On incorporating time-lapse seismic survey data into automatic history matching of reservoir simulations

Laurence R. Bentley

ABSTRACT

Porosity, permeability and other parameters must be specified at every node within a petroleum reservoir simulator. The parameters are developed from sparse and noisy data so that they are not known exactly. After preliminary assignment of porosity and permeability data, simulators are run in the forward mode, but they can seldom match the observed production history adequately. Consequently, a "history matching" step is required in which porosity, permeability and other parameters are adjusted until the model predictions match the production history to a sufficient degree. In "automatic history matching", an objective function is formed by squaring and weighting the difference between observations and the computed predictions. In addition, terms are often added to the objective function to penalize departures from original parameter estimates. This objective function is then minimized with respect to the parameters using solution techniques such as the Marquardt method.

Time-lapse seismic surveys produce images at different times in a reservoir's history. The seismic response of a reservoir may change due to changes in pressure, fluid saturation and temperature. Given appropriate rock physics models, the output of the reservoir simulator can be used to predict the change in seismic response. The difference between the predicted change in seismic response and the observed change in seismic response forms another set of residuals. These residuals are squared, weighted and added to the objective function.

INTRODUCTION

Reservoir simulators are used to approximately solve the mathematical equations that describe the physics of flow of fluids within petroleum reservoirs. The simulators calculate changes in pressure and saturation within reservoirs due to production and the injection of fluids during secondary recovery operations. Reservoir simulators are used to forecast production, assess risk and test conceptual models.

The mathematical models used in reservoir simulators require the specification of reservoir parameters at all locations within the reservoir. In particular, the intrinsic permeability and the porosity need to be specified. The values of these attributes are extremely heterogeneous in geologic formations, and they exert strong control on the flow and storage characteristics of the reservoir. In practice, these parameters are known approximately, because hard data from wells are sparse, correlation to densely spaced seismic data is typically weak and differences exist between measurement and modeling scales.

In general, forward simulations using these values will not adequately match the observed petroleum reservoir production history on the first attempt, because of the uncertainty in the initial estimates of the parameters. Consequently, the parameters

are iteratively adjusted until an adequate match of the production history is obtained. This process is known as “history matching”

Automatic history matching (AHM) is a process in which history matching is performed using optimization techniques. In this approach, an objective function is formed which is a function of the mismatch between the calculated values of pressure and production history and the observed pressure and production history. The objective function is then minimized with respect to the parameters to obtain an updated set of parameters for use in the simulator.

DEVELOPMENT

Notation

H vector of true production history (pressure, water cut, gas oil ratio, etc.)

H^* vector of measured production history

H^C vector of computed production history

H’ vector of computed production history at the optimization point

N_H number of production history vector entries

e_{H^*} vector of errors associated with H^*

$\sigma^2_{H^*}C_{H^*}$ covariance of e_{H^*}

C_{H^*} covariance of e_{H^*} scaled by $1/\sigma^2_{H^*}$

$\sigma^2_{H^*}$ common variance of e_{H^*}

W_H production history weight

r_H production history residual vector

X_H Production history sensitivity matrix

A vector of true seismic attributes

A^* vector of measured seismic attributes

A^C vector of computed seismic attributes

A’ vector of computed seismic attributes at the optimization point

N_A number of seismic attribute vector entries

e_{A^*} vector of errors associated with A^*

Contents

$\sigma^2_{A*}C_{A*}$ covariance of e_{A*}

C_{A*} covariance of e_{A*} scaled by $1/\sigma^2_{A*}$

σ^2_{A*} common variance of e_{A*}

W_A seismic attribute weight

r_A seismic attribute residual vector

X_A seismic attribute sensitivity matrix

P vector of true parameters values

P^* vector of measured parameter values (prior information)

