Benefiting from Variogram Information to Characterize Facies Distribution in History Matching with EnKF*

A. Abadpour[1], P. Bergey[1], R. Piasecki[1], and T. Blanchard[1]

Search and Discovery Article #41315 (2014)
Posted April 7, 2014

*Adapted from extended abstract prepared in conjunction with oral presentation at GEO-2014, 11th Middle East Geosciences Conference and Exhibition, 10-12 March 2014, GEO-2014©2014

[1]Geosciences Research Centre, Total E&P, Aberdeen, United Kingdom

Abstract

Constraining reservoir models to available field observation is an important step in reservoir simulation process especially when models aim at further development decision making. In the past two decades, several assisted history-matching methodologies have been proposed to help in the matter (gradient methods, streamlines, etc. - Oliver and Chen, 2011). More recently, ensemble based methodologies were adapted from other domains to perform reservoir history matching (Evensen, 2007).

Introduction

Ensemble methods rely on establishing correlations between sets of uncertain observations and a priori models and acting based on such correlations in a statistical manner. The approach is Bayesian by essence.

Ensemble based history matching methods have the capacity of dealing with almost any type of uncertain parameters in the geo-model and also to incorporate any sort of observation from the wells and particularly “special” ones like seismic or tracer data. These algorithms are very computationally efficient and comparatively insensitive to the number of parameters to be treated. This opens the possibility of addressing uncertainty on a cell-by-cell basis.

The Ensemble Kalman filter approach best performs in linear cases and Gaussian distribution of uncertainty in the prior models. Its performance remains acceptable in nonlinear cases and non-Gaussian distributions, but it is still recommended to re-parameterize whenever possible toward linear acting models with Gaussian parameters and independent observations.

This paper briefly describes the reformulation of information beneficial both in the model space and in the observation space, mainly to parameterize facies distribution uncertainty adaptable to be handled with ensemble methods and 4D seismic observations to be used to constrain the models.

EnKF and Uncertain Facies Distribution

3D distribution of petrophysical parameters (porosity, permeability, relative permeability coefficient, etc) constitutes one of the key reservoir uncertainties. Facies are commonly used as high order controls for petrophysical parameters. While the petrophysical characteristics of each facies can be evaluated at wells, their spatial distribution is very often poorly known, typically derived only from indirect means (such as 3D seismic). This makes the distribution of different facies type in the reservoir model one of the important (if not the main) sources of uncertainty. Of particular importance are the discrete petrophysical properties such as relative permeability tables that are often directly linked to facies.

Facies are discrete and often non-sort-able parameters. By non-sort-able, we refer to a specific characteristic of petrophysical properties. Different petrophysical properties of different facies can be sorted in different orders making averaging approaches inefficient (e.g. 3 facies A, B and C could have porosities ranked in the ABC order, horizontal permeability ranked in the CBA order but vertical permeability ranked in order BCA; in such set-up B is not intermediate between A and C).

Consequently, adjusting the facies uncertainty with a methodology like ensemble Kalman filter is not straightforward. Simply assimilating on discrete facies values there is a high probability of finding non-discrete ones, which will not be obvious to associate to the reservoir cells.

An appropriate re-parameterization scheme seems required in such a case. Several ideas have been proposed. In the series of studies reported here we focused on a level set methodology proposed in (Lorentzen et. al. 2011) to calculate the closest distance to the boundary of facies type and to use this distance as a parameter to be treated by Kalman filter.

To summarize the level set re-parameterization workflow, it is obvious that when the boundary of the facies is treated, cells have the possibility to change facies type; and they need to be filled with petrophysical properties related to their variable facies type. That is why the approach requires starting with the full cube of properties for every single facies type; petrophysical parameters conditional to facies can be updated in the course of assimilation while facies can be directly used to govern relative permeability tables.

This approach increases the number of uncertain parameters in the initial ensemble matrixes, bearing also in mind that the level set manipulation is computationally costly. Example: from 1 NTG, 1 KX, 1,KVKH, 1 PHI, total of four 3D cubes of uncertain parameters to 4 NTG, 4 KX, 4 KV/KH, 4 PHI and 4 distance to facies in a four facies setup. One needs to make sure that facies are important source of uncertainty to justify addressing them explicitly. One also needs to ensure that the number of facies remains limited; otherwise, the cost increase will be prohibitive.

The main idea stayed intact in this paper but the approach to calculate the distance evolved a lot during the course of our studies, mainly to benefit from variogram information used to populate the facies as well as their initial probability. In the following, we will explain the evolution of the calculation methodology and go through the tests performed to prove the performance of implementation of the approach.

The level set approach in the reference work is applied to calculate the distances in the 2D domain, but to take into account the vertical heterogeneities we definitely needed to adapt it for real 3D calculations especially for dealing with real reservoir models. A publicly available level set toolbox (Mitchelle, 2007) has been used to calculate the distance functions. Such distance calculation was based on counting the number of cells between cell centre and the boundary of facies type and not the exact Euclidian distance. Ideal was to calculate the exact distance specially when not dealing with rather sugar box grid and the structure contains rather sophisticated shapes, bearing in mind that the first aim is to investigate the eventual improvement on the quality of the history match and moreover geological realism in the final realizations.

It worth mentioning such an approach to address facies boundary and then petrophysical properties at the next level already allows to better respect the geological description of the models. One should remember that data assimilation with ensemble-based approaches produces new set of realizations by looking for a linear combination of realizations that produces smaller miss match of observation. So eventually mixing the petrophysical properties of different facies type and associating unacceptable values to them is not only unacceptable from geological point of view but also provokes inconsistency between static (porosity, permeability, etc) and dynamic parameters (relative permeability, compressibility, etc) in such a cell. Those inconsistencies typically introduce many numerical issues for the simulator.

To properly calculate the distances to the facies boundary a first order fast marching algorithm was used which takes the exact reservoir grid into account (here in our case the Eclipse EGRID file), to export as many distance sets as the number of facies in the model. At the same time the new algorithm has the capacity of stretching the grid in any direction, benefitting from this option we could deform the grid differently for every facies type to represent the anisotropy in the variogram of each. This extra step showed very helpful in keeping the continuity of the facies during the assimilation, and improved the smoothness of retrieved bodies.

