}} \sum_{k=1}^{N_w} \left| (q_{o,ts,k}^{i} - q_{o,ts,k}^{j}) \right|$$

As an illustration, the field oil production rate over the entire simulation time will be used, which is shown for the 72 reservoir models in **Figure 2** below.

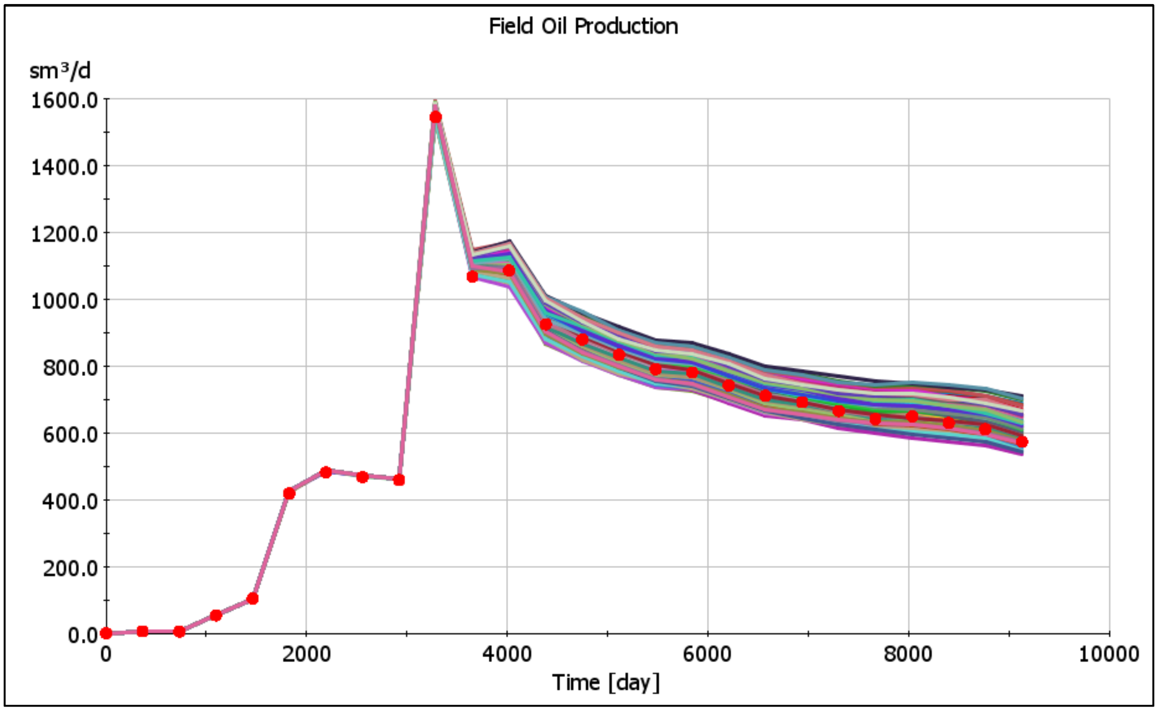


Figure 2: Field oil production rate for 72 reservoir models. The dots indicate the historical data.

The distance measures between pairs of runs form a dissimilarity distance matrix, shown schematically in **Figure 3**. Note that the matrix is symmetric, and the diagonal of the distance matrix is zero by definition (distance between the model and itself is zero). The dissimilarity distance matrix defines the metric space.

A consequence of using a flow response in the distance calculation is that, if history exists for the reservoir, the simulation data can be directly compared to historical data. In this case, an additional distance is calculated between the simulation data and the historical data, adding a new column and row to the distance matrix, representing the unknown, “true” reservoir.

	model 1	model 2	model 3	model 4	...
model 1	0	δ_{12}	δ_{13}	δ_{14}	...
model 2	δ_{21}	0	δ_{23}	δ_{24}	...
model 3	δ_{31}	δ_{32}	0	δ_{34}	...
model 4	δ_{41}	δ_{42}	δ_{43}	0	...
...	...	...	...	...	0

Figure 3: Example of a dissimilarity distance matrix. The diagonal elements of the matrix are zero (the dissimilarity between a model and itself is zero).

