Geophys. J. Int. (2010) **181,** 321–333 doi: 10.1111/j.1365-246X.2009.04498.x

Reducing temperature uncertainties by stochastic geothermal reservoir modelling

C. Vogt,[1] D. Mottaghy,[2] A. Wolf,[1,3] V. Rath,[1,4] R. Pechnig[2] and C. Clauser[1]

[1]*Applied Geophysics and Geothermal Energy, E.ON Energy Research Center, RWTH Aachen University, Aachen, Germany. E-mail: cvogt@eonerc.rwth-aachen.de*
[2]*Geophysica Beratungsgesellschaft mbH, Aachen, Germany*
[3]*Scientific Computing, RWTH Aachen University, Aachen, Germany*
[4]*Departamento de Física de la Tierra, Astronomía y Astrofísica II, Universidad Complutense de Madrid, Spain*

Accepted 2009 December 23. Received 2009 December 18; in original form 2009 September 10

SUMMARY

Quantifying and minimizing uncertainty is vital for simulating technically and economically successful geothermal reservoirs. To this end, we apply a stochastic modelling sequence, a Monte Carlo study, based on (i) creating an ensemble of possible realizations of a reservoir model, (ii) forward simulation of fluid flow and heat transport, and (iii) constraining post-processing using observed state variables. To generate the ensemble, we use the stochastic algorithm of Sequential Gaussian Simulation and test its potential fitting rock properties, such as thermal conductivity and permeability, of a synthetic reference model and—performing a corresponding forward simulation—state variables such as temperature. The ensemble yields probability distributions of rock properties and state variables at any location inside the reservoir. In addition, we perform a constraining post-processing in order to minimize the uncertainty of the obtained distributions by conditioning the ensemble to observed state variables, in this case temperature. This constraining post-processing works particularly well on systems dominated by fluid flow. The stochastic modelling sequence is applied to a large, steady-state 3-D heat flow model of a reservoir in The Hague, Netherlands. The spatial thermal conductivity distribution is simulated stochastically based on available logging data. Errors of bottom-hole temperatures provide thresholds for the constraining technique performed afterwards. This reduce the temperature uncertainty for the proposed target location significantly from 25 to 12 K (full distribution width) in a depth of 2300 m. Assuming a Gaussian shape of the temperature distribution, the standard deviation is 1.8 K. To allow a more comprehensive approach to quantify uncertainty, we also implement the stochastic simulation of boundary conditions and demonstrate this for the basal specific heat flow in the reservoir of The Hague. As expected, this results in a larger distribution width and hence, a larger, but more realistic uncertainty estimate. However, applying the constraining post-processing the uncertainty is again reduced to the level of the post-processing without stochastic boundary simulation. Thus, constraining post-processing is a suitable tool for reducing uncertainty estimates by observed state variables.

Key words: Probabilistic forecasting; Probability distributions; Heat flow; Hydrothermal systems.

1 INTRODUCTION

An increased use of geothermal energy requires reliable estimates of the risk of failure with respect to the exploration and development of geothermal reservoirs. Suitable geothermal reservoirs need to satisfy certain characteristics for an economic production of geothermal heat. The likeliness to fulfil these requirements defines the risk. For instance, the generation of electrical power with geothermal steam requires a minimum flow rate of 50–100 L s^{-1} and a temperature of 150–200 °C (e.g. Clauser 2006). A wrong estimate of this parameters may result in a failure of the project. Consequently, the entire geothermal system needs to be characterized as precisely as possible. Therefore, numerical simulation of the geothermal system and the processes acting within play a crucial role in all stages of the exploration, development, and operation of any given reservoir. The prediction uncertainty of the numerical simulation and moreover the risk estimate depend on the uncertainties associated different reservoir parameters. These are,

in particular, the subsurface rock properties, such as thermal and hydraulic conductivities, and the state variables in a geothermal reservoir, such as temperature and pressure. As a result, risk of failure and cost may be reduced and estimated with less uncertainty. In this study, we focus on the uncertainty reduction of the state variable temperature, which is a cubical parameter in exploration and exploitation of geothermal reservoirs.

Uncertainties can be quantified based on stochastic approaches provided by various geostatistical algorithms presented by, for example, Caers (2005) and Chiles & Delfiner (1999). Various software tools for geostatistical estimation of rock properties are available (GOCAD User's Manual 1996; Remy 2005). They are based on the widely used geostatistical software library GSLIB (Deutsch & Journel 1998). These tools yield spatial distributions of rock properties. State variables must be predicted by a corresponding heat transport and fluid flow simulation (e.g. Eclipse-100, Technical Description Manual 2001; Clauser 2003; Xu *et al.* 2004). Finsterle & Kowalsky (2007) describe an approach which allows stochastic simulation and Bayesian inverse modelling (Tarantola 2004) combined in one tool. Geostatistical algorithms were already applied successfully in research in hydrogeology (Kitanidis 1997; Cooley 2004; Nowak 2005), repositories for nuclear waste (Neuman & Wierenga 2002), and hydrocarbon reservoir characterization (Campos 2002; Li *et al.* 2008). Uncertainty estimation based on Monte Carlo techniques are described by Robert & Casella (2004). Monte Carlo simulation for solving geophysical inverse problems or predicting hydrocarbon reservoir performance are reported by Sambridge & Mosegaard (2002) and Abhulimen & Otubu (2007), respectively. In geothermal reservoir characterization however, geostatistical algorithms have rarely been used although they are promising for uncertainty quantification and hence risk estimation.

The stochastic approach elaborated in this paper—a Monte Carlo study—is explained in Section 2.1 and tested in Section 3. As part of this approach, multiple parameter realizations of the same geometrical reservoir model are generated. Each realization corresponds equally likely to the real situation defined by data. The reservoir state variables are predicted from simulations of fluid flow and heat transport for each realization. This does not yield only average values and error estimates of a target rock property or state variable at any location in the geothermal reservoir, but also their local probability distribution. Both, the geostatistical and the flow simulation are performed within one software tool. This comprises also parallelization to distribute a large number of realizations over several processor cores. Additionally, the approach allows the stochastic simulation of boundary conditions.

As a rule, geothermal reservoirs are usually explored only by few boreholes, as compared to hydrocarbon reservoirs. Therefore, additional information is used here in order to constrain the stochastic results and hence, to minimize the uncertainty further. By a constraining post-processing of the results as presented in Section 3.2, information from state variable observations is added. The method presented in this paper compares state variables obtained from forward transport simulations, such as temperature or tracer concentration, to available data such as bottom-hole temperatures (BHT). By this comparison, realizations corresponding to an unsatisfactory data fit are discarded.

In order to demonstrate our stochastic approach and its advantages, an exploration scenario is simulated (Section 4) for a current geothermal district heating project in The Hague, Netherlands. In this steady-state simulation, we focus on thermal rock properties. In addition, the stochastic modelling of boundary conditions is also applied to this field study (Section 5).

The overall modelling sequence comprises a conditioning of the realization ensemble to rock properties (geostatistical algorithms) and state variables (constraining method). Based on this information, the risk within a geothermal project can be estimated more accurately.

2 METHOD AND NUMERICAL SIMULATION

By computing fluid flow and heat transport, state variables of a reservoir can be obtained from the spatial distribution of subsurface rock properties and boundary conditions, such as basal specific heat flow. Our stochastic approach—a Monte Carlo study—is based on modifying these rock properties. Hence, the modelling sequence is subdivided into different parts: (i) the stochastic part, Sequential Gaussian Simulation; (ii) the flow simulation by forward modelling and (iii) a constraining post-processing.

2.1 Sequential Gaussian simulation

For a set of chosen geological units, ensembles of realizations for a target rock property are generated using the stochastic algorithm Sequential Gaussian Simulation *Sgsim* from the Geostatistical Software Library GSLIB (Deutsch & Journel 1998). *Sgsim* uses a *Kriging* interpolation technique for spatially distributed data (e.g. Deutsch & Journel 1998). *Kriging* stands for a family of least-square regression methods yielding optimal estimates of the target parameters. The estimator $\hat{z}_0$ for a parameter at a certain point x_0 is written as linear combination of the N related values z at the measuring points x_i

$$\hat{z}_0(x_0) = \sum_{i=1}^{N} \lambda_i \, z(x_i). \tag{1}$$

The weights λ_i are calculated by minimizing—in a least-squares sense—the residual r_0 between estimator $\hat{z}_0$ and unknown true value z_0, that is, $r_0 = (\hat{z}_0 - z_0)^2$.

For each node, *Kriging* returns a mean value, the *Kriging mean*, and an average squared error called *Kriging semi-variance*. By approximating the semi-variogram by a weighting of the available data, *Kriging* accounts for the underlying spatial data correlation. The semi-variogram represents a measure for the dissimilarity of spatial data as a function of the distance between pairs of data (Deutsch & Journel 1998).