Furthermore, we used the variogram curvature information to normalize the calculated distances to the local probability of the occurrence of the facies in the initial ensemble, which is a way to insure the cells in distance further than the range of the variograms are not facing any modification, as we do not have any correlated information after this range. Moreover, to keep the initial probability of the occurrence of each facies intact we have chosen to normalize the distances to the local proportion for each facies in the initial ensemble. At the same time this option helps avoiding the association of a facies type to a cell when it is completely absent in the ensemble of realizations for that cell. Attempts were made to also evaluate the benefits could be gained by introducing variable local proportion which changes after each assimilation step, based on updated realizations (not a constant proportions coming from the initial ensemble of realizations). Or static and dynamic global ones (one value per facies type rather than cell by cell values), important to recall that static local option stays the most reasonable one and gives the best result so far.

Parameterization of 4D Seismic Observation

4D seismic observations have differ from more conventional observations at wells over time. They provide spatial measurements over a large frequency bandwidth and globally include more truly independent observations. On the other hand, they are usually coming with a low time frequency (few monitors).

4D Seismic has been introduced to be used in an EnKF in the recent years, (Skjervheim et al. 2007, Harvel et al. 2005, Fahimuddin et al. 2010, Leeuwenburgh et al. 2011). Their inclusion in the assimilation process improves results due to the extremely rich information they contain but also increase in the computational costs. Moreover, it has been concluded in some studies that the large size of observations may provoke certain level of collapse in the posterior distribution. That collapse has a standard remedy: ensemble inflation to reset the spread among the realizations. Localization also proved beneficial (Chen and Oliver 2010, Emrick and Reynolds 2010); it aims at limiting updates to (typically geographic) domains where correlations are a priori expected.

A front arrival time parameterization has been proposed to linearize the problem as it has been demonstrated that in 1D the front arrival time is linearly dependent upon k/Phi.Ct. This parameterization also has the advantage of decreasing the number of observed data therefore decreasing the risk of ensemble collapse. It presents a low direct computational cost (Trani et al. 2012). The drawback of such parameterization is that the simulations has to be run beyond the assimilation time for many model realizations to be able to reach meaningful predictions for each ensemble member on desired positions. This determines some control challenges if using Eclipse but more importantly it is indirectly extremely costly particularly if the model is heterogeneous.

To overtake this difficulty we proposed a parameterization where the 4D seismic information is transformed into the distance of the shortest path linking a given cell centre to the closest boundary of water front when path are constrained by the simulation grid topology (between a cell and its immediate neighbours). This distance can be calculated relatively cheaply for all the cells in all realizations at any desired time by the fast marching method without any need of additional simulations. To distinguish between flooded and non-flooded areas the distance is signed differently in each domain. This normal projection to the front surface can be seen

At this stage, we needed to come up with the definition of a misfit term to describe the difference between observed and predicted fronts and use such a distance measurement in an assisted history matching procedure. The idea of using local Hausdorff distance for such was first suggested by Tillier et al. (2012) to compute the dissimilarity between different images. In this work, the L2 norm of the local map has been used afterwards to calculate the global dissimilarity because of the limitation in the assisted history-matching scheme that has been used. Using EnKF give us the possibility to increase the number of observation data and handle the misfit locally in each reservoir grid cell.

EnKF does not impose any theoretical limit on the number of observation data that can be handled. However, there is still a computational cost involved and such cost might limit our ability to effectively use the approach. Eventually some mathematical tricks might be needed to alleviate the heavy computational cost. Such a trick was developed in our study when we modified the matrix inversion to obtain the Kalman gain below:

$$\psi^a = \psi^f + (C^e_{\psi\psi})^f M^T \left[M (C^e_{\psi\psi})^f M^T + C^e_{\varepsilon\varepsilon} \right]^{-1} (d - M\psi^f)$$

$$(A + rV^T V)^{-1} = A^{-1} - rA^{-1}V^T [I + rVA^{-1}V^T]^{-1} VA^{-1}$$

Another approach was also used to limit the calculation cost. The expression of 4D seismic results was limited to the common flooded area among all the realizations plus the true observation, expanded by a few cells.

While investigating the performance of the proposed methodology by evaluating the post match models to seek any significant violation of the prior geological information an important deterioration has been observed in the quality of reservoir models after history matching. Extreme changes are observed compared to the prior models. Worth mentioning that the facies models obtained after assimilation on the distance to the boundary of the flooded area as observation was not satisfactory in absolute terms, but compared to the one found after assimilation on cell saturations was still extremely promising.

Certainly, the distance measure in one cell of the reservoir model is not independent from the measures in the adjacent cells. This measure is correlated to the same measure for neighbouring cells. An update without localization would surely over fit the models and provokes an underestimation of uncertainty. An adapted localization scheme could help taking advantage of the observations in better accordance with the hypothesis of independence implicit in the Kalman filter approach. This determines more realistic updates to the input model and a better fit to the observation when using ensemble methods. Accordingly a localization strategy has been deployed to mitigate the Kalman gain to make effective the correlation between the difference of distance to front observation at a given cell and the properties of the same cell to provide modifications to just the considered cell.

Such a scheme not only improved the obtained match but also prevented any excessive change in the model, and revealed the progress made preserving the prior information in the model after the history match process.

Uncertainty Characterization for 4D Seismic Observations

The introduction of measurement error is mandatory for the assimilation of any kind of observation data using ensemble Kalman filter methodology. Measurement error is used first to perturb the observations (to come up with as many perturbed observations than available simulated predictions) and to normalize the misfit term. The level of certainty of provided observed data affects the strength of modification derived from calculated mismatch.

Two ideas have been tested to introduce the error associated to the 4D measurements; those ideas are based on perturbing the volume of flooded region:

- The initial proposition was simply to provide multiple stochastic seismic inversion scenarios unconstrained by reservoir simulation. EnKF can indeed easily factor in directly observation realizations as long as those are numbered as the model realizations.
- In a revised proposal, not having the possibility of producing several inversions, just two interpretations of the water-invaded geo-bodies corresponding to “maximum” and “minimum” scenarios would be enough. Those estimates originate from a deterministic inversion from seismic domain to petro-elastic domain to which a cut-off is applied. The minimum scenario corresponds to the coarse simulation cells that have all their corners in the flooded geo-body as defined at fine seismic scale and the maximum one represents the scheme that at least one corner point is located in the interior. It is very likely that such a process grossly underestimates the 4D seismic uncertainty but provides data fairly well spatially regularized.