The dissimilarity distance matrix can be difficult to analyse for a large ensemble of models. A useful tool for visualizing the metric space is multi-dimensional scaling (MDS) [2]. MDS transforms the reservoir models from the metric space into a Euclidean space where the models can be visualized. Below, we denote this Euclidean space as "MDS space". **Figure 4** displays the 72 reservoir models in MDS space.

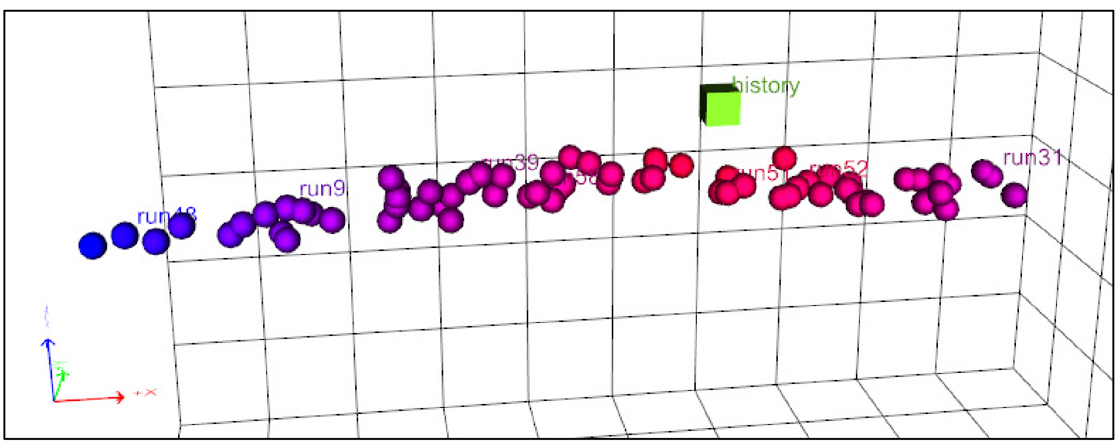


Figure 4: 72 models placed into MDS space using the field-level oil production rate distance measure. Each point in the space represents a reservoir model. Note as well that the historical field oil production rate can also be displayed in the same space (cube), even though no reservoir model can be associated with it.

MDS will reproduce the distances in the distance matrix as close as possible. However, depending on the distance measure, it is not always possible to reproduce the distances exactly. However, the metric space is defined by the distance between elements of the set. The absolute location of the points in the set is unimportant. Consequently, the axes in MDS space have no units and have no intrinsic meaning – what is important is the relative distance between the models. This is an important source of confusion to avoid. For more information on MDS, see Borg and Groenen [2] or in the context of reservoir modelling, Scheidt and Caers [3,4].

Note in **Figure 4** that MDS has in essence reduced 72 high dimensional reservoir models into a low dimensional space, where the distances between each model represent the difference in field-level oil production. **Figure 4** is thus an alternative view of the variability of the reservoir response in **Figure 2**. When the historical data is included in the dissimilarity distance matrix, the location of the history, representing the “true” reservoir can also be displayed in the MDS space (seen in **Figure 4** as a cube), even though no reservoir model can be associated with it.

In addition to MDS, clustering techniques are also employed to group reservoir models which are similar, and separate models which are dissimilar (as measured by the distance). The clustering can be performed directly in metric space (using the k-medoids algorithm, for example), or in MDS space (for example, using the k-means algorithm).

Error! Reference source not found. shows the 72 reservoir models and the historical data placed into 7 clusters using the k-medoids algorithm. Many of the metric space methods employ clustering algorithms to group similar models, and select subsets of dissimilar models (for example, by selecting the centroids of the clusters).

