Groundwater Research Report
WR06R003

APPLICATION OF LSQR TO CALIBRATION OF A REGIONAL MODFLOW MODEL: TROUT LAKE BASIN, WISCONSIN

Mary P. Anderson
Haijiang Zhang
Chris Muffels

2007

Groundwater Research Report
WR06R003

Application of LSQR to Calibration of a Regional MODFLOW Model: Trout Lake Basin, Wisconsin

Principal Investigators
Mary P. Anderson and Haijiang Zhang
Department of Geology and Geophysics
University of Wisconsin-Madison

Research Assistant
Chris Muffels
Department of Geology and Geophysics
University of Wisconsin-Madison

October 1, 2007

Table of Contents

List of Tables

List of Figures

Project Summary

Title: Application of LSQR to Calibration of a Regional MODFLOW Model: Trout Lake Basin, Wisconsin

Project I.D.: WR06R003

Investigator(s): Mary P. Anderson – University of Wisconsin-Madison, Department of Geology and Geophysics
Haijiang Zhang - University of Wisconsin-Madison, Department of Geology and Geophysics

Period of Contract: July 2006 – June 2007

Background/Need: Complex regional groundwater flow models, like the model of the Trout Lake basin in Vilas County, Northern Wisconsin, are used to address a variety of research questions including the impacts of climate, land use changes, and delineation of flow paths. Yet, the level of certainty expected from predictions made by these models is often beyond what most current calibration techniques can provide. Groundwater models are becoming larger and more complex with many more unknown parameters. Current matrix methods that are used in parameter estimation are not able to solve the large matrices associated with these models. For this reason it is important to explore the applicability of new parameter estimation techniques for solving large groundwater problems. LSQR (Paige and Saunders 1982a, 1982b) is a parameter estimation method that uses an iterative subspace inversion technique that is related to the better known singular value decomposition (SVD). SVD is currently used by the groundwater community. LSQR is widely used by the geophysical community to solve large problems in tomography and is more powerful than SVD in solving large problems.

Objectives: The objectives of this research were to 1) use a synthetic groundwater flow model as a test case to develop a strategy for using LSQR to solve groundwater inverse problems, 2) demonstrate that a technique used to calculate the model resolution matrix, which is necessary to quantify uncertainty, for large seismic tomography inverse problems can be applied to groundwater problems; and 3) use the experience gained from the first objective to demonstrate that LSQR gives comparable results to the SVD method in estimating parameters for a regional groundwater flow model of the Trout Lake basin.

Methods: LSQR was tested and compared with the more widely used SVD in terms of computational burden, convergence behavior, and parameter values estimated by both methods. A simple two-dimensional "checkerboard" synthetic model was used for this purpose for which all components of the inverse problem were exactly known. Iterative algorithms like LSQR require stopping rules to decide when an acceptable solution to the inverse problem has been achieved. LSQR provides three such rules and the effects of these were investigated as part of the synthetic exercise. An approach described by Zhang and Thurber (2007), using a modified LSQR, was used to determine the resolution matrix necessary to calculate the uncertainty associated with model predictions. The LSQR algorithm was then used to calibrate the groundwater flow model of the Trout Lake basin and the results were compared to the more traditional SVD. Uncertainty is discussed in the context of using the modified LSQR to determine values for the LSQR truncation criterion. Proper choice of the LSQR truncation criterion is important to minimize prediction uncertainty.

Results and Discussion: The LSQR algorithm was successful in calibrating both the synthetic and Trout Lake models and seems to be well suited to the ill-posed inverse problems often encountered by groundwater modelers. The solution speed of LSQR over the traditional SVD is noteworthy. Several of the LSQR test calibrations of the synthetic model yielded similar reductions in the objective function

compared with the best performing SVD run, but took a third the time to solve the system. For the synthetic model, all the LSQR test cases were successful in reducing the objective function by at least three orders of magnitude, with the most successful LSQR run reducing the objective function from 7.6 to 3.2E-03. For the Trout Lake model, both the SVD and LSQR calibrations produced similar results but LSQR solved the system in a fraction of the time required by SVD. Prediction uncertainties calculated using the modified LSQR exactly matched those calculated using the traditional SVD method. Uncertainties based on pre-calibration estimates were calculated to determine the level of parameterization supported by the observations; the LSQR truncation criterion must be set so that the number of LSQR iterations is commensurate with the number of parameters supported by the observations. Setting the criterion too tight can result in unrealistic solution times and negate the advantage of using LSQR.

Conclusions and Recommendations: The LSQR algorithm shows promise as a robust and versatile algorithm for solving large groundwater inverse models that otherwise are not tractable using inverse methods such as SVD that are currently used in groundwater modeling. As computers become faster with more memory, current methods will be able to solve for more parameters; however, models will continue to increase in size and complexity at the same time and require more parameters for calibration. As such, there will always be a need for alternatives to the traditional approaches. Through this research, LSQR is demonstrated to be a viable means of estimating many thousands of parameters and quantifying the uncertainty associated with model predictions for such large models. Because LSQR is now available as an option within the popular inverse program PEST, modelers throughout the State of Wisconsin and elsewhere can apply LSQR to their problems now and in the future. Additional research is needed to understand the role of the LSQR closure variables ATOL and BTOL and the effects of SVD-Assist on LSQR performance. Furthermore, the use of the modified LSQR, PROPACK-SVD (Larsen 1998), as an alternative to the traditional SVD and LSQR for medium sized problems should be investigated.

Related Publications:

Muffels, C., R. Hunt, J. Doherty, and M. Anderson. 2007. Regularized Inversion of a Groundwater Flow Model of the Trout Lake Basin. Presentation at the 2007 American Water Resources Association Wisconsin Section Meeting, Chula Vista Conference Center, Wisconsin Dells, Wisconsin, March 1-2, 2007.

Muffels, C., H. Zhang, J. Doherty, R. Hunt, M. Anderson, and M. Tonkin. 2006. Incorporating PROPACK into PEST to Estimate the Model Resolution Matrix for Large Groundwater Flow Models. Presentation at the 2006 American Geophysical Union Fall Meeting, Moscone Center, San Francisco, California, December 11-15, 2006. San Francisco, California.

Muffels, C., J. Doherty, M. Anderson, R. Hunt, T. Clemo, and M. Tonkin. 2006. LSQR and Tikhonov Regularization in the Calibration of a Complex MODFLOW Model. Presentation at the Geological Society of America Annual Meeting, Pensylvania Convention Center, Philadelphia, Pennsylvania, October 22-25, 2006. Philadelphia Pennsylvania.

Muffels, C., M. Tonkin, H. Zhang, M. Anderson, and T. Clemo. 2006. Application of LSQR to Calibration of a MODFLOW Model: A Synthetic Study. MODFLOW and More 2006, Managing Ground-Water Systems, International Ground Water Modeling Center, Colorado School of Mines Golden, Colorado, May 22-24, 2006. Vol. 1. 283-287.