We calculate the distance to both of those interfaces for each cell centre. We call them “a” for the distance to the minimum volume and “b” for the maximum one. The ratio a/(a+b) will vary between zero and one for the cells located in between two boundaries, becoming zero when located on the minimum boundary and one when situated on the maximum one.

Several (typically 100) realizations of the front observation are generated by drawing randomly values between 0 and 1 and defining the front location relatively to the comparison of the random value and a/(a+b) ratio of each cell. Every resulting intermediate observation realization is associated to a model realization. Those realizations are strongly spatially correlated; this problem is of no importance in the particular EnKF scheme used as a specific localization scheme is used to limit the influence of distance to front information to the properties of the cell to which it is associated.

Figure 1 illustrates the perturbed realizations of flooded zone produced from minimum and maximum interpretations.

As mentioned before an exact error term also needed to be calculated for integration into the Kalman gain matrix. This has been done by counting the number of cells to the boundary of flooded region and using that as a representative of error knowing that further away less influential the cell’s properties to front location are. In other words if the cell is located far from water front this boundary is governed by so many more cell properties than the one we are dealing with, that’s why we are decreasing the level of updates forced to the actual cell by the calculated far away mismatch.

Test Study

A channelized synthetic model with 70x140x10 dimension containing six producers and four water injectors has been deployed for test purposes. Two producers and one injector are penetrated in the branches of the channel and more or less conditioned in all the realizations. Four facies type including channel are present in the model. Direction, widths, tortuosity and density of channels are the main source of uncertainty in the different realizations. An extra realisation presented in Figure 2 has been created to play the role of the truth case, with the branches of the channels in the middle of the reservoir. Kalman filtering is supposed to move the channel in ensemble of realizations to put it in the middle of the grid where it is located in the truth case.

Porosity, permeability and NTG for the cells as well as the facies boundary should be characterized using EnKF data assimilation on well production data on top of two 4D surveys, before and after water breakthrough in the wells.

Improving the match quality in posterior realizations was the very first objective of this test and more importantly we need to clarify if introducing this extra information helped guiding the water to the flooded areas predicted by the truth model or not. Figure 3 compares some match results with and without 4D assimilation.

Figure 4 summarizes the facies distributions obtained with this way of parameterization and also reveals the benefits of using the facies variogram information with the image on the top being the possibility of finding the channel cell in the initial ensemble and bottom ones in the posterior models, with (left) and without (right) variogram normalization

Figure 5 shows the map of flooded area for one of the realization at the end of data assimilation with (bottom) and without (middle) 4D seismic observations (top). It appears that introducing 4D derived observations helped cleaning up the swept zone to look more like the observed one. Investigating one of the realizations showed promising results but it is also crucial to look at the behaviour of the whole ensemble of posterior models, which is confirmed by Figure 6. Worth mentioning that when all realizations are predicting water flooding in one reservoir cell it is represented in dark red and while none of them are predicting water it would be dark blue.

It is not too difficult to see that improvements have been made to the shape of flooded area in the ensemble of realizations after introducing the 4D seismic observations. It is also important to note that benefiting from a special observation as 4D seismic also improved the characterisation of the reservoir models specially in reducing the uncertainty in the positioning of the facies boundaries.

Another Example (Less Variable Model)

Even though result obtained in the last test case study to position the channel in the middle of the model starting from a huge level of uncertainty was encouraging, but one might argue that the final probability map still does not look satisfactory for decision making processes such as infill drilling after history matching. Accepting above argument and considering that in normal operational studies the level of knowledge represented in prior models should be far more than what illustrated in our synthetic case a similar study with a less variable set of models has been performed.

Figure 7 shows the results obtained incorporating the 4D seismic observation or not. It also presents the evolution of facies proportion in each case that remains almost unchanged during assimilation.

Correlations between petrophysical parameters (e.g. porosity / permeability) are also one of the key features (and underlying key construction mean) of reservoir models. Such laws were often severely transformed in EnKF updated models before using level set parameterisation of facies.

The level set approach and the parallel handling of 3D facies distribution and full cubes petrophysical parameters conditional to facies proved efficient in reducing considerably such effects.

Figure 8 shows the initial and final correlation plot for one of the facies type in the synthetic study of the channelized model with high uncertainty of the channel position. It also compares two posterior distributions using variogram normalisation or not. The use of variogram normalisation appears clearly justified here as well.

Conclusions

Variogram normalized level set functions of distance to interface constitute a practical and efficient re-parameterisation method for discrete non-sort-able parameters such as facies or 4D front. Applied to parameters conditioning others (such as facies that control petrophysical properties), the approach is costly; the size of the problem is increased as many times as there are facies. The approach is however superior to ignoring the high order parameter (e.g. facies) whenever it plays a substantial role in the problem. In order to bring maximal efficiency a number of algorithmic subtleties have to be taken into consideration.

Selected References

Chen, Y., and D.S. Oliver, 2010, Cross-covariances and localization for EnKF in multiphase flow data assimilation: Journal of Petroleum Geosciences and Engineering, v. 66, p. 1-14.

Emrick, A.A., and A.C. Reynolds, 2011, History matching a field case using the ensemble Kalman filter with covariance localization: SPR Reservoir Evaluation & Engineering, v. 14/4, p. 443-452.

Evensen, G., 2007, Data Assimilation: the Ensemble Kalman Filter: Springer, New York, 308 p.

Fahimuddin, A., A. Aanonsen, and J.A. Skjervheim, 2010, Ensemble based 4D seismic history matching – Integration of different levels and types of seismic data: OnePetro, SPE #131453, 15 p.

Leeuwenburgh O., J.H. Brouwer, and M. Trani, 2011, Ensemble-based conditioning of reservoir models to seismic data: Computational Geosciences, v. 15, p. 359-378.

Lorentzen, R.J., G. Naevdal, and A. Shafieirad, 2011, Estimating facies field using ensemble Kalman filter and distance functions - applied to shallow-marine environment: OnePetro, SPE #143031, 14 p.

Oliver, D.S., and Y. Chen, 2011, Recent progress on reservoir history matching: a review: Computational Geoscience, v. 15/1, p. 185-221. DOI 10.1107/s10596-010-9194-2.

