LBL-37611
UC-1240

Automatic History Matching of Geothermal Field Performance

Stefan Finsterle and Karsten Pruess

Earth Sciences Division
Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720

Paper presented at the
17th New Zealand Geothermal Workshop
November 8-10, 1995
Auckland, New Zealand

August 1995

This work was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Geothermal Division, of the U.S. Department of Energy under contract No. DE-AC03-76SF00098.

Automatic History Matching of Geothermal Field Performance

S. Finsterle and K. Pruess

Lawrence Berkeley National Laboratory
Earth Sciences Division
University of California, Berkeley, USA

SUMMARY - We have developed inverse modeling capabilities for the multiphase multicomponent numerical simulator TOUGH2 to facilitate automatic history matching and parameter estimation based on data obtained during exploitation of geothermal fields. The ITOUGH2 code allows one to estimate TOUGH2 input parameters based on any type of observation for which a corresponding TOUGH2 output can be calculated. Furthermore, a detailed residual and error analysis is performed, and the uncertainty of model predictions can be evaluated. This paper focuses on the solution of the inverse problem, i.e. the determination of model-related parameters by automatically calibrating a conceptual model of the geothermal system against data obtained during field operation. We first describe the modeling approach used to simulate fluid and heat flow in fractured-porous media. The inverse problem is then formulated, followed by a brief discussion of the optimization algorithm. A sample problem is given to demonstrate the application of the method to geothermal reservoir data.

1. INTRODUCTION

Predicting the performance of a geothermal field, as well as the design and optimization of field operations, requires reliable numerical modeling techniques. This includes a detailed description of the complex physical processes controlling multiphase fluid flow and heat transport in fractured-porous media. This first step will be referred to as model conceptualization. Furthermore, a site-specific model has to be developed, i.e. the geometry of the reservoir, its hydrogeological properties, as well as initial and boundary conditions have to be determined. After parameter values are assigned in a subsequent calibration process, predictive reservoir simulation can be initiated. It is important to realize that the parameter estimates will always be related to the structure of the model, both conceptually and numerically.

This paper focuses on the solution of the inverse problem, i.e. the determination of model-related reservoir parameters by automatically calibrating a conceptual model of the geothermal field against data obtained during exploitation. We first describe the modeling approach used to simulate fluid and heat flow in fractured-porous media. The inverse problem is then formulated in the framework of maximum likelihood theory, followed by a brief discussion of the optimization algorithm. A sample problem is given to demonstrate the application of the method to (synthetic) field performance data.

2. MODELING APPROACH

2.1 The Forward Problem

We solve the forward problem, i.e. the simulation of fluid and heat flow in a geothermal field, with the TOUGH2 code (Pruess, 1991). Solving the forward problem in an efficient and stable manner is probably the most important step for automatic parameter estimation. TOUGH2 is used here to simulate nonisothermal flow of a single component (water) in two co-existent phases (liquid, vapor).
The mass and energy balance equations for an arbitrary subdomain V_{n} bounded by the surface Γ_{n} can be written in the following form:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{V_{\mathrm{n}}} M \;\mathrm{d}V = \int_{\Gamma_{\mathrm{n}}} \mathbf{F}\bullet\mathbf{n}\;\mathrm{d}\Gamma + \int_{V_{\mathrm{n}}} q\;\mathrm{d}V \qquad (1)$$

The accumulation term M represents mass (m) or internal energy (h) per unit reservoir volume:

$$M_m = \phi\,(S_l\,\rho_l + S_v\,\rho_v) \qquad (2)$$

$$M_h = \phi\,(S_l\,\rho_l\,u_l + S_v\,\rho_v\,u_v) + (1 - \phi)\,\rho_R\,C_R\,T \qquad (3)$$

