PROCEEDINGS, Thirty-Sixth Workshop on Geothermal Reservoir Engineering
Stanford University, Stanford, California, January 31 - February 2, 2011
SGP-TR-191

STOCHASTIC PERMEABILITY ESTIMATION FOR THE SOULTZ-SOUS-FORÊTS EGS RESERVOIR

Christian Kosack[§], Christian Vogt, Gabriele Marquart, Christoph Clauser, Volker Rath*

Institute for Applied Geophysics and Geothermal Energy, E.ON Energy Research Center, RWTH Aachen University, Mathieustr. 6, 52074 Aachen, Germany; [cvogt; gmarquart; cclauser]@eonerc.rwth-aachen.de
[§] now at: E.ON Gas Storage GmbH, Norbertstr. 85, 45131 Essen, Germany; christian.kosack@eon-gas-storage.com
* now at: Departamento de Física de la Tierra, Astronomía y Astrofísica II, Universidad Complutense de Madrid, Spain; vrath@fis.ucm.es

ABSTRACT

Information on the distribution of permeability at depth is of primary concern in geothermal reservoir engineering. Based on a tracer circulation test performed at the European Enhanced Geothermal System (EGS) test site at Soultz-sous-Forêts, France, three different inversion methods are applied and studied with respect to their potential for estimating the permeability distribution: A fully physical, gradient-based Bayesian inversion, a massive Monte Carlo (MC) approach with geostatistical analysis, and an Ensemble-Kalman-Filter (EnKF) assimilation. A limited, high permeability zone is a common feature of all models. It is the main flow conduit for most of the tracer flow and assumed to be associated with a primary fault zone bypassing two of the boreholes. A common finding of all approaches is that permeability in most of the reservoir is on the order of 10^{-15} m^2; between GPK2 and GPK3 there is a good connection with a mean permeability on the order of 10^{-13} m^2, while a barrier exists to GPK4. The Bayesian Inversion returns rock properties yielding the best fit to the observed tracer concentration at GPK3, but shows that permeability and porosity cannot be retrieved independently. The MC approach identifies various alternative flow paths corresponding to tracer concentrations almost fitting the observations at GPK3 equally well. Although the permeability field obtained by the EnKF has a poor spatial resolution, which is a consequence of only three monitoring wells, it is the only method fitting simultaneously the observations at both GPK2 and GPK4.

INTRODUCTION

The use of geothermal energy for an economical generation of electrical energy requires temperatures of 150° C – 200° C and flow rates on the order of 50 L s^{-1} (e. g. Clauser, 2006). For regions outside natural steam systems and enhanced surface heat flow (for example Iceland, Turkey, Indonesia, etc.) these conditions can only be met at depths below 3 km and with engineered heat exchanger surfaces for the flow systems, i.e. Hot Dry Rock (HDR) or Enhanced Geothermal Systems (EGS). In these techniques, a heat exchanger surface of the host rock is engineered or enhanced by hydraulic or chemical stimulation. The first method requires water to be injected under high pressure to create irreversible shearing and opening of fractures, the second one uses acid to dissolve minerals and to increase permeability. While both methods have proven successful for enhancing permeability at depth, it is still a challenge to plan and control the stimulation process.

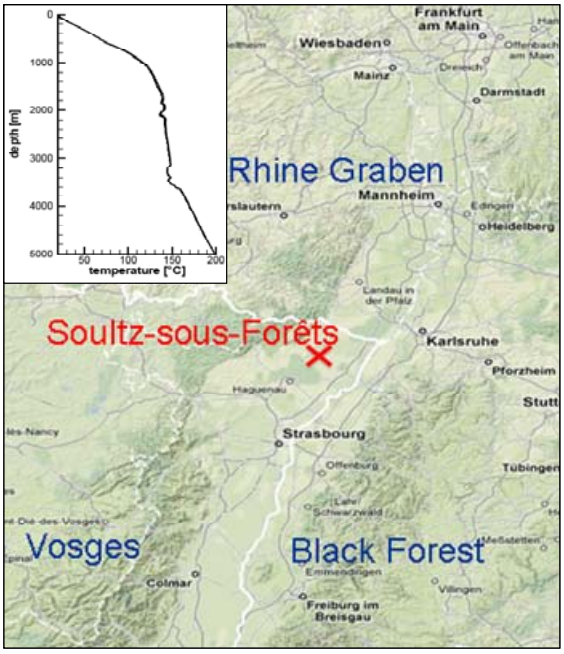

Figure 1: Location of the European EGS test site at Soultz-sous-Forêts, France. Inset: Temperature depth profile at the EGS test site.

However, reliable information on the response of the host rock to stimulation and the uncertainty of the engineered permeability field is vital for the

development and acceptance of the EGS technique and for estimating the technical and economic risk of geothermal projects (Manzella, 2010). Information on the distribution of enhanced permeability at depth can be obtained from interpretation of micro earthquake clouds observed during stimulation (e.g. Shapiro, 2000; Delépine et al., 2003) and from pumping and tracer tests.

In order to study the performance of an EGS reservoir, a European EGS test site was set up in 1987 at Soultz-sous-Forêts, France, located in the Lower Rhine Graben (Figure 1). The Rhine Graben is part of the European Cenozoic Rift System and developed during the Tertiary concomitant to the European Alpine collision that formed the Alps along pre-existing Permian troughs. It has a pronounced Graben structure characterized by normal faults which can be followed in seismic sections down to 4000 m. Faulting is listric in character displacing the Cretaceous sediments and can be traced into the granitic basement where it shallows out (Cloetingh et al., 2006). The Quaternary fluviatile sediments of the river Rhine cover the older faulted sediments.