Reiso, E., M.C. Haver, and M. Aga, 2005, Integrated workflow for quantitative use of time-lapse seismic data in history matching: A North Sea Field Case: OnePetro SPE #94453, 10 p.

Skjervheim, J.A., G. Evensen, S.I. Aanonsen, B.O. Ruud, and T.A. Johansen, 2007, Incorporating 4D seismic data in reservoir simulation models using ensemble Kalman filter: SPE Journal, v. 12/3, p. 282-292, SPE #95789.

Tillier, E., M. Le Ravalec, and S. Da Veiga, 2012, Simultaneous Inversion of Production Data and Seismic Attributes: Application to a Synthetic SAGD Produced Field Case: Oil & Gas Science and Technology – Rev. IFP Energies nouvelles, v. 67/2, p. 289-301.

Trani, M., R. Arts, and O. Leeuwenburgh, 2012, Seismic history matching of fluid fronts using Ensemble Kalman filter: SPE Journal, v. 18/1, p. 159-171, SPE #163043-PA. DOI: 10.2118/163043-PA.

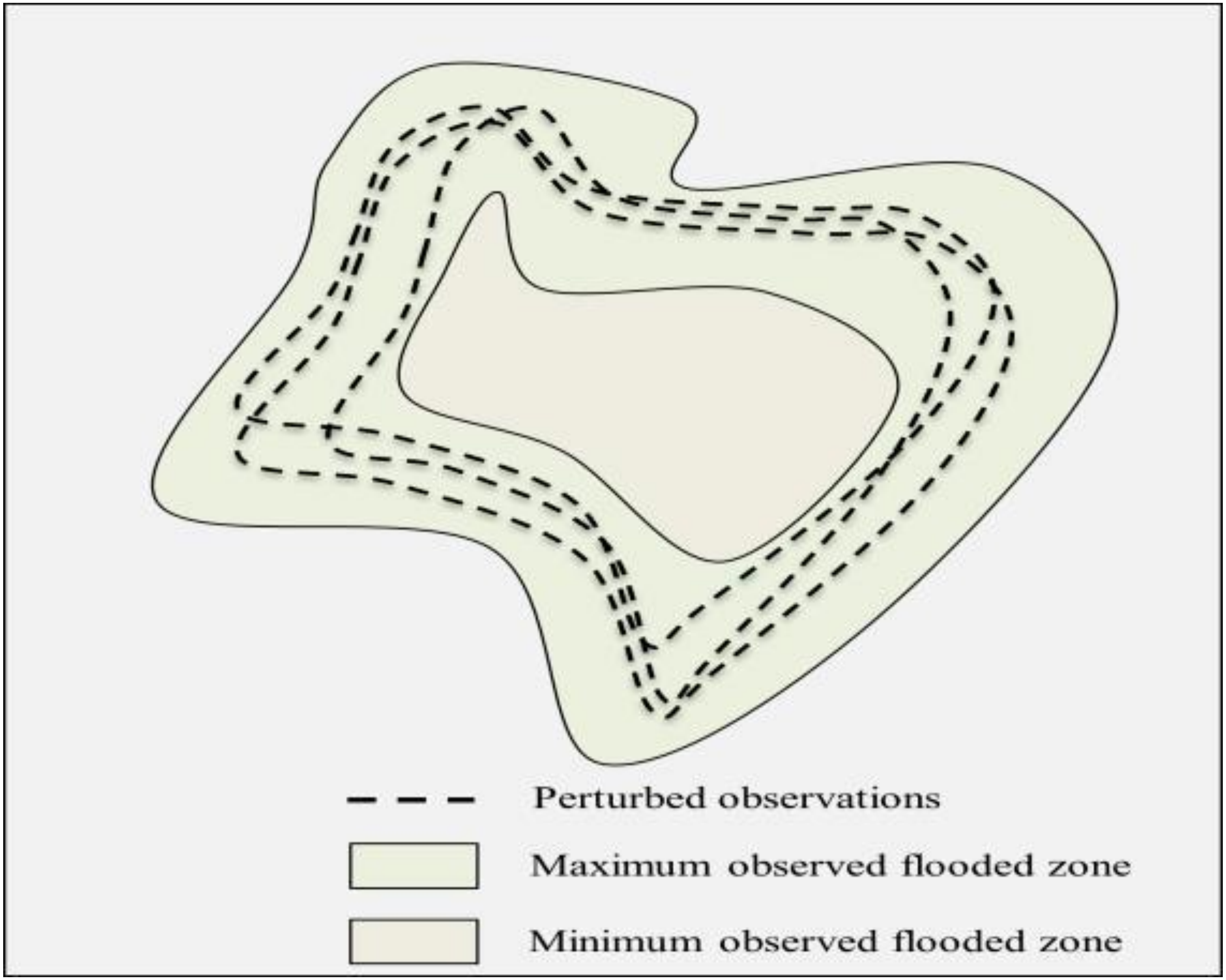

Figure 1. Provided maximum and minimum observed flooded zones used to be perturbed and obtain several realizations of flooded area.

