UCRL-TR-230755

LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Initial Report on the Development of a Monte Carlo-Markov Chain Joint Inversion Approach for Geothermal Exploration

W. Foxall, A.L. Ramirez, S. Carlson, K. Dyer, Y. Sun

May 7, 2007

Disclaimer

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

UCRL-TR-230755

Initial Report on the Development of a Monte Carlo-Markov Chain Joint Inversion Approach for Geothermal Exploration

Bill Foxall, Abe Ramirez, Steve Carlson, Kathy Dyer, and Yunwei Sun

Lawrence Livermore National Laboratory

Introduction

Geothermal exploration and subsequent characterization of potential resources typically employ a variety of geophysical, geologic and geochemical techniques. However, since the data collected by each technique provide information directly on only one or a very limited set of the many physical parameters that characterize a geothermal system, no single method can be used to describe the system in its entirety. Presently, the usual approach to analyzing disparate data streams for geothermal applications is to invert (or forward model) each data set separately and then combine or compare the resulting models, for the most part in a more or less *ad hoc* manner. However, while each inversion may yield a model that fits the individual data set, the models are usually inconsistent with each other to some degree. This reflects uncertainties arising from the inevitable fact that geophysical and other exploration data in general are to some extent noisy, incomplete, and of limited sensitivity and resolution, and so yield non-unique results.

The purpose of the project described here is to integrate the different model constraints provided by disparate geophysical, geological and geochemical data in a rigorous and consistent manner by formal joint inversion. The objective is to improve the fidelity of exploration results and reservoir characterization, thus addressing the goal of the DOE Geothermal Program to improve success in exploration for economically viable resources by better defining drilling targets, reducing risk, and improving exploration/drilling success rates.

Approach

We are developing a joint inversion methodology for application to geothermal exploration by adapting a stochastic Monte Carlo-Markov Chain (MCMC) software package designed and developed at LLNL to jointly invert a wide range of data sets for characterization of complex geological systems. This package, dubbed the "Stochastic Engine" (SE), is described in the Appendix, together with a general overview of the joint inversion problem. The solutions produced by such stochastic methods are in the form of likelihood distributions across alternative models, which form an appropriate basis for developing risk-based drilling and development strategies.

We are adapting the SE by applying it initially to a specific type of geothermal system. In the first stage of development we chose a fault-hosted Basin and Range type system, based on the Dixie Valley, Nevada geothermal field. However, following definition of the conceptual framework for the application, we decided to base further development on rift-type magmatic systems based on the Salton Sea, California field. This change was motivated by the opportunity to integrate our effort with an allied project funded by the California Energy Commission (CEC). The objective of the CEC project (Bonner et al., 2005) is to constrain porosity, permeability, fluid saturation and the relationship of those properties to active faulting by innovative analysis of the abundant seismicity within and surrounding the Salton Sea field. The results of that study, therefore, hold the potential of providing valuable constraints on the joint inversion.

In the second stage of the project we implemented and tested a generalized base representation (see Appendix) for rift-type magmatic systems, based on existing data and interpretations for the Salton Sea field. We have selected four data sets, surface heat flow, shallow temperature gradient, gravity, and electrical resistivity, for the initial development, and written or adapted and interfaced the corresponding forward modeling codes to the MCMC core (see Appendix). At the heart of the computational scheme is a coupled fluid flow-heat transport code. Much of the effort during this second stage has been devoted to adapting and validating a finite difference code that is fast enough to render ~10^4 simulations of the temperature field per inversion run practicable.

The first task in the next stage of the project would be to interface the finite difference code to the SE core. The complete SE implementation would then be tested by inverting synthetic temperature, heat flow, resistivity, and gravity data sets generated from an increasingly complex series of alternative models representing the most important features of the Salton Sea field and the Salton Trough. These tests would be designed to assess the performance of the SE and to quantify the information content of each of the selected data sets, thus enabling optimization of the implementation.

The Salton Sea Geothermal System

The Salton Sea geothermal field occupies an actively subsiding pull-apart basin within the Salton Trough, a major transtensional basin located along the right-lateral transform boundary between the Pacific and North American plates. The Salton Sea field has long been recognized as one of the largest and hottest magmatic-type systems in the world. In a recent investigation, Hulen et al. (2002) used hitherto unavailable data to build upon numerous past studies to produce what is probably the most comprehensive conceptual model of the field to date, as shown in Figure 1.