The decision for this site was based on the high surface heat flow and a very promising thermal gradient observed in shallow drill holes of previous oil exploration (Gérard and Kappelmeyer, 1989; see Figure 1). Although the thermal gradient declines below 1000 m depth, the temperature still reaches about 200° C at 5 km. At Soultz, three wells were drilled in the granitic bedrock down to 4 km - 5 km and were hydraulically stimulated to enhance the hydraulic connectivity between the wells. An overview of the project in general is given by Gérard et al. (2006) and on the brine-rock-interaction during stimulation by Bächler and Kohl (2005).

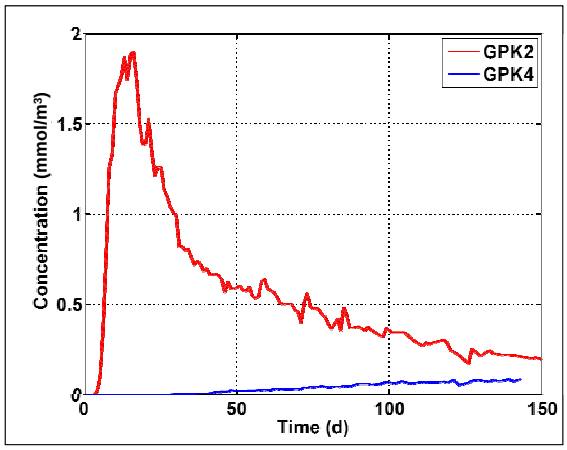

Figure 2: Tracer concentration in GPK2 and GPK4 in response to injection in GPK3.

A tracer test was performed in July 2005 (Sanjuan et al., 2006) in order for estimating the changes of the hydraulic properties between the injection borehole GPK3 and the two production boreholes GPK2 and GPK4,. The outcome of this experiment, the tracer concentration at GPK2 and GPK4 versus time is shown in Figure 2. The experiment showed that GPK2 and GPK3 are well connected. Former studies (e.g. Gessner et al., 2009) proposed a dominant fluid flow through a pre-existing natural fault system already detected by seismic investigation (Cloetingh et al., 2006). The inclined conductive fault is intersected by two of the boreholes but missed by GPK4 which ends above the fault, as indicated in Figure 3.

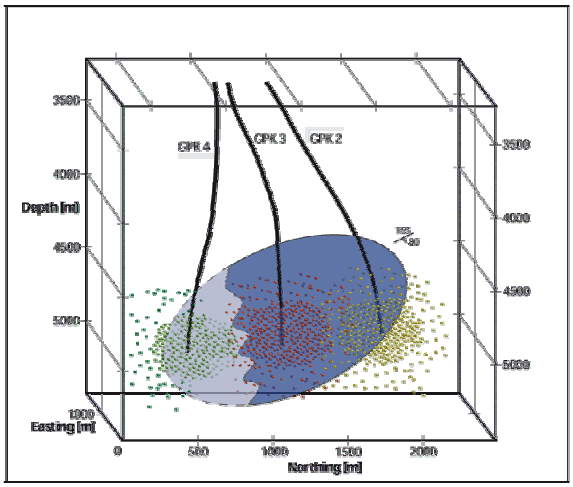

Figure 3: Sketch of the borehole configuration in Soultz-sous-Forêts 2005 adopted from Gessner et al. (2009). The fault is shown in dark blue where it is intersected by the boreholes GPK2 and GPK3, the light blue area illustrates the possible continuation in the direction of GPK4, but GPK4 ends above the fault. The green, red and yellow dots outline the zones affected by the hydraulic stimulation of the boreholes.

Based on the observed tracer concentrations (Figure 2) we used three inversion methods in combination with a numerical simulator of porous flow (*Shemat_suite*; Clauser, 2003; Rath et al., 2006) to reveal the permeability distribution in the Soultz reservoir and to compare the advantages and disadvantages of the different inversion methods.

THE SOULTZ-SOUS-FORÊTS 2005 TRACER EXPERIMENT

A long-term tracer circulation test was performed between the boreholes GPK2, GPK3, and GPK4 for characterizing the effects of hydraulic stimulation on the reservoir at a depth of 5000 m,. From July 2005 to December 2005, a fluid volume of about 209 000 m^3 was injected into GPK3 and 165 000 m^3 and 40 000 m^3 were produced from GPK2 and GPK4, respectively (Sanjuan et al., 2006), yielding a nearly even mass balance. A mass of 150 kg of 85 % pure fluorescein was used as a tracer. The fluorescein was

dissolved in 0.95 m^3 of fresh water which results in a concentration of 0.389 mmol L^{-1}. This fluid was injected into GPK3 over 24 hours, as geochemical fluid monitoring started at GPK2 and GPK4. Figure 2 shows the observed concentration in GPK2 and GPK4 versus time. Fluorescein was first detected in GPK2, 4 days after the injection into GPK3. In GPK4, fluorescein was detected only 24 days after the injection. The average pumping rates were 11.9 L s^{-1} in GPK2, 15 L s^{-1} in GPK3, and 3.1 L s^{-1} in GPK4, already indicating a reduced water supply to GPK4. In combination with the tracer arrival times this shows that the hydraulic connection is very heterogeneous. The maximum tracer concentration measured in the produced fluid was 2 mmol m^{-3} in GPK2 and appeared 9 days to 16 days after the injection (Figure 2). It dropped to approximately 0.2 mmol m^{-3} after 5 months. No clear maximum tracer concentration could be detected in GPK4 even after 5 months of production. The final measured value was approximately 0.09 mmol m^{-3}. The relative uncertainty of the measured tracer concentration was estimated to be on the order of 10 % to 15 %. During the entire time of the experiment only 23.5 % of the tracer was recovered (Sanjuan et al., 2006).

