Beaney from 'Begriffsschrift, eine der arithmetischen nachgebildete Formelsprachen des reinen Denkens', L. Nebert, Halle, 1879.
Geach, P. T. and Black, M. (eds.). 1952. *Translations from the philosophical writing of Gottlob Frege*. Basil Blackwell, Oxford.
Janssen, T. M. V. 1997. Compositionality. In J. van Benthem and A. Ter Meulen (eds.), *Handbook of Logic and Language*, pp. 417-473. Elsevier, MIT Press, Amsterdam, The Netherlands, Cambridge, MA.
Janssen, T. M. V. 2001. Frege, Contextuality and Compositionality. *Journal of Logic, Language and Information*, vol. 10, pp. 115-136.
Harnish, M. (ed.). 1994. *Basic Topics in the Philosophy of Language*. Prentice-Hall, Englewood Cliffs, NJ.
Hodges, W. 2001.Formal features of compositionality. *Journal of Logic, Language and Information*, vol. 10, pp. 7-28.
Landman, F. and Veltman, F. (eds.). 1984. *Varieties of Formal Semantics*. Foris, Dordrecht.
Mihálydeák, T. 2003. *Tarskian and Husserlian models in general type-theoretical semantics*. Manuscript, forthcoming.
Partee, B. 1984. Compositionality. In F. Landman and F. Veltman (eds.) *Varieties of Formal Semantics*, pp. 281-312. Foris, Dordrecht.
Pelletier, F. J. 2001. Did Frege Believe Frege's Principle? *Journal of Logic, Language and Information*, vol. 10, pp. 87-1114.
Rott, H. 2000. Words in Contexts: Fregean Elucidations. *Linguistics and Philosophy* vol. 23, pp. 621-641.
Salmon, N. 1994. Sense and Reference: Introduction. In M. Harnish, (ed.) *Basic Topics in the Philosophy of Language*, pp. 99-129. Prentice-Hall, Englewood Cliffs, NJ.
Szabo, Z. G. 2000. Compositionality as Supervenience. *Linguistics and Philosophy*, vol. 23, pp. 475-505.
Thomason, R. H. 1999. Type Theoretic Foundations of Context, Part 1: Contexts as Complex Type-Theoretic Objects. In P. Bouquet, L. Serafini, P. Brézillon, M. Benerecetti, and F. Castellani (eds.) *Modeling and Using Contexts: Proceedings of the Second International and Interdisciplinary Conference, CONTEXT'99*, pp. 352—374. Springer-Verlag, Berlin.

The Most Frequent Value Method in Groundwater Modeling

Péter Szűcs

University of Miskolc
Department of Hydrogeology
3515 Miskolc-Egyetemváros
www.hidrotanszek.hu
hgszucs@uni-miskolc.hu

The University of Oklahoma
School of Petroleum and Geological Engineering
Norman, OK 73019
http://mpge.ou.edu/
Adviser: Professor Faruk Civan

Abstract
The Most Frequent Value Method (MFV) is applied to groundwater modeling as a robust and effective geostatistical method. The Most Frequent Value method is theoretically derived from the minimization of the information loss called the I-divergence. The MFV algorithm is then coupled with global optimization (Very Fast Simulated Annealing) to provide a powerful method for solving the inverse problems in groundwater modeling. The advantages and applicability of this new approach are illustrated by means of theoretical investigations and case studies. It is demonstrated that the MFV method has certain advantages over the conventional statistical methods derived from the maximum likelihood principle.

1. Introduction

One of the main objectives of groundwater modeling is to determine the properly working earth models in order to adequately explain the hydrogeological observations. From the mathematical point of view, such solutions can be found by optimization (Lee, 1999). Frequently, the inverse methods are used to determine the optimal parameter values of the groundwater models. In Earth science applications, the objective functions may have multiple hills and valleys in the multi-dimensional parameter space. The conventional local search algorithms are usually trapped in one of the local minima instead of approaching the global minimum (Sen and Stoffa, 1995). Such limitations of the conventional methods for hydrogeological problems can be circumvented by the application of the global optimization methods. The calculated or theoretical data can be determined from the solution of mathematical models by assigning a set of prescribed values to the model parameters. This constitutes the forward problem. Besides the forward and inversion problems, the applied statistical principle is also a key factor in successful modeling as the objective or error functions are based on different statistical norms and principles. Unfortunately, the old dogma still exists in that the estimation or the measuring errors are approximately normally (Gaussian) distributed (Huber, 1981). Therefore, the application of the least-squares principles based on the maximum likelihood theory has been widespread even in geosciences. However, the efficiency of such classical algorithms is questionable when the actual error is not a Gaussian distribution. The Most Frequent Value (MFV) procedure (Steiner, 1991, 1997) has been introduced as a robust and resistant method for geostatistical data analysis and processing.

