History Matching in Low-Dimensional Connectivity-Vector Space

Kwangwon Park and Jef Caers

Stanford Center for Reservoir Forecasting

ABSTRACT

History matching is a high-dimensional and severely non-linear optimization problem, often requiring some form of dimensionality reduction. In addition, the history-matched models should honor prior geologic information often integrated within geostatistical algorithms. Traditional optimization methods that can handle these problems are, however, time-consuming and highly-dependent upon the initial realization.

In this study, we generated a large ensemble of equi-probable realizations that honor prior information (training image) by means of multiple-point geostatistical simulations. A history-matched model is obtained by searching amongst the ensemble of initial realizations. In order to do this, we assigned each realization its proper connectivity vector. A connectivity vector is calculated by means of TOF (time-of-flight) between an injector and a producer. The connectivity vector is directly related to the production history of existing producers in the field.

It provides two main advantages to rearrange the set of realizations into connectivity-vector space. First, the new space is very low-dimensional (the number of producers) while the set of realizations originally exists in extremely high-dimensional space (the number of grid-blocks). Secondly, in the new space, the objective function changes very smoothly with distance, which means that we can apply various optimization methods (neighborhood algorithm, gradient-based method, etc) more efficiently.

In the connectivity vector space, we applied a simple search method with Kriging as interpolator to search a history match amongst an ensemble of realizations (SIMPAT generated channel reservoir models). The search method provides a reasonable history match in terms of reproduction of production response, injector-to-producer connectivity, and channel geometry.

INTRODUCTION

Various challenges still remain in using traditional optimization methods for history matching reservoir models. Gradient-based optimization methods, which are usually based on the sensitivity coefficient, are highly dependent upon the initial realization, since we need to optimize a large number of variables with the possible existence of several local minima. Although stochastic methods have been developed to find the global minimum, those methods require a considerable number of forward simulations. Extremely time-consuming forward simulation in history matching makes it difficult to apply stochastic methods to field-scale history matching.

Additionally, history matching should honor prior geologic information often integrated within geostatistical algorithms. Perturbing reservoir variables to match history without any consideration of geologic information is meaningless, because, in that case the history match often can not predict future reservoir performance (Caers, 2005). Recently Suzuki and Caers (2006) developed an innovative technique using similarity distance. They show how a (static) distance between any two realizations that correlates with their difference in flow response can be used to search for history matched models by means of efficient search algorithms, such as neighborhood algorithm and tree-search
algorithm. This method was successfully applied to structurally complex reservoirs (Suzuki et al. 2006)

In order for this method to work, the similarity distance between any two realizations should be reasonably correlated with the production history that we want to match. For that reason, the Hausdorff distance proposed in Suzuki and Caers (2006) may not always be an appropriate distance of choice. In this study, we propose a new, connectivity-based distance between two realizations that correlates well to difference in production data between those two realizations. To calculate this distance one does not require any flow simulation. Since the distance is defined in low-dimensional connectivity vector space, we can employ not only Suzuki and Caers (2006) search algorithms but various other optimization methods efficiently.

In addition to the new distance, the connectivity vector space makes it possible to select a history-matched one from a large ensemble of realizations that are simulated by geostatistical algorithms. If uncertainty exists in the prior geologic information, possibly a major source of uncertainty, we can generate an ensemble that is based on multiple geological scenarios (training images). In order to select a history-matched realization efficiently, we map the ensemble into a dimension-reduced space and try to map the realizations such that the objective function values show spatial continuity in that space. In this paper, we demonstrate how well the ensemble mapping and how efficiently the selection of a history-matched realization can be done in this new space.

In the next section, we define a connectivity vector space and investigate the application of a search algorithm in this space. Lastly, concluding remarks summarize the entire study.