Here ϕ is porosity, S is saturation, ρ is density, u is internal energy, C is specific heat, and T is temperature. The subscripts l, v, and R denote liquid, vapor, and rock, respectively. The mass flux is a sum over the fluxes in the liquid and vapor phase:

$$\mathbf{F}^m = \sum_{\beta=\mathrm{l,v}} -\mathbf{k}\,\frac{k_{r\beta}}{\mu_\beta}\rho_\beta\,(\nabla P_\beta - \rho_\beta\,\mathbf{g}) \qquad (4)$$

where $\mathbf{k}$ denotes the permeability tensor, k_r is relative permeability, μ is viscosity, P_β is the pressure in phase β, and $\mathbf{g}$ is acceleration of gravity. In Eq. 1, $\mathbf{n}$ is the outward unit normal vector. The total heat flux containing conductive and convective components can be written

$$\mathbf{F}^h = -K\nabla T + \sum_{\beta=1,v} (h_\beta \, \mathbf{F}_\beta) \tag{5}$$

with K the thermal conductivity of the rock-fluid mixture and h_β the specific enthalpy, which is a nonlinear function of temperature. Thermophysical properties of liquid water and vapor are calculated using steam table equations given by the International Formulation Committee (IFC, 1967). The continuum equations (1) are discretized in space based on an integral finite difference formulation (Narasimhan and Witherspoon, 1976), and a multiple interacting continua (MINC, Pruess and Narasimhan, 1982, 1985) approach is used to represent fractured-porous media. Time is discretized fully implicitly as a first-order finite difference. Discretization results in a set of nonlinear coupled algebraic equations which are solved simultaneously by means of Newton-Raphson iterations. A conjugate gradient algorithm is used to solve the linear equations arising at each iteration step.

2.2 The Inverse Problem

The determination of reservoir properties from performance data, such as pressures, temperatures, and flow rates, is referred to as the inverse problem. The indirect approach to inverse modeling consists of minimizing the differences between the observed and simulated field responses, which are assembled in the residual vector **r** with elements

$$r_i = y_i^* - y_i(\mathbf{p}) \tag{6}$$

Here y_i^* is an observation (e.g. pressure, temperature, flow rate, etc.) at a given point in space and time, and y_i is the corresponding simulator prediction, which depends on vector **p** of all unknown or uncertain model parameters, including initial and boundary conditions. If the error structure of the residuals is assumed Gaussian and described by a covariance matrix **C**, the objective function to be minimized is simply the sum of the squared residuals weighted by the inverse of the prior covariance matrix (Finsterle and Pruess, 1995):

$$z(\mathbf{p}) = \mathbf{r}^T \mathbf{C}^{-1} \mathbf{r} \tag{7}$$

In maximum likelihood theory, it can be shown that minimizing z is equivalent to maximizing the probability of reproducing the observed system state. Eq. 7 corresponds to the generalized nonlinear least squares estimator.

Due to the strong nonlinearities of multiphase flows, an iterative procedure is required to minimize the objective function. The Levenberg-Marquardt modification of the Gauss-Newton algorithm (Levenberg, 1944; Marquardt, 1963) has been found to be the most robust for our purposes. The basic idea of this method is to move in the parameter space along the steepest descent direction far from the minimum, switching continuously to the Gauss-Newton algorithm as the minimum is approached. This is achieved by decreasing a scalar λ, known as the Levenberg parameter, after a successful iteration, but increasing it if an uphill step is taken. The following system of equations is solved for $\Delta\mathbf{p}$ at an iteration labeled k:

$$(\mathbf{J}_k^T \, \mathbf{C}^{-1} \, \mathbf{J}_k + \lambda_k \mathbf{D}_k) \, \Delta\mathbf{p}_k = -\mathbf{J}_k^T \, \mathbf{C}^{-1} \, \mathbf{r}_k \tag{8}$$

