INVERSE MODELING USING ENSEMBLE KALMAN FILTER (EnKF) APLICATION IN RESERVOIR ENGINEERING

Asist.drd.Bogdan Sebacher
Technical University of Civil Engineering Bucharest, Romania
E-mail: bogdansebacher@yahoo.com
Conf.dr. Mircea Ariciuc
Technical University of Civil Engineering Bucharest, Romania
E-mail: mirceaariciuc@yahoo.com

1. INVERSE MODELING

Time series models can be used to predict the behavior of environmental processes. However these models are derived from the data and do not include physical knowledge of the process. Furthermore, measurements alone do not generally provide a complete picture of the process. Especially in case of processes that vary in space and time it is very hard to reconstruct the spatial and temporal patterns only from data. Physically based models, deterministic or stochastic, produce results that are spatially and temporally consistent. However these models are usually not able to accurately reproduce the measurements that are available. The information provided by the models and by the measurements is often complementary. Therefore it is important to study a methodology for integrating measurements with the physically based mathematical models. This methodology is called **data assimilation**. The data assimilation techniques gives us the possibility of assessment of the system variables based on the measurements, therefore we say that we deal with an inverse modeling procedure. An inverse problem is referring to the determination of the physical properties of the system or information about these properties, given the observed response of the system to some stimulus (measurements).The inverse problem is in most all cases **ill-posed** problem in the sense that at least one of the Hadamard condition is not accomplished. A **well-posed** problem is a problem that satisfies three conditions (Hadamard):

(1) The problem has a solution.
(2) The solution is unique.
(3) The solution is a continuous function of the problem data.

We state that the general inverse problem is related to determine plausible values for important model parameters given the (uncertain) data and assumed that exist a theoretical model relating the observed data to the model. The inverse modeling could be made using variational methods (gradients based optimization methods) or sequential methods (without gradients). One of the non gradient method for inverse modeling is the Ensemble Kalman Filter. Ensemble Kalman Filter (EnKF) (Evensen 1997) is a Monte Carlo technique introduced in the Kalman filtering framework where the probability density of the state is represented by an ensemble of possible realizations that are simultaneously updated. The Kalman filter is an efficient recursive filter that estimates the state of a linear dynamical system from a series of noisy measurements. The method was introduced by Kalman (1960) and is more a theoretical model suited for linear cases.

2. PARAMETERS ESTIMATION WITH ENSEMBLE KALMAN FILTER (EnKF)

Let consider a dynamical model described by a nonlinear system of partial differential equations and assume that these equations with boundary conditions has been discretized in a space and the errors in the errors in the boundary conditions are zero. We denote with $u = u(t) \in \mathbb{R}^{N_u}$ the discretized approximation of the solution at time t. We assume that the model depends on some poorly known parameters $m \in \mathbb{R}^{N_m}$ to be estimated. Then the time equation of the state became $\frac{du(t)}{dt} = F(u,m) + w_m(t), u(t_0) = u_0$ being the initial state and $w_m(t)$ is the model error. The initial state u_0 usually is a random variable representing the uncertainty in the initial condition. In reservoirs engineering characterization u contain the dynamical variables pressure and saturation. We further assume that the model is constrained by some noisy measurements collected at times $t_1, t_2, \dots t_n$ and the measurements are related to the state through the general nonlinear relation $d_i = g_i(u(t_i), m) + v_i^d \in \mathbb{R}^{N_d}$, where $v_i^d \sim N(0, C_{d_n})$ represents the measurements errors. In the parameter estimation problems with EnKF we define an augmented state vector that contains parameters, dynamical variables and the simulated data $x_k = \begin{bmatrix} m & u_k & g(m,u_k) \end{bmatrix}^T \in \mathbb{R}^{N_m + N_u + N_d}$.