Key Words: groundwater flow modeling, parameter estimation, prediction uncertainty, LSQR

Funding: State of Wisconsin Groundwater Research Program through the University of Wisconsin Water Resources Institute, the Department of Geology and Geophysics, UW-Madison, and S.S. Papadopulos and Associates Inc.

Introduction

The Trout Lake basin (Figure 1) is part of the NSF's NTL-LTER (Long Term Ecological Research) program and the USGS WEBB (Water Energy Biochemical Budgets) program. Surface water systems are well connected to the groundwater system, and there is very little surface runoff; therefore, understanding groundwater flow is critical to determine the movement of water and transport of solutes in the system (Walker and Bullen 2000). Recent versions of the Trout Lake basin model by Pint (2002), Pint et al. (2003), Hunt et al. (2003), John (2005), and Hunt et al. (2005) are based on the groundwater flow code MODFLOW2000 (Harbaugh et al. 2000). Calibration of the Trout Lake basin model using 11 parameters is described by Hunt et al. (2005) using UCODE (Poeter and Hill 1998), a universal parameter estimation program that solves the inverse problem using a modified Gauss-Newton method (Cooley and Naff 1990). The Trout Lake basin model is designed to be a general purpose watershed model that is used to address a variety of research questions, including flow path delineation (Pint et al. 2003, Masbruch 2005), climate change (John 2005), and land use change.

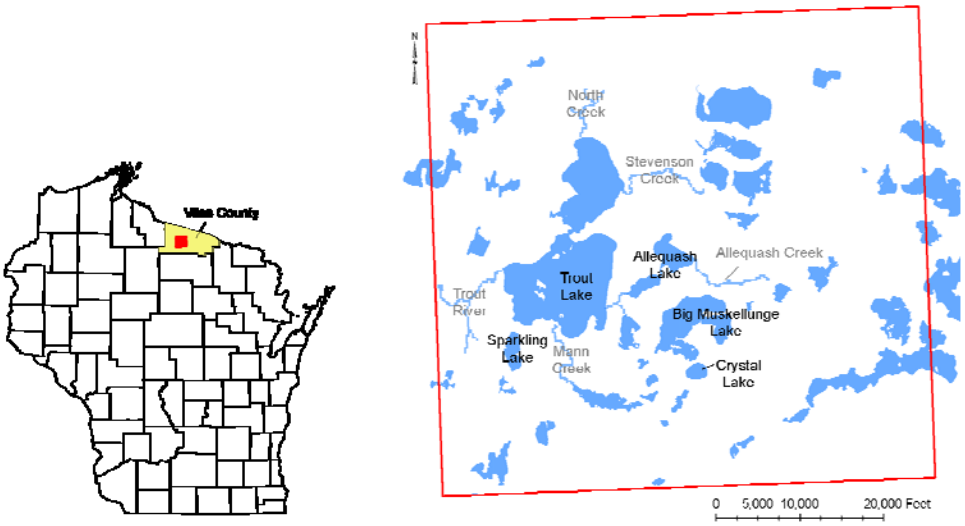

Figure 1. Location map of the Trout Lake basin and the extent of the regional model (red outline).

Calibration of a groundwater flow model involves determining values for *n* unknown system parameters, **x** (often horizontal and vertical hydraulic conductivities and recharge rate), from a set of *m* field observations, **b** (often hydraulic heads and stream/river fluxes). An *m* x *n* sensitivity matrix, **A**, is used to relate the parameters to the observations and inform the inverse problem how the model responds to perturbations in parameter values. Parameter value updates are calculated according the following equation (Cooley and Naff 1990, Doherty 2004),

$$\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b} \qquad (1)$$

The determination of best parameters is an iterative process that starts with an initial estimate of the parameters from which **A** is determined, and the parameter values are recalculated. The iterative process terminates, ideally, when the objective function, Φ, is a global minimum.

$$\Phi = (\mathbf{b} - \mathbf{A}\mathbf{x})^T(\mathbf{b} - \mathbf{A}\mathbf{x}) \qquad (2)$$

The global minimum is not always easy to find, and the process can be complicated if there are local minima. As such, the initial parameter estimates must be a good approximation of the true values (Doherty 2004).

The inverse problem becomes intractable for these methods as the number of parameters increases because it is less likely that the sensitivity matrix can be inverted. This difficulty can be mitigated if the principle of parameter parsimony (Hill 1998) is practiced by parameterizing the model

with zones of piecewise constancy. However, trying to minimize the number of parameters in this way restricts and pre-determines the spatial heterogeneity of the system because the zones are determined before the process begins (Doherty 2003).

A means of stabilizing the inverse process to allow many more parameter values to be estimated is to regularize the inversion. Regularization is a method for producing a close approximation to poorly-posed inverse problems so that they are numerically tractable. Regularized inversion can take many forms, including methods that augment the sensitivity matrix (pre-conditioning methods) and methods that truncate the solution calculations (subspace methods) (Tonkin and Doherty 2005). Pre-conditioning methods provide a means by which the modeler can include their "hydrogeological wisdom". For example, if the aquifer system being modeled is believed to be homogeneous the inverse process can be informed of this preference through the formulation of appropriate regularization equations. With regularized inversion there is no need for pre-calibration zoning of spatial heterogeneity (Doherty 2003), which is a significant advantage since there is often insufficient information for making informed decisions about zonation *a priori*. Prior to this research, calibration of the Trout Lake basin model was achieved with 1000 parameters using PEST (Doherty 2004).

Solving large systems in this manner is common place in many scientific applications and singular value decomposition (SVD) is the subspace regularization method most often employed. The SVD method decomposes any arbitrary *m* x *n* matrix, **B**, according to:

$$\mathbf{B} = \mathbf{U}\mathbf{S}\mathbf{V}^{T} \qquad (3)$$

where **U** (*m* x *m*) and **V** (*n* x *n*) contain the left and right singular vectors of **B**, respectively, and **S** (*m* x *n*) the singular values (Lawson and Hanson 1995, Anderson et al. 1999). The singular vectors represent weighted combinations of parameters that comprise each observation value and the singular values indicate the magnitude of that influence. Truncated singular value decomposition (TSVD) is a mechanism for determining **x** from the *k* most dominant singular vectors according to:

$$\mathbf{x} = \mathbf{V}_k\mathbf{S}^{-1}\mathbf{U}_k{}^{T}\mathbf{b} \qquad (4)$$

Truncation is a form of regularization and is often necessary because the singular vectors associated with small singular values magnify the noise inherent in the observed values and contaminate estimation of **x**.

