Constraining Geostatistical Realizations to Temperature Data with an EnKF

Yevgeniy V. Zagayevskiy and Clayton V. Deutsch

Reservoir characterization mostly depends on the quality of a model that describes performance of the reservoir. Inverse problems are encountered when observations on the results of the process are available; they should be accounted for to constrain the static reservoir description. The Steam Assisted Gravity Drainage (SAGD) method is increasingly popular in Northern Alberta oil sand fields. The technique is employed for a long time and 4-D seismic and surveillance wells provide a large amount of temperature data. The research study aims to integrate temperature observations using the Ensemble Kalman Filter (EnKF) to characterize permeability field. This inverse problem technique is applied to synthetic example and compared to a similar study, where Sequential Self-Calibration (SSC) method was used. Steam, Thermal, and Advanced Processes Reservoir Simulator (STARS) from Computer Modeling Group Ltd. (CMG) is used. EnKF has shown reasonable estimation results with much smaller number of iterations than SSC requires. Initially assumed permeability values are the most crucial part in EnKF-based estimation and should be generated considering all prior information about primary variable.

1. Introduction

The permeability field is a key component of a fluid flow model in reservoir characterization; however, its values at every location are rarely known; hence, they should be estimated using all available primary and secondary variables. Fluid temperature can be considered as a secondary variable that is easily available at SAGD thermal in-situ oil recovery method.

The Steam Assisted Gravity Drainage or SAGD method for bitumen extraction from oil sands is widely employed for many years. The SAGD method is applied to heavy oils which are too deep to be mined, but economically feasible to extract. The SAGD mechanism is as follows. Two wells are drilled horizontally with approximate length of 1 kilometer, where an injector is placed above a producer and close to the base of a reservoir. The separation distance between the wells is often about 5 meters. Steam is generated at a nearby steam generator and is pumped into the injector at a temperature that heats the oil and lowers the viscosity to permit the oil to drain down along boundary of formed steam chamber to the producer by gravity. An emulsion of warmed oil and water condensate is carried up to the surface through the producer. The mechanism is based on the difference in densities of steam and oil. As the oil is extracted, the steam chamber grows upwards and sideward. The process of injection and production should occur continuously and simultaneously. SAGD has several advantages over conventional steamflooding, where the formation fluid is pushed by injected hot liquid with less viscosity causing possibility of fingering to occur (Butler, 1997; Deutsch and McLennan, 2005).

Reservoir temperature data will become available as the SAGD process operates. This data has the potential to improve the reservoir description and help with reservoir management. The relationship between the variables is established in Steam, Thermal, and Advanced Processes Reservoir Simulator (STARS). Using measurements of secondary variable to predict values of primary variable is deemed as a traditional inverse problem. Measurements of primary variables will also be used to lead to better estimates. An inverse problem can be solved using any inverse technique. In this paper Ensemble Kalman Filter (EnKF) is proposed to assimilate temperature observations in order to estimate permeability field. EnKF results are compared to similar work on simple synthetic example, where another inverse technique Sequential Self-Calibration (SSC) was employed (Hassanpour and Deutsch, 2009).

The paper is organized in the following manner. First, theoretical background of both inverse techniques, SSC and EnKF, is presented. Then, they are compared on synthetic 2D example comprising SAGD methodology, where thermal flow simulator STARS is applied. Both techniques show promising result of permeability estimation.

2. SSC Background

SSC is a widely used inverse modeling technique that comprises both geostatistics and optimization (Gómez-Hernánez et al., 1997). The objective of the method is to generate equally plausible realizations of primary variables honoring measurements of primary and available secondary variables, whose relationship is known. Model parameters or static variables are considered as primary variables and model responses (states) or dynamic variables are treated as secondary variables. The solution of inverse problem is achieved by modifying conditional realizations of primary variable with finite-difference approximation of perturbation vector and tuning boundary conditions. A central feature of the SSC method is the derivation of sensitivity coefficients that permit the realizations of primary variables to be altered to get closer to matching the secondary data.

In the context of SAGD, the primary variable is the permeability field and the secondary variable is temperature (Hassanpour and Deutsch, 2009). Their relationship can be expressed through mass conservation and energy conservation equations. According to Butler (Butler, 1991) thermal energy is transferred by two processes of conduction through rocks and convection through fluids (water, oil, and gas) in reservoir, though main source of energy transport in SAGD is thermal convection. Energy conservation equations for both processes in the simplest form can be expressed as follows. Equation (1) describes thermal conduction process, and Equation (2) – thermal convection (Chen, 2007).

$$\vec{u} = -K_T \cdot \nabla T \tag{1}$$

$$\frac{\partial}{\partial t}\left(\phi \cdot \sum_i \rho_i \cdot S_i \cdot C_{Vi} \cdot T + (1-\phi) \cdot \rho_s \cdot C_s \cdot T\right) - \nabla\left(K_T \cdot \nabla T + \sum_i \rho_i \cdot \frac{k_i}{\mu_i} \cdot C_{Pi} \cdot T \cdot \nabla P_i\right) = q_C - q_L, \quad i = o,w,g \tag{2}$$

where $\vec{u}$ is the heat flux vector; K_T is the total thermal conductivity; T is the temperature; ∇ is the gradient operator ($\nabla T = \frac{\partial T}{\partial x}\vec{i} + \frac{\partial T}{\partial y}\vec{j} + \frac{\partial T}{\partial z}\vec{k}$); t is the time; ϕ is the porosity of porous medium; i stands for type of fluid: water (w), oil (o), and gas (g); ρ_i and ρ_s are the density of i^{th} fluid and rock per unit volume; S_i is the i^{th} fluid saturation; C_{Vi} and C_{Pi} is the heat capacities of i^{th} fluid at constant volume and constant pressure; C_s is the specific heat capacity of rock; k_i and μ_i are the permeability and viscosity of i^{th} fluid; P_i is the pressure applied to i^{th} fluid, q_c and q_L are the heat source item and heat loss to overburden and underburden. Exerted pressure to each fluid is found from mass conservation Equation (3) (Chen, 2007).

$$\frac{\partial(\phi \cdot \rho_i \cdot S_i)}{\partial t} = \nabla\left(\rho_i \cdot \frac{k_i}{\mu_i} \cdot \nabla P_i\right) + q_i, \quad i = w,o,g \tag{3}$$

where q_i is the production or injection rate of i^{th} fluid.