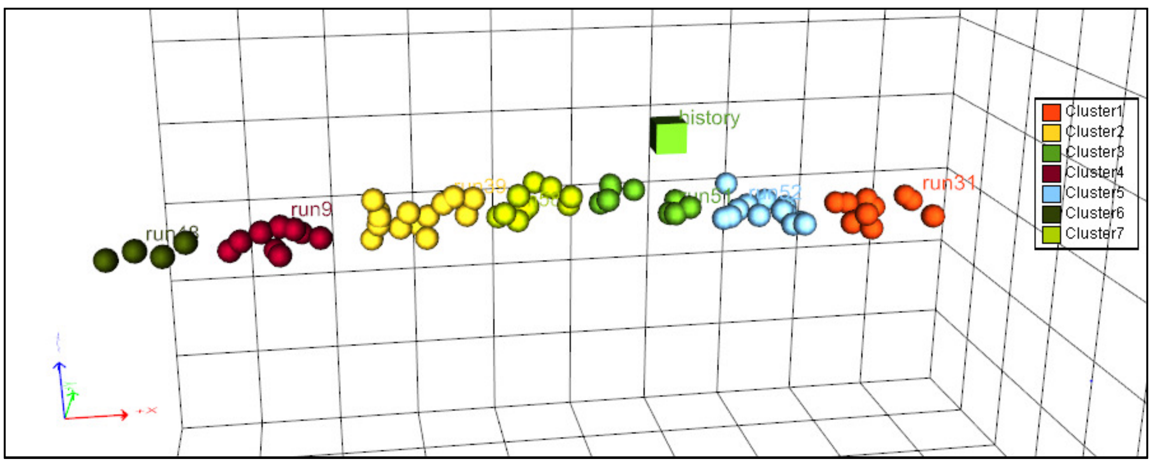


Figure 5: The same reservoir models shown in Figure 4, grouped using k-medoids into 7 clusters. The historical data is in the 3rd cluster, shown as a cube.

For field-level responses, there may not be much advantage in understanding reservoir model ensembles in MDS space compared with field production plots (**Figure 2**). However, there may be a significant advantage when analysing well-level responses. As an example, a metric space is created with the 72 reservoir models using a well-based distance measure (oil production rate), shown in **Figure 6**. Note how this figure shows more of a spread in the distances, than the field-level distance measure of **Figure 5**, a reflection of more variability if flow results on a well-by-well basis. In this case, 8 clusters are created, one of which contains the “true” reservoir**.** For well-level analysis, the reservoir model ensemble using **Figure 6** may be easier than examining production data individually for each well, considering the large number of wells in the model.

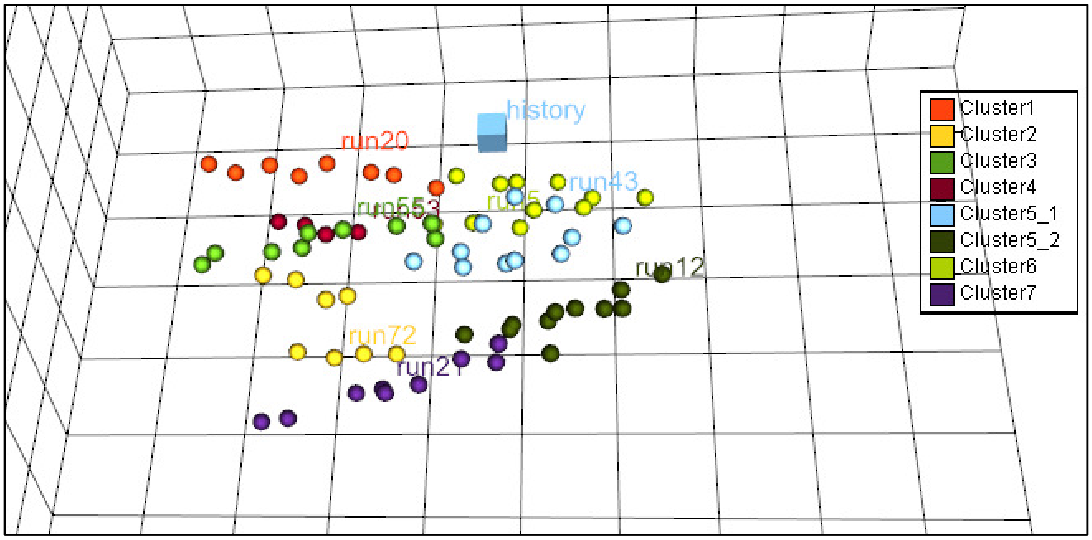


Figure 6: Visualization of 72 reservoir models in MDS space with historical data (cube) clustered into 8 groups. The distance measure is the absolute difference of the well-level oil production rate.