CONNECTIVITY VECTOR SPACE

Definition of a connectivity vector space

Geostatistical reservoir models usually contain 1 to 100 million gridblocks. Considering permeability as the only history matching parameter each geostatistical realization can be mapped in 1 to 100 million Cartesian “realization” space. Equations 1 and 2 define a feature mapping that maps the realizations into a lower-dimensional space (Figure 1).

\[
\Phi : \mathbb{R}^{N_{gb}} \rightarrow \mathbb{R}^{N_{pw}}
\]

\[
x \mapsto \mathbf{x} := \Phi(x)
\]

where, \(N_{gb}\) means the number of gridblocks and \(N_{pw}\) the number of producers. \(x\) represents a reservoir model in \(N_{gb}\)-dimensional space and \(\mathbf{x}\) the connectivity vector that indicates the location of corresponding reservoir model in a \(N_{pw}\)-dimensional connectivity vector space. This space relies on the following definition of a connectivity vector:

\[
(x)_i := \frac{1}{\min_j \left(d(w_i^j, w_i^P)\right)}
\]
where \( (\mathbf{x})_i \) is \( i \)-th element of connectivity vector \( \mathbf{x} \) and \( d(w_j^I, w_i^P) \) a measure of spatial distance between the \( j \)-th injector \( w_j^I \) and \( i \)-th producer \( w_i^P \). Since typically, the nearest injector predominantly affects the production history of a producer, we combined the injector-to-producer distances for each producer into a producer dominant distance by selecting the minimum of the distances.

In order for the connectivity vector to have high correlation with the production history, it is critical to define an appropriate distance between an injector and a producer. Here, we propose a TOF (time of flight, Datta-Gupta and King, 1995) based distance calculation. TOF from an injector to a producer is calculated by streamline simulation (Batycky et al., 1997) in steady state conditions. Typically, steady state TOF calculation requires only one hundredth to one thousandth simulation time equivalent to the usual reservoir simulation time. Equation 4 shows the TOF-based injector to producer distance calculation. We can choose any arbitrary percentile among TOFs of streamlines that arrive at a producer.

\[
d(w_j^I, w_i^P) = P_u \left[ \int_{w_j^I}^{w_i^P} \frac{d\zeta_k}{v(\zeta_k)} \right]_{k=1,...,N_{sl}}
\]

(4)

where, \( P_u[\tau] \) represents the \( u \)-percentile of a set of \( \tau \)'s (TOFs) of corresponding streamlines, \( \zeta_k \) is the coordinate along the \( k \)-th streamline and \( v(\zeta) \) is the interstitial velocity along streamline, \( \zeta \). \( N_{sl} \) means the number of streamlines.

In the mapped connectivity vector space (\( N_{pw} \)-dimensional space), it is easy to calculate the similarity distance between any two realizations. We can calculate the similarity distance by means of a two-norm (Euclidean distance)

\[
D(x_a, x_b) = \| \Phi(x_a) - \Phi(x_b) \| = \| x_a - x_b \|
\]

(5)

between any two realizations \( x_a \) and \( x_b \).
Example calculation of connectivity vector

Consider a facies model (Figure 2) that contains channel sand (red) and background mud (blue), in which three injectors and three producers are installed.

Since we have three injectors and three producers, there can be nine possible injector-to-producer connectivities (Figure 3). “Injector-to-producer connectivity” represents quantitatively how strongly an injector is connected to a producer. In other words, the larger the injector-to-producer connectivity is, the more the injector affects the flow response of the producer. Hence, this connectivity is function of the geological heterogeneity as well as the position of the wells and their production history. In this example model, injector 1 is connected to producers 2 and 3. Injector 2 is connected only to producer 1 and injector 3 is not connected to any producer. In addition, we can expect that the connectivity between injector 1 and producer 2 is the largest and the connectivity between injector 1 and producer 3 will be relatively small because the existence of producer 2 in between them. Therefore, qualitatively the nine connectivities can be summarized as in Figure 3. After Equation 3, the nine connectivities are combined into three connectivities for three producers (Figure 4).