P^C vector of current parameter values derived from the optimization routine

P' vector of parameter values at the optimization point

N_P number of prior information vector entries

e_{P*} vector of errors associated with P^*

$\sigma^2_{P*}C_{P*}$ covariance of e_{P*}

C_{P*} covariance of e_{P*} scaled by $1/\sigma^2_{P*}$

σ^2_{P*} common variance of e_{P*}

W_P prior information weight

r_P prior information residual vector

X_P prior information sensitivity vector

n_p number of decision variables, $n_p \geq N_P$

p reservoir pressure

S_O oil saturation

S_W water saturation

S_G gas saturation

J_K objective function

J_K' objective function at solution

Contents

X total sensitivity (Jacobian) matrix

X' total sensitivity matrix at solution

r augmented residual vector

r' augmented residual vector at solution

G gradient vector

H_e approximate Hessian matrix

λ_M Marquardt parameter

λ_R step size relaxation parameter

s^2 common variance

Formulation

Consider three objective function terms associated with the production history, seismic attributes and prior knowledge of the parameter values. Typically the decision parameters, P^C, will consist of the log-permeability and the porosity. The production values, H^C, are calculated using the simulator and the model parameters P^C. A set of rock physics models, such as the Gassman equation, are used with the pressure and saturation values computed by the flow simulator to calculate the seismic attributes, A^C.

Production History

Assume that production history measurements are related to the real production history by,

$$H^* = H + e_{H^*} \tag{1}$$

Further, the form of e_{H^*} is assumed to be normal with zero mean and covariance $\sigma_{H^*}C_{H^*}$. The matrix C_{H^*} has the form of the covariance of e_{H^*}, but it is scaled by the inverse of the common variance $\sigma_{H^*}^2$. The most common models for C_{H^*} is that the production measurement errors are uncorrelated in space but may be correlated by a one lag autocorrelation function in time. This means that C_{H^*} is a diagonal or block diagonal matrix. The diagonal terms of C_{H^*} would vary if, for example, one set of pressure data was considered less reliable than another set of pressure data.

Contents

The production history objective function term is,

$$\frac{1}{\sigma_{H^*}^2} J_H = \frac{1}{\sigma^2_{H^*}} \left(H^C - H^*\right)^t C_{H^*}^{-1} \left(H^C - H^*\right). \tag{2}$$

The weighting terms $\frac{1}{\sigma_{H^*}^2}$ and $C_{H^*}^{-1}$:

1. Scale the equation for variations in units (e.g. psi versus Kpa),
2. Weight more reliable data more than less reliable data, and
3. Filter the estimates based on the correlation between errors.

Seismic Attributes

In a manner analogous to the production history match, the seismic attribute match leads to,

$$\frac{1}{\sigma_{A^*}^2} J_A = \frac{1}{\sigma^2_{A^*}} \left(A^C - A^*\right)^t C_{A^*}^{-1} \left(A^C - A^*\right). \tag{3}$$

In this case, it is possible the e_{A^*} may be correlated and the optimal C_{A^*} would be a non-diagonal matrix.

Prior Estimates of Parameters

Estimates of porosity have been developed using well log information or well log information in conjunction with seismic attributes. Permeability estimates may be derived from upscaling core permeability, DST analysis, and correlation of porosity with permeability. The information contained in these prior estimates can be represented by the objective function term,

$$\frac{1}{\sigma_{P^*}^2} J_P = \frac{1}{\sigma^2_{P^*}} \left(P^C - P^*\right)^t C_{P^*}^{-1} \left(P^C - P^*\right). \tag{4}$$

In this case, the form of the covariance of e_{Y^*} may be determined from the estimation errors generated in the interpolation step (e.g. kriging or cokriging). This term is sometimes referred to as a regularization term and improves the stability of inverse procedures. In general, e_{P^*} will be correlated, and using the full correlation structure while developing the weighting matrices has been shown to improve estimates (Bentley, 1997).

Optimization

Assuming that the statistical assumptions (normality, etc.) are correct and that near the solution the objective function is nearly linear in the parameters, then the following objective function is optimal in a least squares sense,

Contents

$$\underset{W.R.T.P^C}{MIN}\ J_1 = \frac{1}{\sigma_{H^*}^{\ 2}} J_H + \frac{1}{\sigma_{A^*}^{\ 2}} J_A + \frac{1}{\sigma_{P^*}^{\ 2}} J_P \tag{5}$$

The correct weighting is seen to be the inverse of the covariance of errors in the measurements and prior estimates. Using this form requires a priori knowledge of the covariance of the errors, which are not known. In addition, we have neglected model error in the sense that the flow simulator may not be able to match the true system response due to inadequacies in the model equations or discretization. The assumption of normally distributed errors leads to problems if there are extremely large residuals.