2. Theory of Inverse Procedures

A synthetic data set generated from a mathematical model using a set of assumed values for model parameters is compared with measured data. If the match is acceptable, the model parameter values are accepted as the best estimates. Otherwise, the parameters are modified to generate a new calculated data set and the quality of the match is investigated. This procedure is continued until a satisfactory match between the measured and calculated data is obtained. Therefore, the inverse procedures are usually regarded as optimization. The discrete data used in groundwater modeling is usually composed into a column vector as (Sen and Stoffa, 1995):

$$d_{measured} = [d_1, d_2, d_3, \ldots, d_{ND}]^T \quad (1)$$

where ND is the number of measured data, and *T* denotes a matrix transpose. The parameters of a groundwater model are also given in a column vector:

$$m = [m_1, m_2, m_3, \ldots, m_{NM}]^T \quad (2)$$

where NM is the number of model parameters. The calculated or synthetic data (d_{cal}) can be generated by the solution of the forward problem, namely the g-operator, as:

$$d_{cal} = g(m) \quad (3)$$

Generally, the forward problem operator is not linear in hydrogeology. The objective is to determine the best estimate values

of the model parameters, leading to the minimization of the difference between the measured ($d_{measured}$) and calculated (d_{cal}) data. The least-square norm, referred to as the L_2-norm, is the most common form derived from the L_p norms (Lines and Treitel, 1984):

$$\|e\|_2 = \left[\sum_{i=1}^{ND} |e_i|^2\right]^{1/2} \quad (4)$$

Weighted L_2-norms can also be used when there is additional information about the measurements. The particular type of norm used in modeling determines the effectiveness and accuracy of parameter estimation. As the measured data can be originated from a very wide range of distributions and some errors or outliers can also be expected, the application of the L_2-norm has certain disadvantages in Earth science applications (Sun, 1994). Hence, the use of the robust and resistant L_1-norm is more advantageous under these circumstances. However, the following P_k-norm, based on the Most Frequent Value (MFV) method (Steiner, 1991, 1997), provides additional advantages over the L_1- and L_2-norms, as a robust and resistant measure of the model

$$P_k = \varepsilon \left[\prod_{i=1}^{ND}\left(1 + \frac{(d_i^{measured} - d_i^{cal})^2}{(k\varepsilon)^2}\right)\right]^{\frac{1}{2ND}}$$

where □ denotes the scaling parameter or dihesion of the differences, as determined later. When the relationship between the model parameters and the calculated data is not linear, a suitable linearization method, such as based on the truncated Taylor series expansion, can be resorted to simplify the solution:

$$d_{measured} = d_{cal} + \left.\frac{\partial g(m_0)}{\partial m}\right|_{m=m_0} \Delta m$$

$$\Delta d = G_0 \Delta m,$$

where $\Delta d = {}_{dmeasured} - d_{cal}$. G_0 denotes a sensitivity matrix, including the partial derivatives of the calculated data with respect to various model parameters.

Inverse problems that do not possess uniqueness and stability are called ill-posed inverse problems. Otherwise the inverse problem is called well-posed. Nevertheless, techniques known as regularization can be applied to ill-posed problems to restore their being well-posed. Te L_2-norm yields a solution as:

$$\Delta m_{est} = [G^T G]^{1} G^T \Delta d \quad (8)$$

To characterize the model parameters obtained by the inversion, consider the covariance given by:

$$[\mathrm{cov}\, \Delta m_{est}] = \sigma_d^2 [G^T G]^{1} \quad (9)$$

where σ_d is the root-mean-square difference between the computed and measured data. The measured data can be weighted with a diagonal W-matrix if additional information is available for different observations. Applying the Marquardt – Levenberg algorithm and the weights, Eq. (8) can be modified as following to improve the search properties by an iterative procedure:

, $$\Delta m_{est} = \left[G^T G\right]^{1} G^T \Delta d$$

where α is called the Marquardt parameter, whose value gradually decreases to zero as the iteration progresses. Thus, initially the Marquardt–Levenberg method, frequently named as ridge-regression, operates based on the gradient principle. It then transforms into the Gauss-Newton method to seek an optimal solution. Although the Marquardt – Levenberg calculation can provide more stability, the effective operation still depends strongly on the initial guess assumed for the values of the model parameters for starting the iterative search. If the objective function has several local minima, the above-mentioned local search algorithms cannot provide the global minimum as a solution in case of a "bad" start of the parameter values search. This can be demonstrated by the following simple test problems. For example, consider the two-dimensional sinus cardinalis error function. This error-norm surface has several local minima and one global minimum location (x = 0, y = 0). Their locations are shown in Fig. 1. If the Levenberg-Marquardt algorithm is started from $x = 3.5$ and $y = 0$, the local minimum at $x = 3.53$ and $y = 0$ will be obtained as the solution. The experiment was repeated several times with different starting values to check the effectiveness of the global minimum search. The global minimum solution was obtained by the Levenberg-Marquardt method only when the starting point remained inside the "big pit". In contrast, the simulated annealing algorithm, described later, could easily solve this task without being trapped in the local minima locations. The solution $x = 0$, $y = 0$ was achieved in all cases regardless of the start of the model.