The heat source (or sources) is identified as a cooling felsic intrusion that probably arose from a primitive magma body at the base of the crust. The intrusion may be located as shallow as ~1 km or less below the deepest wells (~2.5 km) in the field. The Salton Trough contains a sequence of low-density, brine-saturated sediments up to 6 km thick overlying an intermediate-density basement composed of metamorphosed basin sediments and/or pre-basin continental crust. The brines enable convective heat transfer, and comprise the geothermal production fluids.

The uppermost sedimentary layers consist of a thin (250-700 m) impermeable caprock overlying 500-600 m of essentially unaltered (or only slightly altered) sediments (Younker et al., 1982). Below this the sediments have undergone extensive hydrothermal alteration. The geothermal reservoir occupies both the unaltered and altered layers (termed the upper and lower reservoirs, respectively, by Younker et al.).

Past studies have indicated that flow -- and hence heat transport -- within the upper reservoir interval is predominantly horizontal, either because thin shale layers hamper vertical flow (Younker et al., 1982) or because of layered thermohaline convection (Oldenberg and Pruess, 1998). Both interpretations imply the absence of large-scale convection cells in the upper reservoir. Younker et al. presented evidence for predominantly porous flow and small-scale convection limited to individual sand beds, perhaps superimposed on large-scale horizontal flow, within the upper ~1 km of the reservoir. They concluded, however, that within the lower reservoir fracture permeability, perhaps maintained by seismicity within this tectonically active zone, and larger-scale convection probably predominate. Whereas the entire lower reservoir may be extensively fractured, none of the past studies has examined in detail the possible role that significant steeply-dipping fault zones might have in localizing high-permeability pathways, particularly deeper in the reservoir. Although such faults are indicated by recent and ongoing analyses of abundant seismicity and are known to be present within the Salton Sea field, Younker et al. (1982) (citing Kendall,1976) discount faults as major pathways for fluid and heat transport. However, this assertion is based on scant evidence limited to the upper ~2.5 km of the reservoir penetrated by production wells. It seems likely that enhanced vertical permeability along deep fault zones may play an important role in convective transport of heat from the depth of the intrusive source to production depths. Therefore, fault zones form a part of the base representation for the joint inversion.

Application of Joint Inversion to Magmatic Geothermal Systems

Base Representation

The base representation used in the MCMC inversion is designed to capture and quantify the possible elements of a generalized rift-type magmatic hydrothermal system, including the subsurface temperature distribution, stratigraphic and structural controls on the permeability and fluid saturation distribution, and the location(s) and temperature of the heat source(s). We implemented a simplified 2D base representation for the Salton Sea system based on the conceptual model of Hulen et al. (2002), together with more detailed shallow stratigraphy from Younker et al. (1982) and other earlier studies. Bounding ranges of parameter values (rock and fluid properties, temperatures, structural characteristics, etc.) available to the base sampler (see Appendix) are specific to the Salton Sea field, and were extracted from past studies.

Five example realizations of the 2D Salton Sea model are shown in Figure 2. The base representation comprises:

1. *Stratigraphy*: The stratigraphy consists of the caprock and the unaltered and altered rock layers. Variable parameters are rock density, thermal and electrical conductivity, and spatial distributions of permeability, fluid saturation, salinity and temperature. (Note that fluid saturation is included in the general base representation, although the Salton Sea reservoir itself is thought be 100% saturated.)

2. *Heat Source*: Variable parameters are location, size, shape, density, and temperature.
3. *Faults*: Variable parameters are number, location, width, dip, along-dip extent, permeability, and fluid saturation, salinity and temperature.

The color scale in Figure 2 corresponds to a range of different rock categories. Each category comprises a combination of rock properties and fluid content and properties within given *a priori* bounds. As shown in Figure 2, strata thicknesses can be varied in the general representation, but are held fixed in the Salton Sea application because the near-surface layer boundaries within the geothermal field are relatively well defined.