Here, **J** is the sensitivity matrix with elements $J_{ij} = \partial y_i / \partial p_j$. **D** denotes a matrix of order n (n being the number of parameters to be estimated) with elements equivalent to the diagonal elements of matrix $(\mathbf{J}_k^T \, \mathbf{C}^{-1} \, \mathbf{J}_k)$. The improved parameter set is finally calculated:

$$\mathbf{p}_{k+1} = \mathbf{p}_k + \Delta\mathbf{p}_k \tag{9}$$

Under the assumption of normality and linearity, a detailed error analysis of the final residuals and the estimated parameters can be conducted (for details see Finsterle and Pruess (1995)). For example, the covariance matrix of the estimated parameter set is given by:

$$\mathbf{C}_p = \frac{\mathbf{r}^T \, \mathbf{C}^{-1} \, \mathbf{r}}{m - n} (\mathbf{J}^T \, \mathbf{C}^{-1} \, \mathbf{J})^{-1} \tag{10}$$

where m is the total number of observations. As a byproduct of calculating the Jacobian matrix **J**, one can qualitatively examine the contribution of each data point to the solution of the inverse problem as well as the total parameter sensitivity (for details see Finsterle (1995)).

The inverse modeling formulation outlined above is implemented in a computer program named ITOUGH2 (Finsterle, 1993). ITOUGH2 has been applied to a number of laboratory and field data (Finsterle and Pruess, 1994).

3. APPLICATION

The purpose of this section is to illustrate the use of the proposed methodology for the characterization of geothermal fields. ITOUGH2 provides the flexibility to take advantage of any type of data usually collected during field exploitation. For the sake of simplicity and reproducibility, we will analyze a synthetic case. Applications of ITOUGH2 to geothermal field data are described in O'Sullivan and Bullivant (1995) and White (1995).

We consider a two-dimensional five-spot production-injection problem previously studied by Pruess (1991) and Pruess and Wu (1993). Due to the symmetry of the large well field five-spot configuration, only 1/8 of the basic pattern needs to be modeled (Fig. 1).

The problem specifications correspond to conditions typically encountered in deeper zones of hot two-phase flow reservoirs. The medium is assumed to be fractured with embedded impermeable matrix blocks in the shape of cubes with side lengths of 50 m. The permeable volume fraction is 2 % with a porosity of the fracture domain of 50 %. Reservoir thickness is 305 m. Water with an enthalpy of 500 kJ/kg is injected at a rate of 30 kg/s. Production rate is also 30 kg/s.

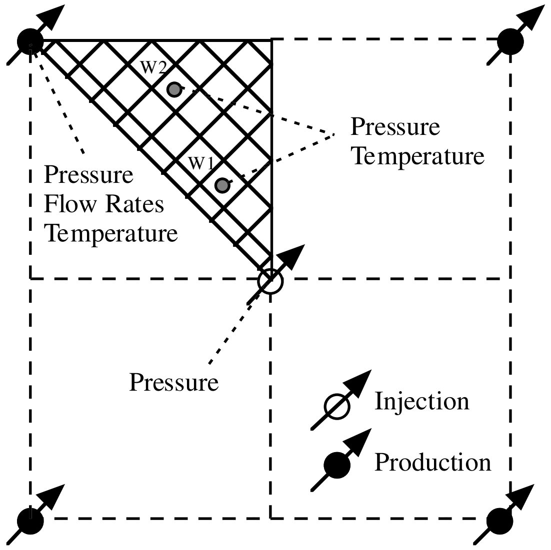

Figure 1 - Five-spot well pattern with grid for modeling 1/8 symmetric domain. Observation points and type of data measured is also indicated.

We assume that temperature and pressure measurements are taken in the injection (Inj) and production well (Pro) as well as in two abandoned wells (W1, W2; see Fig. 1). Furthermore, liquid and vapor flow rates are measured in the production well. Note that temperature and pressure measurements are redundant as long as two-phase conditions prevail. TOUGH2 is run in forward mode to generate data for five years of field performance history, and random noise is added to simulate measurement errors (see Table 1 for standard deviations).

Subsequently, the TOUGH2 model is automatically calibrated against these observations in order to determine certain input parameters considered unknown or uncertain. The parameters include the effective permeability of the fracture system, porosity, heat conductivity, specific heat of the rock grains, fracture spacing *a* (which is a parameter of the MINC preprocessor), and the initial reservoir temperature T_i.

The estimated parameter set is shown in Table 2. The covariance and correlation matrices are summarized in Table 3, and some statistical measures are given in Table 4. The latter need some explanations. The second column of Table 4 contains the standard deviation σ_p of the estimate which is the square root of the corresponding diagonal element in Table 3.