Regardless of whether the parameter estimation technique is that employed by UCODE or by PEST, or another inverse code, at most the number of parameters uniquely estimable is equal to the number of observations. However, in a typical groundwater problem the observations are noisy and inter-dependent and do not always contain unique information and so fewer parameters are actually supported. A distinction between UCODE and PEST is the manner in which these estimable parameters are treated. UCODE, through the use of zones, requires the user to specify these parameters (number, types and locations) *a priori*. In other words, the modeler is required to provide the model space that can be explored during the calibration. The model space is defined by all the possible combinations of parameter values. PEST, through the use of subspace regularization (TSVD), allows the calibration process to determine the model space supported by the observations because it can be well represented by the singular vectors corresponding to the dominant singular values. The PEST program includes a novel technique called SVD-Assist (Tonkin and Doherty 2005) that defines these dominant parameters as "super" parameters and recasts the inverse problem, posed on the basis of many base parameters, to one based on the super parameters. The calibration only explores the model space defined by these super parameters, which requires many fewer forward model runs because only the super parameters are perturbed to calculate the sensitivity matrix. Values for the base parameters, **x**, are determined from the super parameters, $\mathbf{x}_k$, using the linear relationships defined in $\mathbf{V}_k$, according to (Tonkin and Doherty 2005):

$$\mathbf{x} = \mathbf{V}_k\mathbf{x}_k \qquad (5)$$

The advantage of using many parameters is that the model space supported by the observations can be determined more accurately because more possible parameter combinations can be tested. The problem is that the number of parameters the traditional SVD algorithm can estimate (i.e. the size of the matrix it can decompose) is limited by computer memory space and speed (Berryman 2000). As computers continue to get faster and have more memory, the number of parameters that can be estimated

by SVD will continue to increase. For this reason it can be argued that the life-span of alternatives to SVD will be short-lived. However, consider that the computers available 20 years ago when groundwater modeling with computers became wide-spread were vastly inferior in speed and memory to the computers we have today. Despite the vastly superior computers we have today, there are groundwater models constructed with over 4000 parameters and 4000 observations that cannot be calibrated using SVD (Doherty personal communication 2006). As computers become more sophisticated, models become larger and more complex with more parameters. These models will continue to push the bounds of SVD and we will always require alternative parameter estimation techniques.

The most efficient method available to solve large linear seismic tomographic systems is LSQR. LSQR is an iterative inversion technique that uses the Lanczos bidiagonalization process for reducing the matrix **A** to bidiagonal form and then minimizes the least squares norm $\|\mathbf{Ax} - \mathbf{b}\|_2$ using QR factorization (Paige and Saunders 1982a, 1982b). The LSQR is better suited to large problems than SVD because it requires much less storage and can quickly iterate to a solution. However, it cannot directly calculate the model resolution matrix because vectors obtained from the Lanczos process do not retain orthogonality as they iterate due to finite machine precision. The model resolution matrix is important because it is used to evaluate the ability of the real observations to estimate the model parameters (Aster et al. 2005) and to quantify the uncertainty in predictions made with the model (Moore and Doherty 2005). It was originally proposed that a model independent parameter estimation program be developed that uses the LSQR algorithm. Partly owing to early results of this research, LSQR was incorporated into the widely used code PEST (Doherty 2004) and is available free from www.sspa.com/pest.

LSQR has four input parameters to control the number of iterations needed to solve equation 1: ATOL, BTOL, CONLIM, and ITNLIM. Paige and Saunders (1982a) and Doherty (2007, addendum to the PEST manual) give a description of each and recommend values. The ATOL and BTOL parameters represent the user assumed accuracy of the sensitivity matrix **A** and the observation vector **b**, respectively. When using perturbation sensitivities, Doherty (2007) suggests that ATOL and BTOL are of the order 10^{-2} and 10^{-3}, respectively. The variable CONLIM is a limit on the condition number of matrix **A** and is intended to prevent small or zero singular values from affecting the solution (Paige and Saunders 1982a), which is similar to TSVD. If, for each singular value, the ratio of the highest value to the current value is calculated, the variable CONLIM can be thought of as an upper bound on this ratio. That is, when the ratio of the highest estimated singular value to the current value in any iteration exceeds CONLIM then LSQR will terminate. This value is opposite to the PEST variable EIGTHRESH, which is the ratio of the lowest to highest singular value and the cutoff used to determine *k* for TSVD. The ITNLIM variable is an upper bound on the number of iterations that controls termination of LSQR.

A method described by Zhang and Thurber (2007), which makes use of PROPACK-SVD (Larsen 1998), was used to calculate the resolution matrix. The PROPACK-SVD algorithm is based on Lanczos bidiagonalization with partial reorthogonalization and only estimates a user-specified number of singular values rather than all of them, which is a significant time savings and allows it to be used with large models. With the Zhang and Thurber (2007) method, the model is first calibrated using LSQR and then the PROPACK-SVD is used to estimate the *k* dominant singular vectors of the final sensitivity matrix that are needed to calculate the resolution matrix according to:

$$\mathbf{R} = \mathbf{V}_k \mathbf{V}_k^T \qquad (6)$$

where **R** is the model resolution matrix. Moore and Doherty (2005) present the following equation to calculate predictive error variance (synonymous with uncertainty):

$$\sigma^2 = \mathbf{y}^T(\mathbf{I} - \mathbf{R})\mathbf{C(x)}(\mathbf{I} - \mathbf{R})^T\mathbf{y} + \mathbf{y}^T\mathbf{G}\mathbf{C(\varepsilon)}\mathbf{G}^T\mathbf{y} \qquad (7)$$

where σ^2 is the total predictive error variance, **y** is an *n*-row vector of the sensitivity of the prediction (observation) of concern to model parameters (i.e. a row from the sensitivity matrix), **I** is the *n* x *n* identity matrix, **C(x)** is the *n* x *n* parameter covariance matrix, **G** is an *n* x *m* matrix and is equal to $(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ such that $\mathbf{x} = \mathbf{Gb}$, and **C(ε)** is the *m* x *m* covariance matrix of the data noise. The first term on the right-hand side of the equation is the contribution to a prediction's uncertainty due to the inability of the calibration to determine the true parameter set and the second term is the uncertainty due to the inability

of the calibration to determine a parameter set that exactly reproduces the observations due to noise. A more detailed explanation of this equation is beyond the scope of this report, but it is important to note that the two terms in equation 7 compete: increasing the number of parameters reduces the uncertainty due to the first term and increases the uncertainty due to the second term (Figure 2). Typically, prediction uncertainty is calculated after a model has been calibrated; however, calculating the pre-calibration uncertainty for different levels of parameterization can be used to determine the number of parameters that are supported by the observations. The number of parameters that should be used with TSVD or SVD-Assist is the number of parameters that minimizes the prediction uncertainty.