Plotting the reservoir models in MDS space provides interesting insight into the relationship between the reservoir models and the historical data. Ideally, we would like to find the "true" reservoir within the cloud of reservoir models (models "bracket" the data). This would indicate that the parameterization of the model (i.e. the prior model) is sufficient to find one or more models which match history. When the unknown, "true" reservoir is found far from the ensemble of reservoir models, this indicates that the model parameterization is insufficient, or indeed incorrect, strongly suggesting that the reservoir modelling team re-examine the assumptions which go into the creation of the reservoir models. This analysis in metric space may prevent reservoir engineers from undergoing the time-consuming process of well-level history matching. Note that this analysis goes beyond simply examining the values of the objective function for the 72 models – it is the spatial relationship of the reservoir models and the true model in multi-dimensional MDS space which provides the important information. For our example application (**Figure 5** and **Figure 6**), the history is located within the cloud of reservoir models in the two largest principle axes. Although not perfect, this is acceptable for the sensitivity analysis given below.

This section offered a quick overview on how metric spaces are created, and described the basic tools that are used to visualize and analyse reservoir models in metric space. For further details, see References 3-7.

Application of Metric Space Methods in Reservoir Modelling Workflows

In this section, we present some example applications which employ the distance measures, MDS, and clustering methods described above. The common denominator of all these methods is the ability of the distance measure and clustering to identify models which are similar, and models which are dissimilar with respect to the desired response.

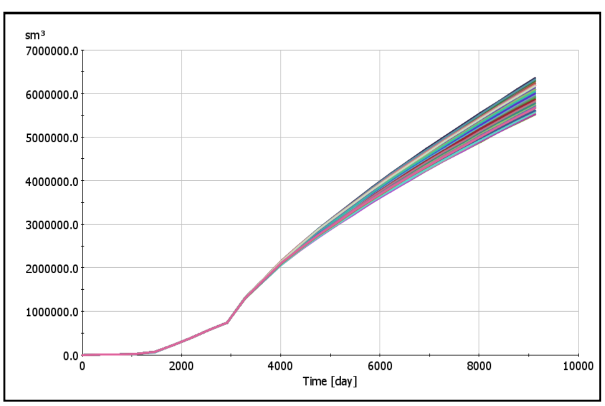
Uncertainty Quantification. Scheidt and Caers [3,4] apply metric space methods for quantifying the variability in flow response. They use streamline simulation to compute the flow-based distance measures, and clustering using kernel k-means. The reservoir models closest to the cluster centroids are selected for evaluation with finite-difference (FD) flow simulators. The response from FD flow simulation estimates the uncertainty of the reservoir response of interest (cumulative oil production at a given time). The weights of the responses are determined by the size of the cluster associated with each model. Comparison with other methods of uncertainty quantification show that using SLS-based distances, metric spaces and clustering provides a more accurate quantification of uncertainty when compared to the uncertainty determined by FD flow simulation of the full set of models.

Scheidt and Caers show that flow-based distance measures from SLS are preferable to static-based measures when possible. When flow-based distances are not feasible static based measurements can be used.An alternative static-based distance is the sum of the absolute difference over the N_{gb} grid block properties (rather than global properties) between reservoir models, such as the local grid block pore volume.

$$\delta_{ij} = \sum_{k=1}^{N_{gb}} \left| (\mathrm{PORV}_k^i - \mathrm{PORV}_k^j) \right|$$

This grid-block pore volume-based distance measure may be appropriate in cases where the reservoir models have local variations in reservoir volume (spatial uncertainty), and for convective displacements such as waterfloods where variation in local properties such as porosity, permeability, and net-to-gross will have a significant impact on production behaviour.

In the example case given here, variation in local grid-block pore volume is an excellent distance measure as demonstrated in **Figure 7**The cumulative oil production for all 72 models is adequately captured by the 7 models that were selected from the centroids of the clustering in metric space based on the distances computed using the above local grid block pore volume



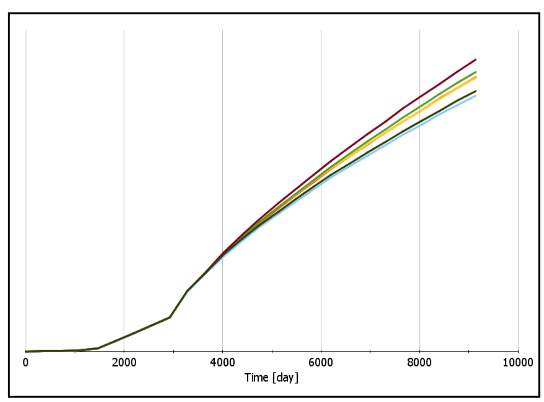


Figure 7: Comparison of cumulative oil production from the 72 models (left) and the 7 models (right) representing the centroids of 7 clusters based on a static distance local pore volume measure, The variation of cumulative oil production is very similar.