Complex groundwater models require

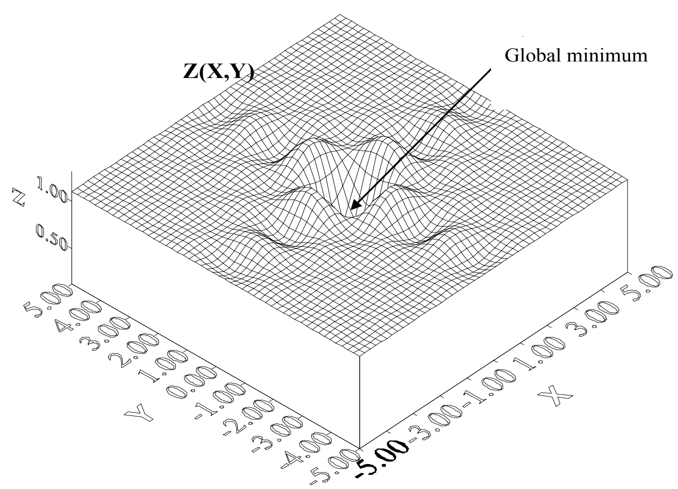


Fig. 1. Two dimensional error surface with several local minima and a global minimum at x = 0 and y = 0.

numerical approaches for evaluation of the partial derivatives in the above-mentioned G sensitivity matrix. Consequently, additional numerical errors would be involved in the inversion process during the inverse matrix calculation of an inverse problem. These drawbacks of the local search algorithms underline the advantages of the global optimization methods for applications not only in hydrogeology but also in different branches of geosciences (Szucs and Civan, 1996).

3. Global optimization and Simulated Annealing Methods

Besides the Genetic Algorithms (GA), the Simulated Annealing (SA) methods have been applied widely to seek for global optimum in different engineering and natural science problems (Sen and Stoffa, 1995). The works by Kirkpatrick et al. (1983) has shown that the model for simulating the annealing of solids, proposed by Metropolis et al. (1953), could be used for optimization problems, where the objective function to be minimized corresponds to the energy states of the solid and the control parameter corresponds to temperature, as defined later in the following. There are several modifications besides the classical Metropolis algorithms. The Very Fast Simulated Annealing (VFSA) method introduced by Ingber (1989) seems to be the fastest and most effective in multi-variable problems. Creating a classical Metropolis algorithm for a given groundwater modeling problem is relatively simple. The initial parameter vector is denoted as m_i. Consider the objective function (or error norm) denoted as $E(m_i)$. First, a new parameter vector (m_j) and the corresponding objective function $E(m_j)$ are generated. Then, the change in the value of the objective function, given as following, is examined:

$$\Delta E_{ij} = E(m)_j - E(m_i) \qquad (11)$$

If $\Delta E_{ij} \leq 0$, then the new m_j parameter vector is always accepted. Contrary, if $\Delta E_{ij} > 0$, then the probability of the acceptance of m_j parameter vector is determined using the Metropolis criterion, given by:

$$P = \exp\left(-\frac{\Delta E_{ij}}{T}\right),$$

where T corresponds to the temperature. This acceptance criterion provides an opportunity for avoiding entrapment in local minima. The temperature is decreased following a cooling schedule. An appropriate cooling schedule guarantees the convergent behavior of the method. Several studies have shown that decreasing temperature may result very rapidly in entrapment in a local minimum of the objective function (Sen and Stoffa, 1995). Typically recommended choice considers a temperature variation proportionally to $1/\ln(n+1)$ at the n-th iteration (Szucs and Civan, 1996). Usually, the model parameters in practical problems may have different finite ranges of variations and may affect the error function differently. Therefore, it is reasonable to allow the various model parameters different amounts of perturbations from their current positions. Hence, Ingber (1989) modified the Metropolis algorithm to elaborate the Very Fast Simulated Annealing (VFSA) method.

4. The Most Frequent Value (MFV) method

Besides a suitable optimization scheme, the formulation of an appropriate objective or error function also has a significant importance during any inverse calculation seeking for the best estimate values of the model parameters. The particular form of the statistical norm determines the performance of the optimization for a given error distribution. As proven previously by several geo-science applications and examples (Steiner, 1972, 1988, Ferenczy et al., 1990, Steiner and Hajagos, 1994, Szucs and Civan, 1996), the application of the MFV procedure provides several advantages over the least-squares or other conventional statistical techniques in hydrogeology and groundwater modeling.