The methodology of SSC is based on four main steps (Gómez-Hernánez et al., 1997):

(1) generate the required number of realizations for permeability variable k conditional to permeability measurements if they are available using a valid variogram model, otherwise generate realizations resembling possible actual geological pattern of permeability field;

(2) calculate temperature realizations T from permeability realizations using their known relationship (Equations (2) and (3));

(3) modify permeability field to honor temperature measurements (optimization process) by finite-difference approach. Use objective function (Equation (4)) as a criterion for goodness of estimation. It is minimized in respect to permeability perturbation Δk. The concept of master points is applied, at whose locations temperature values are brought to their measured ones and are used to modify rest of the field by ordinary kriging and known semivariogram model of permeability. Sensitivity coefficients ($\partial T / \partial k \approx \Delta T / \Delta k$) derived from Equations (2) and (3) are used in linear approximation of temperature realizations in objective function. Since the relationship between perturbation of permeability and temperature is nonlinear and linear approximation of temperature vector is applied, iterative procedure is conducted to achieve minimum of the objective function.

(4) repeat steps (2) and (3) iteratively until the objective function is not minimized.

$$O^{(n)} = \sum_{i=1}^{N_{obs}} w_i \cdot \left(T^{obs} - T^{cal,(n)}\right)^2 = \left(\left\{T^{obs}\right\} - \left\{T^{cal,(n)}\right\}\right)^T \cdot [W] \cdot \left(\left\{T^{obs}\right\} - \left\{T^{cal,(n)}\right\}\right) \tag{4}$$

where { } represents a vector and [ ] represents a matrix, $O^{(n)}$ is the objective or penalty function at iteration step n; T^{obs} is the measured value of temperature; $T^{cal,(n)}$ is the calculated value of temperature at step n; N_{obs} is the number of temperature measurements; w_i is the weight of i^{th} observation proportional to inverse of covariance matrix [R] of temperature observation error, which consists of measurement and estimation errors, [W] α [R]$^{-1}$.

$T^{cal,(n)}$ is linearly approximated through a first order Taylor expansion (Equation (5)) and substituted to Equation (4). Thus, not exactly the objective function, but a linear approximation (Equation (6)) is minimized in respect to perturbation Δk. Because of approximation minimum value of $O^{(n)}$ cannot be achieved at first calculation, iterative procedure is required.

$$\left\{T^{cal,(n)}\right\} \approx \left\{T^{cal,(n-1)}\right\} + \left.\frac{\partial\left\{T^{cal}\right\}}{\partial\left\{\Delta k\right\}}\right|_{\left\{T^{cal}\right\}=\left\{T^{cal,(n-1)}\right\}} \cdot \left\{\Delta k\right\} \tag{5}$$

$$O^{(n)} \approx O^{(n-1)} + \left\{D\right\}^T\left\{\Delta k\right\} + \left\{\Delta k\right\}^T\left[C\right]\left\{\Delta k\right\} \tag{6}$$

where {D} and [C] are some coefficients stored in vector and matrix forms.

Once the objective function reaches its preset minimum value, it is deemed that all corresponding permeability and temperature realizations represent actual nature of variables distribution. Further analysis can be applied and uncertainty can be assessed. Implementation of SSC is shown later on synthetic example and compared to estimation results of EnKF.

For more details of SSC theory and its application in temperature integration the reader is directed to papers (Gómez-Hernánez et al., 1997), (Hassanpour and Deutsch, 2009), and monograph (Wen et al., 2005).

3. EnKF Background

Another inverse problem technique, Ensemble Kalman Filter (EnKF), can be applied to obtain set of realizations presenting plausible permeability fields using measurements of permeability and temperature variables. The method is aimed to find the maximum likelihood estimate of primary variable of dynamic model and is good for large-scale systems and Gaussian variables (Evensen, 2007).

EnKF is a recursive two-step procedure devised for data assimilation and variable forecast. First step is called forecast or forward step, which predicts variable's value for future time step n_{t+1} using updated variable values from previous time step n_t and relationships between variables in a model. Second step is called update or analysis step, where predicted values are modified to honor measurements of both primary and secondary variables. Mathematically two steps can be presented in form of Equations (7) and (8).

$$X_t^f = M\left(X_{t-1}^a\right) + E_{t-1}^{\text{model}} \quad \leftarrow \quad \textit{forecast equation} \tag{7}$$

$$X_t^a = X_t^f + K_t \cdot \left(D - H \cdot X_t^f\right) \quad \leftarrow \quad \textit{analysis equation} \tag{8}$$

$$K_t = \hat{C}_t^f \cdot H^T \cdot \left(H \cdot \hat{C}_t^f \cdot H^T + R\right)^{-1} \quad \leftarrow \quad \textit{Kalman gain} \tag{9}$$

where X^f_t is the matrix of all variables in raw with possible values in columns at forecast step n_t; X^a_t is the matrix of all variables values at analysis step n_t; M is the model operator that establish relationship between all variables, not necessarily linear; E^{model} is the model error, usually is assumed to be zero; D is the matrix with available observations; K is the Kalman gain matrix, whose coefficients are exactly same to kriging weights; H is the observation matrix consisting of 0s and 1s; $\hat{C}_t^f$ is the sample covariance matrix at time n_t calculated from matrix X^f_t.