To date, minimization of the equivalent of the full objective function J_1 has not been reported. Researchers have optimized the equivalent of $J_2 = \frac{1}{\sigma_{H^*}^2} J_H + \frac{1}{\sigma_{P^*}^2} J_P$ (e.g. Gavalas et al. 1976, Chu et al. 1995, Chavent et al. 1975), or the equivalent of $J_3 = \frac{1}{\sigma_{H^*}^2} J_H + \frac{1}{\sigma_{A^*}^2} J_A$ (Huang, et al. 1997, Huang et al. 1998). In addition, Landa and Horne (1997) use time lapse seismic data to estimate the change in saturation within the reservoir and then solved a form of J_2.

An approach that has been used in water resources is to assume that we know the form of the covariance matrix, but not the magnitude of the common variances (Carrera and Neuman, 1986). In this way, information on the relationship between the errors within each residual type is included in the objective function, but the weights associated with the common variances must be selected by some other criteria. These weights determine the relative importance of each type of information. The revised production history objective function term becomes,

$$W_H J_H = W_H \left(H^C - H^*\right)^t C_H^{-1} \left(H^C - H^*\right) \tag{6}$$

Similar expressions are used for A with weighting W_A and P with weighting W_P. In the literature the weights W_K (where K=H, A or P) have been chosen by a variety of methods. The new objective function becomes,

$$J_4 = W_H J_H + W_A J_A + W_P J_P \tag{7}$$

One method that has been used with success is to vary the weights based on balancing the weighted variances as approximated by the current estimate of the parameters. On the assumption that at the solution P', the values of P', H' and A' are close to their true values, the weighted variances of H^*, A^*, and P^* can be approximated by,

Contents

$$Var\left(H^*\right)^i = \frac{W_H^i}{N_H}\left(H'^i - H^*\right)^t C_{H^*}^{-1}\left(H'^i - H^*\right) \tag{8}$$

$$Var\left(A^*\right)^i = \frac{W_A^i}{N_A}\left(A'^i - A^*\right)^t C_{A^*}^{-1}\left(A'^i - A^*\right) \tag{9}$$

$$Var\left(P^*\right)^i = \frac{W_P^i}{N_P}\left(P'^i - P^*\right)^t C_{P^*}^{-1}\left(P'^i - P^*\right) \tag{10}$$

The index i indicates values derived from the solution at iteration i, and N_H, N_A and N_P are the number of pressure measurements, seismic attribute values and prior estimates of parameters, respectively. In order to balance the contributions of the different data types, new weights are chosen to balance the weighted variances,

$$W_K^i = \frac{W_K^{i-1}}{Var\left(K^*\right)^{i-1}} \tag{11}$$

The weights are adjusted at the end of each run until the variances are approximately equal. Two or three iterations are generally required to reach a balanced set of weighted variances (Weiss and Smith, 1998).

Computational Issues

Gradient and Sensitivity Matrix

Several approaches have been used to solve the minimization problem (e.g. equation (5)). In the following, we focus on the Marquardt method (Marquardt, 1963). Each Marquardt iteration, the gradient and the sensitivity coefficients must be computed. Define the residual vectors,

$$r_H = H^C - H^* \tag{12}$$

$$r_A = A^C - A^* \tag{13}$$

$$r_P = P^C - P^* \tag{14}$$

The residual vectors at the solution are denoted r_K', where K= H, A or P. The global residual vector is

$$r^t = \begin{bmatrix} r_H^t & r_A^t & r_P^t \end{bmatrix} \tag{15}$$

The production history residual vector contains N_H entries. These will be arranged as time series data from several to tens of production and injection wells. It is possible for seismic attribute data to exist at each grid block within the simulator, and consequently, the dimension of the vector can theoretically be in the tens of

Contents

thousands. With three permeability values (κ_X, κ_Y, κ_Z) and a porosity associated with each grid block, the number of parameters can be in the tens of thousands.