Table 1 - Observations Used for Model Calibration

Data Type	Location	Standard Deviation
Pressure	Inj/Pro/W1/W2	2.00 bar
Temperature	Pro/W1/W2	5.00 ˚C
Liquid flow rate	Pro	1.60 kg/s (~5 %)
Vapor flow rate	Pro	0.08 kg/s (~5 %)

Table 2 - True, Initial, and Estimated Parameter Set

Parameter	True Value	Initial Guess	Best Estimate
log (permeability [m^2])	-14.22	-13.00	-14.22
fracture zone porosity [-]	0.50	0.30	0.56
specific heat [J/kg˚C]	1000.00	800.00	971.00
heat conduct. [W/m˚C]	2.10	2.50	2.25
fracture spacing [m]	50.00	20.00	50.50
temperature [˚C]	300.00	250.00	300.10

Table 3 - Variance-Covariance Matrix (Main Diagonal and Lower Triangle) and Correlation Matrix (Upper Triangle)

	log(k)	ϕ	C_R	K	a	T_i
log(k)	5E-6	0.21	0.17	-0.25	-0.21	-0.18
ϕ	2E-5	2E-3	0.17	-0.23	-0.18	-0.04
C_R	0.01	0.23	845	0.39	0.53	-0.07
K	-5E-4	-0.01	10.15	0.79	0.98	0.29
a	-4E-3	-0.08	14.52	8.27	89.22	0.19
T_i	-4E-5	-2E-4	0.24	0.02	0.20	0.01

Table 4 - Statistical Measures and Parameter Sensitivity

Parameter	Standard Deviation	σ_p^*/σ_p	Parameter Sensitivity
log (permeability [m^2])	0.002	0.88	3623.1
fracture zone porosity [-]	0.05	0.90	19.6
specific heat [J/kg˚C]	29.10	0.03	64.7
heat conduct. [W/m˚C]	0.89	0.18	50.5
fracture spacing [m]	9.40	0.03	253.6
temperature [˚C]	0.10	0.94	1768.2

Note that this standard deviation refers to the joint probability density function, i.e. it takes into account the uncertainty of the parameter itself and the influence from correlated parameters. The conditional standard deviations σ_p^*, on the other hand, reflect the uncertainty of an estimate provided that all the other parameters are exactly known. Therefore, the ratio σ_p^*/σ_p shown in the third column is a measure of how independently a parameter can be estimated. A value close to one indicates an independent estimate, whereas small values can be interpreted as a loss of parameter identifiability due to its correlation to other uncertain parameters. Finally, we show the total parameter sensitivity (column 4) which is the sum of the absolute values of all sensitivity coefficients, weighted by the inverse of individual measurement errors and scaled by a reasonable parameter variation.

First we note that permeability and reservoir temperature are accurately identified. They are the most sensitive parameters and can be determined almost independently. The estimates of fracture spacing, heat conductivity and specific heat exhibit relatively high standard deviations which is easily explained by the large correlation coefficients among these three parameters (see Table 3). Especially the fracture spacing and heat conductivity have a high positive correlation coefficient, i.e. a larger fracture spacing can be almost completely compensated by an increase in heat conductivity. This statement is true for the type and amount of data available, i.e. the correlation between these two parameters may be reduced by taking additional data. Finally, the fracture porosity can be relatively well determined despite its low overall sensitivity. This is simply due to the fact that fracture porosity is only weakly correlated to the other parameters, resulting in an independent estimate.