A first interpretation of the 2005-tracer experiment was given by Sanjuan et al. (2006) using a simple analytical model which could explain the first 20 days of the measurement. Better results were obtained by Blumenthal et al. (2007) with a 2D numerical model. They revealed the mutual dependence of porosity and permeability inside the reservoir by varying the peak concentration. Nevertheless, the 2D model did not fit the tracer curve over time due to the restriction of the flow in a plane. A 3D model by Kosack (2009), based on a deterministic Bayesian inversion, discriminated only between two flow zones and obtained a nearly perfect fit of the tracer curve between GPK2 and GPK3 (we will discuss this model in the present paper). Even though the tracer concentration fit was optimal, this study showed that different geometrical models were able to fit the tracer curve equally well and only products of parameters (for example porosity and permeability) could be resolved due to their high spatial correlation. Gentier et al. (2010) provided an approach based on a manual fit of permeabilities of an underlying fracture network. The fit quality was not as perfect as the results of Kosack (2009). However, both the GPK2 and the GPK4 tracer curves were matched by a combined model.

Beside these studies a deeper insight into possible flow paths and fracture networks and their statistical significance is required. We therefore used two additional stochastic approaches to characterize fracture networks able to fit the tracer curves: A Monte Carlo (MC) approach assuming a heterogeneous permeability field derived from a characteristic permeability distribution based on lab measurements on rock samples and log observations, and an assimilation filter method (Ensemble Kalman Filter) minimizing successively the error between observed and numerically predicted tracer concentration by adjusting the permeability distribution according to calculated error statistics.

SIMULATION CODE AND MODEL SET-UP

In our study the geothermal system is simulated by the code *Shemat_suite* (Clauser, 2003; Rath et al., 2006) which solves the coupled transient equations for groundwater flow, heat transport, and transport of chemical tracers in a porous rock. Assuming mass conservation, the flow equation in terms of the constant density hydraulic head h_0 is

$$S\frac{\partial h_0}{\partial t} = \nabla \cdot \left[K\left(\nabla h_0 + \rho_r \nabla z\right)\right] + Q_f \qquad (1)$$

where S is the specific storage coefficient, $\rho_r = \frac{\rho - \rho_0}{\rho_0}$ is relative density and ρ_0 is constant reference density, K is hydraulic conductivity, and Q_f a source or sink term, $v = -K\left(\nabla h_0 + \rho_r \nabla z\right)$ is the Darcy velocity with $K = \frac{\rho_f g}{\mu_f} k$ where ρ_f and μ_f are the density and viscosity of the fluid, g is gravity and k is the permeability, which is physically a tensor but treated here as an isotropic space variable $k(\boldsymbol{x})$. The temperature T is determined assuming conservation of energy by the transient heat transport equation

$$\frac{\partial T}{\partial t}\left[\phi(\rho c)_f + (1-\phi)(\rho c)_r\right] = \nabla \cdot \left(\lambda \nabla T - (\rho c)_f T v\right) + Q_h \qquad (2)$$

where λ is the bulk thermal conductivity, (ρc) is the thermal capacity with the indices r for the rock skeleton and f for the fluid; □ is porosity, and Q_h the heat source term. Tracer transport is modeled by a diffusion-advection equation for concentration c

$$\phi\frac{\partial c}{\partial t} = \nabla \cdot \left[D\nabla c - (vc)\right] + Q_c \qquad (3)$$

where D the dispersion coefficient and Q_c is the source term for the dissolved tracer. The dispersion coefficient is physically a tensor and defined by the transversal and longitudinal dispersion length and the molecular diffusion. To keep our approach simple we regard D as a scalar and do not discern between transversal and longitudinal dispersion lengths, and neglect molecular diffusion which is a very slow process. Thus, the dispersion coefficient is given by $D = \alpha_D |v|$, where α_D is the dispersion length. Equations (1) to (3) are expressed in a finite volume notation for rectangular non-equidistant grids and