A sensitivity coefficient is the partial derivative of an individual residual with respect to a decision variable. Each type of residual has a sensitivity matrix associated with it. An element of the production history sensitivity matrix is,

$$X_{Hij} = \frac{\partial H_i^C}{\partial P_j} \tag{16}$$

where each row i=1 to N_H is associated with a production history residual and each column j=1 to n_p is associated with a decision parameter. Similarly, the seismic attribute residual sensitivity coefficients are,

$$X_{Aij} = \frac{\partial A_i^C}{\partial P_j} \tag{17}$$

where i=1 to N_A and j=1 to n_p. Assume that variations in seismic attribute can be calculated from the pressure and saturation,

$$A = F(p, S_O, S_W, S_G) \tag{18}$$

Then, the seismic attribute sensitivities can be calculated using the chain rule,

$$\frac{\partial A}{\partial P} = \frac{\partial F}{\partial p}\frac{\partial p}{\partial P} + \frac{\partial F}{\partial S_O}\frac{\partial S_O}{\partial P} + \frac{\partial F}{\partial S_W}\frac{\partial S_W}{\partial P} + \frac{\partial F}{\partial S_G}\frac{\partial S_G}{\partial P} \tag{19}$$

Consequently, the partial derivative of the pressure and saturations with respect to each parameter is required at each location of a seismic attribute value. These same derivatives are required for the evaluation of equation (16), but only at the locations of the wells used in the production history match. In some cases, temperature would need to be modeled as a state variable, because it can also have a significant effect on the seismic response.

The sensitivity coefficients for the prior information are,

$$X_{Pij} = \frac{\partial P_i^C}{\partial P_j} \tag{20}$$

with rows i=1 to N_p and columns j=1 to n_p. The prior information matrix consists of entries of ones and zeroes, and if every parameter has prior information then the matrix is the identity matrix.

Contents

The total sensitivity matrix is constructed,

$$X^t = \begin{bmatrix} X_H^t & X_A^t & X_P^t \end{bmatrix} \tag{21}$$

and it has $N=N_H+N_A+N_p$ rows and n_p columns. It is possible for the number of parameters to be in the tens of thousands. The huge dimensionality of the parameter space and the limited amount of observational data lead to a problem that is inherently underdetermined. Consequently, uniqueness and identifiability are issues in the parameter optimization problem. In practice, the dimension of the parameter space is reduced by lumping many grid blocks together in zones of equal value or to use interpolation methods such as the pilot point method to populate the reservoir with a smaller number decision variables located ate the interpolation points (pilot points). We are currently investigating the use of cluster algorithms for dynamic zoning.

The calculation of the sensitivity matrix is a major computational burden. Three approaches have been used. In the perturbation method, each parameter is perturbed and a forward simulation is run to compute the new value of the residuals. The sensitivity matrix is then constructed by the finite difference method. This method is only computationally feasible with a small number of parameters. In the adjoint equation method, the equivalent of a separate simulation run is needed for each of the observation locations. This method was useful when the number of observation locations was limited to a small set of production and injection wells. However, the inclusion of spatially dense seismic attribute data requires the sensitivity calculations to be carried out at thousands of locations, and the adjoint equation method is not as attractive for the cases in which time-lapse seismic data is included in the objective function. In the gradient simulation method (Anterion et al. 1989), the equivalent of an extra flow simulation plus for each decision variable one back substitution is required for each Marquardt iteration. At present, this appears to be the best approach for calculating sensitivity coefficients for the cases with time-lapse seismic terms in the objective function.

The global weighting matrix for objective function J_4 is,

$$C^{-1} = \begin{pmatrix} W_H C_H^{-1} & 0 & 0 \\ 0 & W_A C_A^{-1} & 0 \\ 0 & 0 & W_P C_P^{-1} \end{pmatrix} \tag{22}$$

The other objective functions have analogous weighting matrices.

The gradient of the objective function is,

$$G = \frac{\partial J_k}{\partial P} = 2 X^t C^{-1} r \tag{23}$$

where k indicates the form of the objective function (i.e. k=1, 2, 3, or 4). The vector G has a length of n_p.

Contents

If the residuals are small and/or the equation is quasi-linear, then the Hessian matrix can be approximated as,

$$H_e = \frac{\partial^2 J_k}{\partial P^2} \approx \left(X^t C^{-1} X\right) \tag{24}$$

The matrix H_e has dimensions n_p by n_p.

The Marquardt algorithm update is,

$$P^{k+1} = P^k + \lambda_R^k \left(\left(X^t C^{-1} X\right) + \lambda_M^k I\right)^{-1} G \tag{25}$$

where k is the iteration index, λ_M is the Marquardt parameter and λ_R is a relaxation parameter to control the step size.