The methodology of EnKF is as follows (Aanonsen et al., 2009):

(1) Generate initial realizations of the primary variable using all prior information about it. Actual semivariogram model is highly desirable to track geological patterns. Initial values are very important and mostly determine forecasted values, whose covariance matrix is being preserved during assimilation steps;

(2) Use generated realizations in forecast Equation (7) to get initial values for rest of the variables and calculate sample covariance matrix;

(3) Update all variables through analysis Equation (8) where all available measurements are assimilated;

(4) Proceed to next forecast step and compute objective function similar to Equation (4). If acceptable minimum level of objective function is achieved stop EnKF process, otherwise repeat steps (3) and (4) until objective function is not minimized.

An ensemble size of at least 60 realizations, to conduct at least 2 recursive steps, and to use all available data with semivariogram model resembling actual distribution (Zagayevskiy and Deutsch, 2010).

4. Comparison of the Methods on Synthetic Example

The comparison of two inverse modeling techniques is conducted on a synthetic SAGD example consisting of two variable types 1) primary variable type is permeability, and 2) secondary variable type is temperature. While permeability field is considered to be static, temperature field changes in time, therefore, it is dynamic. 5 time steps are examined that altogether lasts 10 years with increment step of 2 years (720 days). So, total number of examined variables is six: one permeability field and five temperature fields. The relationship between variables is established through thermal flow simulator STARS. For most cases only temperature measurements are available, whose initial values before launching an inverse modeling technique are 7^0 Celsius. The objective of current example is to estimate permeability field conditional to temperature data obtained at five time steps using SSC and EnKF methods and compare the results. For this reason base case is used as reference case to assess quality of permeability estimation and temperature prediction.

The settings of this simple 2D example are as follows. Vertical cross section in *X-Z* plane is examined. The model consists of N_X = 13 rectangular blocks in *X* direction and N_Z = 10 rectangular blocks in *Z* direction, whose sizes are 10 meters and 2 meters respectively. Values of variables are assigned to centers of the blocks. Injector well is placed above producer and is in the middle of the model closer to its bottom (Figure 1). Steam is injected at temperature of 223^0 Celsius. Observation locations with total number of 12 are shown in Figure 1 and are extracted from base case. Base cases of permeability, porosity and temperature fields are presented in Figure 2. It is clear that permeability variable is characterized by two regions of high and low values: 1000 mD on the right side of the model and 10 mD on the left side. Base cases of temperature fields are not so simple. Associated porosity field has same distribution pattern as permeability field with values of 0.10 and 0.25. The histograms of base case and initial permeability fields are presented in Figure 3.

Implementation details and results of SSC method can be found in the paper (Hassanpour and Deutsch, 2009). Procedure of getting estimates from EnKF-based inverse technique is discussed below. Ensemble size of 100 realizations is chosen as optimal one (Zagayevskiy and Deutsch, 2010). All negative values of EnKF permeability estimates are set to 5 mD to avoid crash of thermal flow simulator. Five different initialization cases of EnKF are examined in order to conduct small sensitivity analysis and to achieve better results. The cases are as follows.

1) case 1 – initial permeability values are generated unconditional using sequential Gaussian simulation (SGS) and follow normal distribution, where negative values are inversed to positive ones. Later all values are multiplied by coefficient to reach mean values of double value of base case;

2) case 2 – initial permeability values follow uniform distribution with mean close to base case mean and are unconditional;

3) case 3 – initial permeability values are generated unconditional using SGS and correct semivariogram model (Figure 4 and Equation (10)). The values are multiplied by a coefficient to reach mean of base case and follow normal distribution;

4) case 4 – same as case 3, but mean of generated values is doubled and, hence, twice bigger than mean of base case;

5) case 5 – same as case 4, but later permeability measurements are used in EnKF update step sampled at same observation locations.

$$\gamma(h) = 0 * Nugget + 0.1 * Gaussian_{\substack{R_{MAX}=+\infty \\ R_{MIN}=0.1\,m}}(h) + 0.9 * Gaussian_{\substack{R_{MAX}=+\infty \\ R_{MIN}=60.0\,m}}(h) \quad (10)$$

In the context of this example assimilation/recursive step in EnKF has close meaning to iteration step in SSC terminology.

Relative objective function (slightly modified Equation (4)) is used as criterion for goodness of estimation results. Relative objective function is found as division of objective function at i^{th} assimilation/iteration step by objective function at initialization step. Once it reaches value close to zero, the assimilation process in EnKF is stopped. For EnKF 6 assimilation steps are enough to reach stable results, but this number is much higher for SSC and equals to 40 (Figure 5). Histograms of EnKF estimates after sixth assimilation step are shown in Figure 3 from which it is clear that estimates tend to Gaussian distribution. This is main reason of poorer results of permeability estimation by EnKF compared to SSC. The permeability estimation results are presented in Figure 6. Both techniques are able to reproduce zones with high and low permeability values, especially when more information is used for EnKF initializations. Moreover, temperature fields are reproduced reasonably well (Figure 7). But for EnKF initial permeability values should have higher mean than base case mean, since estimates tend to be Gaussian with lower mean than of initial values. History matching of temperature variable at observation locations #1, 3, 8, and 10 is shown in Figure 8 and shows reasonable reproduction.

Overall average difference (OAD) term is used to compare results accuracy for EnKF permeability estimates and predicted temperature values. Its expression is shown in Equation (11) and OAD values of all variables for 5 examined cases are tabulated in Table 1. It is clear that case 5 is the best, where permeability measurements are used and correct semivariogram model of permeability is considered. Beside it, case 4 is second best, where only correct semivariogram model is taken into account.

$$OAD = \sum_{i=1}^{N_{real}} \sum_{j=1}^{N_X \times N_Z} \frac{\left| X_j^{base\ case} - X_{i,j}^{cal} \right|}{N_{real} \cdot N_X \cdot N_Z} \quad (11)$$

where N_{real} is the number of realizations or ensemble size; N_X is the number of blocks in X direction; N_Z is the number of blocks in Z direction; $X^{base\ case}{}_j$ is the variable value at j^{th} block of base case; $X^{cal}{}_{i,j}$ is the variable value at j^{th} block of i^{th} realization.