The *Sgsim* algorithm operates as follows:

(i) The geometry of a geothermal reservoir model is discretized on a specific grid.

(ii) All data is transformed into Gaussian shape with mean equals zero and variance equals one.

(iii) The algorithm follows a random path through the model.

(iv) For each grid node, nearby data and already simulated nodes are used for a *Kriging* of the target property.

(v) A property value drawn randomly from the distribution defined by *Kriging mean* and *variance* is assigned to the node. Consequently, these values take into account (1) assumed or observed property distributions; (2) the correlation length; (3) primary data, such as borehole measurements and (4) secondary data, such as seismic data.

(vi) A realization is completed when property values are assigned to all nodes of the model.

(vii) The model is transformed back from Gaussian into the original space.

(viii) More realizations are created by following other random paths. Each of these realizations is equally likely with respect to the real situation defined by the data.

The *Sgsim* algorithm is integrated as a module into a fluid flow and heat transport simulator (Rath *et al.* 2006). The corresponding equations are described in Section 2.2. This way, the generated realizations can be used directly as input for the fluid flow and heat transport simulations, wasting no time and effort for format conversion. Further, already existing models implemented in the flow simulator can be easily updated with a stochastic simulation.

2.1.1 The integration method

During the construction of the geometrical model for the simulation, groups of cells are combined into units corresponding to geological layers. Each unit is defined by a set of up to 13 rock properties, depending on the particular choice of problem to be solved. When the stochastic module is run for a target property, the corresponding property distribution is simulated with the *Sgsim* algorithm as if the unit would comprise the entire model. This is done in each unit for a specific set of rock properties. The individual cell values—specified rock properties—are assigned according to the generated *Sgsim* model distribution. Thus, cells of a unit no longer have identical properties. As a result, each geological unit in the model is simulated with respect to its measured value distribution. For different rock properties, the stochastic simulation may be performed in a linear or logarithmic mode. For permeability simulation in particular, the logarithmic mode is useful. This property initialization process is repeated for each realization. Then, fluid flow and heat transport are simulated which requires usually the larger part of the computation time.

The grids for stochastic and transport simulation need to be similar to avoid geometric discrepancies. The sums of all cell boundary positions in one dimension are not allowed to differ more than one 1 per cent. However, a grid refinement is possible for the fluid flow and heat transport simulations, for example, around boreholes. Because *Sgsim* works only with equidistant grids, the property values in each node of the coarser grid are assigned to the corresponding node of the finer resolution grid. To avoid upscaling problems, the grid used for stochastic simulation is always the coarser one.

Because all realizations are computed independently, the algorithm lends itself very well for parallelization. Here, a nested parallelization is implemented which means that each individual realization is simulated in parallel by a group of computing threads and, additionally, the different realizations are computed by different groups of threads. As a consequence, the code can be run on several processor cores which allows fast computation, even for large ensembles. Details of the parallelization technique are described in Wolf *et al.* (2008).

The stochastic variation of boundary conditions with *Sgsim* is also implemented in a similar way to the simulation of the rock properties.

2.2 Forward modelling

Fluid flow through a porous medium is commonly described by Darcy's law (Darcy 1856)

$$\boldsymbol{v} = -\frac{\boldsymbol{k}}{\mu_f}(\nabla P + \rho_f g \nabla z)\,, \tag{2}$$

where $\boldsymbol{v}$ is the specific discharge ($m^3\,m^{-2}\,s^{-1}$), $\boldsymbol{k}$ the hydraulic permeability tensor (m^2), μ_f the fluid dynamic viscosity (Pa s), ρ_f fluid density ($kg\,m^{-3}$), g gravity ($m\,s^{-2}$) and P the hydraulic pressure (Pa). Here, the unity vector is pointing upwards.

The equation for fluid flow implemented here is derived from eq. (2) and the equation of continuity using an extended Boussinesq approximation (e.g. Kolditz *et al.* 1998; Diersch & Kolditz 2002)

$$\rho_f\,(\alpha + \phi\beta)\,\frac{\partial P}{\partial t} = \nabla \cdot \left[\frac{\rho_f \boldsymbol{k}}{\mu_f}(\nabla P + \rho_f g \nabla z)\right] + W\,. \tag{3}$$

Here, ϕ is porosity while α and β denote the compressibility (Pa^{-1}) of the rock and the fluid phase, respectively. W corresponds to a mass source term ($kg\,m^{-3}\,s^{-1}$).

The heat transport equation follows in a way analogue from conservation of energy (e.g. Beardsmore & Cull 2001)

$$(\rho c)_e \frac{\partial T}{\partial t} = \nabla \cdot [\boldsymbol{\lambda}_e \nabla T] + (\rho c)_f \boldsymbol{v} \cdot \nabla T + H\,. \tag{4}$$

In this equation, T is the temperature (°C), $(\rho c)_e$ is the effective volumetric heat capacity of the saturated porous medium and the fluid ($J\,m^{-3}\,K^{-1}$), $(\rho\ c)_f$ is the thermal capacity of the fluid ($J\,m^{-3}\,K^{-1}$), $\boldsymbol{\lambda}_e$ is the tensor of effective thermal conductivity ($W\,m^{-1}\,K^{-1}$) and H a heat generation rate source term ($W\,m^{-3}$). Effective values are appropriate averages defined for the fluid-rock system, in particular the arithmetic mean for thermal capacity, and the geometric one for thermal conductivity. In the following, both, thermal conductivity and hydraulic permeability are assumed to be isotropic, that is, they can be treated as scalars.

2.3 Constraining method

State variables are used to find appropriate realizations of an ensemble characterized by different spatial distributions of rock properties. These variables are obtained from the coupled modelling with respect to stochastic simulation of rock property distributions and forward simulation of the fluid flow and heat transport. Chen & Zhang (2006) reported (for an another stochastic technique: ensemble Kalman filtering) that state variable data is more effective for estimating formation properties and transport characteristics than rock property data from limited locations. The details of our approach will be explained and discussed in the following Section 3. Realizations which do not agree with observed state variables, for example, temperature, at certain locations in the model sufficiently well are discarded from the ensemble.

3 TESTING THE METHOD ON A SYNTHETIC MODEL

To test the stochastic approach, a synthetic model is created and later a constraining method applied to select the realizations which fit best the temperature data. Results are compared using *Sgsim* with varying number of wells at the one hand and unconditioned *Sgsim* with constraining post-processing at the other hand.

A small model is generated comprising 12 × 12 × 12 cells and four geological units. This corresponds to a total model size of 1.2 km × 1.2 km × 1.2 km. Using *Sgsim*, a different Gaussian distribution of the thermal conductivity λ is simulated in each unit. In one central permeable layer, a bi-modal distribution of permeability k is modelled with values between 10^{-16} and 10^{-12} m^2. A flow corresponding to a hydraulic head difference of 30 m passes through the model in y-direction. This high value is chosen to emphasize the effects of the hydraulic properties on the temperature field. A

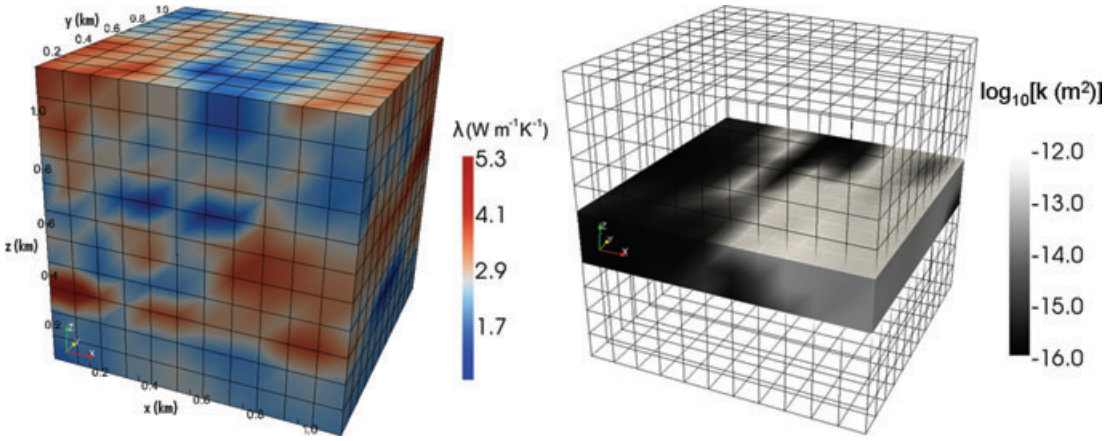

Figure 1. The reference model: Single realization of thermal conductivity λ (left-hand panel) and permeability k (right-hand panel) generated with the *Sgsim* algorithm. The wire frame is used to emphasize the permeable layer, which top is located at a depth of 600 m with a formation thickness of 360 m.

transient fluid flow and heat transport simulation is performed for a period of 10 yr. This realization is called the reference model and is shown in Fig. 1.