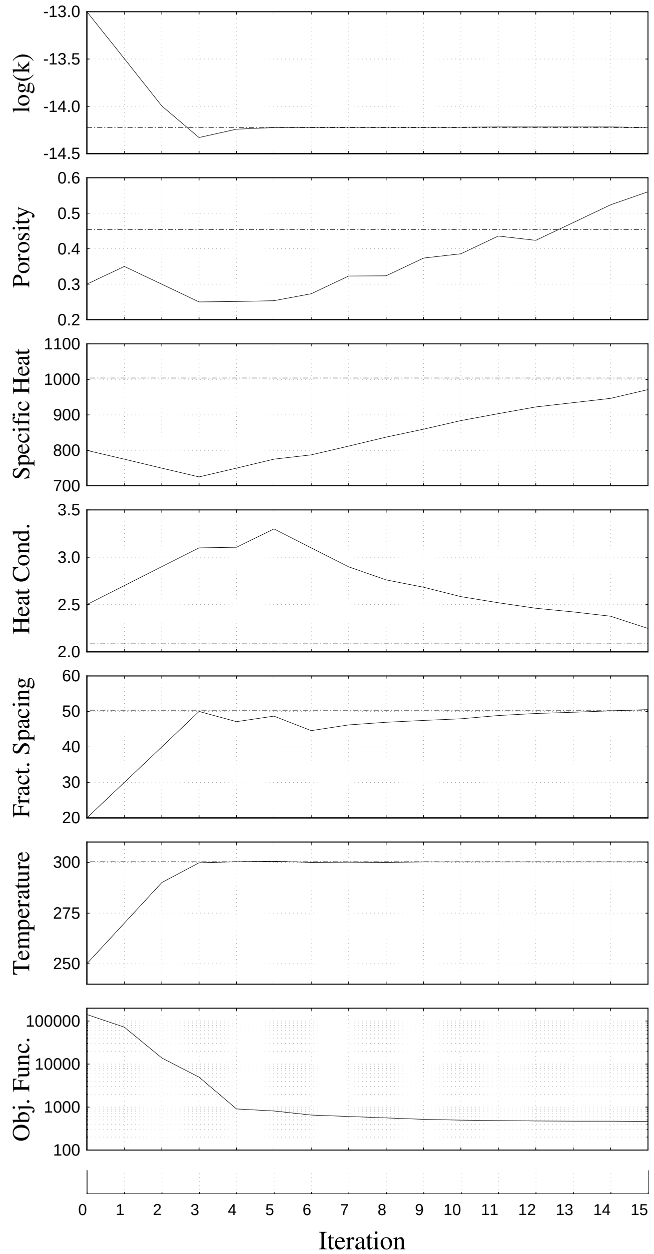

Figure 2 - Optimization path in parameter space (solid lines). The true values are indicated as dash-dotted lines.

Figure 2 depicts the solution path in the parameter space, i.e. the updated parameter set after each Levenberg-Marquardt iteration. The two most sensitive parameters ($\log(k)$ and initial reservoir temperature) quickly converge to the true values (dash-dotted line), reducing the value of the objective function (Eq. 7) by more than two orders of magnitude within four iterations. Fracture spacing is increased from its initial value of 20 m toward the true value of 50 m, forcing the less sensitive parameters to actually go away from their true values according to the correlation rules prevailing at that time. They later recover after the main parameters have stabilized. However, this fine-tuning only leads to minor reductions of the objective function. Fracture zone porosity as the least sensitive parameter does not seem to converge to the true value.

The solid lines are the pressures, temperatures, water and vapor flow rates simulated using the estimated parameter set (Table 2, column 4). For the first 5 years, the deviations between the solid lines and the squares minimize Eq. 7. Beyond 5 years, the solid lines are predictions, i.e. an extrapolation of the system response matched during the calibration period. The model predictions are uncertain due to uncertainties in the estimated parameters. The standard deviation of the calculated system response, i.e. the uncertainty of the predicted temperature in the production well at a certain point in time, is the square root of the corresponding diagonal element of matrix $\mathbf{C}_z$ which is calculated using first-order error propagation analysis:

$$\mathbf{C}_z = \mathbf{J}\,\mathbf{C}_p\,\mathbf{J}^T \qquad (11)$$

Here, matrix $\mathbf{J}$ is the sensitivity matrix for the predicted system response, and $\mathbf{C}_p$ is the covariance matrix of the estimated parameters (Eq. 10). The resulting error bands on the model predictions are shown as dash-dotted lines in Figure 3. They have to be considered optimistic because only the six parameters analyzed in this study are considered uncertain. All the other parameters as well as the model structure are assumed to be exactly known, which of course is only true for a synthetic case. However, it is interesting to note that the true system response (triangles) lies within the estimated error band despite the fact that the parameter set used for the prediction does not exactly correspond to the true one.