Data Sets

We selected four initial data sets, each of which provides constraint on a different set of system parameters:

1. <u>*Borehole temperature and surface heat flow*</u>: Shallow subsurface temperature gradient and heat flow measurements comprise the most fundamental data that can be inverted for the subsurface temperature distribution. Inversion of temperature and heat flow also helps constrain fluid saturation and rock permeability because convection as well as conduction is an important heat transport mechanism.
2. <u>*Gravity*</u>: Surface gravity measurements are inverted for the spatial distribution of rock density. Gravity data, therefore, could potentially constrain the composition, location and shape of the heat source, basin shape, thicknesses and compositions of basin sediments, and the locations and geometries of major fault zones. Gravity measurements generally offer relatively low-resolution of deeply buried features, so they are often most useful when analyzed jointly with passive (earthquake) and/or active seismic data. Although at this initial stage the implementation does not include inversion of seismic data directly, results from the CEC project will be used to constrain the *a priori* distribution bounds on rock properties and structures.
3. <u>*Electrical Resistivity*</u>: Surface resistivity measurements constrain the electrical conductance of buried rock units and the fluids they contain. Fluid conductivity is determined largely by dissolved salts and by temperature. Therefore, inversion of the resistivity data is used chiefly to help constrain the distribution of fluid temperatures, and, because the fluids hosting convection (at least within the Salton Sea system) are most likely brines, the overall distribution of subsurface fluids.

Forward Modeling Codes

1. <u>*Coupled Non-isothermal Fluid Flow and Heat Transport*</u>: During the conceptual stage of development we decided that an adequate characterization of rift-type magmatic systems necessarily includes estimates of the position and temperature history of the heat source, and the balance of the mechanisms -- conduction and convection -- by which heat is transported to production depths. This requires

computing the full evolution of the temperature field from the time of heat source intrusion to the present (as opposed to inverting for just a snapshot of the present temperature field). In addition to synthesizing the shallow sub-surface heat flow and temperature data, the computed temperature field and fluid properties are also used to drive the resistivity inversion.

We had first intended to use LLNL's multi-purpose Non-isothermal Unsaturated Flow and Transport (NUFT) code (Nitao, 1996), but had to develop an alternative for two reasons: (1) The steam tables in NUFT (and in other competing codes) do not extend to the high temperatures and pressures (depths) characteristic of magmatic/instrusive heat sources; and (2) evolving the temperature field for every trial solution in the MCMC inversion using NUFT would impose a very high computational burden.

As a simpler alternative to NUFT, we selected the 2D finite difference code developed by Lau (1980; see also Kasameyer et al., 1985), which employs an Alternating Direction Implicit (ADI) method to solve the energy equation and Gauss-Seidel iteration for the flow. This code is much faster than NUFT but treats the (single-phase) convection problem less rigorously. Therefore, a significant part of the effort in the second stage of work was devoted to modifying Lau's code and validating it against NUFT for temperatures and pressures below the critical point. Validation results are shown in Figure 3. Lau's code produces satisfactory matches to NUFT, although it still suffers from some degree of mismatch near the lateral boundaries of the model. The next stage of work would include completing the validation using temperature gradient data from the Salton Sea field and then interfacing the code to the SE core.

2. *Gravity*: The gravity code, CGRAV, that we developed for this project is specifically designed to interface with the SE. The code calculates the surface gravity field arising from an Earth model composed of a 3D grid of cells, each of which is assigned a density value. Therefore, arbitrary 3D or 2D geometries for strata, basins and discrete rock bodies -- such as the intrusive heat source -- can be represented. We integrated the gravity code to the SE core as an integral part of the code development.

3. *Electrical Resistivity*: We use the LLNL 3D electrical resistance tomography code MULTIBH to invert resistivity data. This is a mature code that has been employed in numerous studies, including inversion using the SE (e.g. Ramirez et al., 2005).

Conclusions

We have developed the generalized framework, base representation and forward modeling codes necessary to apply a Bayesian Monte Carlo-Markov Chain (MCMC) method to joint inversion of data sets commonly collected for geothermal exploration and reservoir characterization. Formal joint inversion provides a rigorous and rational means of optimizing the complementary constraints on the system parameters provided by the disparate data sets. This should lead to improved exploration results and reservoir models, together with full characterization of model uncertainties.

We based development of the geothermal application on current models, data and hypotheses for the rift-type magmatic Salton Sea geothermal field. The four data sets that are initially incorporated in the base representation are shallow temperature gradients, surface heat flow, gravity and electrical resistivity. These data potentially constrain geologic structure and rock and fluid properties, rock and fluid temperature distributions, and the distribution of permeability within the geothermal system. The resistivity forward modeling code had previously been implemented in the MCMC scheme, and validated in controlled experiments. We developed a new 3D gravity modeling code and interfaced it to the MCMC computational core. The greatest challenge during this stage of development was selection and adaptation of a coupled fluid flow-heat transport code capable of dealing with the high temperatures and pressures encountered in magmatic systems and fast enough to make the $\sim 10^4$ forward simulations typical of an inversion run practicable. We have successfully adapted a fast finite difference scheme and are ready to complete final validation and interfacing to the MCMC core.