The optimization objective of a groundwater modeling problem requires some kind of a norm of the residuals to be minimum. In most cases, the principle of the least-squares is applied. The classical statistics is based on this well-known principle, which can be easily formulated by the X_i residuals. Hence, the best model parameters set fulfils the minimum condition stated by:

$$\sum_{i=1}^{ND} X_i^2 = \text{minimum}.$$

Although this minimum condition is commonly used, it has several disadvantages concerning the effectiveness and outlier sensitivity. Steiner (1965) alleviated this difficulty by introducing a principle of maximum reciprocals as:

$$\sum_{i=1}^{ND} \frac{1}{X_i^2 + S^2} = \text{maximum},$$

where S is a scaling parameter, characterizing the measurement error. A comparison of the above-defined principles reveals that the outliers heavily influence Eq. (13). Large measuring errors associated with one or more X_i–s may lead to unreal or misleading results in some cases. The property of the proposed statistical procedure is referred to as resistance. Therefore, in this sense, the least-squares principle is not a resistant statistical procedure.

Applying the principle of the maximum reciprocals in a geostatistical analysis leads to the Most Frequent Value (MFV) technique (Steiner, 1988, 1990, Hajagos and Steiner 1991, Steiner, 1991, 1997). A statistical method is called an "MFV" technique if the X_i residuals are most frequently small (or even near zero) values. The condition stated by Eq. (14) forces the X_i residuals to be as small as possible in the overwhelming majority and it does not matter if therefore some X_i values become eventually very large. For example, also the following condition results in an MFV technique:

$$\prod_{i=1}^{ND} (X_i^2 + S^2) = \text{minimum} .$$

The MFV procedures are sometimes called "modern statistical methods" (Steiner, 1997), based on the idea first introduced by Steiner (1965). The hegemony of classical statistics even today can be perhaps excusable with the acceptance of the old dogma that "error distributions are always Gaussian" (Steiner and Hajagos, 1995). Szucs (1994) showed that the frequently used statistical hypothesis tests, like the chi-square test (the χ-test), might lead to greatly misleading results. Monte Carlo simulations proved that the χ-test could not be recommended for the normality tests of different distributions occurring in practice. Even when the data samples significantly deviate from the Gaussian distribution, the χ-test accepts them as normally distributed with high probabilities at the most frequently used significance levels. Assuming the observations to be normally distributed, the classical estimations are based on the maximum likelihood principle (MLE – Maximum Likelihood Estimators). The MFV algorithm follows a completely different theoretical approach. The MFV method tends to achieve minimization of the I-divergence (Steiner, 1997). I-divergence (information divergence) can be called as relative entropy and Kullback-Leibler distance, or a measure of information loss (Huber, 1981).

Dutter (1987) showed that this probability density would be the most representative type in the geosciences. Instead of the classical L_p-norms, the so-called P_k-norms (Eq. 5) can be defined based on the Most Frequent Value method. On the other hand, the MFV method is a very robust statistical procedure. This means that the choice of the k value in Eq. 5 affects the statistical efficiency very slightly. Therefore, only three different k values are proposed here. The value of k = 2 is recommended if no previous information exists about the type of the actual distributions. If short flanks are expected, then k = 3 should be used. If the actual distribution is of the Cauchy type, then k = 1 provides the best statistical efficiency (Steiner, 1991, 1997). It was further proven by Steiner (1991, 1997) that the MFV procedures are not only resistant but also robust. The attribute denoted as being "robust" generally indicates the efficiency of the statistical procedure is not very sensitive to the type of change. The estimates of T have a finite asymptotic variance, i.e. the law of large numbers is always fulfilled for the most frequent value calculations. The least-squares method does not satisfy this law if for example the error distribution is of the Cauchy type (Steiner, 1991, 1997). Steiner (1991, 1997) proved that L_1–methods (based on the median principle) have a general robustness (efficiency) of 50.1 % for the expected geosciences error distributions.

This is much higher than the general robustness of the classical statistics (based on the L_2-norm). The latter does not have a higher efficiency than 7.8 %. The general robustness of the MFV–methods are, however, always significantly greater than that of the L_1-procedures. The general robustness is higher than 90 % for the P_k-norms. The theoretically most adequate definition of the efficiency of an arbitrary statistical procedure is given by:

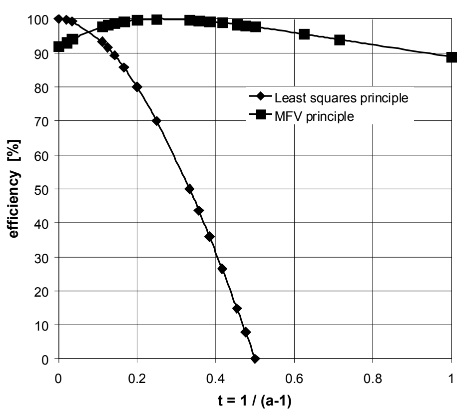
Statistical efficiency = 100 (extracted information / total information) %