The Ensemble Kalman filter is a Monte Carlo method where multiple plausible models are simultaneously updated instead of a single model as in the traditional history matching. In the first step, an ensemble of N_e states $\{x_1, x_2, \dots x_{N_e}\}$ is generated to represent the uncertainty in the initial state $x_0^t = x(t_0)$. In the second step, named the forecast step, the stochastic model propagates the distribution of the true state through the model equations according to $x_i^f(t_k) = M(x_i^a(t_{k-1})) + w_i(t_k)$ where $w_i(t_k)$ represents a realization of the noise process of the model error. For the forecasted state we calculate the mean $\overline{x}^f(t_k) = \frac{1}{N_e}\sum_{k=1}^{N_e} x_k^f(t_k)$ and based on the mean the covariance $C^f(t_k) = \frac{1}{N_e - 1} E^f(t_k) E^f(t_k)^T$, where $E^f(t_k) = \begin{bmatrix} x_1^f(t_k) - \overline{x}^f(t_k) & x_2^f(t_k) - \overline{x}^f(t_k) & \dots & x_{N_e}^f(t_k) - \overline{x}^f(t_k) \end{bmatrix}^T$. When the measurements become available values of each ensemble member are adjusted based on the Kalman equation $x_i^a(t_k) = x_i^f(t_k) + K(t_k)[d_{obs}(t_k) - H(t_k)x_i^f(t_k) + v_i(t_k)]$, where $K(t_k) = C^f(t_k)H(t_k)^T[H(t_k)C^f(t_k)H(t_k)^T + R(t_k)]^{-1}$ is the Kalman gain, $H(t_k)$ is the observation operator, $R(t_k)$ is the covariance matrix of the measurements error and $v_i(t_k)$ is the realization of the noise added to observed measurements. At the end of the assimilation period we will have an estimator for each parameter, defined by the ensemble mean together with his uncertainty given by the forecasted covariance matrix.

4. APLICATION: ESTIMATION OF THE FACIES DISTRIBUTION IN A RESERVOIR DOMAIN

4.1 INTRODUCTION

A facies is a body of rock with relatively uniform characteristics such as porosity or permeability and there are significant differences in these petrophysical characteristics between facies types. Consequently, the spatial distribution and the location of the different facies types presents in the reservoir have significant impact on fluid flow. Therefore the estimation of the facies distribution into a reservoir domain has two goals; one is related with the high predictive capacity of flow measurements of a reservoirs that are geologic realistic and the other one is related with the need of accurate properties to predict the consequences of changing condition in the reservoir lifetime (e.g. the response to enhance recovery).The first step in the assessment of the facies distribution is to define a geological model to simulate plausible facies maps that are consistent with the prior knowledge about subsurface geology (numbers of facies type that occur, facies proportion, core information, type of patterns etc).The geological model can be construct based on a parameterization of the facies maps with some uncertain variables(parameters)through which we are able to adjust them such that these will honor the underlying knowledge of subsurface geology. These parameters can be calibrated (hence, as well the maps) using additional informations about flow measurements into a process named in reservoir engineering History Matching (HM) (Oliver et al. 2008). In our study, for simulation of the facies maps we used a truncation plurigaussian method (Galli et al. 1994) (data assimilation in general).The method consists in a projection from a continuous space (the space determined by the Gaussian Random Fields) into a discrete space (the facies maps space) through a map designed in the GRF space and defined by intersection of some threshold (values for one GRF, curves for two GRF, surfaces for three GRF, etc.). In this paper we investigate the reservoirs having a complex geology defined by three facies type that occur each two could having direct contact.

4.2 THE GEOLOGICAL SIMULATION MODEL FOR FACIES DISTRIBUTION

In every well (production or injection) of our reservoir we are sure (perfect observations) about the type of facies present there, so we can use that information for the whole grid where the well is situated. Let be (i, j) a grid block of our reservoir domain and suppose that in this grid facies type 1 occur. Consider $A_k^{i,j}$ the event in which in the grid (i,j) the facies type k occur, then, in terms of probabilities we have $P(A_1^{i,j})=1, P(A_2^{i,j})=0, P(A_3^{i,j})=0$. Hence, every facies type generates a field defined on the reservoir domain whose values in grids cells are 0 or 1 depending on whether the facies type occurs or not in those locations. These fields have binary values (discrete fields) and to estimate them we will use probabilities fields which are defined as random fields with values in interval [0,1] having spatial correlations. The probabilities fields are modeled through projection in [0,1] of some Gaussian Random Fields defined on the reservoir domain with a truncation function(projection function).The truncation function used is $\varphi_m(t)=\begin{cases}-\frac{|t|}{m}+1 & if \quad |t|\le m\\ 0 & if \quad |t|>m\end{cases}$, where m represents a truncation parameter.