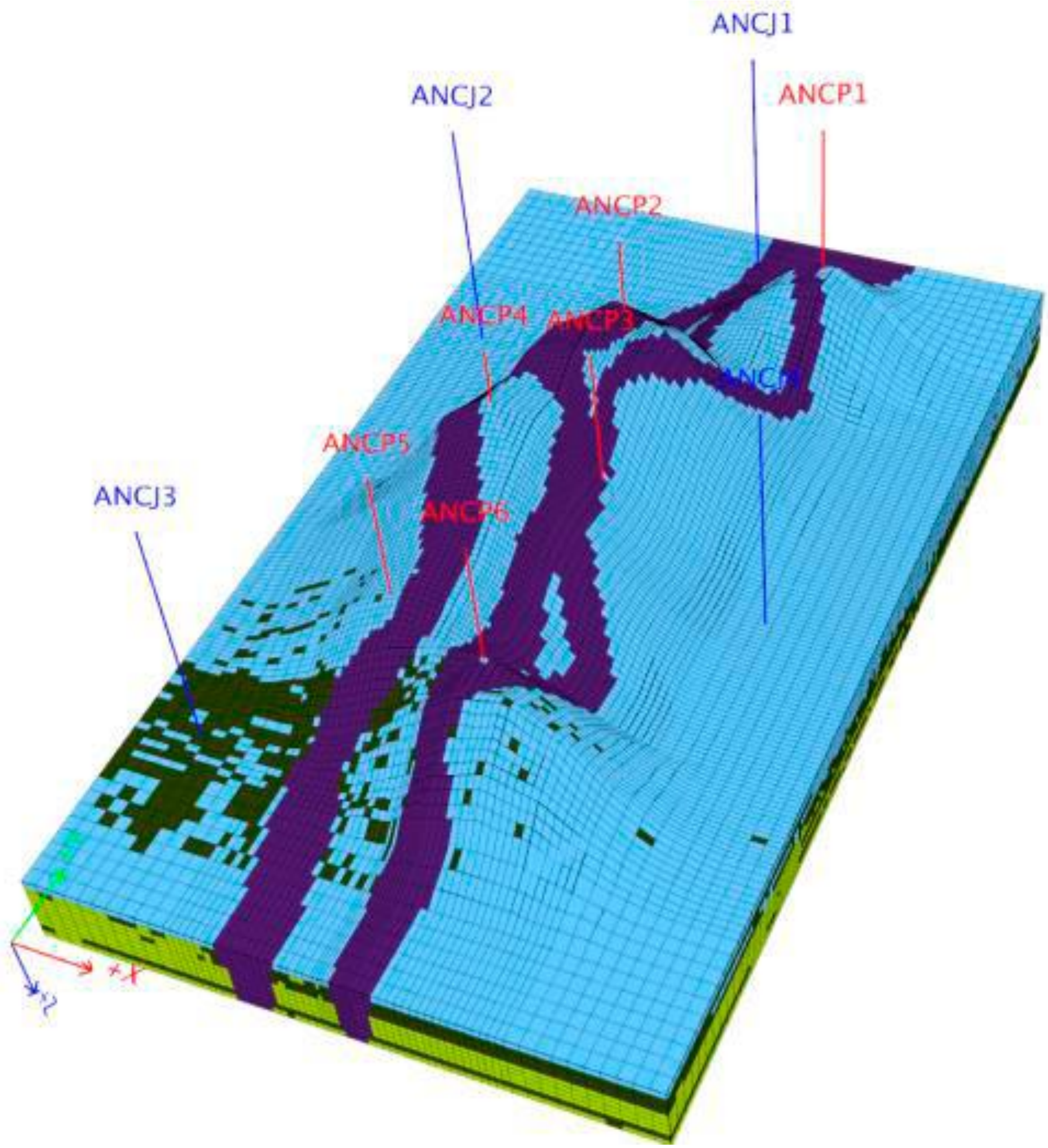

Figure 2. Truth case for channelized model, case study to characterize facies boundaries and to improve the shape of water flooded zones by assimilation on 4D seismic data.

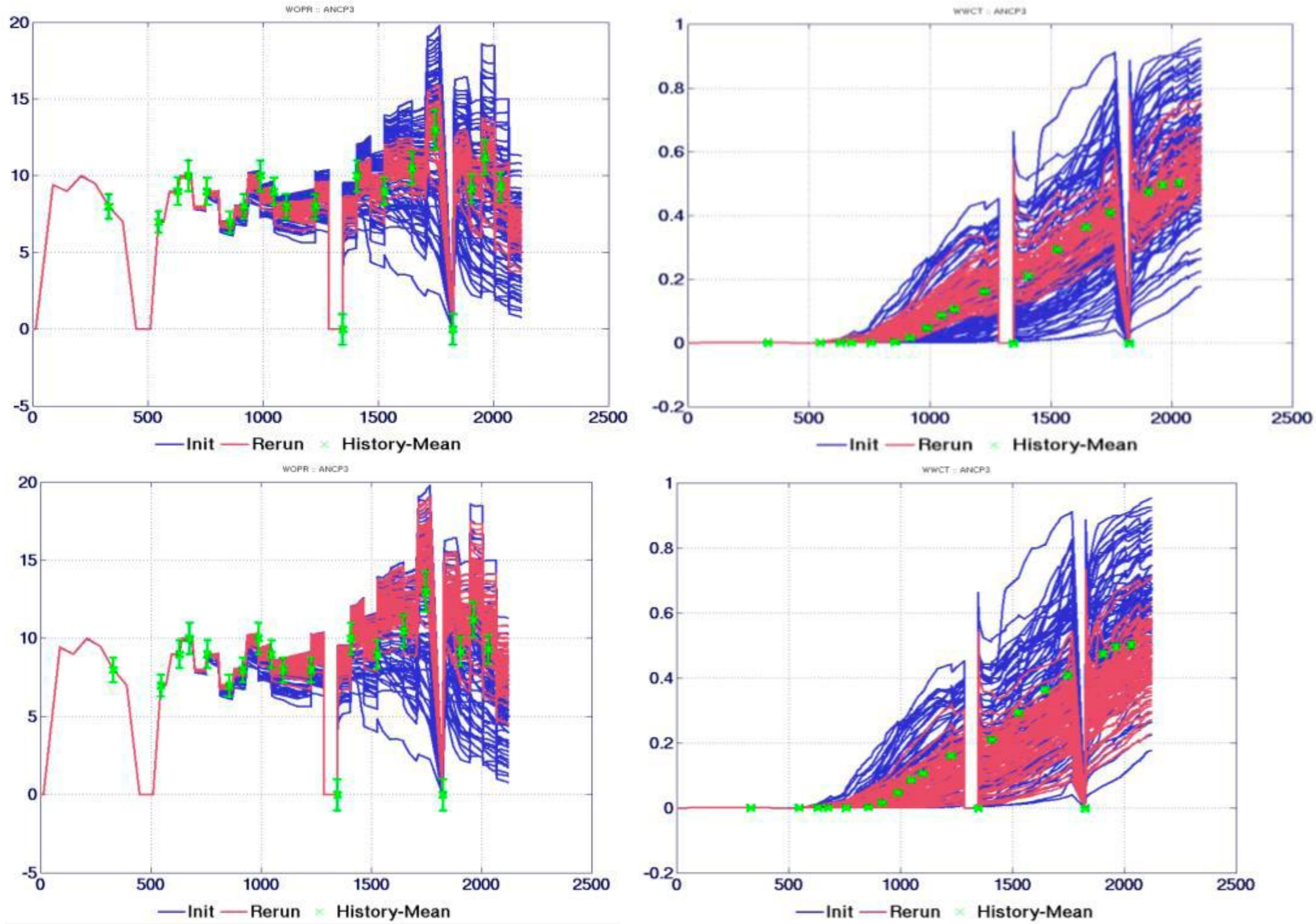

Figure 3. Improvement on the quality of obtained match by assimilation on 4D seismic observations (top) oil production rate and water cut for one of the producers.