The entire joint inversion package is now ready for validation, which we would begin by inverting synthetic data generated in a model based on existing knowledge of the Salton Sea field. Within this base representation, it will be possible to explore ranges of parameter values and to assess the sensitivity of the inversion results to given data sets. The performance of the present implementation of the method, together with the constraints provided by the data sets presently incorporated, would be evaluated based on the sensitivity study in order to identify improvements that could be made before applying the method to data collected at Salton Sea.

References

Bonner, B., L. Hutchings, and P. Kasameyer, A strategy for interpretation of microearthquake tomography results in the Salton Sea Geothermal Field based upon rock physics interpretations of State 2-14 borehole logs, UCRL-PROC-222141, Lawrence Livermore Natl. Lab., 2005

Hulen, J.B., D. Kaspereit, D.L. Norton, W. Osborn, and F.S. Pulka, Refined conceptual modeling and new resource estimate for the Salton Sea geothermal field, Imperial Valley, California, *Geotherm. Resources Coun. Trans.*, **26**, 29-36, 2002.

Kasameyer, P., L. Younker, and J. Hanson, Inversion approach for thermal data from a convecting hydrothermal system, *Jo. Geodynamics*, **4**, 165-181, 1985.

Kendall, C., Petrology and stable isotope geochemistry of three wells in the Buttes area of the Salton Sea Geothermal Field, Imperial Valley, California, U.S.A., (unpublished) M.S. thesis, Univ. Calif. Riverside, 1976.

Lau, K.H., Effect of permeability on cooling of a magmatic intrusion in a geothermal reservoir, UCRL 52888, Lawrence Livermore Natl. Lab., 1980.

Nitao, J. J., Reference Manual for the NUFT Flow and Transport Code version 1.0, UCRL-ID-113520, Lawrence Livermore Natl. Lab.,1996.

Oldenburg, C.M., and K. Pruess, Layered thermohaline convection in hypersaline geothermal systems, *Transport in Porous Media*, **33**, 29-63, 1998.

Ramirez, A.L., J.J. Nitao, W.G. Hanley, R. Aines, R.E. Glaser, S.K. Sengupta, K.M. Dyer, T.L. Hickling, and W.D. Daily, Stochastic inversion of electrical resistivity changes using a Markov Chain Monte Carlo approach, *Jo. Geophys. Res.*, **110**, B02101, doi:10.1029/2004JB003449, 2005.

Younker, L.W., P.W. Kasameyer, and J.D. Tewhey, Geological, geophysical and thermal characteristics of the Salton Sea geothermal field, California, *Jo. Volcanology Geotherm. Res.*, **12**, 221-258, 1982.