The convergence characteristics of estimated realizations of every variable for case 4 at observation location #1 are shown in Figure . Average of realizations for most temperature variables reaches its actual value at second assimilation step, but permeability average is far from its true value, since permeability measurements are not used in this case.

Table 1 – Comparison table for overall average differences of 5 cases after six assimilation steps.

Variable	OAD for Permeability		OAD for 720 day temperature		OAD for 1440 day temperature		OAD for 2160 day temperature		OAD for 2880 day temperature		OAD for 3600 day temperature	
Case #	Mean	St.dev.	Mean	St.dev.	Mean	St.dev.	Mean	St.dev.	Mean	St.dev.	Mean	St.dev.
(1) *Random initialization*	778.0	65.6	2.214	0.673	5.108	2.449	7.240	3.114	8.819	2.826	10.937	4.009
(2) *Uniform initialization*	416.4	13.9	7.423	1.477	3.353	0.832	5.070	1.220	10.961	1.619	13.534	2.526
(3) *Normal initialization with correct semivariogram structure*	334.7	18.9	2.940	0.135	7.098	0.382	8.654	0.370	11.608	0.444	9.950	0.649
(4) *Normal initialization with higher mean and*	445.6	34.5	0.659	0.227	0.977	0.217	1.603	0.330	2.259	0.672	2.903	1.102

correct semivariogram structure											
***(5)** Normal initialization with higher mean, correct semivariogram structure and conditional to permeability measurements*	292.4	21.0	1.330	0.181	0.931	0.127	1.335	0.326	2.182	0.882	3.214	1.722

5. Conclusion

Two inverse modeling techniques, SSC and EnKF, are compared on simple 2D synthetic example featuring SAGD implementation. Although SSC method has shown better estimation results for permeability field, it required 40 iteration steps to reach minimum of relative objective function, what is much higher compared to EnKF, where only 6 iteration steps are enough. So, EnKF is computationally faster than SSC, but is not so accurate and heavily depends on prior information (initial values). However, both techniques are able to distinguish high and low permeability zones, especially its boundary. Also, EnKF temperature prediction results are as good as of SSC. Moreover, EnKF is much simpler than SSC and easily handles large-scale systems.

References

Aanonsen, S.I., Nævdal, G., Oliver, D.S., Reynolds, A.C. and Valles, B., 2009, The Ensemble Kalman Filter in Reservoir Engineering - a Review, *SPE Journal*, 393-412.

Butler, R.M., 1991, *Thermal Recovery of Oil and Bitumen*, Prentice Hall, Englewood Cliffs, New Jersey, 524 pp.

Chen Z., 2007, *Reservoir Simulation: Mathematical Techniques in Oil Recovery*, CBMS-NSF Regional Conference Series in Applied Mathematics; 77, SIAM, Philadelphia, 218 pp.

Deutsch, C.V. and Journel, A.G., 1998, *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, New York, 2nd Ed., 369 pp.

Deutsch, C.V. and McLennan, J.A., 2005, *Guide to SAGD (Steam Assisted Gravity Drainage) Reservoir Characterization Using Geostatistics,* Centre for Computational Geostatistics, Guidebook Series Vol. 3, 129 pp.

Evensen, G., 2007, *Data Assimilation: The Ensemble Kalman Filter*, Springer Verlag, Berlin, 305 pp.

Gómez-Hernánez, J.J., Sahuquillo, A. and Capilla, J.E., 1997, Stochastic Simulation of Transmissivity Fields Conditional to Both Transmissivity and Piezometric Data – I. Tehory, *Journal of Hydrology* 203, 162-174.

Hassanpour, R.M. and Deutsch, C.V., 2009, Framework for Surveillance Temperature Data Integration, *Centre for Computational Geostatistics* 11, 208-1 – 208-10.

Wen, X.H., Deutsch., C.V., Cullick, A.S., and Reza, Z.A., 2005, Integration *of Production Data in Generating Reservoir Models*, Centre for Computational Geostatistics, Monograph Series Vol. 1, Edmonton, Alberta, 219 pp.

Zagayevskiy, Y.V. and Deutsch, C.V., 2010, Characteristics of EnKF for Geostatistical Inverse Problems, *Centre for Computational Geostatistics* 12, 125-1 – 125-10.

Figures

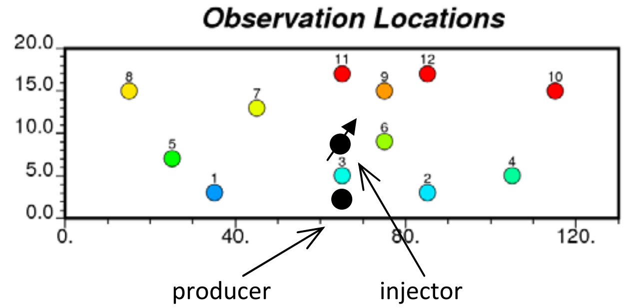

Figure 1 – *X-Z* cross section of 2D model on which SAGD method is applied. Dots represent observation locations with their order number beside. Black dots represent producer and injector wells.