Undoubtedly this definition provides a real measure of the statistical efficiency. However, a practically usable definition for numerical calculation of the statistical efficiency, denoted by e, is given by:

e=100 (minimum possible asymptotic variance / asymptotic variance) %

The denominator can be calculated for the actual applied statistical procedure. The nominator is the so-called Cramer-Rao bound, which can be found in almost every handbook of mathematical statistics, such as in Huber's (1981) famous book about robust statistics. Fig. 2 describes these efficiency relationships as a function of t = 1/ (a-1). This simple parameter transformation has a particular advantage in mapping the semi-infinite range of the supermodel parameter "a" value, varying from 2 to ∞, to the finite range of 0 to 1 for variation of the "t" value. The least-squares procedure works with 100 % efficiency if the error distribution is Gaussian. This is not surprising because the least-squares estimate is the best when the distribution is Gaussian. Unfortunately, its efficiency diminishes sharply to zero for error distributions having longer tails. Therefore, the least-squares principle should not be applied for any type of distribution other than the Gaussian type. In contrast, the MFV procedure is very highly efficient (> 90 %) regardless of the distribution type. The MFV procedure performs the best statistics for the geostatistical distribution (a = 5), where the efficiency value is 100 %. Therefore, the general high robustness of the MFV procedure is unarguable.

Fig. 2. Efficiency curves for the least squares and



MFV procedures for the f_a(X) supermodel

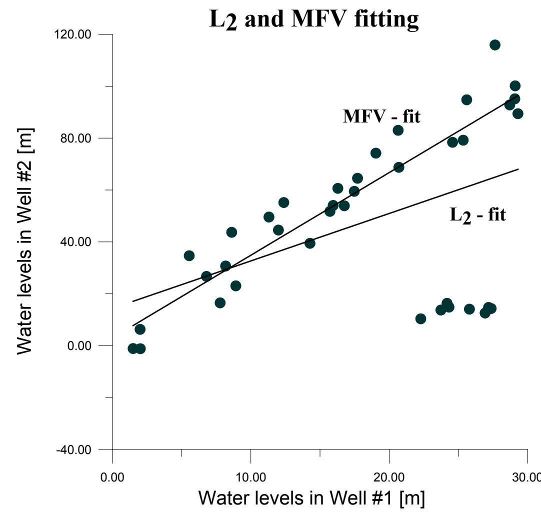
Besides their robustness, the MFV methods are also resistant. It can be seldom guaranteed that data are outlier-free. The appearance of the outliers may be very different (even rhapsodic). Statistical algorithms should be tested about their sensitivity or insensitivity (resistance) against the outliers. The outliers are due to measurement as well as model errors (Valstar et al., 2004). The above-mentioned double iteration process of the MFV methods guarantees a convergent solution to find the most frequent value and the dihesion or the inverse problem solution independently from the initial model parameter estimates

5. Applications for model and field problems

Although a natural science approach, that a groundwater reservoir is a general geologic medium, is very important in hydrogeology (Tóth, 1999), sophisticated mathematical and statistical methods are also inevitable to increase the accuracy and efficiency of the relevant interpretation processes. Marsily at al. (2000) present an outstanding review about the inverse problems in hydrogeology.

The MFV procedures may be facilitated for many problems in hydrogeology. Fig. 3 shows an example of a simple linear fitting of the water level data derived from a thick Pleistocene aquifer. Water levels were measured in two different wells, where there was a strong correlation between the levels because of hydraulic communication between the screened layers. This strong relationship was also given by the generalized and robustly correlation factor (Steiner, 1997). The traditional (Pearson-type) linear correlation factor showed only a weak relationship due to presence of the outliers. Fig. 3 indicates that the linear relationship based on the least-squares principle can be heavily influenced by the presence of outliers (produced artificially by human error in this case). Whereas, the MFV procedure clearly avoids the misleading bunch of data and provides a realistic linear physical relationship instead of a statistically distorted one.

Fig. 3. Linear regression for water level data



using the least-squares and MFV principles

Szucs (2002), and Szucs and Ritter (2002) applied the MFV procedures successfully for interpretations of several pumping tests. A well-defined geo-statistical method based on the MFV concept has been elaborated to determine the hydraulic parameters and their uncertainties. These are the necessary input data for a reliable groundwater modeling.

The suggested algorithm is well posed from the point of existence, uniqueness, stability, and robustness. This new evaluation method is proven and validated for different pumping test interpretation methods (Kruseman and Ridder, 1990). The main advantage of the suggested inverse procedure is that the uncertainty or the reliability of the hydraulic model parameters can also be determined by the MFV procedure and Monte Carlo simulations using one set of measured field data (Fig. 4). Artificial measuring errors are superimposed to the original measurements, and the inversion process is repeated several times. The applicability and usefulness of the introduced procedure have been demonstrated by means of several case studies on the ground water modeling problems concerning the Northern Hungarian region (Szucs, 2002, Szucs and Ritter, 2002).