If we define the injector-to-producer connectivity as an inverse distance of the shortest path between an injector and a producer, we may not obtain a proper quantification of connectivity. A non-weighted distance is not affected by any other wells located on this path. In order to get an appropriate measure for connectivities, we propose to calculate a flow-based injector-to-producer distance; TOF is one possible way to represent these paths. Figures 5 and 6 show the calculated connectivities of the shortest path and minimum TOFs.
Spatial continuity in connectivity vector space

As an ensemble of equi-probable realizations, we used the 405 realizations (80x80) of Suzuki and Caers (2006), which had been generated by the SIMPAT algorithm (Arpat, 2005) based on 81 training images (250x250) of channel sand distributions with varying geometrical characteristics. The details on the reference field, training images and the realizations can be found in Suzuki and Caers (2006).

First of all, in order to check that a pair of similar connectivity vectors means similar injector-to-producer connectivity, we have listed the realizations whose percentile in distance from the base realization (S71R5) in the connectivity vector space is the 0-percentile (itself), the 1-percentile (the closest one), the 10-percentile, the 25-percentile, the 50-percentile, and the 100-percentile in Figure 7. As the distance from the base case becomes larger, the injector-to-producer connectivities become more different.

Figure 7-Variation of injector-to-producer connectivity according to the connectivity vector.

In order to verify the correlation between the similarity distance of Equation 5 and the objective function (difference between production history, specifically watercut), we select 5 realizations among the ensemble of 405 realizations as reference cases (Figure 8). Each one of them has different injector-to-producer connectivities.

Figure 8-Five reference cases.

Figure 9 represents 405 realizations mapped into the new connectivity vector space. The connectivities are calculated using the minimum TOFs as a measure. Since we have three producers, the connectivity space is three-dimensional.
Moreover, the objective function values obtained from 405 exhaustive simulations show spatial correlation in this space (Figure 10). Without any forward simulation, we mapped the realizations into the new space. Figure 10 also displays the variograms of the objective function in the new space. Objective function has spatial correlations in the connectivity vector space.

Figures 12 and 13 exhibit the production history (watercut) of 5 or 10 realizations that are nearest to the reference in the connectivity vector space. Compared with the watercut curves for all 405 realizations, which show a large range of distribution, the watercut curves of 5 or 10 realizations show relatively small range and sometimes look very similar. As the distance between two realizations in connectivity vector space is smaller, the two realizations show more similar production history. (Figures 14 and 15 show the production history of 5 or 10 realizations that have the smallest Hausdorff distance to the references.)

In summary, the proposed mapping rearranges the realizations in a low-dimensional space such that their distance in this space is correlated with the difference in production history.
We have developed a new space based on the connectivity vector, in which the ensemble of equi-probable realizations is mapped such that spatial correlation in the objective function is achieved. In connectivity vector space, we can apply various optimization methods efficiently because of its low dimensionality. As a first try, a simple technique has been employed for the previous five reference cases.

**Greedy search**

First of all, we select a small number of realizations uniformly in connectivity vector space. Then we evaluate the objective functions for this small set through reservoir simulation. Figure 11 shows the locations of 27 realizations selected and their objective function.

Second, we interpolate the sample objective function over the entire space through simple Kriging. Spherical variograms (range = 0.35, which is equivalent to one third of the space size) are applied for simple Kriging. Since a small number of sample realizations are used, variogram with relative large range makes it possible to find a minimum point which locates in between sample points. Figure 11 also shows the Kriging results. The interpolated objective function represents the distribution of objective function in the space reasonably well when compared with Figure 10. Figure 11 additionally depicts the scatterplot between the interpolated objective functions and the actual objective functions that are calculated through exhaustive reservoir simulation.

Third, we find the location with minimum objective function among the interpolated points. Next, we search for reservoir models whose locations are nearest to the minimum objective function location. For example, we select the nearest 10 models and evaluate the objective functions. Then, we need to simply select the one that has the minimum objective function.

Figures 16 to 30 display the greedy search results for five reference cases. We change the number of the nearest realizations that are selected; 5, 10, and 20 realizations. In general, all models searched exhibit reasonable history matches. In the breakthrough curves, red circles represent the reference curves, blue dotted lines the curves for a history matched model, and the others the breakthrough curves of the realizations that are searched during the selection of history matched one.