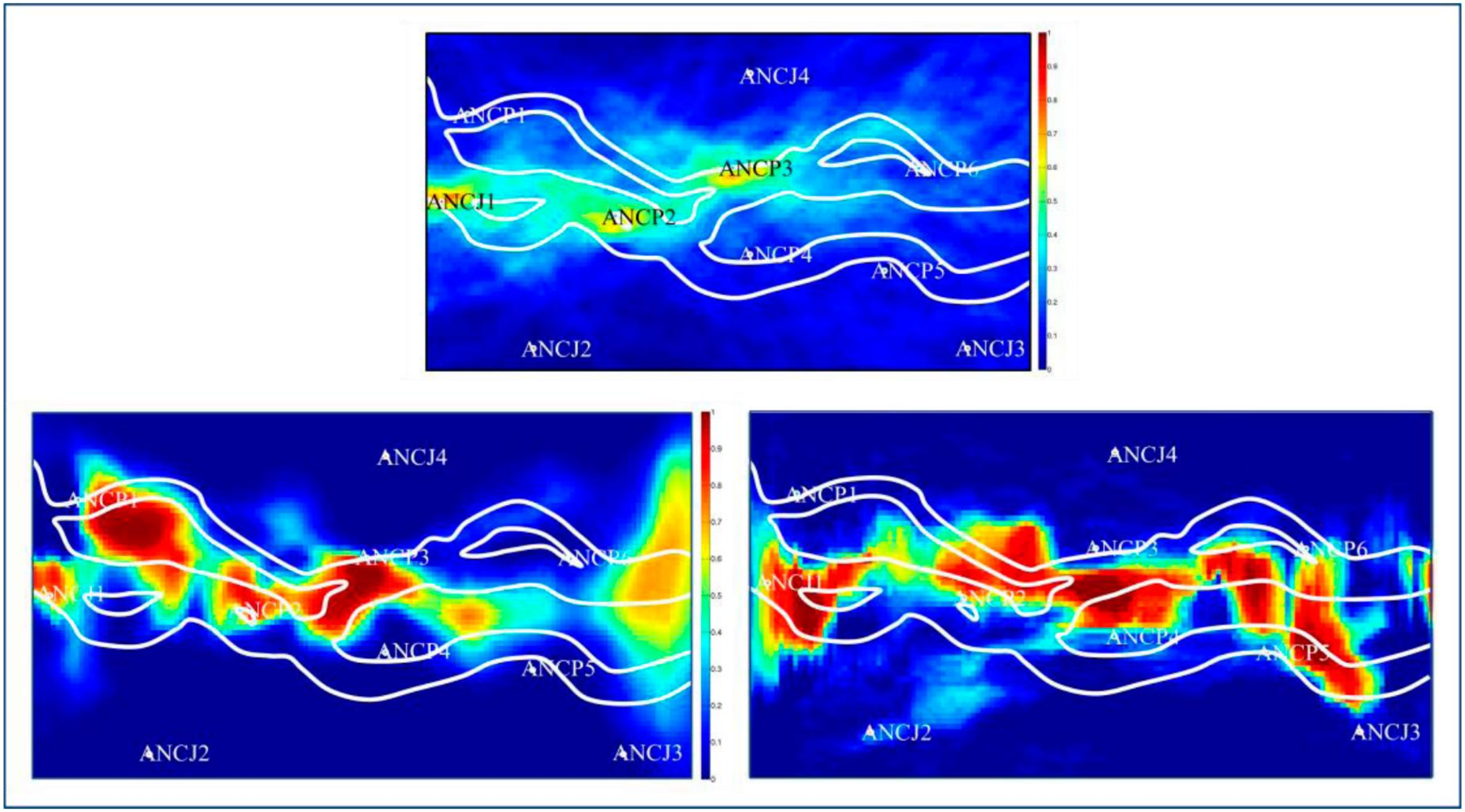

Figure 4. Probability of finding the channel cell in the initial ensemble (Top) and posterior models, with (Bottom left) and without (Bottom right) variogram normalization.