3.1 Model reconstruction with varying well number

Considering boundary conditions, probability distributions, and correlation lengths to be well known, the *Sgsim* algorithm is run to reproduce the reference model. The simulation is conditioned by data: thermal conductivity and permeability from a varying number of synthetic wells inside the reference model. The positions of the wells are randomly chosen. The realizations using the same number of wells form a group. Each group comprises three realizations. Target parameters for this test are thermal conductivity and permeability as well as temperature. The latter is of special interest because it is a state variable which can be measured directly. Additionally, the temperature is a parameter of major significance in the field of geothermal energy.

Fig. 2 illustrates the differences in the temperature fields ΔT between the reference model and simulated realizations from a group using data from three wells. It shows a horizontal cross-section in the centre of the model. The three patterns are different, but indicate a similar level of uncertainty. In total, the simulation runs yield comparable patterns within each group and very different ones between the groups. Of course, the differences decrease with increasing number of used wells. Interestingly, some simulated realizations, in this case model (a), reflect the reference model more successful, although they are based on a relative small number of wells.

Typical realizations for a different number of wells are shown in Fig. 3. As expected, the quality of the fit to the reference model increases with the number of wells. Interestingly, simulation using only a small amount of wells (Figs 3b and c) does not yield better results than an unconditioned simulation (Fig. 3a). Therefore, an unconditioned simulation appears preferable in the case of only few available wells or a low exploration borehole density. This holds particularly for projects for geothermal energy use as mentioned in Section 1. However, some successful members of the ensemble such as the one shown in Fig. 3(d) fit the reference distribution quite well, even though they are constrained by relatively few wells. The spatial permeability distribution is supposed to be the major parameter for the shape and quantity of the temperature field because in operated reservoirs fluid flow affects temperature stronger than conductive heat flow. Thus, different realizations feature large ranges of temperature differences ΔT, varying from less than 1–20 K. Considering the absolute temperature mean in the analysed layer of the reference model of 30 °C, the uncertainty in the temperature prediction is as very large. Successful realizations, in terms of a successful fit of the temperature in the reference model, result mostly from a successful fit of the corresponding permeability distribution.

3.2 Constraining method

A simple constraining technique is proposed for identifying the best fits of the reference model and discarding all others. To this end, a larger ensemble of 50 realizations is generated using unconditioned *Sgsim*. A large ensemble size is essential when applying constraining methods to obtain a sufficient number of successful members. Unfortunately, generating large ensembles is at the price of large computing time. However, for the small synthetic model in this case, 50 realizations can be generated in short time on a single core. Again, information is used from three synthetic wells ($i = 1, 2, 3$) inside the reference model. One single temperature value $T^i{}_{\mathrm{ref}}$ from the permeable layer is recorded for each well instead of the rock properties. This corresponds to measured BHT in boreholes. These temperatures are compared to the corresponding temperatures of a simulated realization T^i_{sim}. This yields three fitting parameters c_i

$$c_i = \frac{|T^i_{\mathrm{sim}} - T^i_{\mathrm{ref}}|}{T^i_{\mathrm{ref}}}. \tag{5}$$

To select the three best-fitting members of the ensemble, realizations are discarded which yield one of the fitting parameters above an appropriate threshold. This technique is repeated for each realization. The temperature difference ΔT of the three residual realizations is shown in Fig. 4 to access the quality of the fit. The figure should be considered in comparison with Fig. 2, which illustrates the results using three wells for conditioning *Sgsim* without constraining post-processing.

Even though the *Sgsim* algorithm is run in unconditioned mode with respect to thermal conductivity and permeability, constraining post-processing based on temperature data reduces the maximum misfit between simulation and reference model from 20 K to 0.8 K. Obviously, constraining post-processing is a major improvement in uncertainty minimization. Therefore, it is performed

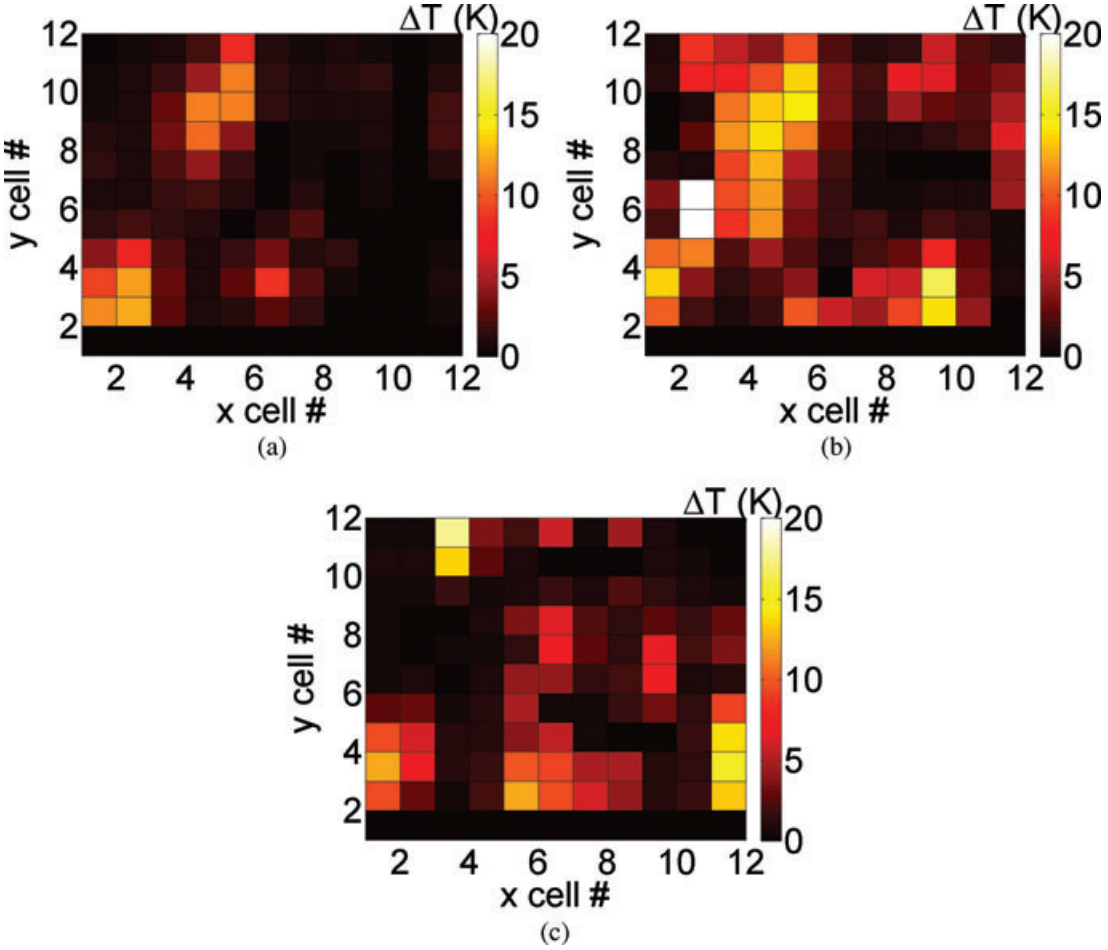

Figure 2. Temperature differences ΔT between reference model and three simulated realizations (a–c) from the same set using data from three wells. The mean temperature in the displayed layer is 30 °C. The flow source corresponds to the origin of the y-axis.