- **Case 1:** A reasonable history match was obtained. But the connectivity between the producer 1 and the injector 1 of the history matched model is lower than that of the reference, so the history of the producer 1 is not perfectly matched. Additionally, we observe that the injector 3 in the history matched model is connected to all producers, while the injector 3 in the reference is not connected to any producer. Note that in our technique we do not consider injector dominant connectivity but producer dominant connectivity.

- **Case 2:** We observe some improvement with an increased number of searched realizations. Although we need more forward simulations, selecting larger number
of realizations raises the probability to find a better history match. S25R3 (in 20 realizations case) has more similar injector-to-producer connectivity than S23R3 (in 5 or 10 realizations cases). S25R3 also has a correct orientation of channels.

- **Case 3**: S25R3 (in 10 realizations case) shows almost exactly same injector-to-producer connectivity for all wells. As a result, the production history is matched reasonably well. Even though the training images are different, the width of channel and orientation are similar.

- **Case 4**: S70R3 (reference) has almost the same connectivity with S27R3 (history match). Although the geology of the two models is completely different, the connectivity is very similar so the production history is similar.

- **Case 5**: The history matches (S12R3 and S21R2) look similar to the reference, such as channel distribution and width, etc. Also, the production history is well matched. Although S12R3 has some different orientation from the reference, S21R2 shows similar orientation.

### Global search

The Kriging map in Figure 11 has several local minima. If one uses only the minimum point of the kriged points, we may not obtain a good history match. Therefore we propose a global search which can be accomplished as follows. After Kriging, we search for all local minima and choose from which we select a few with low objective function. Then, we search a small number of realizations nearest to those multiple minima. This search makes it possible to select multiple history matches and / or increase the possibility to find a better history match.

*Figures 31 to 45* represent the global search results. We used the best five local minima and from those points searched the nearest 1, 3, or 5 realizations. Therefore, we may obtain at most 5 final history matches. (The number of final history matches may be smaller than the number of local minima used, because the search ranges of different local minima may overlap.) Multiple history matches show good connectivity matches and similar channel geometry.

- **Case 1**: The nearest 1 and 3 realizations cases found the same history matched realization as the greedy case. However, the nearest 3 realizations cases selected exactly same realization. A global minimum was located near a local minimum point that was different from the local minimum found in the greedy case. This is a clear motivation to use a global search.

- **Case 2**: We found multiple history matches, all of which show good history match and good connectivity match. Moreover, channel geometry is quiet similar (S26R1, S25R2, and S23R4 are based on similar training images).

- **Case 3**: As with the previous cases, we are able to select multiple history matches. The injector-to-producer connectivity is similar, too.
• **Case 4**: We found adequate history matches, but the channel geometry is completely different from the reference. The lack of geologic information to constrain the reservoir geometry (resulting in the existence of multiple training images) makes this possible (S74R2, S25R2, and S27R3).

• **Case 5**: A good history match and good connectivity match is achieved. Also, the channel geometry is very similar, especially in S75R1.

**Performance**

Figures 46 and 47 depict the performance of the techniques proposed in this study. We have applied 405 greedy and global searches where of all 405 realizations were taken as reference. The performance in Figures 46 and 47 is the experimentally observed frequency of selecting the best history match in the ensemble (this is 1 in the legend, which means the rank), or at least the second best history match (this is 2 in the legend). In other words, for the purple line (10th best match), if we evaluate the objective function 47 times, we can expect to select at least the tenth best history match with a probability of around 0.75.

For greedy search, we selected 5, 10, 20, 30, and 50 nearest realizations and select the one that has the smallest objective function. Therefore the number of objective function evaluations is 32 (27+5), 37 (27+10), 47 (27+20), 57 (27+30), and 77 (27 + 30), because we evaluate the objective functions of 27 realizations for Kriging interpolation and 5, 10, 20, 30, and 50 nearest realizations. This technique is different from traditional history matching in that the number of objective function evaluations is determined prior to history matching.