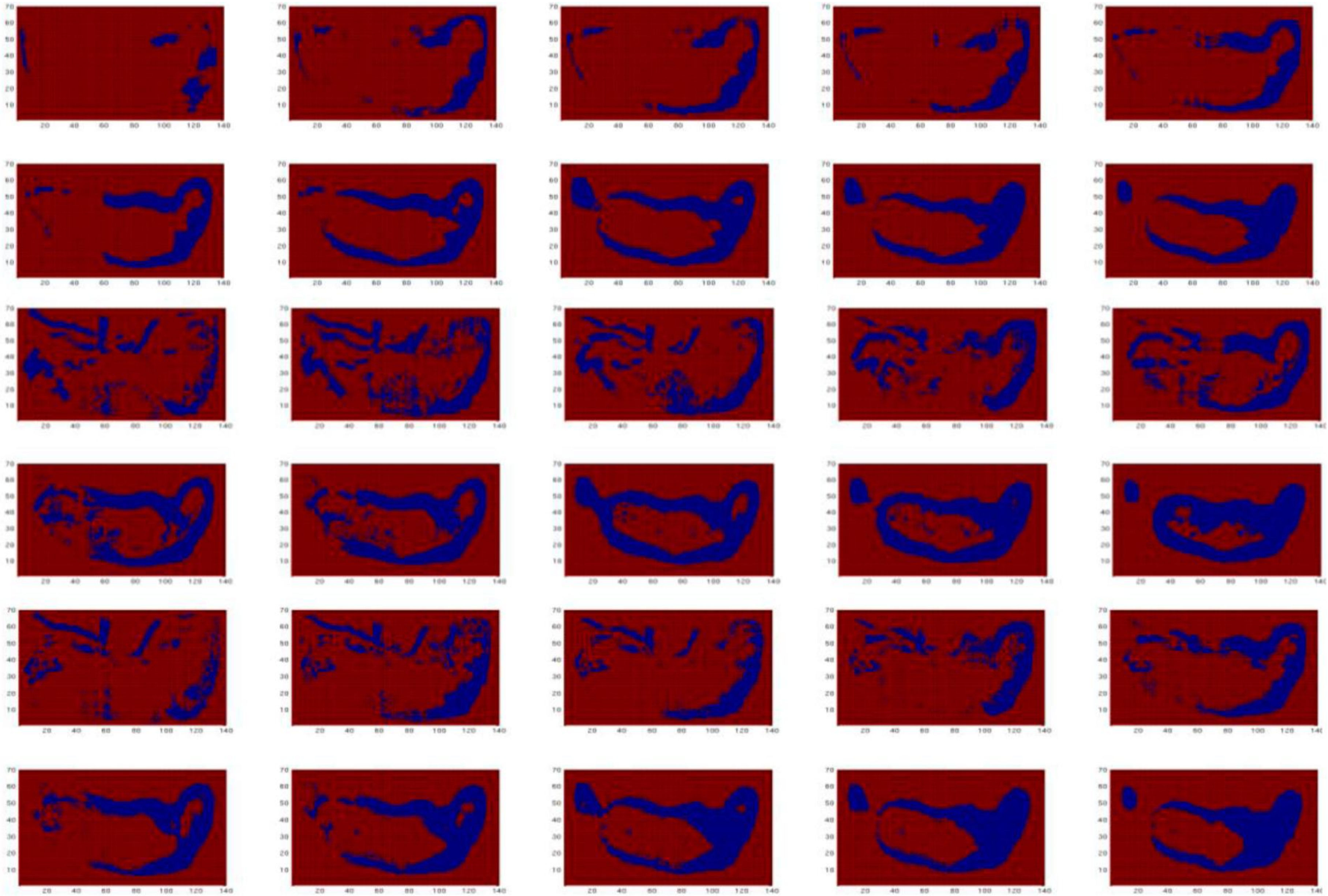

Figure 5. Map of flooded area for one of the realizations at the end of data assimilation with (bottom) and without (middle) 4D seismic observations (top).

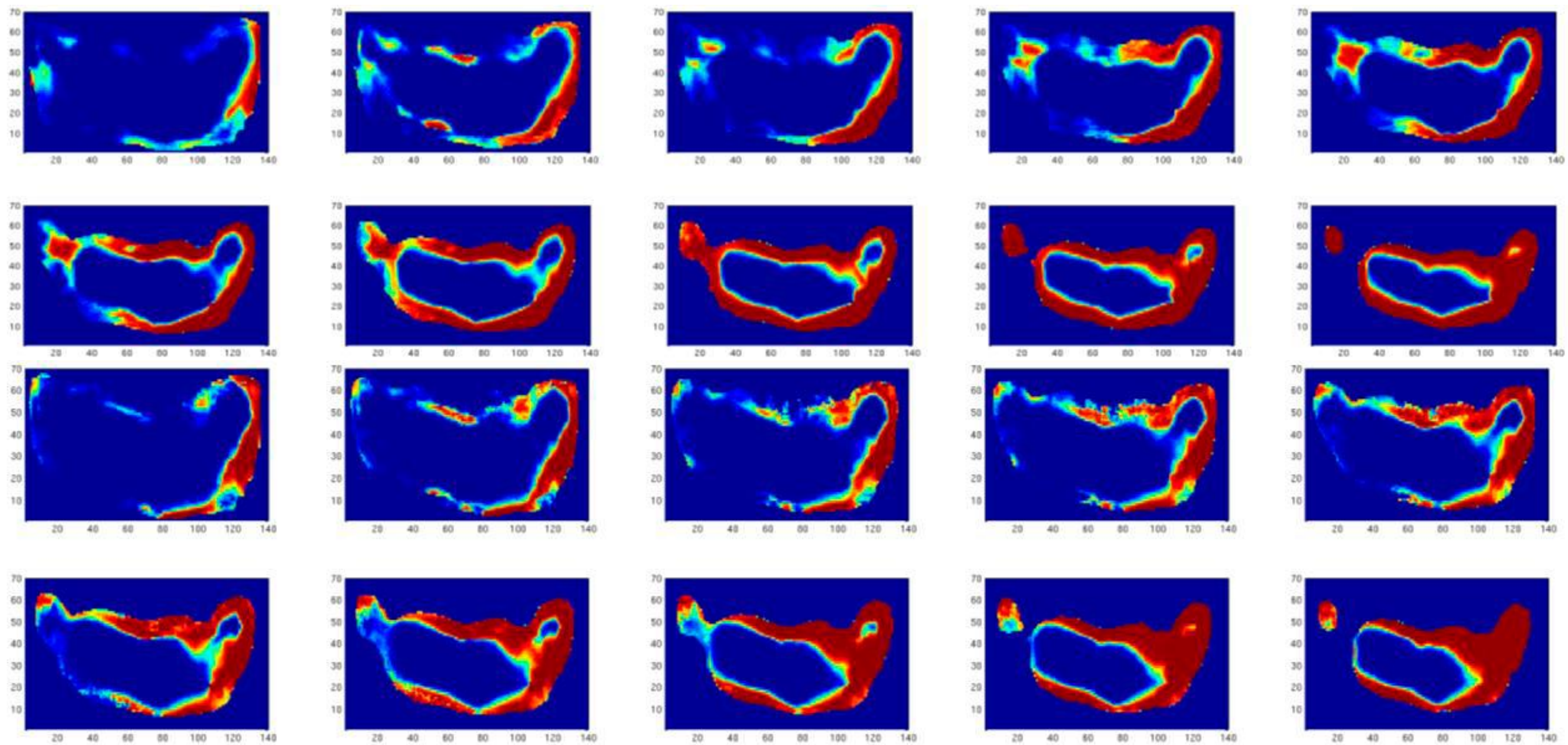

Figure 6. Map of flooded area for all of the realizations. Result of data assimilation with (bottom) and without (top) 4D seismic observations (Figure 5).