Sensitivity Analysis. Sensitivity analysis studies the variation in the output of a reservoir model response which is attributable to variations in the model inputs. Standard methods for sensitivity analysis are very powerful [8,9], yet traditionally have difficulty analysing sensitivities to parameters which have only discrete values, such as different facies proportion cubes, structural interpretations (different fault scenarios), as well as responses which have a stochastic component (spatial uncertainty). Metric space methods offer a complementary approach to standard sensitivity analysis methods which avoid such limitations. This suggests that metric space methods for sensitivity analysis may be useful upstream of standard methods.

A common use of streamline simulation is to measure reservoir connectivity. One measure which quantifies reservoir connectivity is the flux-pattern map (FPmap), shown for two reservoir models in **Figure 8**. The FPmap quantifies injector/producer pairs in terms of reservoir fluxes between the well-pairs and is used for reservoir flood surveillance [10]. The thickness of the line indicates the volume flux of the injector/producer pair. In an example using the connectivity distance, we analyse the sensitivity of reservoir flux for the five parameters in the set of 72 reservoir models. In this case, we looked at the differences in FPmaps of all models, just using the latest timestep. One additional advantage of this distance measure is that although it is flow-based, it does not require the expensive transport step of a flow simulation. Only the computation of the streamline paths and bulk flow between well-pairs for the timesteps of interest are required, giving fast surveillance flow simulations.

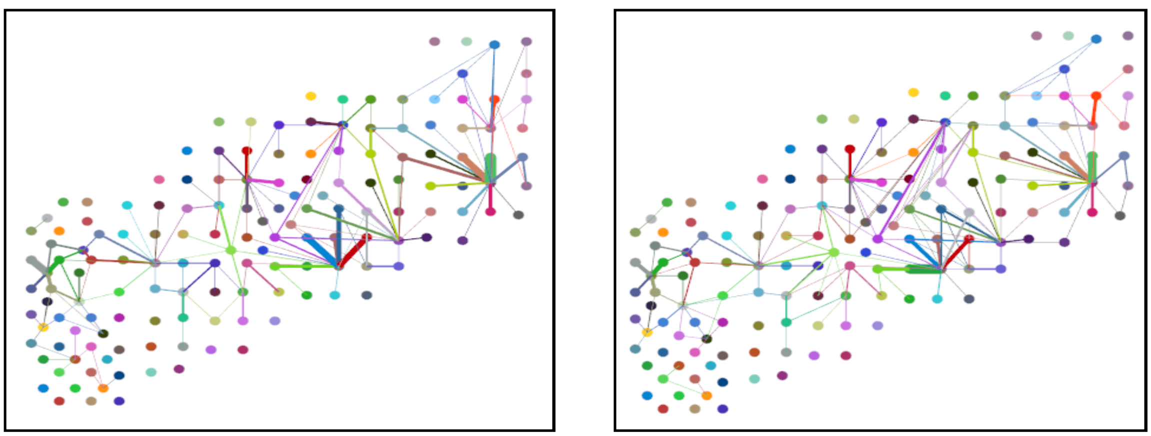


Figure 8: The flux-pattern maps (FPmaps) at a given time step of two reservoir models in the example set of 72 models.

After calculating the connectivity-based distance matrix for the 72 reservoir models, MDS is performed and displayed in **Figure 9**. The separation of the colours in **Figure 9** indicates the impact of the parameter on the variability of the reservoir model connectivity, hence giving a qualitative understanding of the impact of various parameters on the model variability (as defined by the distance). In this case, the variogram angle has a strong impact on the injector/producer well connectivity, compared to the Kv/Kh ratio, which is consistent with expectations.