The parameter m can be initially chosen based on geological prior knowledge about the facies proportions in the certain case and estimated in the process of history matching. We start with two Random Gaussian Fields y_1 and y_2 defined on the entire reservoir domain and α_1 and α_2 will be the projection of the Gaussian fields in [0,1]. Then $\alpha_k^{i,j}=\varphi_{m_k}(y_k^{i,j})$, $k\in\{1,2\}$, represent an estimator for the probability of occurrence of the facies k at the location (i,j), where m_k represent the truncation parameter for the random field y_k. In order to appoint the estimator for the probability of the third facies we will use the following rules: $\alpha_3^{i,j}=\begin{cases}1-(\alpha_1^{i,j}+\alpha_2^{i,j}) & if\ \alpha_1^{i,j}+\alpha_2^{i,j}<1\\ 0 & otherwise\end{cases}$. At a certain location (i,j) we assign facies of type $k\in\{1,2,3\}$ if $\alpha_k^{i,j}=\max\{\alpha_r^{i,j}, r=1,2,3\}$ with the convention that if $\alpha_1^{i,j}=\alpha_2^{i,j}\ge\alpha_3^{i,j}$ we assign facies type 1 and if $\alpha_2^{i,j}=\alpha_3^{i,j}>\alpha_1^{i,j}$ we assign facies type 2(maximization criterion). Also, if we represent the point $(y_1^{i,j}, y_2^{i,j})$ in the Cartesian plain $(y_1 O y_2)$ we can assign the facies type to the grid (i,j) based on the region where the point $(y_1^{i,j}, y_2^{i,j})$ falls in that two dimensional space(see Figure 1).

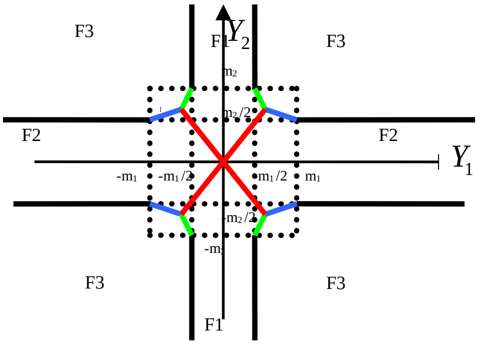

Figure 1: Truncation map for Gaussian Random Fields

4.3 ENSEMBLE KALMAN FILTER IMPLEMENTATION FOR FACIES UPDATE

The state vector for the jth ensemble member at the kth assimilation step is:
$x_j^k = \begin{bmatrix} y_1 & y_2 & m_1 & m_2 & p & s & BHP & q_w & q_o & \alpha_1^w & \alpha_2^w \end{bmatrix}_j^T$,where y_1 and y_2 represents the two Random Gaussian Fields, m_1 , m_2 are the truncation parameters of the random fields , p is the pressure, s is the saturation, BHP is the pressure measured at the injector, q_w, q_o are the water and oil rates measured at the producers and α_1^w, α_2^w represents the simulated facies measurements at the wells locations. The facies measurements at the wells location are written in probability terms. If facies type 1 occur then $\alpha_1^w = 1, \alpha_2^w = 0$, if facies type 2 occur then $\alpha_1^w = 0, \alpha_2^w = 1$and if facies type 3 occur we have $\alpha_1^w = 0, \alpha_2^w = 0$.We generate an ensemble of 120 replicates where the values for pressure and saturation are kept constant for each member of the ensemble. The uncertainty in the initial ensemble is given by the choice of the two Gaussian Random Fields and the choice of the truncation parameters. The Random Gaussian Fields y_1 and y_2 are generated with sequential Gaussian simulation method specifying the Geostatistical properties (isotropy or anisotropy, principal directions and the range correlation) and with constraint given by the type of facies find in the grids where the wells are situated. If in a grid with a well located we have observation about the existence of facies type 1 then the value in this grid for y_1 is 0 and of course if we have observation about the existence of facies type 2 then we generate y_2 with value 0 in this grid.