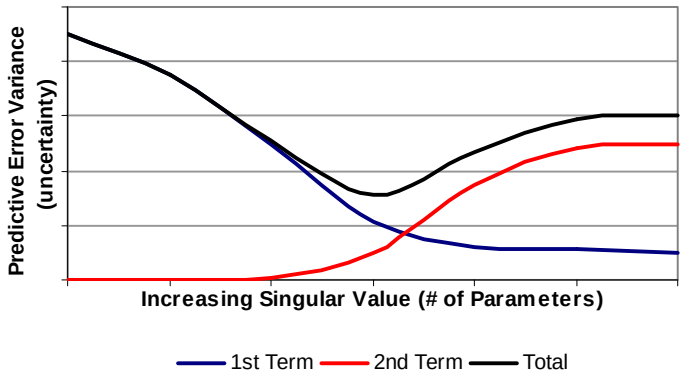

Figure 2. Idealized prediction uncertainty curves demonstrating trade-off between the two terms on the right-hand side of equation 8.

Procedures and Methods

Synthetic Problem. To test the LSQR, a synthetic groundwater flow model was constructed in MODFLOW2000, which is a widely used code developed by the U.S. Geological Survey. The model was calibrated with PEST using SVD, as well as LSQR, to provide a base from which to draw conclusions regarding the convergence and computational burden of LSQR. Using the experience gained from this synthetic exercise the flow model of the Trout Lake basin was calibrated using LSQR and SVD. The synthetic model was two-dimensional with 32 rows and 32 columns; each cell had dimensions of 1 ft by 1 ft. Constant-head boundaries of 40 ft and 32 ft were specified on the left and right sides of the model with no-flow conditions along the top and bottom. Two zones (1ft/day and 100 ft/day) of hydraulic conductivity were evenly distributed throughout the model in a checkerboard pattern (reference 5, Appendix A). To test LSQR, the model was calibrated to the model calculated heads (for the true hydraulic conductivity distribution) using hydraulic conductivity as the only variable parameter. Hydraulic conductivity was assumed to be isotropic and was assumed unknown in each model cell; the starting value for each parameter was calculated as a random fraction (up to ±50 %) of the true value. In total there were 1024 parameters and 1024 observations. The model was calibrated with PEST using both SVD and LSQR. Different levels of truncation were used with the SVD calibration to illustrate the effects of subspace regularization, TSVD. The LSQR was tested by first using the default LSQR settings for ATOL, BTOL and CONLIM and then varying those numbers to evaluate effects on the final calibration (Table 1).

Trout Lake Basin Model. There are 230 rows and 240 columns and 6 layers in the Trout Lake model. Each cell is 75 x 75 ft. The boundaries of the model are specified fluxes, which are calculated by an analytic element screening model (Hunt et al. 1998) that encompasses a larger regional area. The boundary conditions are implemented using MODFLOW's Well Package. Details on model design and calibration are contained in Muffels (reference 1, Appendix A).

For the purposes of this research, the model was calibrated using 1501 parameters and 120 observations. Parameter types included horizontal and vertical hydraulic conductivity, bed conductance for lakes and streams, which were represented using MODFLOW's River Package, and porosity, which is used in particle tracking simulations. Recharge was not used as a parameter; its distribution and values were taken from a previous calibration by Hunt (personal communication 2007). The parameters were represented as pilot-points; nodal values for each cell in the model were interpolated from these pilot

Table 1. Summary of SVD truncation levels, k (with corresponding EIGTHRESH values in parentheses), and LSQR stopping criteria (ATOL, BTOL, CONLIM, and ITNLIM) used to test SVD and LSQR with the synthetic model.

Run		ATOL	BTOL	CONLIM	ITNLIM	Truncation Level, *k* (EIGTHRESH)
SVD	1	-	-	-	-	960 (1E-14)
	2	-	-	-	-	512 (1E-08)
	3	-	-	-	-	200 (1E-04)
	4	-	-	-	-	380 (1.5E-05)
	5	-	-	-	-	600 (1.0E-10)
LSQR	1	1.0E-06	1.0E-06	1.0E+06	5000	-
	2	1.0E-03	1.0E-03	1.0E+06	5000	-
	3	5.0E-05	5.0E-05	1.0E+06	10000	-
	4	1.0E-06	1.0E-06	1.0E+04	5000	-
	5	1.0E-12	1.0E-12	2.0E+06	30000	-
	6	1.0E-12	1.0E-12	1.0E+06	600	-

points using kriging (Doherty 2003). Observation types included hydraulic head, horizontal and vertical head differences, groundwater flux estimates to and from the lakes, baseflow estimates for each stream, and travel time and depth of flow paths determined from isotope analyses for flow paths from Crystal Lake to Big Muskellunge Lake, and from Big Muskellunge Lake to Allequash Lake. The particle-tracking program MODPATH (Pollack 1994) was used to determine the simulated equivalent to travel time and depth of flow path. The objective function was calculated as a weighted difference between simulated and observed values (equation 2). Each observation was normalized using weighting functions (Moore and Doherty 2005) so that each observation type contributed equally to the objective function, Φ.

The calibration was performed using both SVD-Assist (Tonkin and Doherty 2005) with a preferred-value regularization scheme for each parameter, and LSQR. The preferred values chosen for each parameter were taken from the values estimated by Pint (2002). For demonstration purposes for this report the prediction of primary interest was arbitrarily chosen to be the groundwater discharge to Big Muskellunge Lake. The Trout Lake basin model was then calibrated with PEST using both LSQR and SVD. A method described by Zhang and Thurber (2007), which uses the PROPACK-SVD algorithm to determine the singular vectors necessary to calculate the resolution matrix, was used to calculate prediction uncertainty. To do so, a utility called PREDVAR1 (Doherty 2007) was modified to call the PROPACK-SVD routine rather than that of the more traditional SVD. PREDVAR1 calculates prediction uncertainty according to equation 7. Uncertainty was calculated to determine truncation levels for LSQR to demonstrate the modified LSQR works and could be used for a more rigorous post-calibration uncertainty analysis if need be.

Results and Discussion

Synthetic Problem. A value for the variable EIGTHRESH (Table 1) must be selected to determine the truncation level, *k*, for the TSVD. Because this is a synthetic exercise with no "real" prediction of concern, a simple procedure was used for this research to determine the truncation level for SVD. A plot of the ratio of each singular-value to the highest value for every parameter (Figure 3) was constructed. The shape of the resulting curve is different for each problem, but the ratio will always decrease with increasing parameter number. The procedure involves choosing a truncation level based on obvious breaks or inflection points in the curve. From Figure 3, obvious breaks are at EIGTHRESH values of 1E-14 and 1E-08 or 960 and 512 parameters respectively. The break at 960 parameters is caused by the specified head boundary conditions. The 64 parameters associated with the constant-head cells on the