Transformation and Scaling

Improving the conditioning of the objective function will improve the solution and reduce the computational effort required to reach a solution. Parameters and residuals are often scaled or transformed in order to provide a better conditioned objective function. In addition, transformed variables often have better statistical properties than the primary variables. For example, permeability is generally transformed to log-permeability for geologic and statistical reasons. Appropriate weights will depend on the form of the scaled or transformed variables, so transformation choices are related to weighting choices.

Often the decision variables are normalized to one in order to improve the conditioning of the objective function. If the current value of the parameters is P^C, then the scaling takes the form,

$$H_{eij}^S = H_{eij} P_i^C P_j^C \tag{26}$$

$$G_i^S = G_i P_i^C \tag{27}$$

$$P_i^S = \frac{P_i}{P_i^C} \tag{28}$$

Error Estimates

At the solution, the common variance of the weighted errors can be estimated by linear error analysis,

$$s^2 = \frac{J'_4}{(N_H + N_A + N_P) - n_P} = \frac{J'_4}{N_D} \tag{29}$$

where N_D is the difference between the number of residuals and the number of decision variables.

Contents

An estimate for the estimation variance of the parameters is,

$$Var(P') = s^2 \left(X'^t \, C^{-1} X' \right)^{-1} \tag{30}$$

where the prime indicates that the values are evaluated at the optimal solution. A linear approximation of the confidence regions is given by,

$$(P - P')^t \, X'^t \, C^{-1} X' (P - P') \le n_P s^2 F_\alpha (n_P, N_D) \tag{31}$$

where F_∞ is an F distribution with n_P and N_D degrees of freedom.

CONCLUSION

The goal of automatic history matching (AHM) is to improve the parameter selection for petroleum reservoir simulators in order to improve their reservoir forecasting capability. An approach for including time-lapse seismic data into AHM has been presented and some practical issues have been discussed.

AHM is still not routinely used, even after more than twenty years of development. The difficulties are related to the extreme nonlinearity of the objective function, the high dimensionality of the decision variable set, problems associated with uniqueness and identifiability of parameters, and difficulties associated with computing the sensitivity matrices. AHM based solely on production data observed at a limited number of well locations suffers because of the sparsity of the spatial sample. Seismic data has the inherent advantage of being a spatially dense data set, and preliminary indications are that including time-lapse seismic data will lead to improved AHM.

Although the routine application of AHM remains an elusive goal, studying time-lapse seismic survey results in the context of AHM is valuable. The mathematical formulation of AHM gives us insight into the history matching process and helps analysts identify the underlying structure of the history match problem. Study of the sensitivity of seismic attributes to changes in reservoir parameters, objective function gradients and estimation variance will lead to insights into the effectiveness of integrating time-lapse seismic data into reservoir simulation and reservoir forecasting. In turn, these insights will guide time-lapse seismic acquisition and processing choices.

REFERENCES

Anterion, F., Eymard, R., Karcher, B., 1989, Use of parameter gradients for reservoir history matching, SPE 18433.

Bentley, L.R., 1997, Influence of the regularization weighting matrix on parameter estimates, Adv. Wat. Resour., 20(4), 231-247.

Carrera, J. and Neuman, S.P., 1986, Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood estimation incorporating prior information, Wat. Resour. Res., 22(2), 199-210.

Chavent, G., Dupuy, M. and Lemonnier, P., 1975, History matching by use of optimal theory, Soc. Pet. Eng. J., 15(1), 74-86.

Chu, L., Reynolds, A.C. and Oliver, C.S., 1995, Reservoir description from static and well-test data using efficient gradient methods, SPE 29999, 467-482.

Gavalas, G.R., Shah, P.C. and Seinfeld, J.H., 1976, Reservoir history matching by Bayesian estimation, Soc. Pet. Eng., 16(6), 337-350.

Huang, X., Meister, L. and Workman, R., 1998, Improvement and sensitivity of reservoir characterization derived from time-lapse seismic data, SPE 49146.

Huang, X., Meister, L. and Workman, R., 1997, Reservoir characterization by integration of time-lapse seismic production data, SPE38695.

Landa, J.L. and Horne, R.N., 1997, A procedure to integrate well test data, reservoir performance history and 4-D seismic information into a reservoir description, SPE 38653.

Marquardt, D.T., 1963, J. Soc. Ind. Appl. Math., 11, 431-441.

Weiss, R. and Smith, L., 1998, Parameter space methods in joint parameter estimation for groundwater flow models, Wat. Resour. Res., 34(4), 647-661.