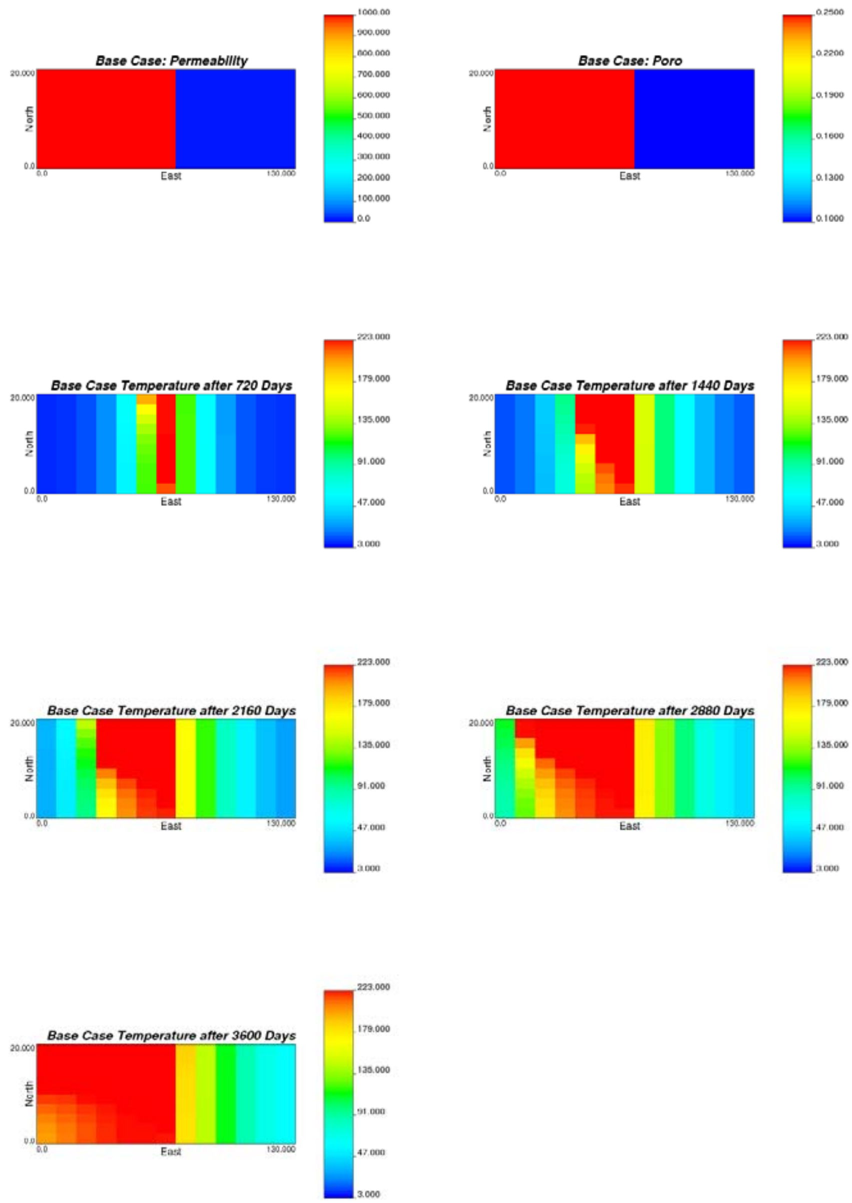

Figure 2 – Base cases of permeability, porosity and temperature at different time steps.

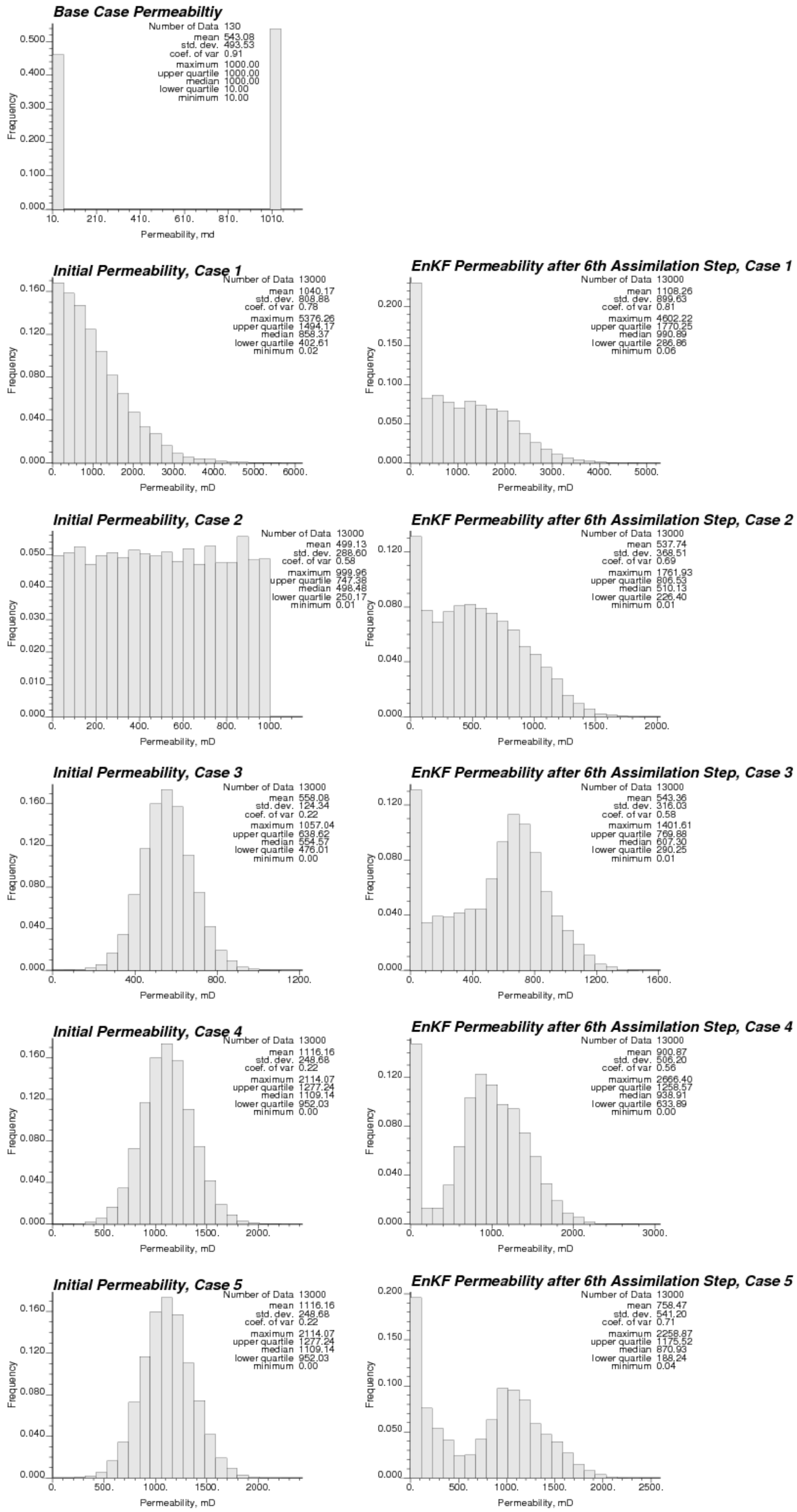

Figure 3 – Histograms of permeability base case and EnKF permeability estimates at initialization and step and after 6 iterations.