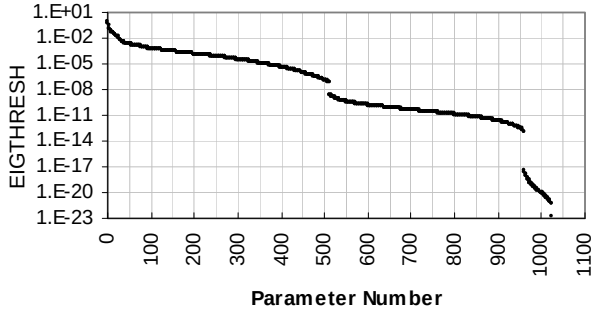

Figure 3. Ratio of each singular value to the highest value (defined as variable EIGTHRESH in PEST) for each of the 1024 parameters in the synthetic model.

left and right boundaries were included in the calibration but the sensitivities of these parameters are very low and contribute little to the calibration. The cause of the break at 512 parameters is unknown. A third cutoff was chosen arbitrarily at 200 singular values (EIGTHRESH ~ 1E-04) to demonstrate the effect of using too few parameters. A summary of the results for the SVD runs (Table 2) shows that, of these three runs (SVD runs 1, 2, and 3), the lowest objective function value, phi, was attained by SVD run #2 that used 512 singular values. The results of SVD runs 1 and 3 demonstrate the effect of truncating too early or too late as they achieved a phi an order of magnitude higher than SVD run 2. The results of SVD runs 4 and 5 are discussed below because they were performed after the LSQR runs were completed.

For the LSQR runs, the CONLIM values (2E+13, 1.8E+07, and 1E+04) were chosen because they correspond roughly to the inverse of the EIGTHRESH values used. The results for the CONLIM value of 2E+13 are not presented because the solution time was too long and negated the advantages of LSQR. Results for LSQR runs (Table 2) show that the lowest phi values were achieved by LSQR runs 5 and 1 as these runs had the tightest convergence criteria with respect to ATOL, BTOL and CONLIM (Table 1). These two runs also took the most time to solve the system of equations; in fact they took longer than the complete decomposition of the sensitivity matrix using SVD. So, even though a better fit to the observations was achieved using LSQR, it took too much time to solve the system of equations.

Table 2. Summary of calibration results for SVD and LSQR runs.

	Run	Average Time to Solve System (seconds)	# Optimization Iterations	# Model Runs	Lowest PHI
SVD	1	45	9	9217	3.1E-02
	2	45	7	7169	3.2E-03
	3	45	8	8193	2.1E-02
	4	45	8	8193	4.8E-03
	5	45	11	11265	7.2E-04
LSQR	1	210	10	10241	3.5E-04
	2	3	11	11265	4.0E-03
	3	3	4	4097	3.4E-03
	4	15	6	6145	3.2E-03
	5	900	8	8193	2.2E-04
	6	3	6	6145	1.4E-02

The success of LSQR for this problem is made evident in the results of LSQR runs 2, 3, and 4. Each of these runs achieved an objective function similar to that of SVD run 2 (512 singular values), but took much less time to solve the system to achieve these results (Table 2) (note these results are not compared to SVD run 5 because there is no obvious reason to choose an EIGTHRESH cutoff at 600 singular values; see discussion below). Also, LSQR runs 3 and 4 required significantly fewer optimization iterations (and correspondingly fewer forward model runs). This result can be important if the most time intensive aspect of the calibration process is generating the sensitivity matrix. However, depending on the inverse problem, this result may not always be the case.

The variable CONLIM had a value of 1E+04 for LSQR run 4, which corresponded to 360 LSQR iterations. For early LSQR iterations, the number of singular values (and vectors) estimated roughly corresponds to the number of iterations. So for LSQR run 4, terminating at iteration 360 corresponded to estimating the first 360 singular values which, according to Figure 3, corresponds to EIGTHRESH of about 1.5E-05. To test this result, an additional TSVD run, SVD run 4, was completed with EIGTHRESH set to 1.5E-05. The lowest phi achieved was similar to that of LSQR run 4 (Figure 4). The

number of singular values used in each solution with the TSVD method is about the same as the number of iterations required by the LSQR method (Figure 4). However, this is only true for very early LSQR iterations as the Lanczos vectors used to approximate the singular vectors quickly lose orthogonality (Zhang and Thurber 2007). As a test, an LSQR run was conducted with the iteration limit, ITNLIM, set to 600 and the results compared with an additional SVD run that truncated the solution after 600 singular values. The SVD run 5 (600 singular values) yielded a much lower phi (on a scale commensurate with LSQR runs 1 and 5) than the LSQR run, which indicates that the Lanczos vectors had lost orthogonality and so contaminated the estimation of the parameter values.

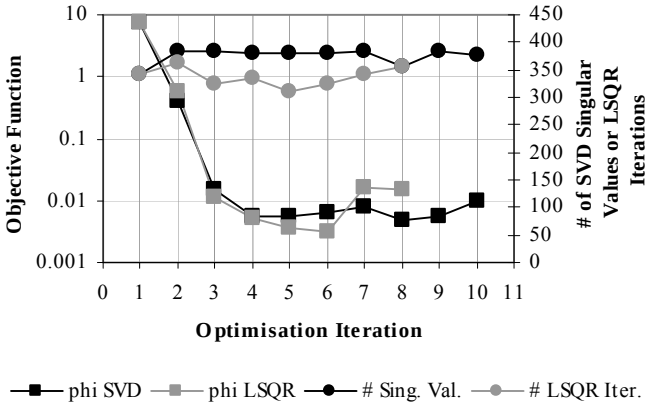

Figure 4. Objective function (phi) values and the number of singular values used or estimated for each optimization iteration for SVD run 4 and LSQR run 4.

<u>Trout Lake Basin Model.</u> Moore and Doherty (2005) outline a process that is incorporated in utility PREDVAR1 of PEST to estimate the uncertainty associated with a model prediction, which is used to determine the level of parameterization necessary to minimize that uncertainty. The source code of this utility was altered to call the modified LSQR algorithm, PROPACK-SVD, to use with the LSQR calibration. The pre-calibration uncertainty estimated using the modified and the original PREDVAR1 executables for the prediction of groundwater flux to Big Muskellunge Lake for different degrees of parameterization is shown in Figure 5 (only the pre-calibration uncertainties were calculated for the purposes of this report as the objective was to demonstrate that the PROPACK-SVD works). From this figure the level of parameterization supported by the observations is about 35 parameters and so 50 super parameters were used with SVD-Assist (a few extra were chosen to give the calibration process some freedom to explore more or fewer parameters during the calibration process). In using LSQR the limiting closure variable was ATOL; iterations terminated after about 40, which was commensurate with the number of parameters supported by the observations (Figure 5). The SVD run was set with a very low EIGTHRESH value and the number of singular values used capped at about 40 corresponding to the number of LSQR iterations. Both the LSQR and SVD runs reduced the objective function from 800 to 300. The reduction is relatively small because this calibration started with results from a previously calibrated model by Hunt (personal communication 2007). The reduction is significant because we show that the calibration is improved by these methods. Furthermore, the LSQR algorithm took just a few seconds to find a solution for each optimization iteration, while SVD took several minutes. This result is significant because it is a 90% speed improvement over SVD. A synthetic test was conducted by Muffels et al. (2006) with 4000 parameters and 8000 observations. The SVD took over an hour to decompose the corresponding sensitivity matrix while LSQR took only a few minutes to return a comparable solution.