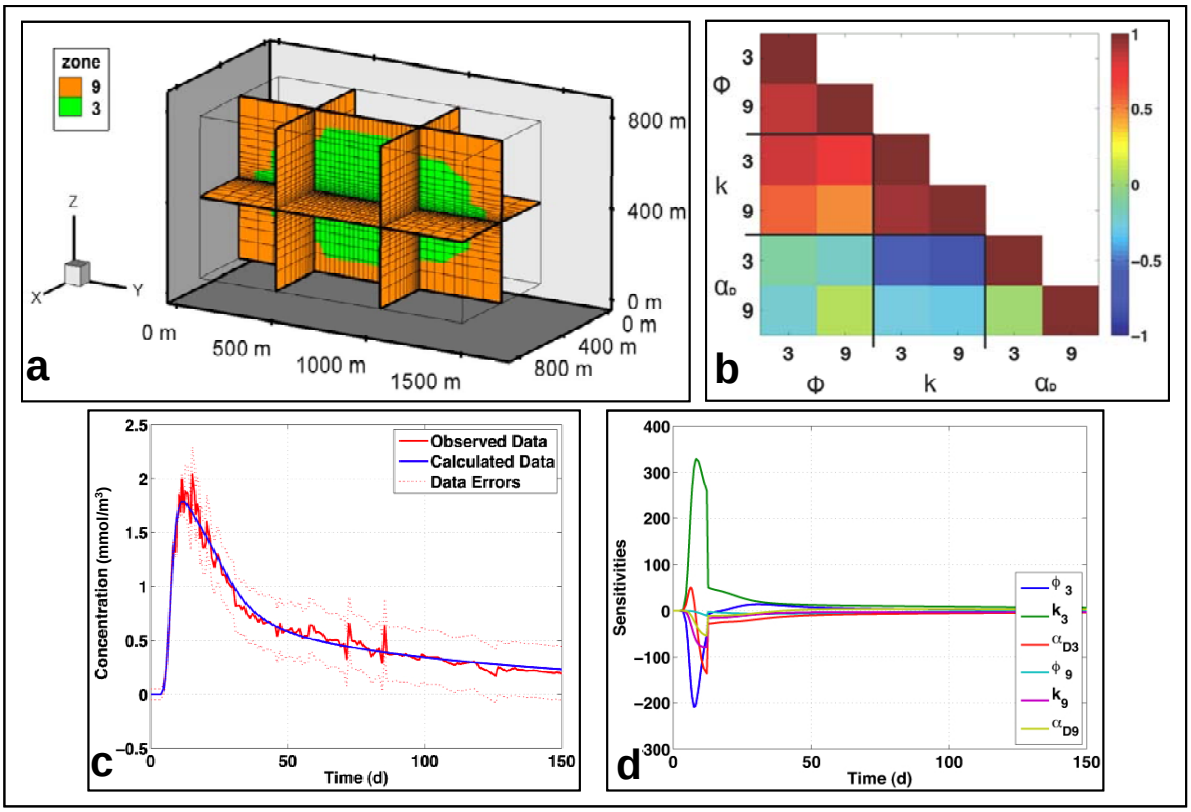

Figure 4: a) Model set up for the Bayesian Inversion; only GPK2 and GPK3 are modeled, located at the intersections of the shown grid planes; b) correlation matrix for the porosity □, permeability k, and dispersion length α_D; c) fit to the observed tracer concentration at GPK3; d) Sensitivities versus time for the various parameters (for the first days a lower data error was assumed).

solved with a Picard iteration method for transient problems.
The injection and production of fluid in the drill holes are implemented by an internal Neumann boundary condition with prescribed flow at the four corner nodes of the rectangular cell.

INVERSION METHODS AND RESULTS

As mentioned above we tested three different inversion methods with their different advantages and disadvantages to infer the permeability distribution at depth.

Bayesian Inversion

A fit to the tracer observation is obtained by minimizing both the quadratic error between data and modeled data, and the distance between the fitting parameter and an a priori parameter. Thus, the objective function is given by (Tarantola, 2004)

$$\Theta = (d - g(p))^T C_d^{-1} (d - g(p)) + (p - p_a)^T C_p^{-1} (p - p_a) = Min \quad (4)$$

p is the vector of system parameters, $g(p)$ the predicted system response (generally non-linear) and, d is the data vector. C_d and C_p are the data and parameter a priori covariance matrices, respectively, given beforehand. Linearization and differentiation of the objective function yields an iteration scheme (with k as the iteration index) for the parameter vector p

$$p^{k+1} = p^k + [J^T C_d^{-1} J + C_p^{-1}]^{-1} \cdot [J^T C_d^{-1} (d - g(p^k)) - C_p^{-1} (p^k - p_a)] \quad (5)$$

The Jacobian $J_{ij} = \frac{\partial g(p)_i}{\partial p_j}$ is the derivative of the model prediction with respect to the various parameters and is determined in our code by automatic differentiation (AD-method; Bischof et al., 2008).
Based on the Jacobian and the a-posteriori covariance matrix $\tilde{C}_p = \left((C_d^{-\frac{1}{2}} J)^T (C_d^{-\frac{1}{2}} J) + C_p^{-1} \right)^{-1}$, a number of resolution matrices can be defined, to study the reliability of the inverted parameters.
Our Bayesian inversion model has a physical dimension of approximately 814 m × 1490 m × 814 m and is discretized into 29 × 36 × 29 orthogonal cells in x, y, and z direction respectively, and refined in the center which is assumed to be the fault zone

(Figure 4a). We consider the flow between GPK3 and GPK2 only, thus, the same injection and production rates are used for GPK3 and GPK2. The boreholes are represented by cells in the fault zone. This appears to be an adequate approach for an inclined conductive fault that is intersected by two boreholes. In reality, the main inflow will occur where the boreholes intersect the fault zone. However, since reliable information of the hydraulic connection between fault zone and boreholes is not available, we model the injection directly in the fault plane.

The hydraulic boundary condition for the top cells of the model is a constant hydraulic head of 4600 m. Constant temperatures of 185 °C and 215 °C are specified at the top and bottom of the model. The tracer can be reasonably well modelled with 2000 time steps, for the 160 days of the experiment.

During the circulation test, the produced fluid was cooled in a heat exchanger prior to re-injection into GPK3. The injection temperature, varying with time, is implemented as a time-dependent boundary condition. At the injection well a concentration of 0.389 mmol L^{-1} of fluorescein is assumed during the first day and injected constantly; later the produced water is re-injected.

Observation data is the tracer concentration with time at GPK2. The inverted parameters are porosity, permeability, and dispersion length in the fault region and the surrounding. This very simple model already yields a nearly perfect fit to the observations (Figure 4c). However, the model parameters (rock properties) cannot be resolved independently. A measure for the mutual dependence is the normalized a-posteriori covariance matrix, i.e. the correlation matrix (Figure 4b), which clearly indicates high correlation and anti-correlation between the various parameters, in particular porosity and permeability. The Jacobian, weighted by the data and parameter errors, is a direct measure for the sensitivity of the resolved parameter on the observed data. One can study the spatial distribution of the sensitivity to obtain information on areas of high and low resolution or one can study the sensitivity variation for various parameters with time at a particular point (Figure 4d). For the tracer experiment at Soultz, the sensitivity is high for all parameters only during the first 20 to 40 days which provide a maximum of information.