In the forecast step at time k+1 the state vector is

$$x_j^{k+1,f} = \begin{bmatrix} y_1^{k+1,f} & y_2^{k+1,f} & m_1^{k+1,f} & m_2^{k+1,f} & p^{k+1,f} & s^{k+1,f} & BHP^{k+1,f} & q_w^{k+1,f} & q_o^{k+1,f} & (\alpha_1^w)^{k+1,f} & (\alpha_2^w)^{k+1,f} \end{bmatrix}^T$$

where

$$\begin{bmatrix} y_1 \\ y_2 \\ m_1 \\ m_2 \end{bmatrix}_j^{k+1,f} = \begin{bmatrix} y_1 \\ y_2 \\ m_1 \\ m_2 \end{bmatrix}_j^{k+1,f} \tag{1}$$

$$\begin{bmatrix} p \\ s \end{bmatrix}^{k+1,f} = M(\begin{bmatrix} p \\ s \end{bmatrix}^{k,a}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}^{k,a}) \tag{2}$$

$$\begin{bmatrix} BHP \\ q_w \\ q_o \end{bmatrix}^{k+1,f} = g_{pred}(\begin{bmatrix} p \\ s \end{bmatrix}^{k+1,f}, \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}^{k+1,f}) \quad (3)$$

$$\begin{bmatrix} \alpha_1^w \\ \alpha_2^w \end{bmatrix}_j^{k+1,f} = \begin{bmatrix} \alpha_1^w \\ \alpha_2^w \end{bmatrix}_j^{k,a} \quad (4)$$

Where M is the forward model, g_{pred} is the prediction operator for the production data. Hence, there are no changes from time k to k+1 for the values of the random Gaussian fields. But, their values have an impact in the changes of the pressures and saturation at time k+1$ a one can see in equation (2). The random Gaussian fields are changed in the update step of the EnKF when the production data is assimilated according to

$$x_j^{k+1,a} = x_j^{k+1,f} + C^f_{x^{k+1,f}} H^T_{k+1}(H_{k+1}C^f_{x^{k+1,f}}H^T_{k+1} + C^{prod}_{obs,k+1})^{-1}(d^{prod}_{obs,k+1} - H_{k+1}x_j^{k+1,f}) \quad (5)$$

where H represents observation operator for the production data, C^{prod}_{obs} represents the error covariance matrix for the production data C_{x^f} represents the forecasted error covariance of the ensemble, and d^{prod}_{obs} are the observed production data. In order to avoid the improper weighting of the pressure and saturation fields we separate part of the state vector as in:

$$\tilde{x}_j^{k+1} = \begin{bmatrix} y_1 & y_2 & m_1 & m_2 & \alpha_1^w & \alpha_2^w \end{bmatrix}_j^{T,k+1} \quad (6)$$

and assimilate de facies data.

If, after the assimilation of the facies data, some of the ensemble members violate the position of the facies type 3 in the grids where we have observations about their existence, we perform an iterative enforcement on facies observation (because the random fields y_1 and y_2 are generated with value 0 in grids where facies type 1 and type 2 respectively occur, during the HM process the position of the facies type 1 and type 2 are not violated). The iterative process will stop when the value of α_3 in the grids where we have observation about existence of the facies type 3 is greater then α_1 and α_2 in those grids. In the iterative process explained above the uncertainty for the facies observations is represented by the error covariance matrix of 0.0001I. The target is to find a field of facies distribution which is an estimator for the "truth" field (or reference field). In the previous studies using truncated Gaussian method for facies simulations, an estimator of the truth field could not be directly presented because the conditional mean of the random fields could not generate a plausible geological model. We calculate the mean of the probabilities, not the mean of the Gaussian Fields and this conditional mean of the probabilities fields of each facies type, generate a field(named estimated probability field) that is an estimator of the binary field of the probability of the occurrence of that facies type in the reservoir domain. Using these three estimated probability fields we define the estimated field with the maximization criterion, meaning that the assignment of a facies type in a grid is made for that facies type with the greater value of probability in that grid.