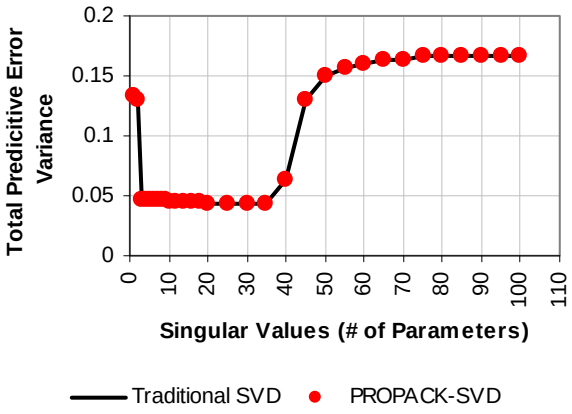

Figure 5. Total predictive error variance per singular value for the discharge to Big Muskellunge Lake prediction calculated by the traditional SVD and PROPACK-SVD.

The kriged fields that resulted from the two different calibration techniques (Figure 6) are shown for the first model layer. A comparison of all the parameter values is presented in Figure 7. In Figure 6, the values of specific parameters are not shown because they are not as important as the net effect of the kriging of these parameters in the delineation of heterogeneities. Any parameter values estimated between 8 and 10 m/day are not colored to indicate no change in value during the calibration. Lower conductivity features are colored brown while higher conductivities are colored green. The circled area encompasses the area where observations are present; parameters in this area are the most sensitive and so the heterogeneities in this area are the focus of our attention. The numbers indicate the major heterogeneity features delineated. These features correspond very closely between the two methods demonstrating that LSQR was able to calibrate the model and return a set of parameter values comparable to that of SVD for layer 1.

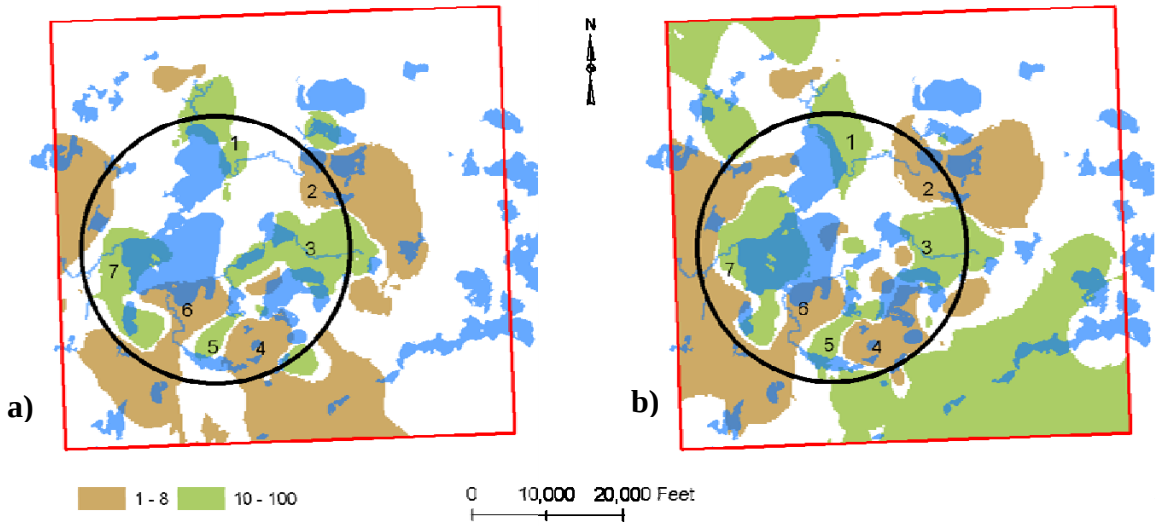

Figure 6. Hydraulic conductivity field estimated by a) SVD and b) LSQR for the first layer of the Trout Lake model. Colors represent hydraulic conductivity ranges in m/day (not colored is the 8 – 10 m/day range).

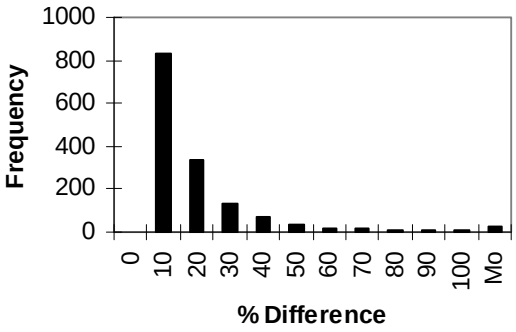

Figure 7. Frequency plot of the percent difference of values estimated using LSQR from those for the same parameter estimated using SVD, calculated as abs[(x_{SVD}-x_{LSQR})/x_{SVD}].

The similarity in results for the other parameters is summarized in Figure 7. It shows that 1165 of the 1501 parameter values estimated by LSQR were no more than 20% different from the value estimated for the same parameter using SVD. Values estimated for porosity were the most different between the two calibrations. The values estimated by LSQR were more reasonable as they ranged between 0.11 to 0.2 compared with a range of 0.005 to 0.2 for SVD.

Conclusions and Recommendations

Alternative methods to estimate parameter values for large groundwater flow models are needed because currently the limit on the number of parameters that can be estimated is determined by the computing power available to the modeler. Hence, large complex models cannot be adequately calibrated to minimize the uncertainty in model predictions. Large geophysical models are limited in the same manner but this community has explored alternative parameter estimation techniques and determined that LSQR is the most efficient method currently available to solve their large problems. This research demonstrated that LSQR is a viable option for large groundwater models.

The LSQR method successfully calibrated both the synthetic and Trout Lake basin models. The sub-space regularization that is implicit in the iterative nature of LSQR is well suited to groundwater inverse models as these are often ill-posed and severely under-determined. Because the algorithm is now an option in the widely used and free parameter estimation program PEST (Doherty 2004), modelers throughout the state of Wisconsin and elsewhere, can apply it to a wide variety of calibration problems. When using LSQR it is important that the number of LSQR iterations be commensurate with the number of parameters supported by the observations. This number can be determined in a number of ways, but the PEST utility PREDVAR1 provides a means of determining it in the context of minimizing the uncertainty associated with key model predictions (e.g., Figure 5). The LSQR stopping criteria, ATOL, BTOL and CONLIM, should be set so that LSQR terminates at about the desired number of iterations. Caution should be exercised as setting these variables too tight can cause long solution times, which might be an indication that the vectors used to approximate the singular vectors are no longer orthogonal and may be contaminating the resulting parameter values. If thousands of parameters are being estimated as part of a calibration process and LSQR is being used it is likely that a technique such as SVD-Assist will be necessary so that calculation of the perturbation sensitivities is tractable. This technique was successfully used with LSQR and the Trout Lake basin model, but more research into its effect on LSQR performance is needed and is a topic for future study.