The values which we obtained for porosity and permeability are 0.0015 and 3.5 10^{-14} m^2 for the fault zone, and for the host rock 0.007 and 3.16 10^{-17} m^2, respectively. Diffusivity is about one order of magnitude higher in the fault zone relative to the host rock. However, the values vary slightly depending on the a-priori assumptions.

Geostatistic MC Modelling

In contrast to the Bayesian approach which is limited to a homogeneous permeability field in each involved zone, we used a stochastic approach and applied a Monte Carlo (MC) method. We assumed a heterogeneous permeability field for characterizing an equivalent fracture network which causes fluid flow fitting the tracer curves. The different permeability distributions for the MC realizations are created following the Sequential Gaussian Simulation method (SGSim) (Deutsch & Journel, 1998). This method is based on a probability density distribution for the rock property, in our case permeability (Figure 5a). The histogram follows two combined Gaussian distributions. The left peak represents the combined effect of background permeability on grain size scale and low fracture density of the solid rock. The smaller peak on the right corresponds to the fractured cells of the reservoir. Using this approach it is possible to capture fractures of very different scales: from major fracture zones cutting through the granite to intra-crystalline micro-fractures inducing weakness in the rock matrix.

Table 1: Model parameters for the geostatistic and EnKF modelling

Parameter	Value
Porosity	$5 \cdot 10^{-4}$
Background permeability	10^{-17} m^2
Rock compressibility	10^{-8} Pa^{-1}
Dispersion length	5 m
Thermal conductivity	2.5 W m^{-1} K^{-1}
Heat capacity	2.06 10^6 J kg^{-1} K^{-1}
Radiogenetic heat generation	10^{-10} W m^{-3}
Hydraulic head at top	4600 m
Temperature at top	185 °C
Specific heat flow at bottom	80 mW m^{-2}

The permeability fields for the MC realizations are created in the following way: (1) The geometrical structure of the reservoir model is discretized on a specific grid. (2) The rock property distribution is transformed into a Gaussian shape, with zero mean and unit variance. (3) The algorithm follows a random path through the model. (4) For each grid node, nearby data and already simulated nodes are used for a Kriging interpolation of the target property. (5) A property value drawn randomly from the distribution defined by Kriging mean and variance is assigned to the node. (6) A realization is completed after property values have been assigned to all nodes of the model. (7) The rock property distribution assigned to the model is transformed back from a Gaussian distribution into the original space. (8) More realizations are created by following other random paths, each of them equally likely.

Additionally, the permeability was set to 10^{-12} m^2 in the grid cells directly adjacent to the boreholes corresponding to the highly stimulated region around the wells. The smoothness of the permeability field is controlled by the correlation length, which in geostatistical modelling is a measure for the distance