The high accuracy of the model prediction can only be achieved by a combined inversion of all available data. It is obvious that the temperature decrease in observation well W1 could not have been predicted by relying only on temperature data during the calibration phase. In fact, the contribution of temperature measurements to the determination of the parameter set is minor. This is mainly due to the fact that a temperature change of 1 ˚C leads to a vapor pressure change of about 1 bar which can be easier detected given the accuracy of pressure measurements. As mentioned earlier, temperature and pressure are not independent in a single-component two-phase flow system.

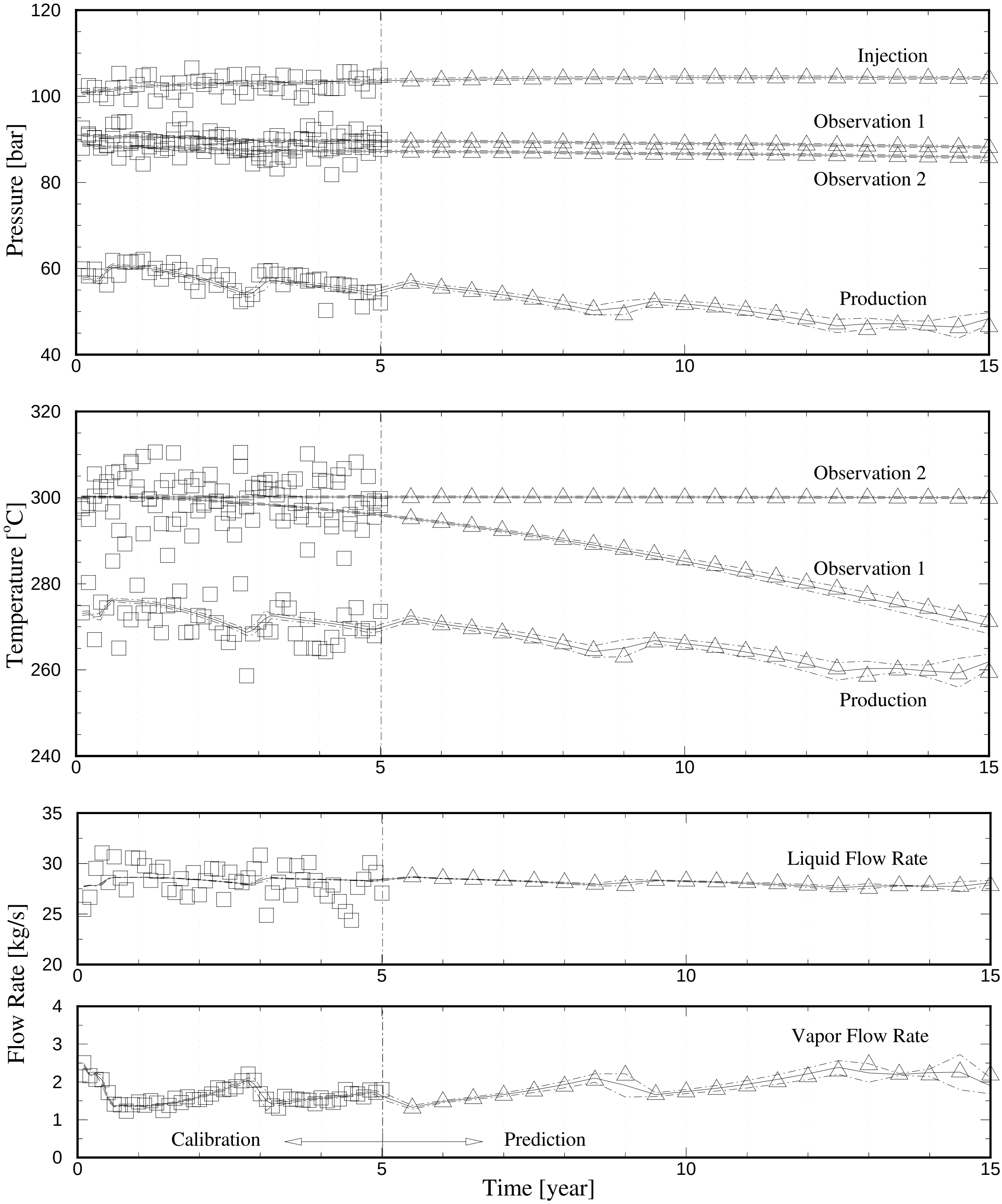

Figure 3 - Calibration and prediction of pressures, temperatures, water and vapor flow rates. Squares are synthetic data points used for calibration. Triangles represent the true system response. Simulation results based on the estimated parameter set are shown as solid lines. Error bands (dash-dotted lines) are calculated using linear error propagation analysis.

This is reflected by correlation coefficients close to one, calculated from matrix $\mathbf{C}_Z$.

Provided that the expected measurement errors (see Table 1, column 3) are reasonable, the bulk of the information about the parameters of interest is contained in the accurate vapor flow rate measurements and the pressure data in the production well. The contribution of a certain observation (e.g. flow rate data of a given accuracy taken over the entire measurement period) to the solution of the inverse problem can be evaluated by adding up all the absolute values of the corresponding sensitivity coefficients, weighted by the expected measurement error and scaled by the inverse of the parameter variation. This qualitative measure is summarized in Table 5, column 2. Comparing total sensitivities of individual observations, one can conclude that accurate measurements of vapor flow rates and pressures and temperatures in the injection and production wells would be sufficient to solve the inverse problem, i.e., data from the observation wells are less sensitive in our example. Note that this kind of an analysis can be performed without actually taking the data, i.e. it can be used to design and optimize monitoring systems. Details of such a procedure are described in Finsterle (1995).

The system response as observed in the injection, production and observation wells is shown in Figure 3. The squares are the synthetically generated and perturbed data points used to calibrate the model. The triangles represent the future system response for the true parameter set (see Table 2, column 2).