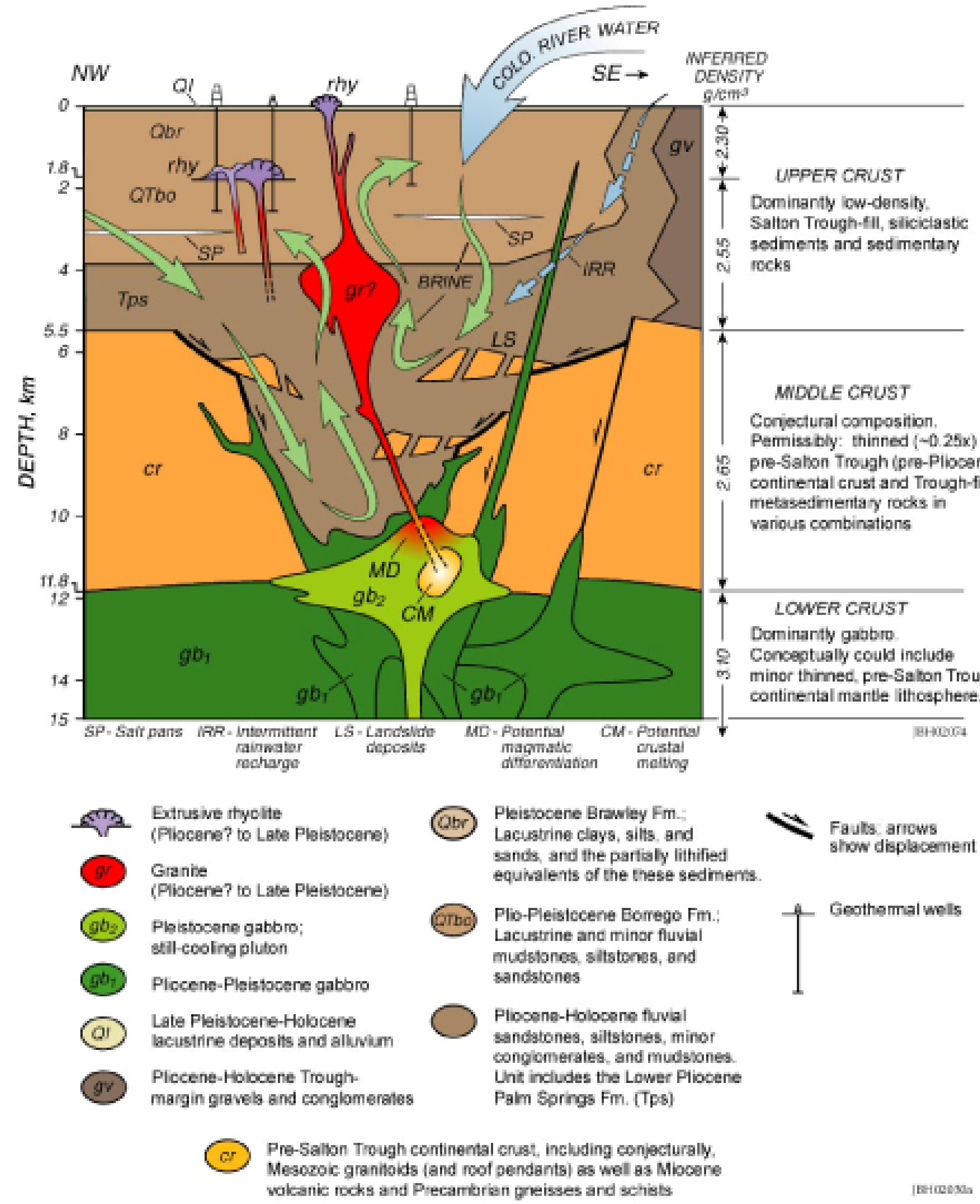

Figure 1: Conceptual Model of the Salton Sea Hydrothermal System. [Reproduced from Hulen et al. (2002)]