4.4 SYNTHETIC EXAMPLE

The simulation model is a 5-spot water flooding 2D-reservoir, black oil model with $50 \times 50 \times 1$active grid blocks. The dimension of each grid block was set at $30 \times 30 \times 20$ ft and there are one injector situated at the center of the reservoir domain and 4 producers situated in the corners. The values of the permeability (k) and porosity (ϕ), corresponding to each facies type, are: for facies type 1 k=174 md, $\phi = 0.18$ for facies type 2 k=372 md, $\phi = 0.25$ and for facies type 3 k=80 md, $\phi = 0.14$.In the next figures we present the reference field(the “truth”) ,

the initial fields and the estimated fields. The blue color represent facies type 1, the green color the facies type 2 and the red color the facies type 3. In the Figure 2 the light blue dots represents the wells positions. For the generation of the Gaussian fields we used isotropic geostatistics characteristics with length correlation of 17 grid blocks and the truncation parameters are generated with mean $\sqrt{2}$ and standard deviation 0.2. In the Figure 3 we present the estimated probability fields for each facies type and the estimated field and in the Figure 4 we present the initial mean of the probabilities fields and the initial mean field.

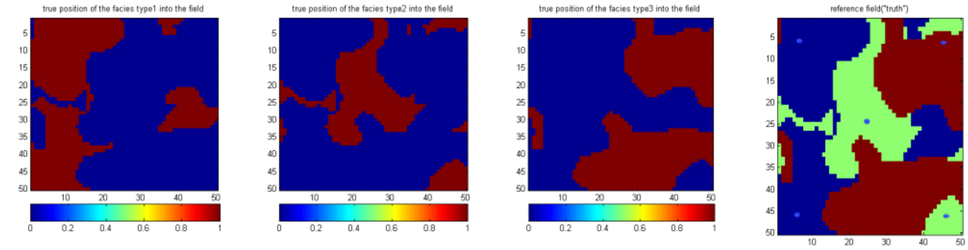

Figure 2: The binary fields defined by each facies type and the reference field

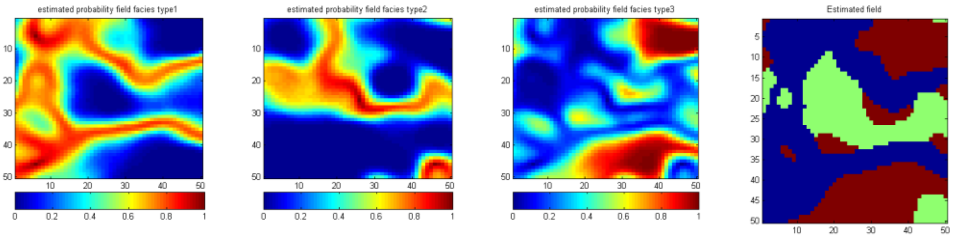

Figure 3: The estimated probability fields of each facies type and the estimated field

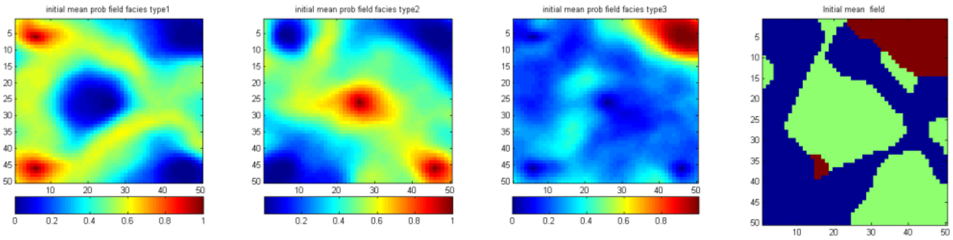

Figure 4: The initial mean probability fields of each facies type and the initial mean field

The water rate production in the initial ensemble and in the updated ensemble is presented in the next two Figures.