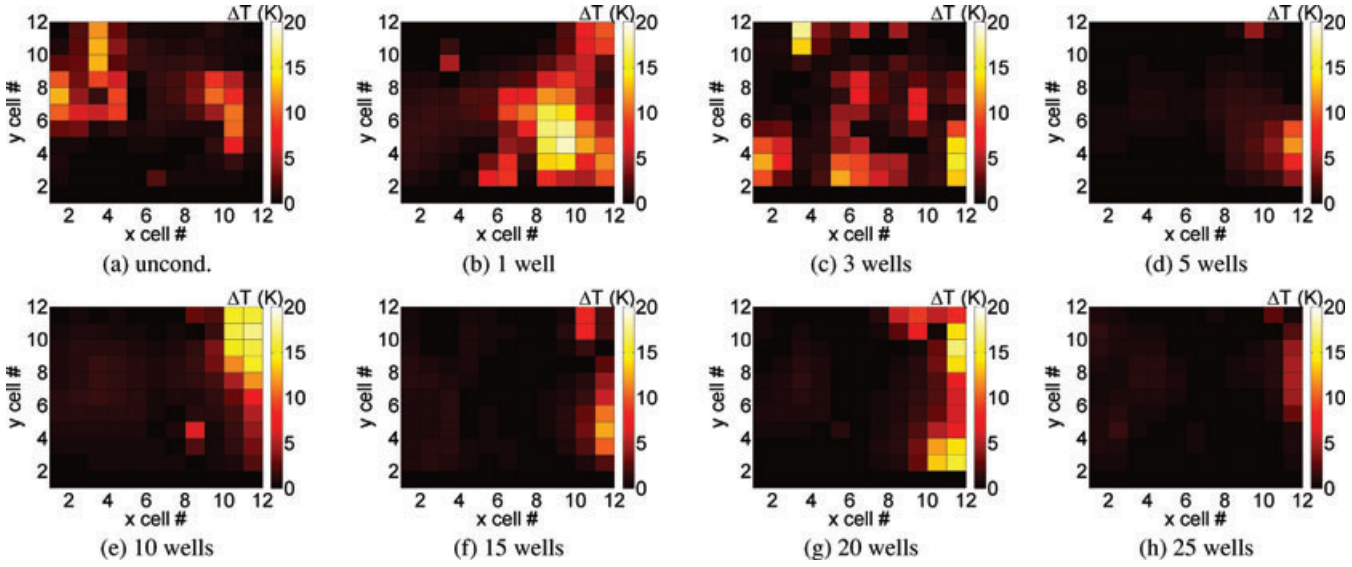

Figure 3. Temperature differences ΔT between reference model and one simulated realizations from different sets: (a) unconditioned and (b)–(h) conditioned to an increasing number of wells. The flow source corresponds to the origin of the y-axis.

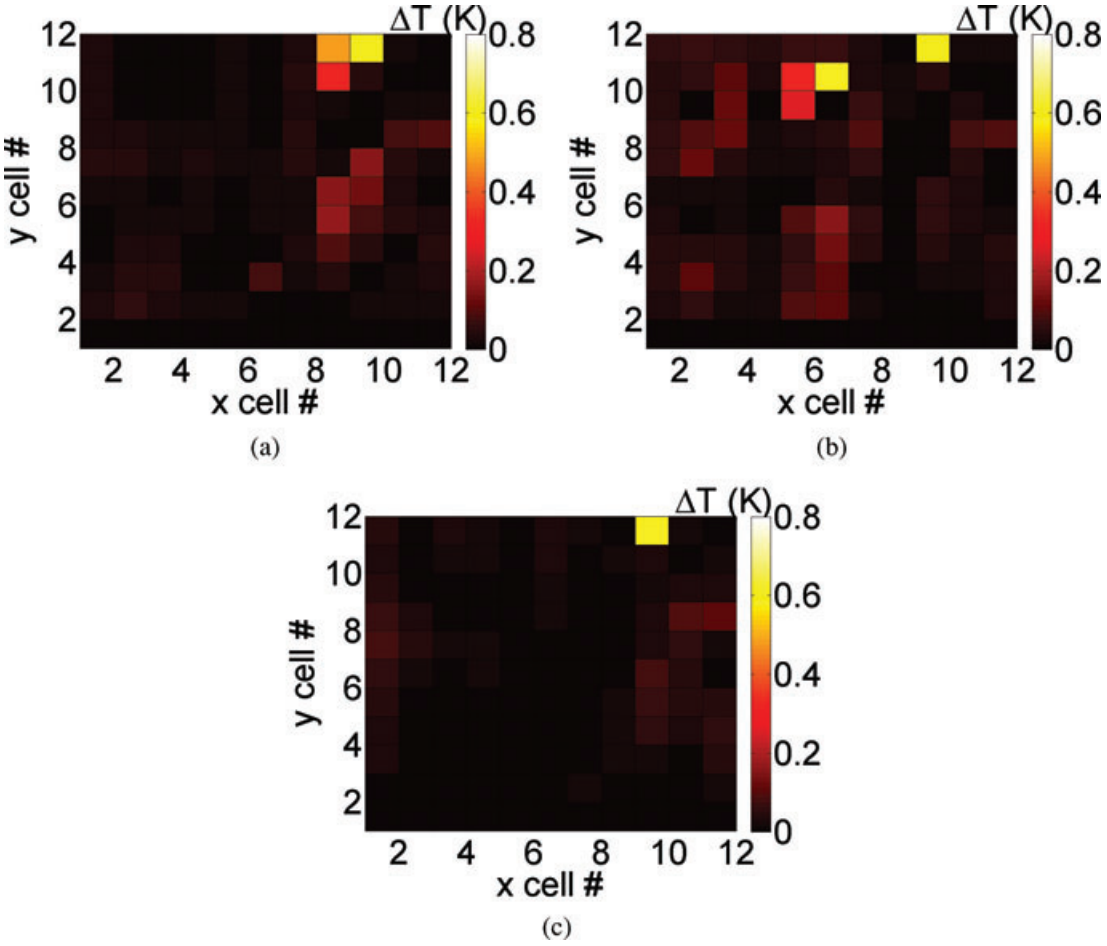

Figure 4. Temperature differences ΔT between reference model and residual simulated realizations after performing constraining post-processing. The flow source corresponds to the origin of the *y*-axis.

during stochastic modelling of real geothermal reservoirs, as described in the following Section 4. As stated above, in the case of exploited reservoirs, hydraulic properties affect the temperature field stronger than thermal properties and their spatial distribution. Consequently, constraining techniques are particularly effective in selecting realizations after stochastic simulations of hydraulic properties in systems where fluid flow affects significantly the temperature.

4 FIELD STUDY FOR A GEOTHERMAL RESERVOIR

A geothermal district heating project in The Hague, the Netherlands is chosen to demonstrate the stochastic modelling sequence as outlined above. In this study, the temperature in the target area is predicted and its uncertainty is minimized.

A geothermal doublet will be installed at a depth of 2300 m to supply up to 6000 households with geothermal heat. The Delft sandstone (van Balen *et al.* 2000) at the transition between Jurassic and Cretaceous formations is identified as target reservoir. Its permeability varies between 10^{-12} and 10^{-13} m^2. Technical requirements for formation temperature and flow rate are 75 °C and 42 L s^{-1}, respectively. Model parameters such as reservoir geometry are reported by Simmelink & Vandeweijer (2008). A map of the studied area is shown in Fig. 5.

Rock properties such as porosity, thermal conductivity and radiogenetic heat production are calculated by us from laboratory measurements on cuttings and logging data from a calibration borehole. To that end, the shale content and the free water content are computed for the key wells. In turn these data are used to compute continuous thermal conductivity profiles (Hartmann *et al.* 2005, 2007). The log derived thermal conductivity profiles were calibrated with the laboratory measurements (Fig. 6). Using these profiles, statistical values of effective thermal conductivity are calculated for the stratigraphic units incorporated in a 3-D model. In addition, radiogenetic heat production profiles were produced by the use of the gamma-ray logs (Rybach 1988).

Consistent with available geological information (Simmelink & Vandeweijer 2008), there is neither significant thermal free convection nor topological driven fluid flow. Thus, conductive heat flow is dominant. The temperature prediction is obtained via steady-state simulation of heat transport. For comparison of model and reality, ten corrected BHT are available distributed evenly in the study area. The modelling process aims for a temperature prediction at the proposed target location near the centre of the model. This target location may change during the further planning progress of the project.

The 3-D geometric model (Pechnig *et al.* 2008; Mottaghy *et al.* 2009) represents a volume of 22.5 km × 24.3 km × 5 km. It is

Figure 5. Map of the studied area containing the proposed target location and borehole labels, and the dimensions of our model.

discretized into 150 × 162 × 100 grid cells. It comprises nine geological layers identified and implemented as separate units. Each unit is characterized by a different saturated thermal conductivity and heat production rate. The mean and the standard deviation of the thermal conductivity in each layer characterizes a unit.

A first estimate for the possible range in the predicted temperature is obtained by characterizing each unit by its mean and standard deviation of the thermal conductivity according to the geometric modelling process. In particular, a temperature range is obtained by using on the one hand always minimum values (mean minus standard deviation) and on the other hand always maximum values (mean plus standard deviation) of the thermal conductivity in each layer for the modelling.

The result of the prediction at the locations of the BHT data is illustrated in Fig. 7. Obviously, the mean reproduces every single BHT successfully, but the uncertainty is over-estimated.

4.1 Quantifying and reducing uncertainty

The stochastic modelling sequence explained in Section 3 is applied to provide better uncertainty estimates. Instead of considering only minimum and maximum values defined by mean and standard deviation, the original probability distribution of the saturated thermal conductivity in the calibration borehole is simulated in six of the nine geological units using the *Sgsim* algorithm. As an example for the different input histograms to be reproduced using *Sgsim*, the measured distributions of Layer 3 (Lower Cretaceous Supergroup) and Layer 4 (Jurassic Supergroup) are illustrated in Fig. 8. The original distributions were calculated by us using different logs, as stated above (Fig. 6).