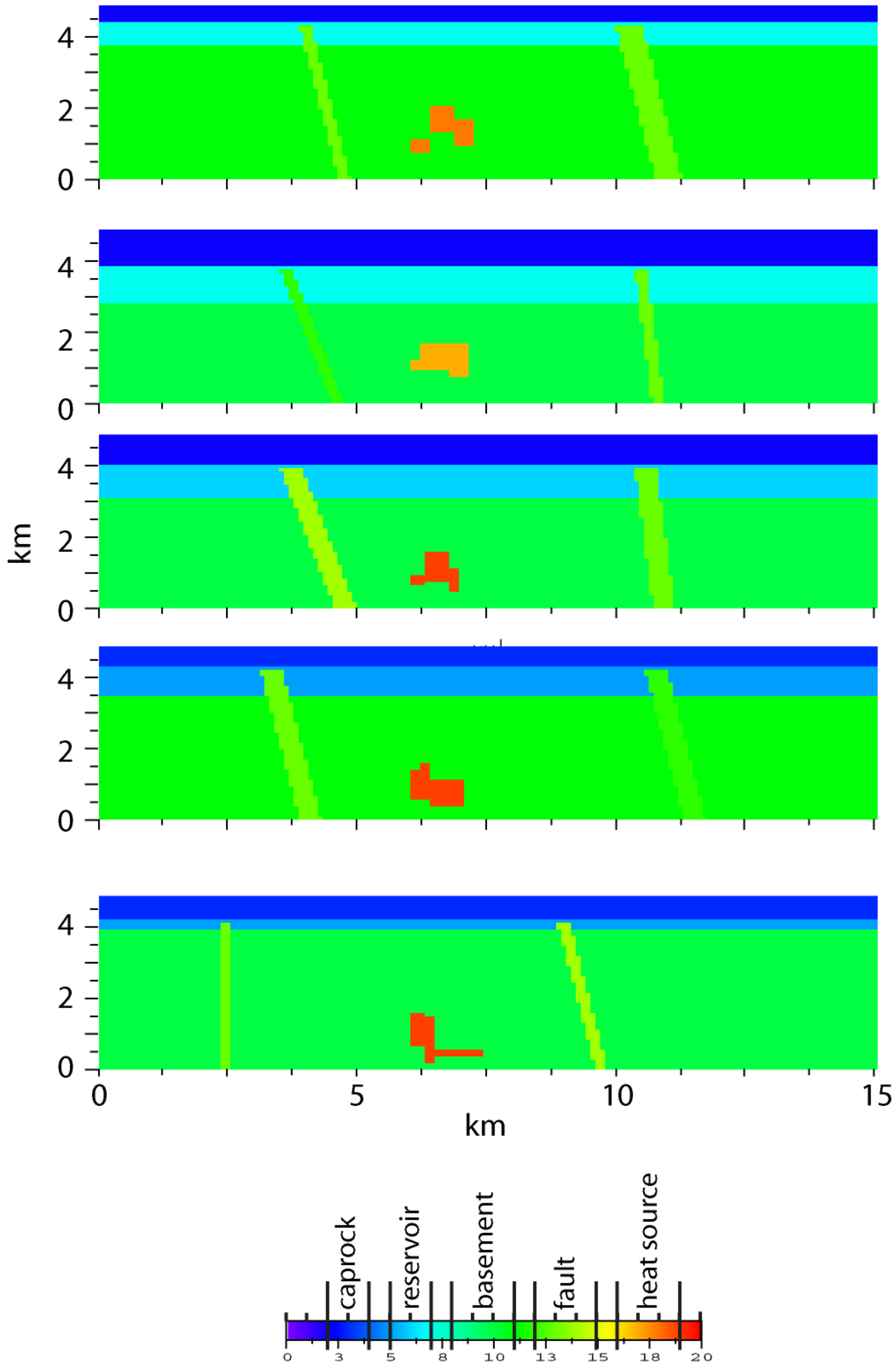

Figure 2: Five realizations of the Salton Sea geothermal field model.