For global search, we select 1, 3, and 5 nearest realizations to the multiple local minima (in this case, 5 points) and select one that has the smallest objective function. In this case, the number of objective function evaluations is 32 (27 + 5), 42 (27 + 15), and 52 (27 + 25), because we evaluate the objective functions of 27 realizations for Kriging interpolation (this is same as greedy search) and one realization for each local minimum.

In both cases, the expected probability to find 10th best history matched one amongst 405 realizations is larger than 75% when 50 objective function evaluations are applied.

Comparing greedy search results with global search results, we can conclude that the probability to have a better history match is slightly higher in the greedy search than in the global search. These results are caused by the fact that in global search the number of realizations to be selected is too small. If we increase the number to 10 or 20, the global search may show a better performance. However, in the cases of large number of searching, the number of objective function evaluations is also increased to 77 (27 + 50) or 127 (27 + 100) (this is too large and meaningless, because the ensemble size is 405 in this case and if we can simulate 127 reservoir models among them, we may find the history match simply by random selection.) Also, for greedy and global searches, the goal is different. In greedy search, the goal is to find a history matched one that is the best. While, in global search, the goal is to find many possible history matched models that are match well enough.
CONCLUDING REMARKS

We have defined a new feature mapping based on injector-to-producer connectivity. This mapping rearranges an ensemble of equi-probable realizations in high-dimensional realization space (the number of gridblocks) into a low-dimensional connectivity vector space (the number of producers). It was shown that the objective functions of the mapped realizations are spatially correlated in the connectivity vector space, a presumption required for the method to work.

In the connectivity vector space, we applied a simple search method with Kriging as interpolator to search a history match amongst an ensemble of realizations. The search method provides a reasonable history match in terms of reproduction of production response, injector-to-producer connectivity, and channel geometry.

This history matching technique is not dependent upon the initial model, because we do not use any initial model. On the other hand, this technique can be applied to determine an initial model for traditional history matching, or alternatively, to eliminate certain geological scenarios not compatible with production data (at least in terms of connectivity). Additionally, all forward simulations need not be done sequentially hence high efficiency can be obtained with parallel computing.

ACKNOWLEDGEMENTS

Satomi Suzuki provided the data of 405 realizations and codes of neighborhood and tree-search algorithm. We appreciate Satomi Suzuki’s supports for this study.

REFERENCES

Figure 10-405 realizations in connectivity vector space.
Figure 11-27 sample realizations in connectivity vector space and their scatterplots between the krigged values of 405 realization points (x-axis) and those of 27 sample points (y-axis).
Figure 12-Production response of five realizations that is near to the reference.
Figure 13-Production response of 10 realizations that is near to the reference.
Figure 14-Production response of five realizations that is near to the reference based on Hausdorff distance.
Figure 15 - Reduction response of 10 realizations that is near to the reference based on Hausdorff distance.
CASE I + nearest five

REFERENCE: S15R1

HISTORY MATCH: S26R1

Figure 16-Greedy search: S15R1 (search 5 reals).
CASE I + nearest 10

REFERENCE: S15R1

HISTORY MATCH: S26R1

Figure 17-Greedy search: S15R1 (search 10 reals).
CASE I + nearest 20

REFERENCE: S15R1

HISTORY MATCH: S26R1

Figure 18-Greedy search: S15R1 (search 20 reals).
CASE II + nearest five

REFERENCE: S22R4

HISTORY MATCH: S23R4

Figure 19-Greedy search: S22R4 (search 5 reals).
CASE II + nearest 10

REFERENCE: S22R4

HISTORY MATCH: S23R4

Figure 20-Greedy search: S22R4 (search 10 reals).
CASE II + nearest 20

REFERENCE: S22R4

HISTORY MATCH: S25R3

Figure 21-Greedy search: S22R4 (search 20 reals).
CASE III + nearest five

REFERENCE: S45R1

HISTORY MATCH: S26R1

Figure 22-Greedy search: S45R1 (search 5 reals).
CASE III + nearest 10

REFERENCE: S45R1

HISTORY MATCH: S25R3

Figure 23-Greedy search: S45R1 (search 10 reals).
CASE III + nearest 20

REFERENCE: S45R1

HISTORY MATCH: S81R3

Figure 24-Greedy search: S45R1 (search 20 reals).
Figure 25-Greedy search: S70R3 (search 5 reals).
CASE IV + nearest 10