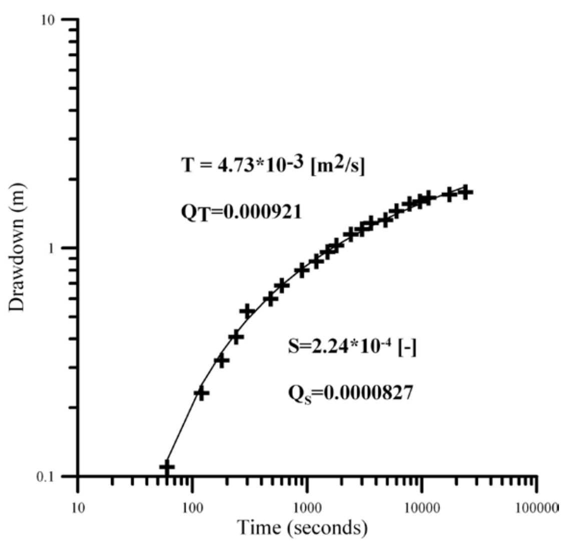


Fig. 4. Results of a pumping test interpretation using the MFV method. The uncertainties of the hydraulic parameters (QT and QS) are also determined.

The following case studies give simple examples of the MFV inverse applications to improve groundwater modeling calibration results. The hydraulic head prediction that comes from a flow model is commonly used as the basis for model calibration. Calibration is a process of adjusting the model parameters to achieve a satisfactory match between the predicted (or calculated) and measured hydraulic heads (Hill, 1998). Practically, calibration is an inverse process. Frequently, the objective function used as a calibration criterion is based on the mean error, the mean-absolute error (L_1 norm), or the root-mean square error (RMSE error, L_2 norm) (Anderson and Woessner, 1992).

Szucs and Ritter (2002) introduced the above-mentioned P_k-norm for groundwater model calibration purposes. Because the real data, and the error or residual distribution can never be known in advance, the usage of $P_{k=2}$-norm is most favorable for groundwater modeling. To demonstrate the advantage of the MFV procedure and global optimization in groundwater modeling, two main examples are provided here.

5.1 Test problem

A simple one-layer unconfined steady-state groundwater model has been facilitated to describe and investigate the behavior of the proposed global optimization (SA) method and the MFV procedure. The x-y dimension of the test model is 1 km by 1 km. The top of the model layer is on 25 m. The bottom of the model layer is 0.0 m. The basic grid size is 20 m. A constant recharge rate at 0.0003 $m^3/(m^2day)$ was applied on the top of the grid system. Four polygons were delineated to represent the layer heterogeneity in the aquifer. The horizontal hydraulic conductivity is assumed to be constant in each polygon. Specified head boundary conditions were introduced on the west and east borders to simulate the natural groundwater flow from west to east. One production well was seated in each of polygon I (-400 m^3/s), II (-500 m^3/s) and III (-300 m^3/s). There is no well in polygon IV. As over-determined systems are preferred for any statistical interpretation, 12 observation points were stationed in the model for the groundwater calibration.

Creating a flow model based on the actual model parameters is called a forward solution. The water levels could be derived exactly for the 12 observation points. To simulate real measured water level data at the observation points, 2 % percent random geostatistical error was superimposed on the exact water levels. Having a pre-defined hydrogeological model and the "measured data set", the inverse investigations could be started. The GMS 4.0 system provides three built-in possibilities for automated inverse parameter estimation. These are the PEST (Doherty, 2000), the UCODE (Poeter and Hill, 1998), and the MODFLOW-2000 PES (Hill at al., 2000) procedures. The MODFLOW-2000 PES method has been selected for comparison of the investigations with the present MFV based inverse algorithm using a global (Metropolis Simulated Annealing) optimization (noted as MFV – SA). The MFV – SA inverse method has been also linked to the popular MODFLOW-2000 package, which provides the forward solution. In addition to the well-described error functions (RMSE and P-norm), the relative model distance (RM) has also been used to characterize the accuracy of the compared inversion procedures.

The present application of the MFV procedure utilized the classical Simulated Annealing global optimization search.. For illustration, the Metropolis (SA) algorithm was applied with the parameter values given as follows. The initial temperature is T_0 = 1.0. The final temperature is T_f = 0.0001. The temperature reduction constant is α = 0.975. The number of iterations at each temperature is R(t) = 300. Table 1 gives a summary of the most important results obtained by the MODFLOW-2000 PES and MFV+SA algorithms. The results clearly indicate a great difference in the relative model distance (RM) values although the objective function values (RMSE and P-norm) are not far from each other. The relative model distance (RM=0.58, MODFLOW-2000 PES) reduces by half when the MFV based inverse procedure is applied (RM=0.27). Fig. 5 also indicates the advantage of the Most Frequent Value approach. Fig. 5 shows nearly the same flow pattern as that of the original model. Note that even the MFV – SA method was not able to give back the original model parameters. This is truly understandable because a complication in groundwater problems arises when the information about the head distributions is incomplete (Anderson and Woessner, 1992).

Table 1. The main results of the inverse proce-

dures carried out by MODFLOW-2000 PES and MFV-SA methods in case of 2 % geostatistical distribution error added to the theoretical heads at the observation points.