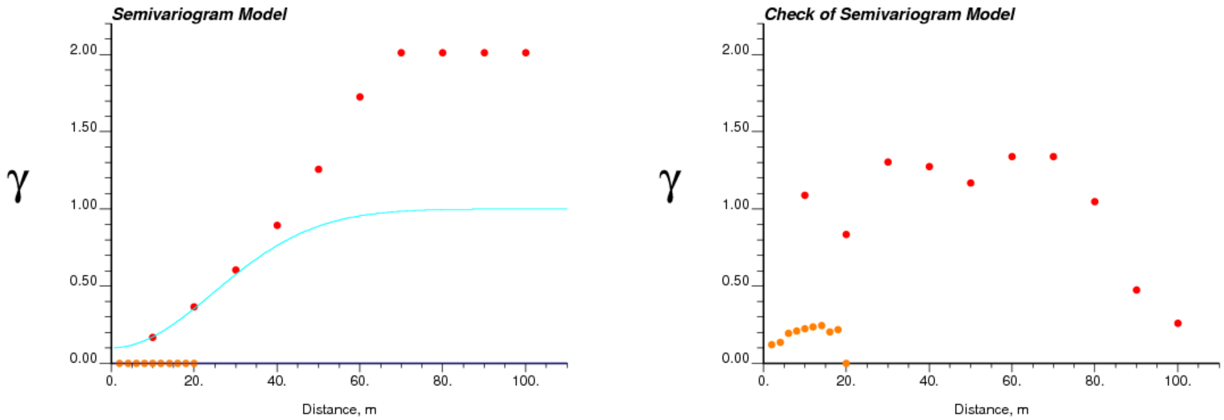

Figure 4 – Semivariogram model of permeability base case (left) and experimental semivariogram of permeability initial values generated using the semivariogram model (right). Orange dots represent maximum continuity direction, red dots stand for minimum continuity direction.

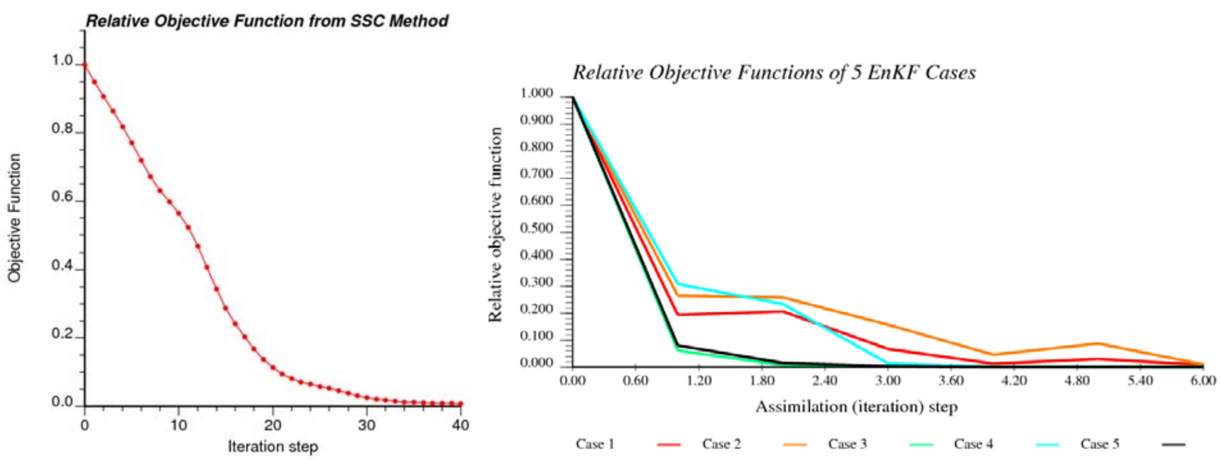

Figure 5 – Comparison of relative objective functions of temperature variable from SSC and EnKF methods.