Although more than 50 exploration boreholes exist in the studied area, there are only BHT measurements. Calculated thermal conductivity data was available for just one calibration borehole with high vertical data density in this study. With this single borehole in an area of about 550 km^2, the borehole density is low, so that unconditioned *Sgsim* has to be performed. This can be done without restricting the results as shown before in Section 3.1. However, the logs in the other boreholes indicate similar lithoglocial conditions within the stratigraphic layers.

Correlation lengths in the vertical direction are calculated using semi-variograms (see Section 2.1). The data required for the semivariograms are obtained from the calibration borehole. The semivariograms yield correlation lengths between 50 and 450 m for the individual geological layers.

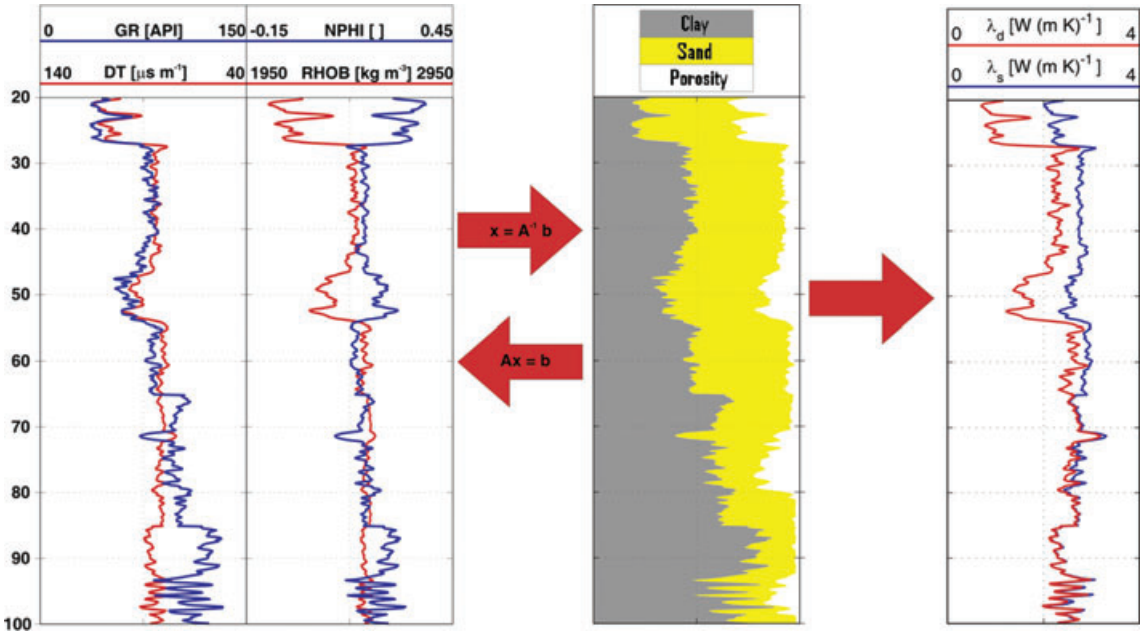

Figure 6. Scheme for calculating dry and saturated thermal conductivity (λ_d and λ_s) profiles from logs. The correlation of shale contend and log response is used to calculate the volumetric fractions of the rock components from log data. Knowing these volumetric fractions and the thermal conductivity of the single rock components from laboratory measurements, the thermal conductivity in the subsurface can be determined using an appropriate mixing law. (See Hartmann *et al.* 2005, 2007, for more details).

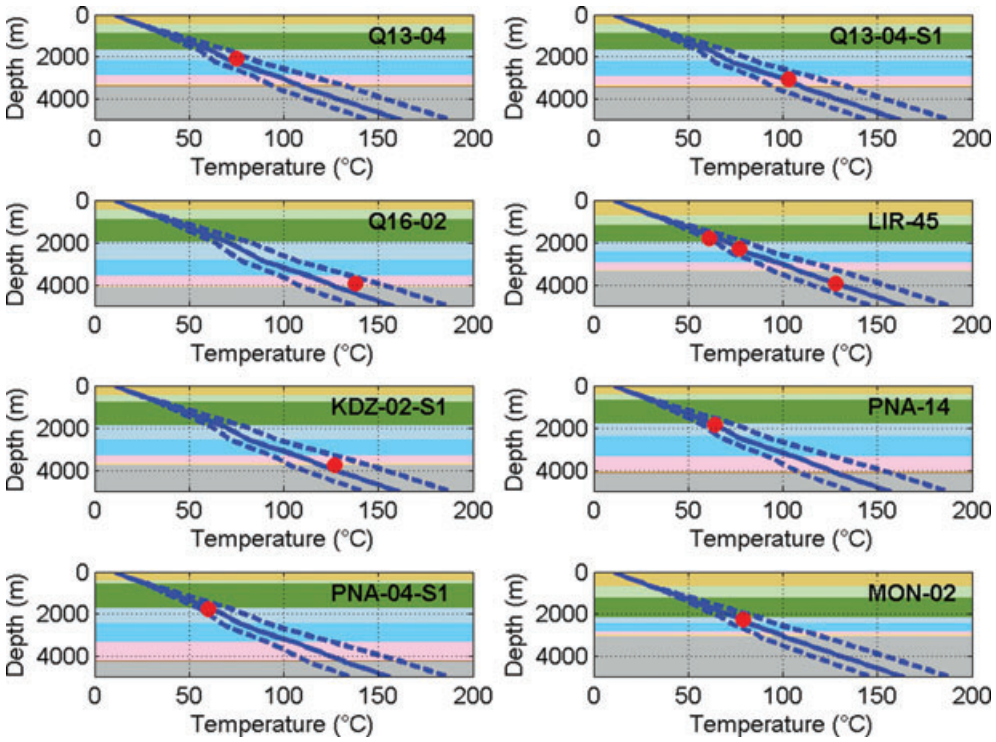

Figure 7. Temperature prediction (blue line) at the positions of the BHT and BHT values (red dots) as well as uncertainty estimate (blue dashed lines). The labels refer to boreholes showed in Fig. 5.

Due to the nature of sedimentation there is an anisotropy between vertical and horizontal correlation lengths (Caers 2005). Vertical correlation lengths are usually much shorter than horizontal ones. For the horizontal directions, the borehole distribution with just one borehole is insufficient to calculate reliable semi-variograms at all. Fortunately, the horizontal direction is less important for predominantly vertical conductive heat flow. However, to obtain a maximum fluctuation in the ensemble of realizations, large correlation lengths are implemented in the horizontal directions by multiplying the corresponding vertical correlation length in each unit. Deutsch & Journel (1998) use a maximum factor of 16 between vertical and horizontal correlation length which is adopted in this study, too.

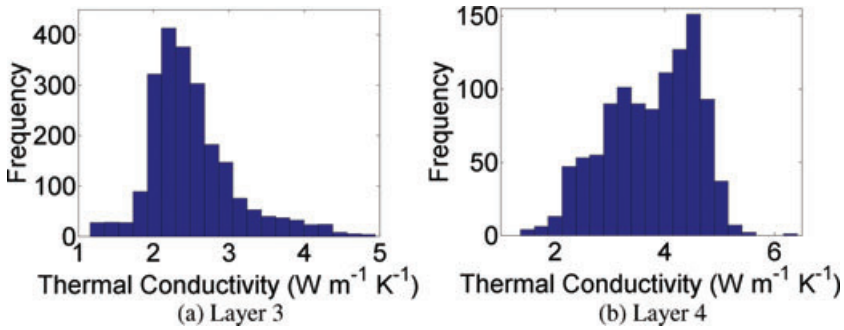

Figure 8. Original distributions of the thermal conductivity in the Lower Cretaceous Supergroup (a) and the Jurassic Supergroup (b) obtained from the logging data. The average thickness of these layers is approximately 1100 and 550 m, respectively.

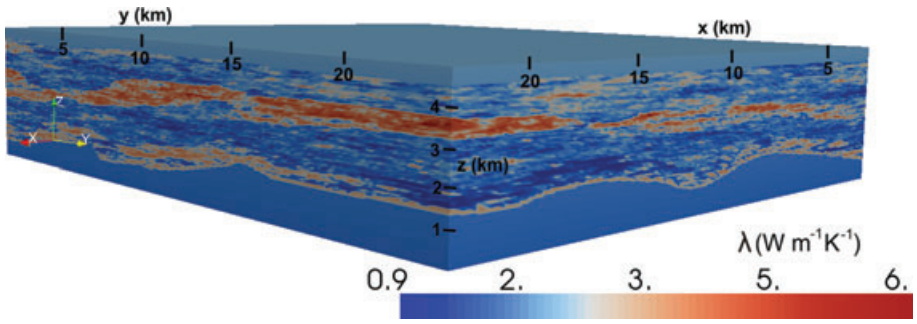

Figure 9. One realization of the thermal conductivity in the model.