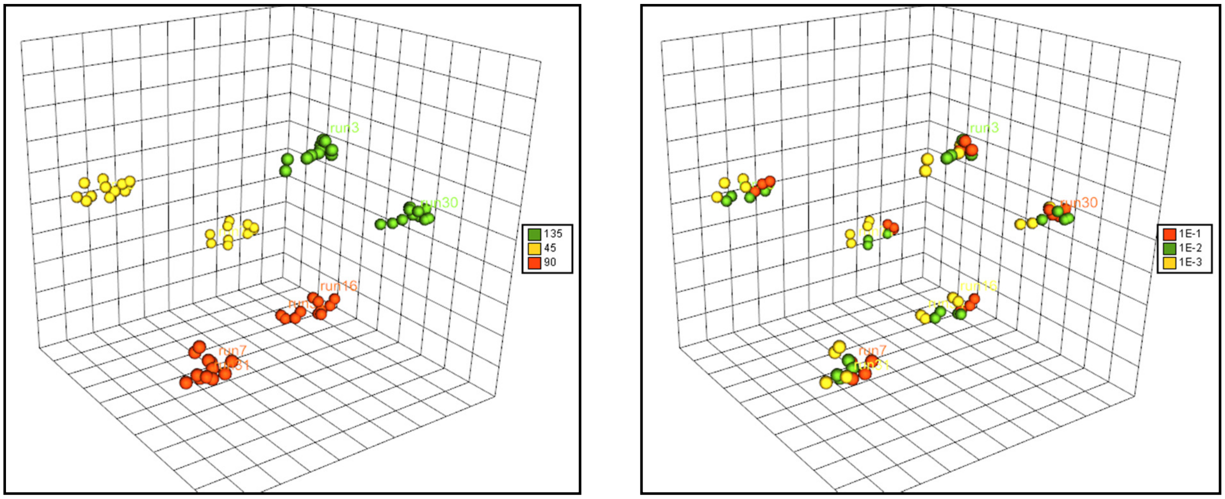


Figure 9: MDS space plots of 72 reservoir models, coloured by variogram angle (left), and Kv/Kh ratio (right), derived from a metric space using the FPmap-based distance measure.

More quantitative sensitivity analyses are available when the models are clustered in metric space or MDS space. Below, two examples are presented in the context of history matching.

When historical data is included in the distance matrix, clustering identifies a subset of reservoir models close to the “true”, unknown reservoir. Analysing the reservoir models in the subset leads to understanding which parameters have an impact on the history match, and which parameter values offer an improved history match.

In the example using the field production, the MDS space for the field-level oil production distance measure was previously shown in **Figure 5**, where the 3rd cluster contains the historical data. Examining only the runs in this cluster, we see in **Figure 10** that the models in the cluster match the oil and water production very well. Interestingly, when examining the parameters for these runs, we see a diversity of values for the models in the cluster (right side of **Figure 10**). Aside from KvKh, there does not appear to be significant influence of a particular parameter for the field-level match. This also implies that we could have a diversity of forecasts, preserving uncertainty in the forecast process.