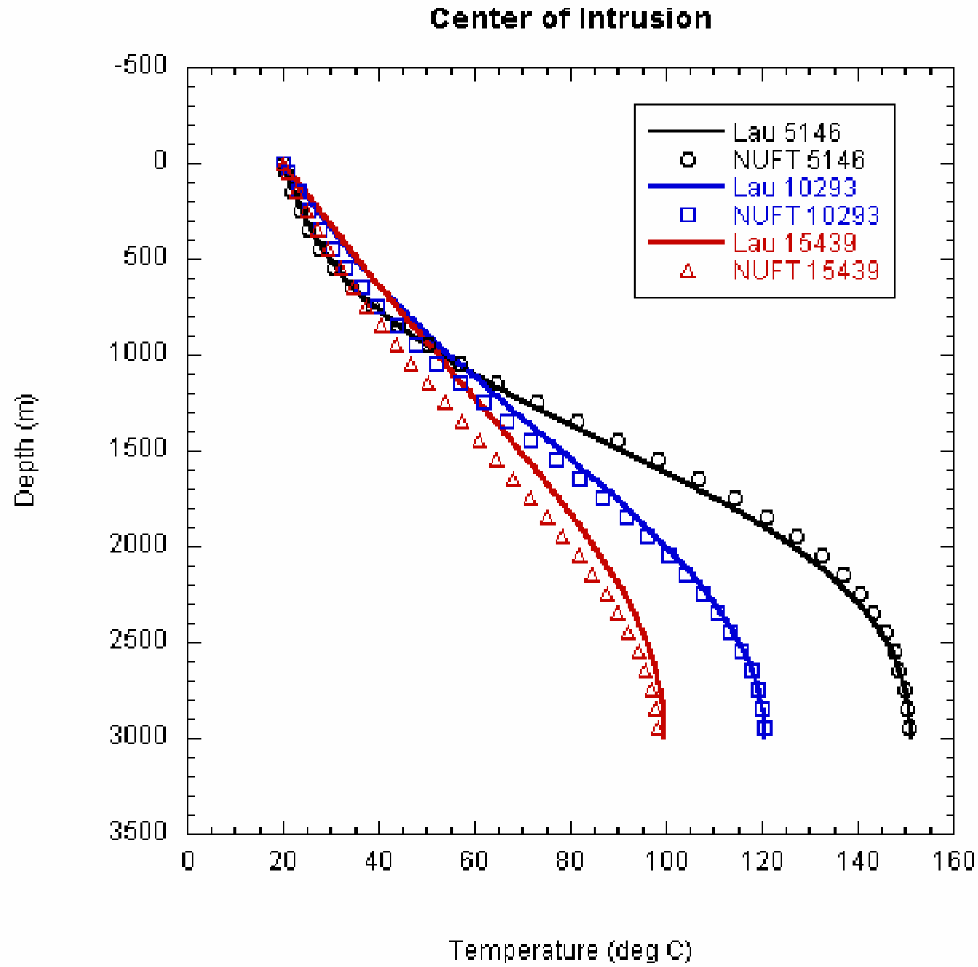

Figure 3: Comparison of temperature-depth profiles computed by the ADI finite difference code (Lau, 1980) and NUFT. Heat source is a 1 km wide by 1.5 km deep intrusion embedded in a uniform porous medium underlying a 200m-thick impermeable caprock. The top of the intrusion is at a depth of 1.5 km. Intrusion temperature is 175°C above surrounding rock.

APPENDIX

Joint Inversion to Characterize Geological Systems

The Inverse Problem

Measured data $\mathbf{d}$ (e.g. geophysical, geological, geochemical or hydrological observations) can be expressed as a function of the parameters $\mathbf{m}$ of a subsurface geological system (model) by:

$$\mathbf{d} = \mathbf{g}(\mathbf{m})$$

where $\mathbf{g}$ is the functional form of a forward (physical) model with which the data can be computed from the model parameters. The inverse problem is to determine the model parameters given the data. In essence, the inverse problem integrates the general knowledge represented by the forward model with specific knowledge represented by data. The model space $\mathbf{m}$ can comprise both fixed, usually spatially-varying parameters, which might include, for example, stratigraphy and structure, that are prescribed *a priori* based on a geologic model, and varying parameters, such as temperature, fluid saturation, porosity and permeability, that are solved for in the inversion. The simplest case is a problem in which the vector $\mathbf{d}$ of observations is linearly related to a vector $\mathbf{m}$ of model by $\mathbf{d} = \mathbf{G}\mathbf{m}$, where $\mathbf{G}$ is a matrix of analytic or semi-analytic response/transfer functions involving the Earth properties and (fixed) source-observation geometry.

Joint Inversion

The present state-of-the art in analyzing multiple data streams to detect and characterize subsurface geological systems is to invert (or forward model) each data set separately and then combine or compare the resulting models, for the most part in a more or less *ad hoc* manner. However, while each inversion may yield a model that fits the individual data set well, the models are usually inconsistent with each other to some degree. This reflects uncertainties arising from the inevitable fact that the data in general are to some extent noisy and incomplete, and imperfectly constrain the model parameters. Each data set offers constraint on only part of the model space, or strong constraint on some of the model parameters but weak or no constraint on the rest.

Joint inversion, on the other hand, formally combines the constraints supplied by the different data types to produce a model (or distribution of models - see below) that is consistent with all of the data simultaneously. In addition to making more complete use of the data by improving the span on the model space, combining the data sets can also convert an underdetermined inverse problem (number of data less than number of model parameters to be solved for; no unique solution) to an overdetermined one (for a given discretization of the model) by increasing the number of measurements (e.g. Pritchard et al., 2002). Another way of viewing this is that adding data permits finer discretization of the model and hence higher resolution. Joint inversion can be carried out on disparate data sets that provide constraint on quite distinct parts of the parameter space. For example, surface deformation, constraining subsurface fluid volume changes, can be inverted with electrical resistivity, constraining change in fluid saturation, to jointly constrain a model of fluid saturation and pressure.

Deterministic and Stochastic Inversion Approaches

Joint inversion can be carried out using either deterministic or stochastic (probabilistic) approaches. Deterministic joint inversion utilizes a variety of conventional linear inversion techniques employing, for example, gradient-based algorithms such as constrained, weighted least squares to search the model space for a solution that provides the best fit (minimum misfit) to the data. The advantage of these methods is that they are usually computationally relatively undemanding. The objective of a deterministic inverse is to produce a single model that provides the "best fit" to the data together with estimates of the uncertainties (variance, covariance) on the parameters of that solution. The ill-posedness of the most geophysical inverse problems requires regularization of the optimization based on *a priori* or assumed information; for example, a common assumption is that the solution should be smooth to some degree.