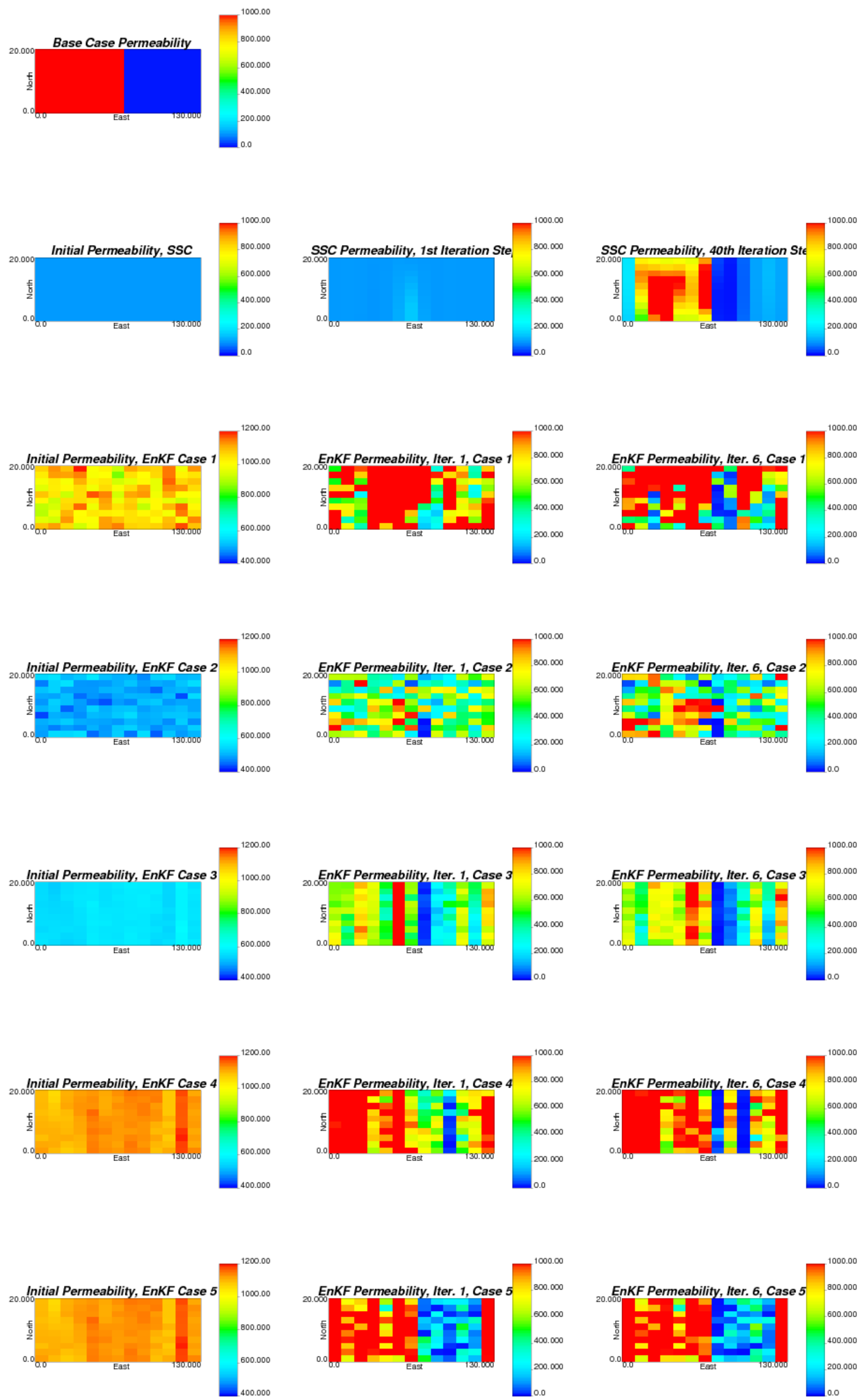

Figure 6 – E-type maps of permeability estimates derived from SSC method after 40 iteration steps and 5 cases of EnKF method after 6 iteration steps.

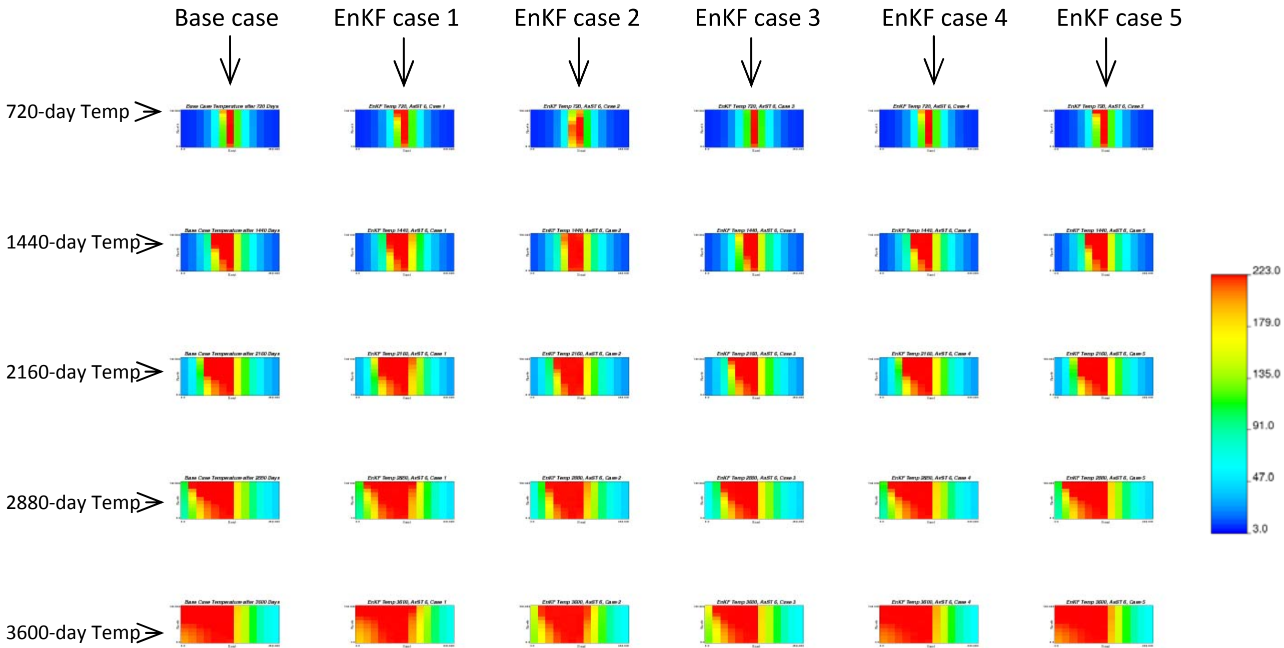

Figure 7 – E-type maps of temperature estimates after 6 iteration steps from EnKF inverse modeling method.