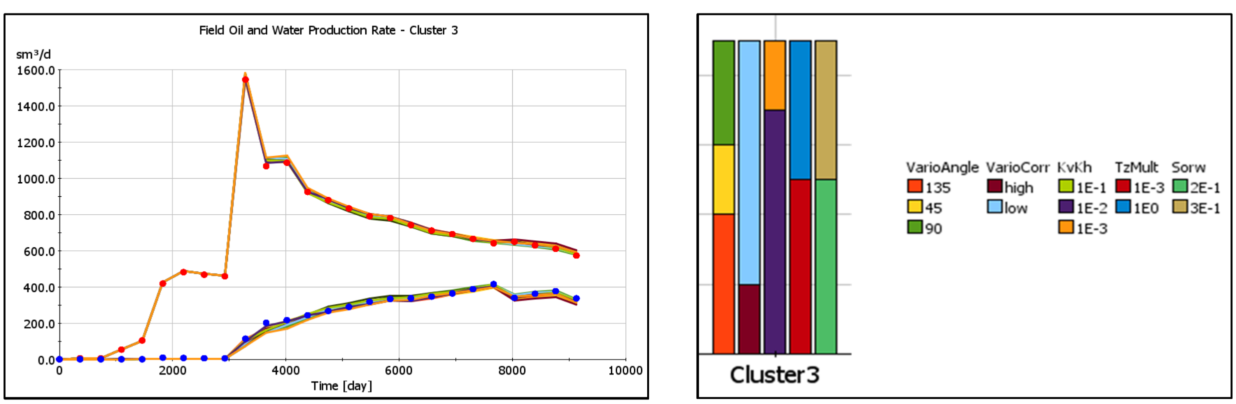


Figure 10: The oil and water production for the 9 models in the cluster containing the historical oil production data (left), and (right) the histogram showing the frequency of parameter values for the five parameters in the cluster (Cluster 3). The MDS space is shown in Figure 5.

Recall that the MDS analysis gives different clusters for field-level vs. well-level distances. Similarly the clustering in metric space based upon the well-level distance (**Figure 6**) results in a different distribution of parameter values, shown in **Figure 11**. A metric space constructed from well responses will carry greater amount of information compared to field response-based metric spaces.

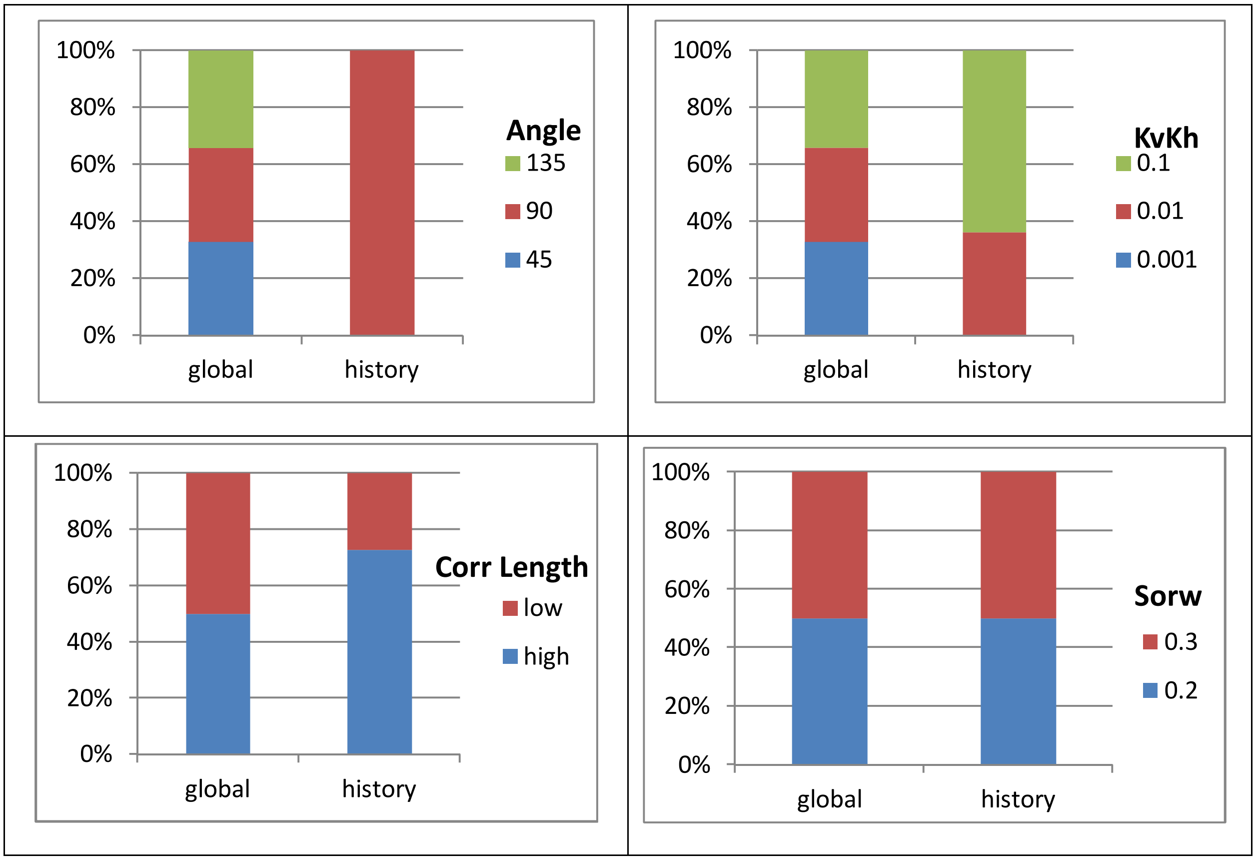


Figure 11: Comparison of the normalized distributions for 4 of the 5 parameters in the example set of 72 models (global) and the cluster containing the “true” reservoir (history).