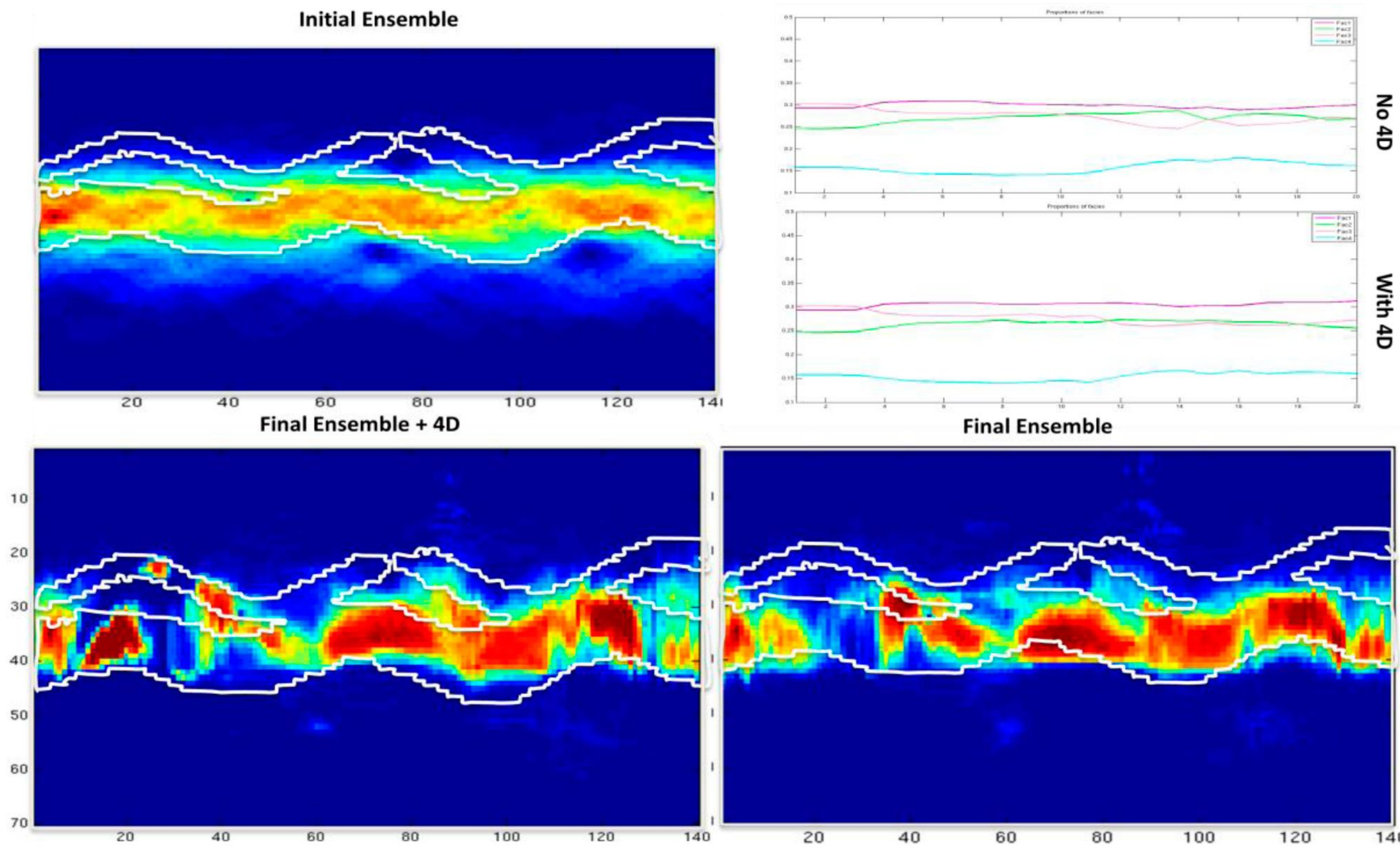

Figure 7. Probability of finding the channel cell in the initial ensemble (Top left) and posterior models, with (Bottom left) and without (Bottom right) incorporating 4D seismic observations. Evolution of four facies proportion in the 3D model during the assimilation process (Top right).

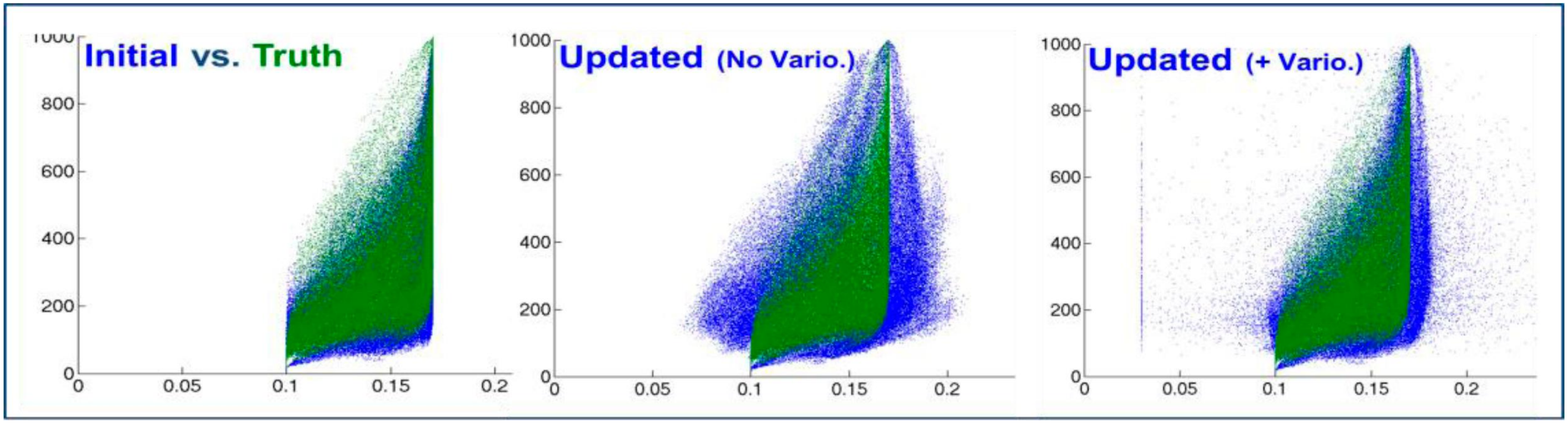

Figure 8. Initial and final correlation plot (Blue) versus the truth (Green) one, respecting initially in place correlation laws using variogram information.