The PROPACK-SVD routine was added as an option to PREDVAR1 and successfully used to calculate the prediction uncertainty for the Trout Lake basin model. (However, at the time of this report, the modified version of the utility was not yet available for download with the PEST suite of programs at www.sspa.com/pest.) Beyond calculating the model resolution matrix, this algorithm could be used as an alternative to both the traditional SVD and LSQR algorithms for medium sized problems. Its advantage over the traditional SVD algorithm is that it does not have to decompose a matrix entirely; the user can specify the number of singular values and vectors desired, which is likely to be a significant time savings. This idea is similar to terminating LSQR after a specified number of iterations; however, the PROPACK-SVD method does not suffer from a loss of orthogonality in the same manner and will accurately calculate the singular values and vectors rather than approximate them. It is recommended that a future study examine the use of the PROPACK-SVD routine as a solution technique.

References

Anderson, E., Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammerling, A. McKenney, and D. Sorenson, 1999. LAPACK Users' Guide Third Edition, 407 pp, SIAM Publications, Philadelphia, PA, (http://www.netlib.org/lapack/index.html)

Aster, R.C., B. Borchers, and C.H. Thurber, 2005. *Parameter Estimation and Inversion Problems*, 301 pp. Elsevier, New York.

Berryman, J.G., 2000. Analysis of approximate inverses in Tomography. II. Iterative inverses. *Optimization and Engineering* 1, 437-473.

Cooley, R.L. and R.L. Naff, 1990. Regression Modeling of Ground-water flow: U.G. Geological Survey *Techniques in Water-Resources Investigations*, Book 3, Chap, B4, 232 p.

Doherty, J., 2003. Groundwater Model Calibration using Pilot Points and Regularization. *Ground Water* 41, no. 2: 170-177.

Doherty, J., 2004. PEST, Model Idenpendent Parameter Estimation. User Manual: 5th Edition. Watermark Numerical Computing, Brisbane, Queensl., Australia.

Doherty, J., Senior Research Fellow, University of Queensland, Brisbane, Australia; and Director, Watermark Numerical Computing Pty. Ltd. Verbal communication 2006.

Doherty, J., 2007. *PEST Manual*. Watermark Numerical Computing, Brisbane, Queensl., Australia.

Harbaugh, A.W., E.R. Banta, M.C. Hill, and M.G. McDonald, 2000, MODFLOW-2000, the U.S. Geological Survey modular ground-water model -- User guide to modularization concepts and the Ground-Water Flow Process: U.S. Geological Survey *Open-File Report* 00-92, 121 p.

Hill, M.C., 1998, Methods and Guidelines for Effective Model Calibration: U.S. Geological Survey *Water-Resources Investigations Report* 98-4005, 90 p.

Hunt, R. J., M.P. Anderson, and V.A. Kelson, 1998. Improving a Complex Finite-Difference Groundwater Flow Model Through the Use of an Analytic Element Screening Model. *Ground Water* 36, no. 6: 1101-1117.

Hunt, R. J., C.D. Pint, and M.P. Anderson, 2003. Using diverse data types to calibrate a watershed model of the Trout Lake Basin, Northern Wisconsin. In: *Proceedings of MODFLOW 2003*, IGWMC, Golden Colorado, 600-604.

Hunt, R.J, D.T. Feinstein, C.D. Pint, and M.P.Anderson, 2005, The Importance of Diverse Data Types to Calibrate a Watershed Model of the Trout Lake Basin, Northern Wisconsin, USA, *Journal of Hydrology* 321(1-4): 286-296, published online 28 Sept. 2005, doi: 10.1016/j.jhydrol.2005.08.005.

Hunt, R.J., Research Hydrologist, Wisconsin Water Science Center, USGS, Middleton WI; and Adjunct Professor, University of Wisconsin-Madison, Madison WI. Written communication 2007.

John, R. 2005. *Simulating Response Rates and Historical Transience of Surface Water and Groundwater: Trout Lake Basin, Northern Wisconsin*. M.S. thesis, Department of Geology and Geophysics, University of Wisconsin-Madison.

Larsen, R. M.,1998, Lanczos bidiagonalization with partial reorthogonalization for large sparse SVD and linear least squares calculations, *Tech. Rep*., Computer Science Department, Aarhus University.

Lawson, C. L. and J, R. Hanson, 1995. Solving Least Squares Problems, *SIAM Classics in Mathematics*, pp 337, Philadelphia, Pennsylvania.

Moore, C., and J. Doherty, 2005. Role of the calibration process in reducing model predictive error, Water Resources Research. VOL. 41, W05020, doi: l0.1029/2004WR00350.

Masbruch, M.D. 2005. *Delineation of Source Areas and Characterization of Chemical Variability using Isotopes and Major Ion Chemistry, Allequash Basin, Wisconsin*. M.S. thesis, Department of Geology and Geophysics, University of Wisconsin–Madison.

Muffels, C., J. Doherty, M. Anderson, R. Hunt, T. Clemo, and M. Tonkin. 2006. LSQR and Tikhonov Regularization in the Calibration of a Complex MODFLOW Model. *Abstract.* Proceedings, Geological Society of America Annual Meeting, Philadelphia, Pennsylvania.

Paige, C.C. and M.A. Saunders, 1982a. LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares. *ACM Trans. Math. Softw*. vol. 8: 43-71.

Paige, C.C. and M.A. Saunders, 1982b. LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares. *ACM Trans. Math. Softw*. vol. 8: 192-209.

Pint, C.D. 2002. *A groundwater flow model of the Trout Lake of the Trout Lake Basin, Wisconsin: Calibration and lake capture zone analysis*. M.S. thesis, Department of Geology and Geophysics, University of Wisconsin-Madison.

Pint, C.D., R.J. Hunt, and M.P. Anderson, 2003. Flow Path Delineation and Ground Water Age, Allequash Basin, Wisconsin. *Ground Water* 41, no. 7: 895-902.

Poeter, E.P. and M.C. Hill, 1998, Documentation of UCODE, a Computer Code for Universal Inverse Modeling: U.S. Geological Survey *Water-Resources Investigations Report* 98-4080, 116 p.

Pollock, D.W., 1994, User's Guide for MODPATH/MODPATH-PLOT, Version 3: A particle tracking post-processing package for MODFLOW, the U.S. Geological Survey finite-difference ground-water flow model: U.S. Geological Survey *Open-File Report* 94-464, 6 ch.