Comparing the distribution of parameter values for the global distribution with just the cluster containing history identifies high-impact parameters giving models close to historical production. This is done in **Figure 11**, and shows that the variogram angle (90 degrees) and a "high" correlation length give results closer to history. **Figure 11** also suggests a small probability for a low value of Kv/Kh, and a higher probability for a high correlation length, which is interestingly the opposite of what is found for the field-level clustering. This type of information may be very useful for a reservoir modelling team which can then construct additional models with the parameter values of interest if so desired.

History Matching. For additional examples of the use of metric space methods for history matching, we refer to Scheidt *et al*. [7], and Suzuki and Caers [11].

Concluding Remarks

This paper offers an overview of metric space methods and the novel approaches it offers for analysis and understanding of reservoir modelling, uncertainty, and sensitivity analysis. The common element in these methods is the requirement of a distance measure which is related to the purpose of the study. For a large number of reservoir models, the distance measure must be rapid to calculate, and correlated to the response of interest. We have presented above several different distance measures, emphasizing that when analysing flow-based responses, using a flow-based distance such as streamline simulation is likely to be more efficient.

It should be noted that work in metric space methods is still very new in the oil industry. Further research, and experience in more applications will no doubt lead to better understanding of the advantages, and drawbacks, of these methods. Finally, for a more general understanding of reservoir modelling and uncertainty, including a discussion on metric space methods, we refer reader to Caers [12].

References

1. See http://en.wikipedia.org/wiki/Metric_space for more information and references.

2. Borg, I., Groenen, P.: Modern Multidimensional Scaling: Theory and Applications. Springer, New York (1997)

3. Scheidt, C., Caers, J.: Representing spatial uncertainty using distances and kernels. Math. Geosci. 41(4), 397–419 (2008). doi:10.1007/s11004-008-9186-0

4. Scheidt, C., Caers, J.: A new method for uncertainty quantification using distances and kernel methods. Application to a deepwater turbidite reservoir. SPEJ (2009). SPE-118740-PA. doi:10.2118/118740-PA

5. Park, K., and Caers, J.: History Matching in Low-Dimensional Connectivity-Vector Space, Proc., EAGE Petroleum Geostatistics, Cascais, Portugal, 10-14 September (2007)

6. Caers, J., Park, K., Scheidt, C.: Modeling Uncertainty of Complex Earth Systems in Metric Space, in Handbook of Geomathematics. Springer Berlin Heidelberg (2010).

7. Scheidt, C., Caers, J., Chen, Y., Durlofsky, L.: A multi-resolution workflow to generate high-resolution models constrained to dynamic data. Computational Geosciences (2011). doi: 10.1007/s10596-011-9223-9

8. Dejean, J.-P. and Blanc, G. 1999. Managing Uncertainties on Production Predictions Using Integrated Statistical Methods, Paper SPE 56696 presented at the SPE Annual Technical Conference and Exhibition, Houston, Texas, 3-6 October.

9. White C. D., Willis, B. J., Narayanan, K and Dutton, S. P. 2001. Identifying and Estimating Significant Geologic Parameters With Experimental Design, SPE Journal, September, 311-324.

10. Batycky, R.P., Thiele, M.R., Baker, R.O., Chugh S. H.: "Revisiting Reservoir Flood-Surveillance Methods Using Streamlines," SPEREE April 2008, Vol 11, No 2, pp 387-394 (SPE95402-PA).

11. Suzuki, S., Caers, J.: History matching with an uncertain geological scenario. Paper SPE 102154 presented at the SPE ATCE, San Antonio, Texas, 24–27 September (2006)

12. Caers, J.: Modeling Uncertainty in the Earth Sciences. Wiley-Blackwell, UK (2011)