This yields horizontal correlation lengths between 800 and 7200 m. However, these values can be considered as more or less arbitrary due to the dominantly vertical direction of heat flow. Nevertheless, simulating fluid flow in the reservoir, e.g. for performance prediction during the operation of the doublet, will require reasonable values in the future.

Using these correlation lengths, 1000 temperature fields are simulated based on 1000 realizations of thermal conductivity in the model. One example is shown in Fig. 9. The ensemble is simulated in parallel on two quadcore processors with a clock rate of 3 GHz each. Thus, a computing time of 127 hr is needed to generate the full ensemble of 1000 realizations. Constraining post-processing is performed in order to reduce the uncertainty further discarding unsuitable ensemble realizations.

4.2 Constraining post-processing

Constraining methods are useful to further reduce the uncertainty significantly. Therefore, we perform the constraining post-processing in this case, too, although we suppose constraining techniques to be less effective in systems lacking significant fluid flow, as stated above in Section 3.2. In this post-processing, the temperatures of all 1000 realizations of the ensemble are compared to the measured BHT. Several studies about the reliability of corrected BHT data and its error are available (e.g. Deming 1989; Hermanrud *et al.* 1990; Förster 2001). Förster (2001) and Hermanrud *et al.* (1990) reported that BHT values corrected with less advanced methods [such as the Horner plot method (Horner 1951)] underestimate the formation temperature by 8 K ± 8 K in their studies. This underestimation is not confirmed by our simulations (Fig. 7). However, according to Hermanrud *et al.* (1990) advanced methods for correcting BHT usually give values close to the formation temperature with a standard deviations of ±9 K. Therefore, this standard deviation of ±9 K is used as a constraining parameter for the post-processing. A realization is discarded from the ensemble, if the difference between the temperature simulated $T^i{}_{\rm sim}$ at the position of the BHT measurement and the corrected BHT value BHT^i exceeds ±9 K for at least one BHT $i = 1, .., 10$)

$$|T^i_{\rm sim} - BHT^i| > 9\,K \qquad \text{discard.} \tag{6}$$

4.3 Results

The probability distribution of the temperature at the proposed target location is shown for 1000 original realizations in Fig. 10(a) and for the 488 realizations which survived the constraining post-processing in Fig. 10(b). The full distribution width of the temperature is reduced from 16 K to 12 K as a consequence of the post-processing. Assuming a Gaussian distribution shape, the standard deviation σ is reduces slightly from 2.0 K to 1.8 K. In particular, realizations with very high and low temperature predictions are removed from the ensemble. Thus, too optimistic and pessimistic assumptions suggested by the stochastic modelling sequence without constraining the ensemble are avoided. The distribution mean $\mu = 79\,^{\circ}$C stays constant at the proposed target position.

As in Fig. 7, the result of the temperature predictions at the locations of the BHT is illustrated in Fig. 11(a) for the 488 realizations. The BHT are reproduced with an appropriate uncertainty.

In total, the final temperature prediction obtained from the 488 surviving realizations is illustrated in Fig. 13(a) in comparison with the initial uncertainty. Obviously, the uncertainty has been minimized significantly by applying the stochastic modelling sequence. The full distribution width is reduced from 25 K to 12 K, making the temperature prediction much more accurate. In addition, a standard deviation of $\sigma = 1.8$ K is can be stated as another

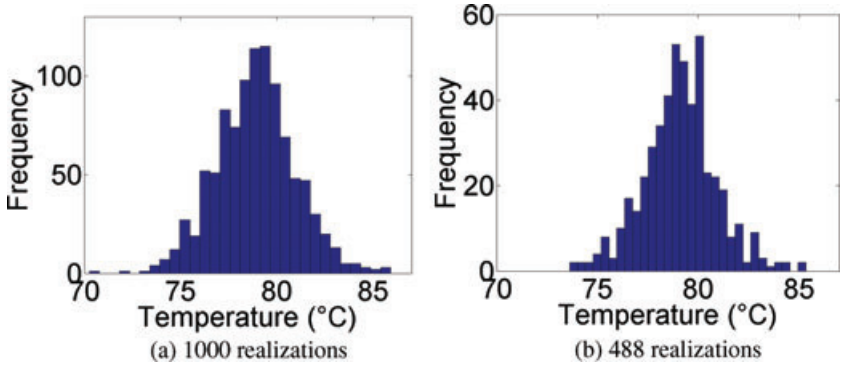

Figure 10. Temperature at proposed target location (a) for the original ensemble and (b) for the ensemble after constraining post-processing. In (b), the uncertainty is further reduced. A minimum temperature of 77 °C is predicted.

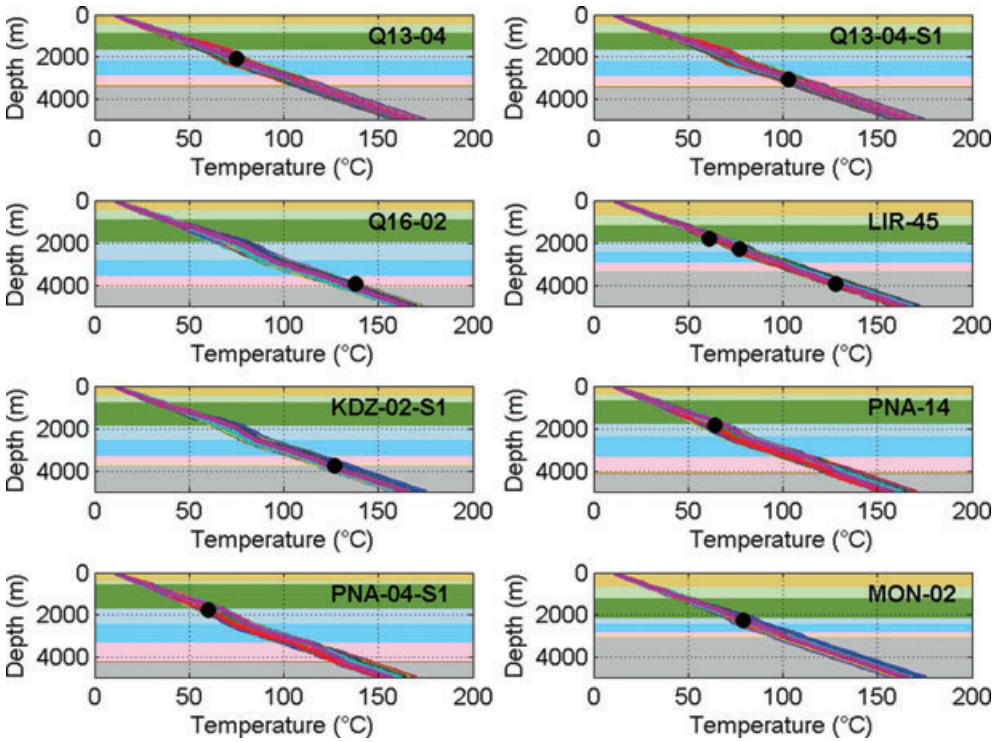

Figure 11. Temperature prediction (coloured lines) at the positions of the BHT and BHT values (black dots) after applying constraining post-processing in comparison with the original uncertainty estimate (blue dashed lines).

uncertainty measure. As further positive result in terms of project feasibility, according to the uncertainty quantification applied in this study, the likelihood for the formation temperature exceeding the required temperature of 75 °C is 97 per cent (without constraining post-processing), respectively, 99 per cent with constraining post-processing.

5 ADDITIONAL STOCHASTIC SIMULATION OF BASAL SPECIFIC HEAT FLOW

Taking advantage of the possibility for stochastic modelling of boundary conditions, an additional stochastic simulation of the basal specific heat flow is integrated in the modelling sequence. The objective is to consider this major factor of uncertainty in the modelling process. The mean of $q = 63\,\mathrm{mW\,m^{-2}}$ is varied with a standard deviation of $\pm 3\,\mathrm{mW\,m^{-2}}$. This small fluctuation of the basal specific heat flow is chosen due to the good agreement of the final corrected BHT with the predicted temperatures obtained from the simulated realizations illustrated in Fig. 7. Moreover, the need of a statistically relevant amount of realizations surviving the constraining selection demands the small but realistic value of $\pm 3\,\mathrm{mW\,m^{-2}}$.

The simulated value is assigned to each grid node at the bottom boundary of the model. Again, an ensemble of 1000 realizations is generated. Now, with an expected larger distribution width due to basal specific heat flow fluctuation, constraining post-processing is essential to identify reliable realizations.