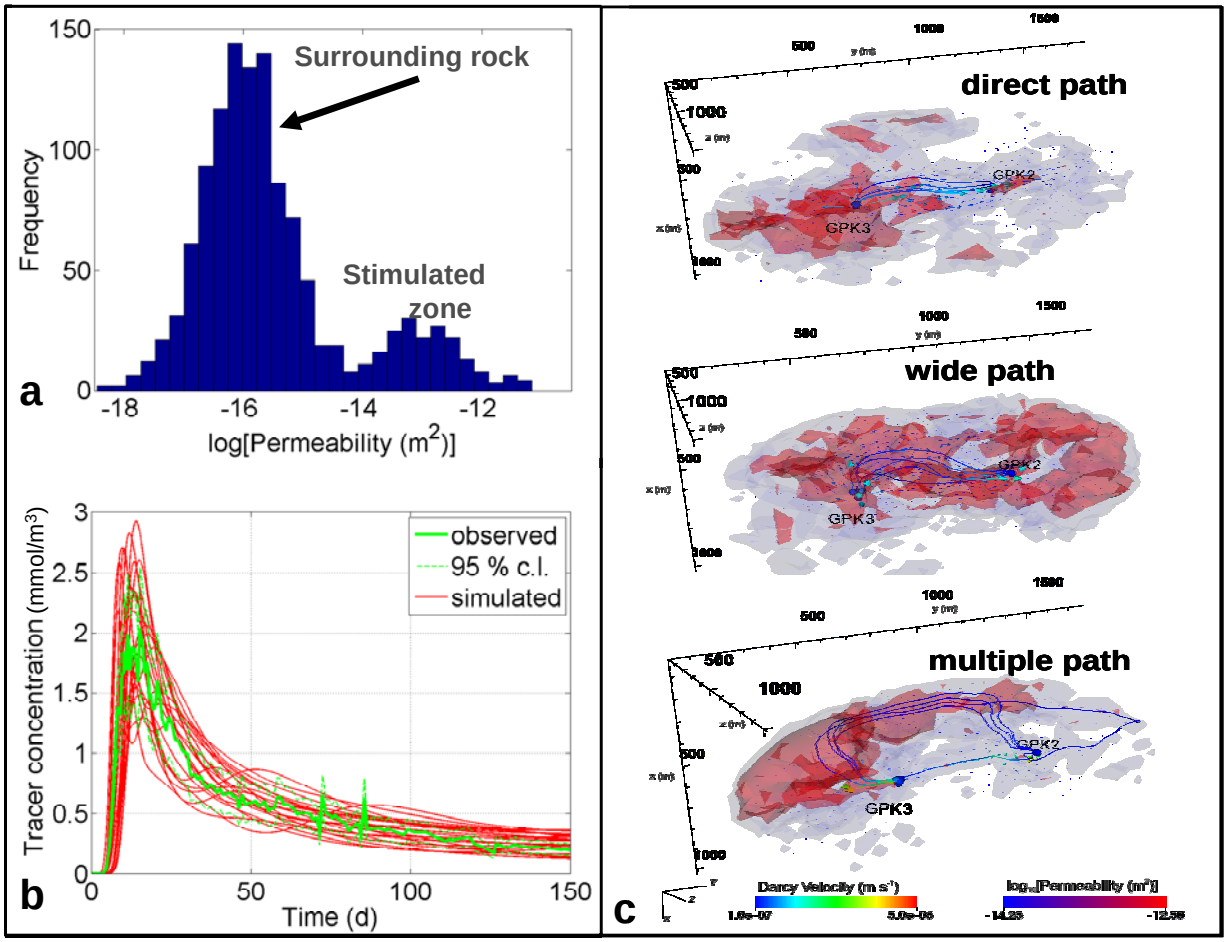

Figure 5: a) Log permeability distribution for the stimulated zone; b) Fit to the tracer concentration curve of the 25 best models; c) 3D view of three typical models for alternative flow paths (direct path, wide path, multiple paths). Indicated are the 10^{-13} m^2 (red) and 10^{-14} m^2 (light grey) permeability isoplanes and the streamlines; along the streamlines the Darcy velocity is color-coded. The locations of GPK3 and GPK4 are also indicated.

within two simulated or observed data points are considered as similar.

To evaluate a reliable value for the correlation length, we used the induced micro seismicity around GPK3 during a hydraulic stimulation. According to Delépine et al. (2003), the total number of micro-seismic events drops to e^{-1} of the original value at a distance of 350 m. This distance is supposed to be a minimum correlation length in all directions.

Two geometrical models were set up: One with a good connectivity between GPK2 and GPK3, and one for the entire reservoir including also GPK4. The reservoir is considered at a depth range between 4500 m and 5500 m. The corresponding models comprise 21 × 36 × 21 and 21 × 48 × 21 grid cells, respectively, on regular grids with a side length of 50 m in each dimension. The values for all other parameters are given in the Table 1.

In this study, we identified groups of realizations with similar permeability distributions yielding optimal fits of the tracer data and reveal different possible pathways. We calculated 5000 forward models and picked the hundred best fitting ones according to a combination of a root mean square error and a Kolmogorow-Smirnow-criterion (Frank and Massey, 1951), which considers the fit of the overall shape of the curve. The tracer concentration curves of the 50 best fitting models are shown in Figure 5b. All fitting models can be grouped in three possible reservoir scenarios as shown in Figure 5c.

Further inspection of these groups suggests the following conclusions: (1) A fit of the tracer curves requires the boreholes to be connected through a main path with permeabilities of at least 10^{-14} m^2. If the permeability exceeds about 10^{-13} m^2, however, the tracer arrives too early, with too large an amplitude or without distinct tailing. (2) If the connection zone between the two boreholes is not wide enough, water of the surrounding of GPK2 is also produced, corresponding to a high permeable zone with $k \geq 3$

10^{-12} m^2 around GPK2. This has no influence on the tracer curve however. It merely shows that also pre-existing water may be produced, which was also found in other studies (Sanjuan et al., 2006). (3) In all simulations parts of the tracer is stored in dead end of flow paths. Thus, some tracer is not retrieved during the circulation test. (4) The periphery of the model in 500 m distance of the main flow path has little influence on the tracer curve and the rock permeability cannot be resolved. The same applies for high-permeability zones unconnected with the boreholes. (5) The tracer is transported on pathways with no larger lateral extent than 300 m, which is additional evidence that the flow does not pass through the entire stimulated area.

As mentioned before, none of the MC realizations could fit the tracer curves at GPK2 and GPK4 simultaneously. We explain this deficiency by a still insufficient number of MC realisations. Considering that permeability values are randomly assigned to about 2 10^4 grid points, the probability to match the observations at GPK2 and GPK4 simultaneously is small.

EnKF Data Assimilation

I Sequential data assimilation methods such as the Ensemble Kalman Filter (EnKF; Evenson, 1994; 2007) provide an alternative to the numerically demanding MC simulations if monitoring data is available. The Kalman Filter is essentially a sequential assimilation of observation data into a numerical simulation of a transient system with Gaussian error statistics. An error estimate is calculated from measured data whenever it becomes available for improving the prediction about the state of the system, i.e. the information on observed data is propagated in the numerical simulation by the dynamic of the system.

The state of the system is described by a state vector Ψ which may comprise various kinds of state variables and parameters. The transient behaviour is described by a generally non-linear transition function F which in our case is the *Shemat_suite* forward simulator. The propagated system is then given simply by

$$\Psi_k^f = F\ (\Psi_{k-1}^a) + \varepsilon_s \qquad (6)$$

where k is the time step number. Independent of the type of transition function, the Kalman Filter is a two-step procedure where the first step is a forecast or prediction (denoted by the superscript f) and the second step is the assimilation or update step (indicated be superscript a); ε_s is the system noise. At the same instant in time observation data d_k might be available which can be represented by

$$d_k = H\,\Psi_k^f + \varepsilon_r \qquad (7)$$

where H is a measurement operator linking data d to the system state vector Ψ and ε_r is the measurement error. The basic idea in Kalman Filter is that the improved state of the system is given by the state vector forecast and a weighted difference between the measured data d and the predicted data $H\Psi$

$$\Psi_k^a = \Psi_k^f + K_k(d_k - H\Psi_k^f) \qquad (8)$$

where the matrix K, the Kalman Gain, determines the weights for the new observations as they are incorporated into the updated state vector. Since we assume a stochastic system, the state vector is given by a probability density which in turn is conditioned by the available observations. Since we assume a Gaussian distribution, the probability density is characterized by a covariance matrix $C_{p,k}$. The expression for the Kalman gain K is found by minimizing this error covariance matrix after the assimilation step

$$K_k = C_{p,k}^f\, H^T \cdot (HC_{p,k}^f\, H^T + C_{d,k})^{-1} \qquad (9)$$

$C_{d,k}$ is the data error covariance and $C_{p,k}$ is the system error covariance obtained from the variance of an ensemble of system realizations which converges during repeated data assimilation steps.