There are two major drawbacks to deterministic methods. The first is that they cannot deal with more than weakly nonlinear systems, since beyond that the optimization problem becomes too complex to solve without strong simplifications to the model and strong damping to remedy numerical instabilities and ensure convergence. The second drawback is that they perform a localized rather than global search of the parameter space and so can become "stuck" in a local minimum, rather than the global minimum representative of the true model. This problem becomes increasingly acute as the nonlinearity of the model increases.

These drawbacks have led to increasing use over the past decade of probabilistic inversions using Monte Carlo methods (see Sambridge and Mosegaard, 2002). The key to the success of Monte Carlo approaches is that they perform a *global* search of the model space by running a large number of forward models to identify the ensemble of acceptable models, according to some threshold criterion. The probability of each of these models is computed based on the fit of data predicted by the model to the observations. The solution to the inverse problem, therefore, is the (often multimodal) probability distribution of alternative solutions that spans the entire model space; i.e. a full description of uncertainty. One approach to assessing the distribution employs Bayesian inference, which combines prior information with the data to produce the posterior probability distribution. Practical Monte Carlo methods employ efficient random but non-uniform searches of the model space. These include Monte Carlo-Markov chain (MCMC) methods employing Metropolis importance sampling algorithms (e.g. Sambridge and Mosegaard, 2002), simulated annealing and genetic algorithms.

Monte Carlo methods can handle highly nonlinear models, since all that is required is the ability to solve the forward problem. Stochastic methods are computationally more demanding than conventional methods because adequate searches of the model space typically involve large numbers of forward calculations. The increasing popularity of these methods reflects the exponential growth in computing power. Although many practical applications can now be run on computer workstations, stochastic methods are ideally suited for parallel computers such as Linux clusters, which render even large nonlinear inverse problems feasible.

Joint Inversion using the Stochastic Engine

Monte Carlo methods are particularly suited to joint inversion. LLNL has developed a flexible Bayesian/MCMC implementation - dubbed the "Stochastic Engine" (SE) - that can be applied to a wide variety of joint inversion problems (Aines et al, 2002; Ramirez et al., 2005, Pasyanos et al., 2005). The SE comprises a Bayesian/MCMC core (the engine) to which existing or purpose-built forward modeling codes are interfaced. Prior knowledge is incorporated on the model space in the form of a "base representation", which is populated by both fixed parameters and parameters that can vary within bounds fixed *a priori*. A "base sampler" proposes trial models to be run through the forward modeling codes by assigning specific parameter values randomly selected from the parameter ranges to the base representation. Joint inversion is accomplished by running the problem as a linked sequence of forward modeling stages, each stage utilizing one of the codes to model the corresponding data set. Any trial model must be accepted by every stage to be included in the posterior model distribution, so that the final solution contains models that are consistent with all of the data. The computational burden is greatly reduced by running the least computationally demanding forward models first, so that far fewer candidate models are passed to later, more computationally intensive stages.

References

Aines, R, et al., The Stochastic Engine Initiative: Improving Prediction of Behavior in Geologic Environments We Cannot Directly Observe, UCRL-ID-148221, 58 p, Lawrence Livermore Natl. Lab., 2005.

Pasyanos, M. E., G. A. Franz, and A. L. Ramirez, Reconciling a geophysical model to data using a Markov Chain Monte Carlo algorithm: An application to the Yellow Sea – Korean Peninsula region, submitted to *Jo. Geophys. Res.*, 2005.

Pritchard, M.E., M. Simons, P.A. Rosen, S. Hensley, and F.H. Webb, Co-seismic slip from the 1995 July 30 M_w=8.1 Antofagasta, Chile, earthquake as constrained by InSAR and GPS observations, *Geophys. Jo. Int.*, **150**, 362-376, 2002.

Ramirez, A.L., et al., Stochastic inversion of electrical resistivity changes using a Markov Chain Monte Carlo approach, *Jo. Geophys. Res.*, **110**, B02101, doi:10.1029/2004JB003449, 2005.

Sambridge, M., and K. Mosegaard, Monte Carlo methods in geophysical inverse problems, *Rev. Geophys.*, **40**, no. 3, 1-29, 2002.