The standard deviations of the final residual (Table 5, column 3) are on the order of the measurement errors, indicating that no significant systematic errors are present. Finally, the contribution of each observation type to the final value of the objective function (Table 5, column 4) is evenly distributed among the measurements, confirming the choice of the weighting factors in matrix $\mathbf{C}^{-1}$.

Recall that this study was made using synthetic data with known error structure, and that no systematic errors are made since the conceptual model is correct. In field applications, the proper conceptual model and the structure of the random errors are not exactly known. Note, however, that the relative weighting of data points can easily be adjusted and is partly automated in ITOUGH2 following the suggestions by Carrera and Neuman (1986). While the problem of systematic errors is not directly addressed by inverse modeling, the automation of the calibration step makes it possible to examine a number of alternative conceptual models. The extensive residual analysis performed by ITOUGH2 provides a means to identify aspects of the model that need to be refined. Moreover, model identification criteria (Carrera and Neuman, 1986) are evaluated which help select the model that most likely represents field conditions. Successful application of ITOUGH2 to a variety of laboratory and field data has been demonstrated (Finsterle and Pruess, 1994, 1995).

Table 5 - Total Sensitivity of Observations, Standard Deviation of Residuals, and Contribution to Objective Function.

Observation	Sensitivity	Std. Dev.	COF
Pressure Inj.	789	1.9	9.7
Pressure Pro.	1500	2.0	10.3
Pressure W1	426	2.2	12.5
Pressure W2	358	2.1	11.6
Temperature Pro.	680	4.6	8.9
Temperature W1	107	5.4	12.2
Temperature W2	100	5.2	11.3
Water flow rate	87	1.6	10.5
Vapor flow rate	1735	0.1	13.0

COF: Contribution to Objective Function [%]

4. CONCLUDING REMARKS

The purpose of this study was to demonstrate the flexibility of an inverse modeling approach for automatic history matching and the estimation of simulation model parameters by performing a joint inversion of all available data. In addition to automatic model calibration, the ITOUGH2 code provides a number of semi-quantitative measures to study parameter sensitivities, correlations between parameters and observations, prediction uncertainties, total parameter sensitivities, potential benefits from taking measurements of a certain kind and in a certain location. This information is useful for the design and optimization of reservoir delineation and monitoring programs.