Tonkin, M. J. and J. Doherty, 2005. A Hybrid Regularized Inversion Methodology for Highly Parameterized Environmental Models, *Water Resour. Res*., 41, W10412. doi:10.1029/2005WR003995.

Walker, J.F. and T.D. Bullen, 2000. Trout Lake, Wisconsin A Water, Energy, and Biogeochemical Budgets Program Site. U.S. Geological Survey *Fact Sheet*-161-99.

Zhang, H. and C.H. Thurber (2007). Estimating the model resolution matrix for large-seismic tomography problems based on Lanczos bidiagonalization with partial reorthogonalization, *Geophys. J. Int*. 170, p. 337-345, 10.1111/j.1365-246X.2007.03418.x

Appendix A. Publications

1) Muffels, C. (in preparation, expected May 2008). A highly parameterized calibration of a groundwater flow model of the Trout Lake basin, Vilas County, Wisconsin. Master's thesis, Department of Geology and Geophysics, University of Wisconsin-Madison.

2) Muffels, C., R. Hunt, J. Doherty, and M. Anderson. 2007. Regularized Inversion of a Groundwater Flow Model of the Trout Lake Basin. Presentation at the 2007 American Water Resources Association Wisconsin Section Meeting, Chula Vista Conference Center, Wisconsin Dells, Wisconsin, March 1-2, 2007.

The Trout Lake basin model is a three-dimensional general purpose MODFLOW-based watershed model used to address a variety of research questions including flow path delineation, and climate and land use change. The flow system is dominated by groundwater flow that is well connected to the surface water system; thus, a good groundwater model is critically important in determining the movement of water and transport of solutes in the system. The latest versions of the model are quite sophisticated, allowing for dynamic interaction with 30 lakes and 5 streams. The model has been historically calibrated using a small number of parameter values; however, recent calibration using many more parameters has been employed to attempt to capture more of the spatial system heterogeneity. To constrain model calibration with a large number of parameters, a "regularized inversion" approach was used in this research. Regularization can take many forms and is a means of stabilizing large problems by using subjective information about the parameters based on the current hydrogeologic understanding of the system in addition to field measurements. In this paper, the results of using different regularization schemes including Tikhonov, truncation, and damping to calibrate the Trout Lake model are presented. Results show that using a preferred homogeneity Tikhonov scheme improved the model fit to field measurements.

3) Muffels, C., H. Zhang, J. Doherty, R. Hunt, M. Anderson, and M. Tonkin. 2006. Incorporating PROPACK into PEST to Estimate the Model Resolution Matrix for Large Groundwater Flow Models. Presentation at the 2006 American Geological Union Fall Meeting, Moscone Center, San Francisco, California, December 11-15, 2006. San Francisco, California.

Regularized inversion of groundwater flow models can be used to determine heterogeneities using subspace methods like the singular value decomposition (SVD). To better characterize the heterogeneity of the model, thousands of system parameters and, with appropriate regularization, thousands of observations may be necessary. The SVD is not practical because it requires significant memory space and is time consuming. We have demonstrated (Muffels et al. 2006a, b) the LSQR can be used to estimate the many unknown parameters in large groundwater flow inverse problems. However, in doing so, a resolution analysis is needed to characterize how reliable the resulting model parameters are. We adopted an approach described by Zhang and Thurber (2006) for large seismic tomography problems and incorporate the PROPACK package developed by Larsen (1998) into PEST, a model independent parameter estimation program. PROPACK is able to efficiently and accurately estimate singular values and vectors for large matrices based on the Lanczos bidiagonalization, the core of LSQR, with partial reorthogonalization. Compared with other LSQR-based resolution approaches, this PROPACK-based approach calculates the full resolution matrix. We estimate the model resolution matrix for a synthetic approximation to the real-world regional MODFLOW model of the Trout Lake Basin, Wisconsin, and compare it with that of the more common SVD.

4) Muffels, C., J. Doherty, M. Anderson, R. Hunt, T. Clemo, and M. Tonkin. 2006. LSQR and Tikhonov Regularization in the Calibration of a Complex MODFLOW Model. Presentation at the Geological Society of America Annual Meeting, Pensylvania Convention Center, Philadelphia, Pennsylvania, October 22-25, 2006. Philadelphia Pennsylvania.

Employing a-priori parameter parsimony to solve the inverse problem in groundwater modeling is conceptually appealing, but has several downfalls – not least of which is that the solution of the inverse problem is restricted to a very limited subspace of the true parameter space. The alternative is to estimate a very large number of parameters and allow the calibration process to determine where heterogeneity may exist. Estimating a large number of parameters during model calibration requires rapid and memory-lean matrix solution techniques that can accommodate highly parameterized models. We demonstrated (Muffels et al, 2006) that although the LSQR - an iterative matrix solution technique similar to the methods of conjugate gradients - can solve the large matrix equations that are produced when thousands of parameters are estimated, convergence can be disappointing. Here, we explore the role of Tikhonov regularization, when employed together with LSQR, for stabilizing the inverse problem and for improving the convergence of a local-approximation to the true inverse problem. The model application is a synthetic approximation to the real-world regional MODFLOW model of the Trout Lake Basin, Wisconsin. Since the number of parameters greatly exceeds the number of observations, parameter sensitivities are calculated using the adjoint-state method. We compare the results to the most commonly employed subspace method, the truncated singular valued decomposition (TSVD). Discussions focus on the advantages – or otherwise – of the Tikhonov regularization, and on the improvement(s) in inference gained through the use of a large number of parameters.

5) Muffels, C., M. Tonkin, H. Zhang, M. Anderson, and T. Clemo. 2006. Application of LSQR to Calibration of a MODFLOW Model: A Synthetic Study. MODFLOW and More 2006, Managing Ground-Water Systems, International Ground Water Modeling Center, Colorado School of Mines Golden, Colorado, May 22-24, 2006. Vol. 1. 283-287.

The inverse problem in groundwater modeling is often made numerically tractable and computationally practical by estimating only a small fraction of the many unknown system parameters. However, this parsimonious approach restricts the solution of the inverse problem to a pre-determined subspace of the true parameter space. To reflect detailed local variations in hydraulic conductivity or recharge, it may be desirable to estimate a very large number of parameters during calibration, which requires an inversion technique that can accommodate highly parameterized models. The least-squares QR decomposition (LSQR) is an iterative subspace solution method that can solve for many hundreds or thousands of parameters. LSQR has been used successfully in seismic tomography inversion problems. As an iterative method, LSQR can solve sparse and dense inverse problems of the form AX=B using significantly less computer storage than direct solution methods. We test the applicability of the LSQR method for solving the inverse problem for groundwater flow using a synthetic model, and compare results with those obtained using the most commonly employed subspace method, the truncated singular valued decomposition (TSVD). Parameter sensitivities are calculated using the adjoint-state method. Both over-determined and under-determined synthetic problems are investigated.