The EnKF method has been proven to recover the permeability field if observations of chemical tracer concentration are available (Marquart et al., 2010). Here we used the same model set-up as for the MC realisations and prescribed the initial permeability fields for 250 realisations by the Sequential Gaussian Simulation method. The permeability field was then assimilated inside an ellipsoid comprising all three wells and held constant at 10^{-17} m^2 outside. For the assimilation, we used the concentration data recorded at both production wells, GPK2 and GPK4, at 40 equally spaced points in time of the 100 first days of the experiment. When the assimilation was completed, we re-run the entire ensemble as well as the ensemble mean to study the resulting tracer concentration curves at GPK2 and GPK4 (shown in Figure 6c and d). While the tracer curve of the ensemble mean fits well the observation (Figure 6c), this approach also predicts the tracer concentration at GPK4 within the observed range (Figure 6d). Concerning the permeability field, the EnKF method, though still preliminary, confirms our previous results. The permeability in most of the reservoir is on the order of 10^{-15} m^2. There exists a good connection between GPK2 and GPK3 with a mean permeability on the order of 10^{-13} m^2. In contrast there is a flow barrier towards GPK4 (Figure 6a).

In EnKF, the standard deviation of the ensemble can be used to define the quality of the estimated property. For the borehole configuration at Soultz the

uncertainty of the permeability estimate is within a factor of 3 in between the wells, increasing to about a factor of 20 towards the periphery.

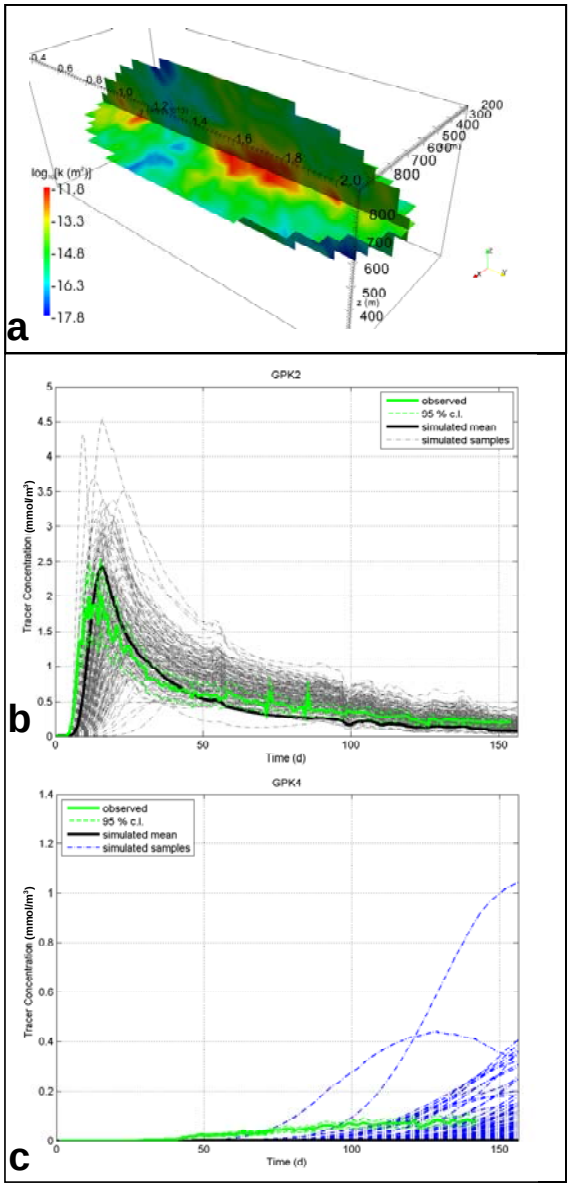

Figure 6: Simulation of the 2005 tracer experiment in Soultz sous Forêts using EnKF: a) Estimated permeability distribution in an ellipsoid containing the wells GPK2, GPK3, and GPK4; b) chemical concentration at GPK2; c) chemical concentration at GPK4; the green line shows observed data, the black line is the ensemble mean, and grey and blue dashed lines denote the ensemble members.

DISCUSSION

All three methods tested for estimating the permeability field in the EGS geothermal reservoir at Soultz-sous-Forêts have proven successful, while each method has advantages and disadvantages.

With the Bayesian Inversion we estimated a parameter set consisting of porosity, permeability, and dispersivity which fitted the measured tracer concentration at GPK3 nearly perfectly. The underlying reservoir geometry is extremely simplified, consisting of a nearly vertical high-permeability fault connecting GPK3 and GPK2, in line with Occam's razor that a model should not be more complicated than necessary. As a by-product of the non-linear inversion, optimal a-posteriori parameter estimates are complemented by an analysis of parameter interdependence and uncertainty. However, the implementation of the Bayesian inversion is numerically demanding, uses considerable computation time, and the obtained results are sensitive to a-priori information and inversion parameters.

Both ensemble methods, MC and EnKF yield a discrete cell-by-cell variation of permeability and a-priori information is only necessary to define bounds or histograms for permeability. Massive Monte Carlo is simple in implementation but very demanding in computation time since the vast majority of the models predict tracer concentrations far outside the range of observation. However, since a large volume of the parameter space is studied, a number of alternative models can be discriminated which fit the observed data nearly equally well.

The EnKF can only be applied to monitoring data recorded over time. Both the numerical implementation and the computational effort are on a moderate scale. As for Massive Monte Carlo only very coarse a-priori information is required and the standard deviation of the ensemble is a measure for the quality of the estimate. However, for very few monitoring locations as in case of Soultz, the spatial resolution of the resolved permeability is poor and the convergence of the ensemble during assimilation is slow. However, the EnKF was the only model approach which fitted the tracer concentration at GPK2 and GPK4 simultaneously.