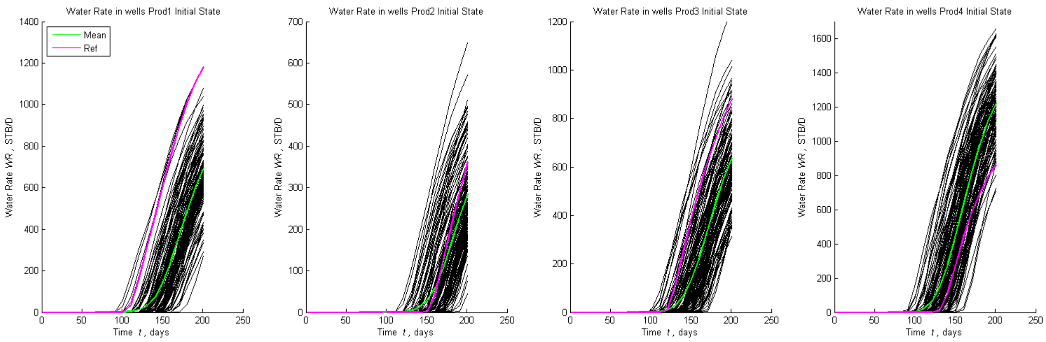

Figure 5: WR in initial state

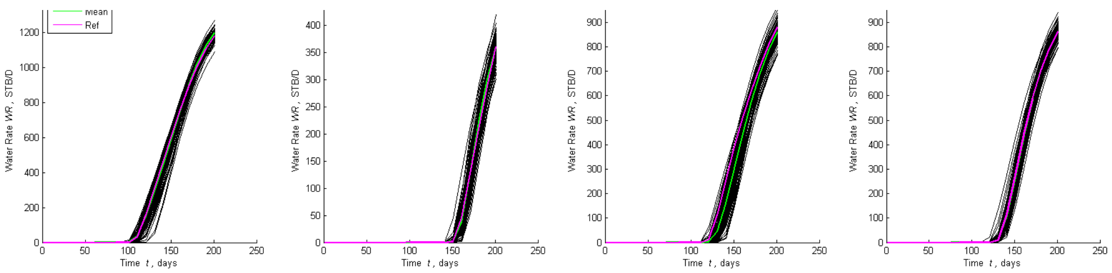

Figure 6: WR in updated state

In the Figure 5 we can observe the initial variability and in the Figure 6 the reduction in trhe variability due the assimilation of the production data and the facies data.
In the next Figure we present the prediction for the next 100 days for water rate production (from 201 days to 301 days).

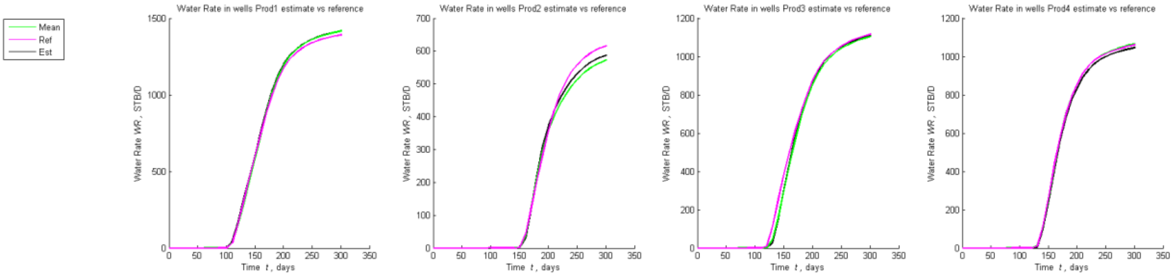

Figure 7: WR prediction for the next 100 days for estimated field and mean of the ensemble

REFERENCES

[1] Agbalaka C. C. , Oliver D. S. , Application of the EnKF and localization to automatic history matching of facies distribution and production data, Mathematical Geosciences, 40, 353374, 2008.
[2] Evensen, G., Data Assimilation: The Ensemble Kalman Filter, Springer, 2006
[3] Galli A, Beucher H, Le Loc'h G, Doligez B, Heresim Group (1994),The pros and cons of the truncated Gaussian method. In: Geostatistical simulations. Kluwer Academic, Dordrecht, pp 217–233
[4] Jazwinski AH (1970), Stochastic processes and filtering theory. Academic, San Diego, California
[5] Liu N. and Oliver D. S., Ensemble Kalman Filter for automatic history matching of geologic facies, Journal of Petroleum Science and Engineering, 47(3-4), 147161, 2005b.
[6] Oliver D, Reynolds A, Liu N (2008), Inverse theory for petroleum reservoir characterization and history matching. Cambridge University Press, Cambridge
[7] SGeMS The Stanford Geostatistical Modeling Software. http://sgems.sourceforge. net/.