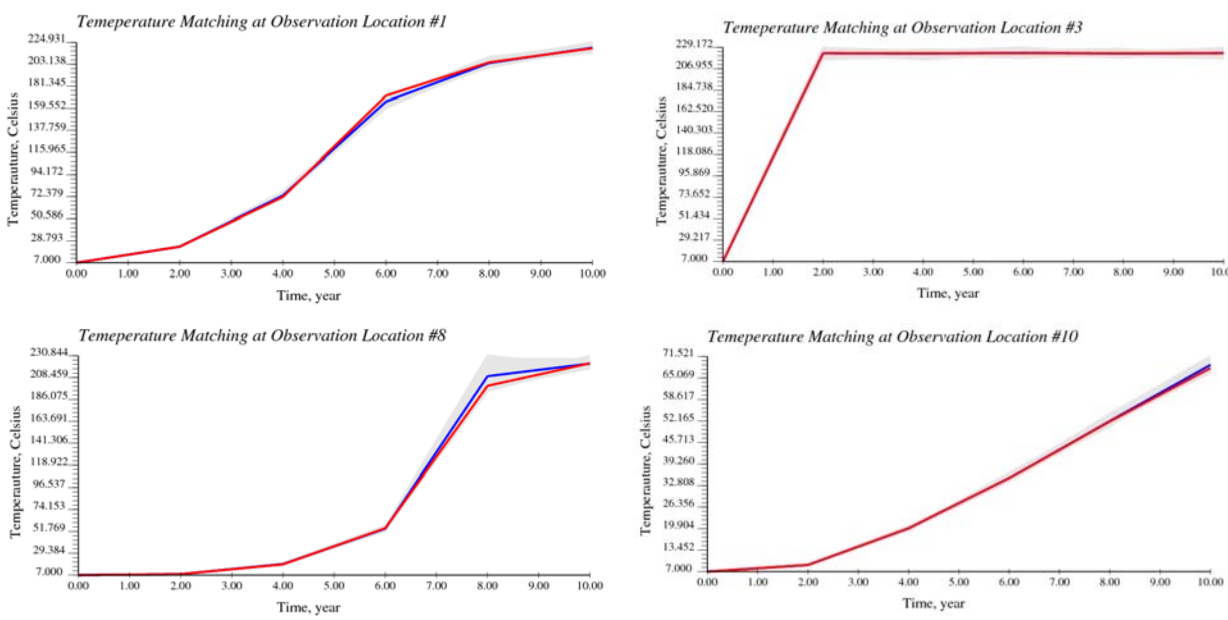

Figure 8 – History matching of temperature variable for observation locations #1, 3, 8, and 10. Red line shows actual values, grey lines stand for possible realizations from EnKF after 6 iteration steps and blue line is their average.

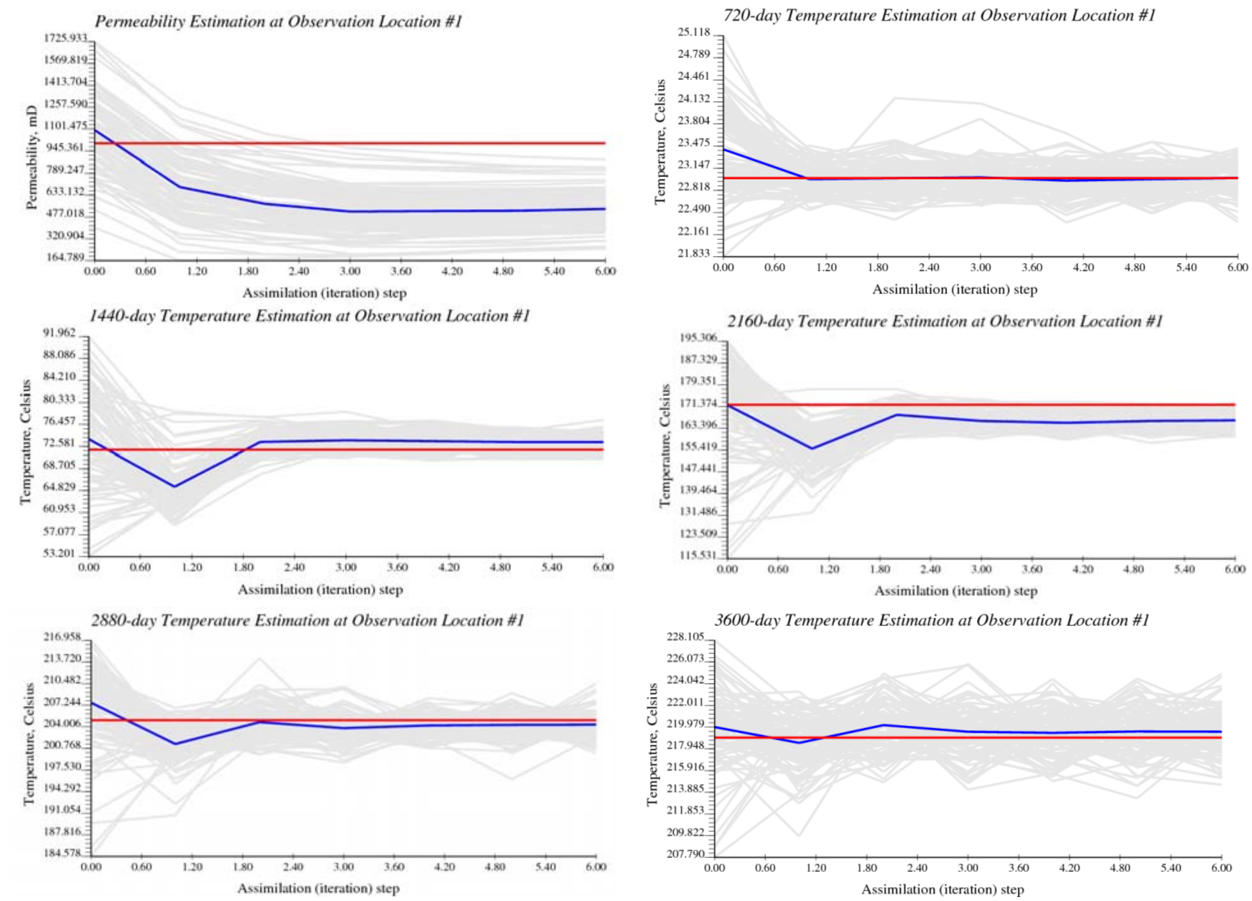

Figure 9 – Convergence of EnKF permeability and temperature realizations at observation location #1 for case 4, where no permeability measurements are used for estimation. Red line represents actual variable value, grey lines stand for possible variable realizations and blue line is their average.