Test problem investigated by different inversion methods			
Model Polygon	Prescribed model parameters	Model parameters from inversion	
		MODFLOW-2000 PES	MFV - SA
I.	25 m/day	11.52 m/day	18.72 m/day
II.	35 m/day	27.65 m/day	32.14 m/day
III.	15 m/day	6.46 m/day	10.92 m/day
IV.	10 m/day	1.90 m/day	7.38 m/day
Error function		RMSE = 0.203 m	P-norm = 0.172 m
Relative model distance		RM = 0.58	RM = 0.27

5.2 Field problem

In general, the field experts prefer to use the commercially available professional groundwater modeling packages for hydrogeological evaluation and interpretation, such as the above-mentioned Groundwater Modeling System (GMS 4.0) or the Processing Modflow (Chiang and Kinzelbach, 2001). Although these packages have built-in inverse modules like PEST, UCODE or MODFLOW-2000 PES, the trial-and-error calibration is still preferred in many cases because the modeler's expertise and experience can be involved in the process easily. In the following, the advantage of the MFV-based inverse groundwater modeling is demonstrated by a field example.

There is an ongoing national project supported by the Hungarian government to delineate the wellhead protection zones for vulnerable groundwater resources. The particle tracking MODPATH module (Pollock, 1994) enables the delineation of the wellhead protection zones around the investigated production wells. During each step of the trial-end-error calibration, the MFV weights can provide very visible and useful information for every observation point about the actual groundwater model condition concerning the strength of matching. The closer the MFV weight is to 1.0, the better the match between the measured and calculated head data for the actual observation points. Besides the individual weight interpretation, the histogram of the MFV weights can also give useful insight about the state of calibration. Fig. 6 shows that the histogram has high relative frequency values at small MFV weights during the beginning of the calibration when the model parameter values are far from their real values. If the calibration is carried out successfully and the measured data are reliable, the histogram should reflect highly on the relative frequency at the greater intervals of MFV weights.

6. Conclusions

It has been demonstrated that the Most Frequent Value method can be applied successfully for effective solution of various problems involving groundwater modeling under certain conditions, such as when the measurement errors are not Gaussian and the model concept errors are insignificant. This robust and resistant geostatistical procedure provides a high general efficiency.

The application of the P-norms based on the MFV principle has been shown to be advantageous over the other types for inverse parameter estimation calculations. The automated parameter estimation method facilitating the Most Frequent Value method and linked to the MODFLOW – 2000-reference flow code has been shown to be effective for deriving the groundwater model parameters. The use of the Most Frequent Value weights of the head residuals readily improves the groundwater interpretation results during traditional trail-and-error calibration processes. The present study has proven that the MFV method provides certain advantages over the conventional statistical methods derived from the maximum likelihood principle. Consequently, the application of the MFV method coupled

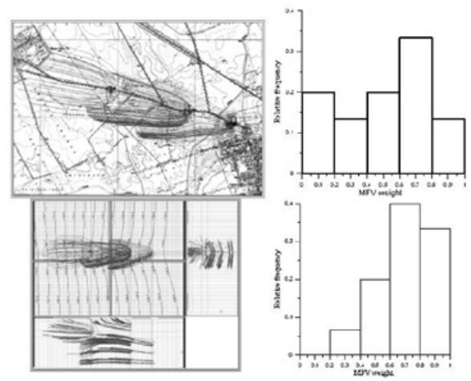


Fig. 6 Histograms of the MVF weights during the calibration process. The upper histogram shows an early stage and the lower histogram reflects the end of the trial-end error calibration.

with global optimization is expected to become a more widespread practice in groundwater modeling. The author gratefully acknowledges the Fulbright Scholarship Program, the Bolyai Janos Research Scholarship of the Hungarian Academy of Sciences, and the Mewbourne School of Petroleum and Geological Engineering at the University of Oklahoma for support of this work.

Bibliography

Anderson M. P. and Woessner W.W., 1992: Applied Groundwater Modeling. Academic Press, San Diego, Calif., 381 p.

Chiang W.H. and Kinzelbach W., 2001: 3D-Groundwater modeling with PMWIN. A simulation system for modeling groundwater flow and pollution. Springer-Verlag, 346 p.

Doherty J., 2000: PEST, Model-Independent Parameter Estimation, fourth edition, program documentation, Watermark Numerical Computing, p. 249.

Dutter R., 1987: Mathematische Methoden in der Montangeologie. Vor-lesungsnotizen. Manuscript, Leoben.

EMRL, Environmental Modeling Research Laboratory of Brigham Young University (2002): Groundwater Modeling System (GMS 4.0), Tutorial Manual.

Ferenczy L., Kormos L. and Szucs P., 1990. A new statistical method in well log interpretation, Paper O, in 13th European Formation Evaluation Symposium Transactions: Soc. Prof. Well Log Analysts, Budapest Chapter, 17pp.