5.1 Results

The results of the complete stochastic modelling sequence including boundary simulation are illustrated in Fig. 12 and Fig. 13(b). As expected, the full distribution width without constraining post-processing increases significantly from 16 K to 24 K at target depth

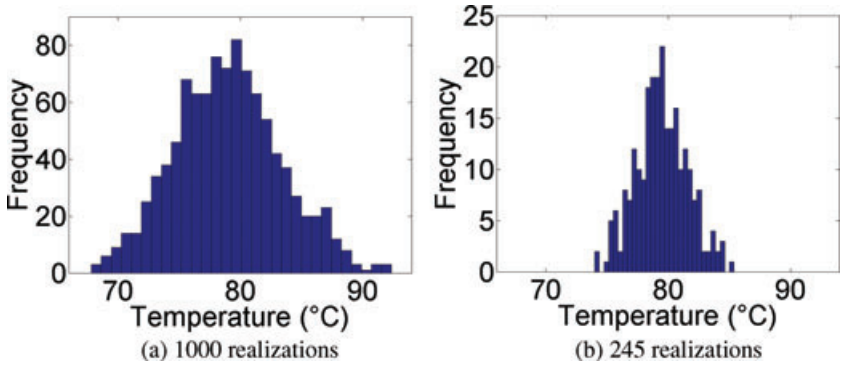

Figure 12. Temperature at proposed target location (a) for the original ensemble and (b) for the ensemble after constraining post-processing including a stochastically simulated basal specific heat flow.

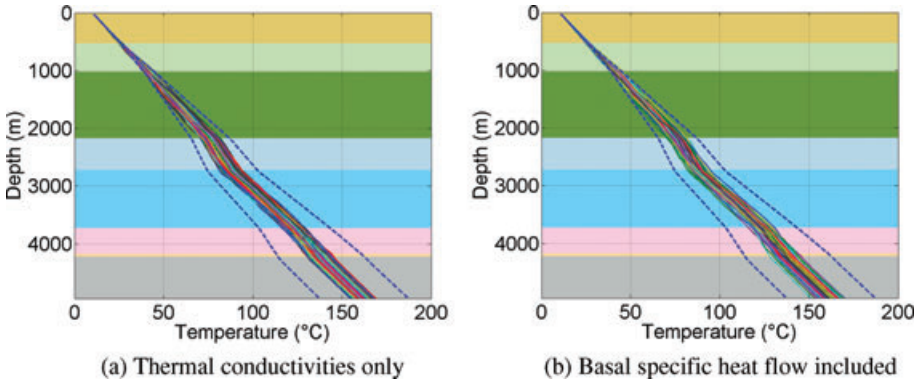

Figure 13. Temperature prediction (coloured lines) at proposed target location in comparison to the original uncertainty estimate (blue dashed lines). In (b), basal specific heat flow is additionally simulated stochastically. Note that the original uncertainty estimate is based on an average value for the basal specific heat flow.

compared to the distribution discussed in Section 4.1. The mean of 79 °C does not change, but the standard deviation increases to 4.3 K.

However, performing the constraining post-processing yields results very similar to the results in Section 4.3 obtained without specific heat flow variation. The standard deviation of the distribution is $\sigma = 2.1$ K. The full distribution width of about 11 K after performing the constraining technique is actually lower than for a constant basal specific heat flow. This is due to the smaller size of the statistical sample. The fraction of realizations, which survived the constraining post-processing decreases from 488/1000 to 245/1000 because of the larger distribution width. Our results are summarized in Table 1.

The uncertainty of the basal specific heat flow before performing constraining post-processing yields a lower likelihood for the formation temperature exceeding the required temperature of 75 °C. Here, this likelihood is 80 per cent without constraining post-processing. However interestingly, this likelihood is increased to 99 per cent again by applying the constraining post-processing, even when the basal specific heat flow is varied.

A variable basal specific heat flow is not considered for the original uncertainty estimate without any stochastic approach. Nevertheless, the temperature profiles of all realizations fall into the estimated original range. In this particular case, constraining post-processing reduces significantly the large effects on the uncertainty estimate by simulating basal specific heat flow. Thus, uncertainty estimates obtained with and without stochastic basal specific heat flow simulation yield similar results. Even though the system is not dominated by fluid flow and thus no hydraulic properties are simulated, the Monte Carlo approach combined with constraining post-processing proves very efficient in this case of stochastic simulation of thermal rock properties and boundary conditions.

Table 1. Results of the stochastic modelling sequence: temperature mean μ, standard deviation σ and full distribution width.

Method	μ (°C)	σ (K)	Full width (K)
Original estimate	77	–	25
Before post-processing without basal specific heat flow variation	79	2.0	12
After post-processing without basal specific heat flow variation	79	1.8	16
Before post-processing with basal specific heat flow variation	79	4.3	24
After post-processing with basal specific heat flow variation	79	2.1	11

6 CONCLUSIONS AND OUTLOOK

Geothermal reservoir exploration and modelling requires sophisticated quantification of uncertainty. The Monte Carlo approach of Sequential Gaussian Simulation combined with constraining post-processing helps to quantify and reduce significantly the uncertainty in the prediction of state variables in geothermal reservoirs. However, there is a price to pay in terms of increased computing power. Multiple realizations are generated conditioned to measured rock properties and equally likely to reflect reality. Moreover, the realizations are conditioned to observed state variables via constraining post-processing. Thus, uncertainty is quantified stochastically much more precisely than basing merely on a minimum–maximum estimate (value range) of the involved rock properties. Hence, it is very effective in reducing uncertainties and therefore eventually reducing risk of failure and cost. In addition, the stochastic simulation of boundary conditions allows a more comprehensive approach to quantify uncertainty in geothermal reservoirs.

Constraining techniques are particularly effective in filtering realizations after stochastic simulation of hydraulic properties in systems where advective heat transport is significant. This has been demonstrated by us in an other field study. To that end, stochastic permeability modelling was applied for transient simulations of the operated geothermal doublet in The Hague. In particular, the temperature development at the target position and the propagation of the cold water front from the injection point is predicted. The possibility to use different grid sizes for the stochastic and the flow simulation was used here. The results will be published soon. Furthermore, as soon as first measurements of the state variables become available after drilling, a constraining approach in a fluid flow dominated temperature field will be performed.

Unfortunately, there is one drawback applying the introduced constraining technique when discussing it in a probability framework. Due to the low number of realizations used in this study, only a raw approximation of the *a posteriori* probability distribution is obtained. Nevertheless, we the results thus obtained to be useful and justifiable. This is supported by the reasonable peaked shapes of the histograms in Fig. 10. In the study presented here, prior knowledge is comprehensive. Therefore, we sample from a well constrained prior model, making use of rather narrow statistical assumptions. This partly explains that we are able to produce useful results even with a relatively small number of samples.

Methods such as Markov Chain Monte Carlo sampling (MCMC) (e.g. Fox & Nicholls 2001) could have been used quantifying the uncertainty more accurately. However, methods of this type are computational expensive due to the large model size (2.4 millions of cells), because they need large numbers of forward simulations. This situation is further aggravated when fluid flow has to be taken into account, and models have to be time dependent. Therefore, more effective sampling strategies will be applied in the future for data assimilation. These include variants of the ensemble Kalman filter (EnKF) (Evensen 2003; Chen & Zhang 2006; Gu & Oliver 2006), or the Kalman ensemble generator (KEG) Nowak (2009).

This study represents the initial step of a strategy for iterative improvement of an ensemble of information forming an geothermal reservoir model. This ensemble will be updated as different kinds of observations, such as logging data, well tests, or even tracer experiments may become available. This iterative process will be supplemented by sophisticated methods such as the ensemble Kalman filter and supported by virtual reality visualization techniques (Wieskopf & Erlebacher 2004; Wolter *et al.* 2009).

ACKNOWLEDGMENTS

Our research was funded by the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU), grant FKZ 0327563, and by the German Ministry of Education and Science (BMBF), grant 03SF0326A. VR was funded by the Ramón y Cajal Program of the Spanish Ministry for Science and Education. We thank IF Technology (Arnheim, the Netherlands) and TNO Built Environment and Geosciences (Delft, the Netherlands) for providing data and granting permission to publish results, as well as E.ON Benelux (Rotterdam, the Netherlands) for initiating and funding the districted heating project in The Hague, from which these data are derived. Comments from two anonymous reviewers were gratefully appreciated.

REFERENCES

Abhulimen, K.E. & Otubu, O.A., 2007. Predicting Reservoir Performance Using Stochastic Monte Carlo Simulation, *E-J. Reservoir Eng.,* **1**(1), http://petroleumjournalsonline.com/journals/index.php/reservoir/.