The advantage of inverse modeling procedures is that they overcome the time and labor intensive tedium of trial-and-error model calibration. Effective, model-related parameters are automatically determined on the scale of interest. This ensures that the reliability of subsequent predictions can be greatly improved if they are based on the same or a similar conceptual model of the geothermal reservoir. Matching data obtained during exploitation of a geothermal field and predicting the future reservoir performance is an excellent example of this philosophy.

5. ACKNOWLEDGMENT

This work was supported by the Assistant Secretary for Energy Efficiency and Renewable Energy, Geothermal Division, of the U.S. Department of Energy under contract No. DE-AC03-76SF00098. We would like to thank Curt Oldenburg and Marcelo Lippmann for thoughtful reviews.

6. REFERENCES

Carrera, J. and S.P. Neuman (1986). Estimation of Aquifer Parameters Under Transient and Steady State Conditions: 1. Maximum Likelihood Method Incorporating Prior Information, *Water Resour. Res.*, *22*(2), 199-210.

Finsterle, S. (1993). *ITOUGH2 User's Guide, Version 2.2*, Report LBL-34581, Lawrence Berkeley National Laboratory, Berkeley, Calif.

Finsterle, S. (1995). *Design of a Welltest for Determining Two-Phase Hydraulic Properties. Report* LBL-37448, Lawrence Berkeley National Laboratory, Berkeley, Calif.

Finsterle, S. and K. Pruess (1994). Estimating Two-Phase Hydraulic Properties by Inverse Modeling, *Proceedings. of the Fifth International Conference on High Level Radio-active Waste Management*, 2160-2167, Las Vegas, Nevada.

Finsterle, S. and K. Pruess (1995). Solving the Estimation-Identification Problem in Two-Phase Flow Modeling, *Water Resour. Res.*, *31*(4), 913-924.

International Formulation Committee (1967). *A Formulation of the Thermodynamic Properties of Ordinary Water Substance*, IFC Secretariat, Düsseldorf, Germany.

Levenberg, K. (1944). A Method for the Solution of Certain Nonlinear Problems in Least Squares, *Q. Appl. Math.*, 2, 164-168.

Marquardt, D.W. (1963). An Algorithm for Least-Squares Estimation of Nonlinear Parameters, *J. Soc. Ind. Appl. Math.*, *11*(2), 431-441

Narasimhan, T.N. and P.A. Witherspoon (1976). An Integrated Finite Difference Method for Analyzing Fluid Flow in Porous Media, *Water Resour. Res.*, *12*(1), 57-64.

O'Sullivan, M.J. and D.P. Bullivant (1995). Inverse Modelling of the Wairakei Geothermal Field Using ITOUGH2, *paper presented at the 17th New Zealand Geothermal Workshop*, Auckland, New Zealand, Nov. 8-10.

Pruess, K. and T.N. Narasimhan (1982). On Fluid Reserves and the Production of Superheated Steam from Fractured Vapor-Dominated Geothermal Reservoirs, *J. Geophys. Res., 87*, (B11), 9329-9339.

Pruess, K. and T.N. Narasimhan (1985). A Practical Method for Modeling Fluid and Heat Flow in Fractured Porous Media. *Soc. Pet. Eng. J.*, *25*(1), 14-26.

Pruess, K. (1991). *TOUGH2 - A General-Purpose Numerical Simulator for Multiphase Fluid and Heat Flow*. Report LBL-29400, Lawrence Berkeley National Laboratory, Berkeley, Calif.

Pruess, K. and Y.-S. Wu (1993). A New Semi-Analytical Method for Numerical Simulation of Fluid and Heat Flow in Fractured Reservoirs, *SPE Advanced Technology Series*, Vol. 1, No. 2, 63-72, Richardson, Texas.

White, S.P. (1995). Inverse Modelling at Kawerau, *paper presented at the 17th New Zealand Geothermal Workshop*, Auckland, New Zealand, Nov. 8-10.