Hajagos B. and Steiner F., 1991. Different measures of the uncertainty. Acta Geod. Geoph. Mont. Hung., 26: 183-189.

Hajagos B. and Steiner F,. 1995. Symmetrical stable probability distributions nearest lying to the types of the supermodel $f_a(x)$. Acta Geod. Geoph. Hung., Vol. 30., (2-4), pp. 463-470.

Hill M.C., 1992: A computer program (MODFLOWP) for estimating parameters of a transient, three-dimensional ground water flow model using nonlinear regression. U.S. Geological Survey, Open-File Report 91-484.

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

Hill M.C., Banta E.R., Harbaugh A.W., and Anderman E.R., 2000: MODFLOW-2000, The U.S. Geological Survey Modular Ground-Water Model – User Guide to the Observation, Sensitivity, and Parameter-Estimation Processes and Three Post-Processing Programs. U.S. Geological Survey, Open-File Report 00-184.

Huber P.J., 1981. Robust statistics. Wiley, New York, NY, 308 pp.

Ingber L., 1989: Very fast simulated reannealing. Mathl. Comput. Modeling, 12, 8, pp. 967-993.

Kirkpatrick S., Gelatt C.D., Jr., and Vecchi M.P., 1983: Optimization by simulated annealing. Science, 220, 671-680.

Kitandis P.K., 1997: Introduction to geostatistics: Applications to hydrogeology. Cambridge University Press, p. 249.

Kruseman G.P. and Ridder de N.A., 1990: Analysis and Interpretation of Pumping Test Data. Publication 47, International Intsitute for Land Reclamation and Improvement, Wageningen, The Netherlands, p. 1-375.

Lee T-C., 1999: Applied Mathematics in Hydrogeology. Lewis Publishers and CRC Press LLC, ISBN 1-56670-375-1.

Lines T.R. and Treitel S., 1984. Tutorial: A review of least squares inversion and its application to geophysical problems. Geophysical Prospecting, 32: 159-186.

Marsily de Gh., Delhomme J.P., Coundrain-Ribstein A. and Lavenue A.M., 2000: Four Decades of Inverse Problems in Hydrogeology. Paru dans Geophysical Society of America, Special paper 348, pp. 1-28.

Marquardt D.W., 1970: Generalized inverses, Ridge regression, biased linear estimation, and nonlinear estimation. Techometrics 12: pp. 591-612.

Metropolis N., Rosenbluth A., Rosenbluth M., Teller A., and Teller E., 1953: Equations of state calculations by fast computing machines. J. Chem. Phys., 21, 1087-1092.

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

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.

Sen M., Stoffa P.L., 1995: Global Optimization Methods in Geophysical Inversion. Advances in Exploration Geophysics 4. Elsevier.

Steiner F,. 1965. Interpretation of Bouguer-maps (in Hungarian). Dissertation. Manuscript. Miskolc, 80-94.

Steiner F., 1972. Simultane interpretation geophysikalischer messdatensysteme. Review Pure and Applied Geophysics, 96: 15-27.

Steiner F., 1990. The Bases of Geostatistics (in Hungarian). Tankonyvkiado, Budapest, Hungary, 363pp.

Steiner F., (Editor), 1991. The Most Frequent Value. Introduction to a Modern Conception Statistics. Akademia Kiado, Budapest, Hungary, 314pp.

Steiner F. and Hajagos B., 1994. Practical definition of robustness. Geophys. Trans., 38: 193-210.

Steiner F. and Hajagos B. 1995. Determination of the parameter errors (demonstrated on a gravimetric example) if the geophysical inversion is carried out as the global minimization of arbitrary norms (demonstrated by the P_c norm). Magyar Geofizika, Vol. 36: 261-276.

Steiner F., (ed.) 1997: Optimum methods in statistics. Akademia Kiado, Budapest.

Sun N-Z., 1994: Inverse Problems in Groundwater Modeling. Kluwer Academic Publishers.

Szucs P., 1994. Comment on an old dogma: 'the data are normally distributed'. Geophysical Transactions 38: 231-238.

Szucs P., Civan F., 1996: Multi-layer well log interpretation using the simulated annealing method. Journal of Petroleum Science and Engineering, 14, pp. 209-220.

Szucs P., 2002: Inversion of pumping test data for improved interpretation. microCAD 2002, International Scientific Conference, University of Miskolc, A: Geoinformatics, pp. 107-112, March 7-8.

Szucs P. and Ritter Gy., 2002: Improved interpretation of pumping test results using simulated annealing optimization. ModelCARE 2002, Proceedings of the 4th International Conference on Calibration and Reliability in Groundwater Modeling. Prague, Czech Republic, 17-20 June 2002. Acta Universitas Carolinae – Geologica 2002, 46 (2/3), pp. 238-241.

Tóth J., 1999: Groundwater as a geologic agent: An overview of the causes, processes, and manifestations. Hydrogeolgy Journal (7), pp. 1-14.