Acknowledgement. We gratefully thank Marion Schindler, Albert Genter (GEIE), Bernard Sanjuan (BRGM), Clement Baujard and Thomas Kohl (Geowatt AG Zürich) for providing the data from the Soultz-sous-Forêts geothermal project for this study. We thank Darius Mottaghy for helpful discussions and Andreas Wolf for invaluable support in programming and executing *Shemat_suite*. This project was funded by the German Federal Ministry for Education and Research (BMBF) under the grant 03SF0326A.

REFERENCES

Bächler, D. and Kohl, T. (2005), "Coupled thermal-hydraulic-chemical modelling of enhanced

geothermal systems", *Geophysical Journal International*, **161**, 533– 548.

Bischof, C. H., Bücker, H. M., Hovland, P. D., Naumann, U. and Utke, J. (2008), "Advances in Automatic Differentiation", 64, LNCSE, Springer.

Blumenthal, M., Kühn, M., Pape, H., Rath, V. and Clauser, C. (2007), "Numerical modelling of hydraulic and thermal processes in the deep heat reservoir of the EGS test site Soultz-sous-Forêts, France", *Proceedings of the Euro-Conference of Rock Physics and Geomechanics on Natural Hazards: Thermo-Hydro-Mechanical Processes in Rocks*, Erice, Italy.

Clauser, C. (ed.) (2003), "Numerical Simulation of Reactive Flow in Hot Aquifers, SHEMAT and processing SHEMAT", Springer, Heidelberg-Berlin.

Clauser, C. (2006), "Geothermal Energy", *in K. Heinloth (ed.), Landolt-Börnstein, Group VIII "Advanced Material and Technologies", Vol. 3 "Energy Technologies",Subvol. C, „Renewable Energies"*, 480–595, Springer Verlag, Heidelberg-Berlin.

Cloetingh, S., Cornu, T., Ziegler, P. and Beekman, F. (2006), "Neotectonics and intraplate continental topography of the northern alpine foreland", *Earth Science Reviews*, **74**, 127–196.

Delépine, N., Cuenot, N., Rothert, E., Parotidis, M., Rentsch, S. and Shapiro, S.A. (2003), "Characterization of fluid transport properties of the hot dry rock reservoir Soultz-2000 using induced micro seismicity", *Journal of Geophysics and Engineering*, **1**, 77-83.

Deutsch, C. V. and Journel, A. G. (1998), "GSLIB. Geostatistical software library and user′s guide", Oxford University Press, New York.

Evensen, G. (1994), "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics", *Journal of Geophysical Research*, **99**, 10143-10162.

Evensen, G. (2007), "Data Assimilation. The Ensemble Kalman Filter", Springer, Berlin-Heidelberg.

Frank J. and Massey, J. (1951), "The Kolmogorov-Smirnov Test for Goodness of Fit", *Journal of the American Statistical Association*, **46**, 253, 68-78.

Gentier, S., Rachez, X., Ngoc, T. D. T., Peter-Borie, M. and Souque, C. (2010), "3D flow of the medium-term circulation test performed in the deep geothermal site of Soultz-sous-Forêts (France)", *in R. Horne (ed.), Proceedings of the World Geothermal Congress*, International Geothermal Association.

Gérard, A. and Kappelmeyer, O. (1989), "Project Européen Roches Chaudes Séches de Soultz-sous-Forêts. Résultats majeurs au 15 février", Soultz-sous-Forêts, France.

Gérard, A., Genter, A., Kohl, T., Lutz, P., Rose, P., and Rummel, F. (2006), "The deep EGS (enhanced geothermal system) project at Soultz-sous-forêts, Alsace, France", *Geothermics*, **35**, 473–483.

Gessner, K., Kühn, M., Rath, V., Kosack, C., Blumenthal, M. and Clauser, C. (2009), "Coupled process models as a tool for analysing hydrothermal systems", *Surveys in Geophysics*, **30**, 133–162.

Kosack, C. (2009), "Inverse Modelling of the 2005 Tracer Circulation Test at Soultz-sous-Forêts, France", Diploma thesis, RWTH Aachen University, Aachen.

Marquart, G., Wolf, A., Rath, V. and Vogt, C. (2010), "Stochastic Estimate of Permeability in Geothermal Reservoir Simulation Using the Ensemble Kalman Filter Method", *submitted*.

Manzella, A. (2010), "Technological challenges in exploration and investigation of EGS and UGR", *in R. Horne (ed.), Proceedings of the World Geothermal Congress*, International Geothermal Association.

Rath, V., Wolf, A. and Bücker, M. (2006), "Joint three-dimensional inversion of coupled groundwater flow and heat transfer based on automatic differentiation: sensitivity calculation, verification, and synthetic examples", *Geophysical Journal International*, **167**, 453–466.

Sanjuan, B., Pinault, J., Rose, P., Gérard, A., Brach, M., Braibant, G., Crouzet, C., Foucher, J., Gautier, A. and Touzelet, S. (2006), "Tracer testing of the geothermal heat exchanger at Soultz-sous-Forêts (France) between 2000 and 2005", *Geothermics*, **35**, 622–653.

Shapiro, S. A. (2000), "An inversion for fluid transport properties of three-dimensional heterogeneous rocks using induced microseismicity", *Geophysical Journal International*, **143**, 931-936.

Tarantola, A. (2004), "Inverse Problem Theory and Methods for Model Parameter Estimation", SIAM, Philadelphia, ISBN-13: 9780898715729, pp 358.