Beardsmore, G.R. & Cull, J.P., 2001. *Crustal Heat Flow,* Cambridge University Press, New York.

Caers, J., 2005. *Petroleum Geostatistics,* Society of Petroleum Engineers, Richardson TX.

Campos, O.M., 2002. Spatial distribution of calcite concretions inferred from borehole image data and their use in reservoir modeling: an application to Breitbrunn field, *Master's thesis*, The University of Texas, Austin TX.

Chen, Y. & Zhang, D., 2006. Data assimilation for transient flow in geologic formations via ensemble Kalman filter, *Adv. Water Resour.,* **29,** 1107–1122.

Chiles, J.P. & Delfiner, P., 1999. *Geostatistics: Modeling Spatial Uncertainty,* Wiley & Sons, New York.

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 *Landolt-Börnstein, Group VIII. Advanced Material and Technologies,* Vol. 3: Energy Technologies, Subvol. C: Renewable Energies, pp. 480–595, ed. Heinloth, K., Springer Verlag, Heidelberg-Berlin.

Cooley, R.L., 2004. A theory for modeling ground-water flow in heterogeneous media, Professional Paper 1679, U.S. Geological Survey (USGS), Reston VA.

Darcy, H., 1856. *Les fontaines publiques de la ville de Dijon,* Dalmont, Paris.

Deming, D., 1989. Application of Botton-hole temperature corrections in geothermal studies, *Geothermics,* **18**(5/6), 775–786.

Deutsch, C.V. & Journel, A.G., 1998. *GSLIB. Geostatistical Software Library and User's Guide,* Oxford University Press, New York.

Diersch, H.J.-G. & Kolditz, O., 2002. Variable-density flow and transport in porous media: approaches and challenges, *Adv. Water Resour.,* **25,** 899–944.

Eclipse-100, Technical Description Manual, 2001. Schlumberger GeoQuest, Houstin TX.

Evensen, G., 2003. The Ensemble Kalman Filter: theoretical formulation and practical implementaion, *Ocean Dyn.,* **53,** 343–367.

Finsterle, S. & Kowalsky, M.B., 2007. *iTOUGH2-GSLIB User's Guide,* Earth Sciences Division, Lawrence Berkeley National Laboratory, University of California, Berkeley CA, http://www-esd.lbl.gov/TOUGH+/manuals/iTOUGH2_GSLIB_Users_Guide.pdf.

Förster, A., 2001. Analysis of borehole temperature data in the Northeast German Basin: continuous logs versus bottom-hole temperatures, *Petrol. Geosci.,* **7**(3), 241–254.

Fox, C. & Nicholls, G.K., 2001. Statistical estimation of the parameters of a pde, *Can. appl. Math. Quater.,* **10,** 277–810.

GOCAD User's Manual, 1996. Association Scientifique pour la Géologie et ses Applications, Ecole Nationale Supérieure de Géologie, Nancy.

Gu, Y. & Oliver, D.S., 2006. The ensemble Kalman filter for continuous updating of reservoir simulation models, *J. Energy Resour. Technol.,* **128**(87).

Hartmann, A., Rath, V. & Clauser, C., 2005. Thermal conductivity from core and well log data, *Int. J. Rock Mech. Mining Sci.,* **42,** 1042–1055.

Hartmann, A., Pechnig, R. & Clauser, C., 2007. Petrophysical analysis of regional-scale thermal properties for improved simulations of geothermal installations and basin-scale heat and fluid flow, *Int. J. Earth Sci.,* doi:10.1007/s00531-007-0283-y.

Hermanrud, C., Cao, S. & Lerche, I., 1990. Estimates of virgin rock temperature derived from BHT measurements: bias and error, *Geophysics,* **55**(7), 924–931.

Horner, D., 1951. Pressure build-up in wells, *Proc. Third World Petrol. Cong.,* **34**(316), 503–521.

Kitanidis, P.K., 1997. *Introduction to Geostatistics, Applications in Hydrogeology,* Cambridge University Press, Cambridge.

Kolditz, O., Ratke, R., Diersch, H.-J.G. & Zielke, W., 1998. Coupled groundwater flow and transport: 1. Verification of variable density flow and transport models, *Adv. Water Resour.,* **21,** 21–46.

Li, S., Zhang, C., Yin, Y., Yin, T. & Yan, S., 2008. Stochastic modeling of reservoir with multi-source, *Earth Sci. Front.,* **15**(1), 196–201.

Mottaghy, D., Pechnig, R., Willemsen, G. & Simmelink, E., 2009. Geothermal Project Den Haag—3-D models for temperature prediction and reservoir characterization, in *Geophysical Research Abstracts,*Vol. 11, European Geoscience Union General Assembly (EGU2009), Vienna, http://meetingorganizer.copernicus.org/EGU2009/EGU2009-11862.pdf.

Neuman, S.P. & Wierenga, P.J., 2002. A comprehensive strategy of hydrogeologic modeling and uncertainty analysis for nuclear facilities and sites, Tech. Rep. NUREG/CR8605, University of Arizona for U.S. Nuclear Regulatory Commission, Office off Nuclear Regulatory Research, Washington DC.

Nowak, W., 2005. Geostatistical Methods for the Identification of Flow and Transport Parameters in the Subsurface, *Doctoral dissertation,* Institut für Wasserbau, Universität Stuttgart.

Nowak, W., 2009. Best unbiased ensemble linearization and quasi-linear kalman ensemble generator, *Water Resour. Res.,* **45**, doi:10.1029/2008WR007328.

Pechnig, R., Mottaghy, D., Willemsen, G. & Simmelink, H.J., 2008. The Den Haag Geothermal District Heating Project—3-D Models for Temperature Prediction, *Schriftenreihe der Deutschen Gesellschaft für Geowissenschaften,* **60,** 309, http://www.dgg.de/pub/schriftenreihe/.

Rath, V., Wolf, A. & Bücker, M., 2006. Joint three-dimensional inversion of coupled groundwater flow and heat transfer based on automatic differentiation: sensitivity calculation, verification, and synthetic examples, *Geophys. J. Int.,* **167,** 453–466.

Remy, N., 2005. S-GeMS: The Stanford Geostatistical Modeling Software: A tool for new algorithms development, in *Geostatistics Banff 2004,* eds Leuangthong, O. & Deutsch, C.V., Springer, Dordrecht.

Robert, C.P. & Casella, G., 2004. *Monte Carlo Statistical Methods,* Springer, New York.

Rybach, L., 1988. Determination of heat production rate, in *Handbook of Terrestrial Heat-Flow Density Determination,* pp. 125–142, eds Haenel, R., Rybach, L. & Stegena, L., Kluwer Academic Publishers, Dordrecht.

Sambridge, M. & Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems, *Rev. Geophys.,* **40**(3), doi:10.1029/2000RG000089.

Simmelink, H.J. & Vandeweijer, V., 2008. Geothermie Den Haag Zuid-West, 2e fase geologisch onderzoek, Tech. Rep. 034.72157, Netherlands Organisation for Applied Natural Science Research (TNO), Utrecht.

Tarantola, A., 2004. *Inverse Problem Theory. Methods for Model Parameter Estimation,* Society for Industrial and Applied Mathematics (SIAM), Philadelpha PA, http://www.ipgp.jussieu.fr/tarantola/Files/Professional/Books/index.html.

van Balen, R.T., van Bergen, F., de Leeuw, C., Pagnier, H., Simmelink, H., van Wees, J.D. & Verweij, J.M., 2000. Modelling the hydrocarbon generation and migration in the West Netherlands Basin, the Netherlands, *Netherlands J. Geosci.,* **79**(1), 29–44.

Wieskopf, D. & Erlebacher, G., 2004. Overview of flow visualization, in *The Visualization Handbook,* pp. 261–278, Elsevier Academic Press, Burlington, MA.

Wolf, A., Rath, V. & Bücker, H.M., 2008. Parallelisation of a geothermal simulation package: a case study on four multicore architectures, in *Parallel Computing: Architectures, Algorithms and Applications, Proceedings of the International Conference ParCo 2007,* Vol. 15 of Advances in Parallel Computing, IOS Press, Amsterdam, The Netherlands, Co-published with NIC-Directors as Volume 38 of NIC Series, 2007.

Wolter, M., Tedjo-Palczynski, I., Hentschel, B. & Kuhlen, T., 2009. Spatial Input for Time Navigation in Scientific Visualizations, *IEEE Comput. Graph. Applicat.,* **29**(6), 54–64.

Xu, T., Sonnenthal, E., Spycher, N. & Pruess, K., 2004. *TOUGHREACT User's Guide: A Simulation Program for Non-isothermal Multiphase Reactive Geochemical Transport in Variably Saturated Geologic Media,* Lawrence Berkeley National Laboratory Report LBNL-55460, Berkeley CA.