REFERENCE: S70R3

HISTORY MATCH: S27R3

Figure 26-Greedy search: S70R3 (search 10 reals).
CASE IV + nearest 20

REFERENCE: S70R3

HISTORY MATCH: S27R3

Figure 27-Greedy search: S70R3 (search 20 reals).
Figure 28-Greedy search: S75R2 (search 5 reals).
Figure 29-Greedy search: S75R2 (search 10 reals).
CASE V + nearest 20

REFERENCE: S75R2

HISTORY MATCH: S21R2

Figure 30 - reedy search: S75R2 (search 20 reals).
CASE I + nearest one + five local minima

REFERENCE: S15R1

HISTORY MATCH: S26R1

Figure 31-Global search: S15R1 (search 1 real for 5 local minima).
CASE I + nearest three + five local minima

REFERENCE: S15R1

HISTORY MATCH: S26R1

Figure 32-Global search: S15R1 (search 3 reals for 5 local minima).
Figure 33-Global search: S15R1 (search 5 reals for 5 local minima).
CASE II + nearest one + five local minima

REFERENCE: S22R4

HISTORY MATCH:
S61R2/S26R1/
S25R2/S23R4

Figure 34-Global search: S22R4 (search 1 real for 5 local minima).
CASE II + nearest three + five local minima

REFERENCE: S22R4

HISTORY MATCH:
S61R2/S26R1/
S25R2/S23R4

Figure 35-Global search: S22R4 (search 3 reals for 5 local minima).
CASE II + nearest five + five local minima

REFERENCE: S22R4

HISTORY MATCH: S61R2/S17R5

Figure 36-Global search: S22R4 (search 5 reals for 5 local minima).
CASE III + nearest one + five local minima

REFERENCE: S45R1

HISTORY MATCH: S74R2/S26R1

Figure 37-Global search: S45R1 (search 1 real for 5 local minima).
CASE III + nearest three + five local minima

REFERENCE: S45R1

HISTORY MATCH: S74R2/S26R1

Figure 38-Global search: S45R1 (search 3 reals for 5 local minima).
CASE III + nearest five + five local minima

REFERENCE: S45R1

HISTORY MATCH:
S74R2/S26R1/S15R1

Figure 39-Global search: S45R1 (search 5 reals for 5 local minima).
Figure 40-Global search: S70R3 (search 1 real for 5 local minima).
CASE IV + nearest three + five local minima

REFERENCE: S70R3

HISTORY MATCH: S74R2/S25R2/S27R3

Figure 41-Global search: S70R3 (search 3 reals for 5 local minima).
CASE IV + nearest five + five local minima

REFERENCE: S70R3

HISTORY MATCH: S74R2/S25R2/S27R3

Figure 42-Global search: S70R3 (search 5 reals for 5 local minima).
CASE V + nearest one + five local minima

REFERENCE: S75R2

HISTORY MATCH: S74R2/S26R1/S12R2

Figure 43-Global search: S75R2 (search 1 real for 5 local minima).
CASE V + nearest three + five local minima

REFERENCE: S75R2

HISTORY MATCH: S75R1/S26R1/S12R3

Figure 44-Global search: S75R2 (search 3 reals for 5 local minima).
CASE V + nearest five + five local minima

REFERENCE: S75R2

HISTORY MATCH: S75R1/S26R1/S12R3

Figure 45-Global search: S75R2 (search 5 reals for 5 local minima).
Figure 46-Performance of greedy search.

Figure 47-Performance of global search.