Comparison of parameter estimation algorithms in hydrological modelling

R. S. BLASONE[1], H. MADSEN[2] & D. ROSBJERG[1]

1 *Environment & Resources, Technical University of Denmark, Bygningstorvet, Building 115, Room 156, DK-2800 Kongens Lyngby, Denmark*
rsb@er.dtu.dk

2 *DHI Water & Environment, Agern Allé 5, DK-2970 Hørsholm, Denmark*

Abstract Local search methods have been applied successfully in calibration of simple groundwater models, but might fail in locating the optimum for models of increased complexity, due to the more complex shape of the response surface. Global search algorithms have been demonstrated to perform well for these types of models, although at a more expensive computational cost. The main purpose of this study is to investigate the performance of a global and a local parameter optimization algorithm, respectively, the Shuffled Complex Evolution (SCE) algorithm and the gradient-based Gauss-Marquardt-Levenberg algorithm (implemented in the PEST software), when applied to a steady-state and a transient groundwater model. The results show that PEST can have severe problems in locating the global optimum and in being trapped in local regions of attractions. The global SCE procedure is, in general, more effective and provides a better coverage of the Pareto optimal solutions at a lower computational cost.

Keywords automatic calibration; hydrological modelling; MIKE SHE model; optimization algorithms; parameter estimation

INTRODUCTION

Local gradient-based search methods have been widely applied in calibration of groundwater models. In more complex models the increased nonlinearity in the parameter-model response provides a more complex shape of the response surface and local search procedures can be trapped in local optima, and thus be unable to reach the global optimum. Global population-evolution based search algorithms, such as genetic algorithms and the Shuffled Complex Evolution (SCE) algorithm, have been demonstrated to perform well for these types of models, although at a more expensive computational cost.

The main purpose of this study is to investigate the performance of local and global parameter optimization algorithms when applied to hydrological models with different degrees of complexity. Two different search algorithms, a local and a global procedure, are applied to a steady-state model and a transient-state groundwater model and their efficiency and effectiveness are compared.

STUDY CATCHMENT

The study site is the Karup catchment in the western part of Denmark, which is drained by the Karup River and about 20 tributaries. The catchment has an area of 440 km^2 and is characterized by a quite homogeneous geology of predominantly sandy soils with high permeability. The aquifer is mainly unconfined with a variable thickness ranging from 90 m at the upstream part of the catchment, to 10 m in the western and central areas.

The Karup catchment has been extensively investigated in previous studies (Refsgaard, 1997; Madsen, 2003), and a comprehensive hydrological database of rainfall from nine stations (daily values), runoff at the river outlet (daily values), groundwater elevation data from 17 wells (recorded every 15 days) and temperature (daily values) is available. The data used for the calibration of the models cover a 6-year period, from 1 January 1969 to 1 January 1975.

MODEL SET-UP AND CALIBRATION PARAMETERS

The software used in the study is the integrated MIKE SHE modelling system (Refsgaard & Storm, 1995), which has been used to model both the steady-state and the transient model of the Karup catchment.

The horizontal computational grid is defined with a spatial scale of 1 × 1 km, and the geology of the saturated zone is represented with vertical and horizontal scales of, respectively, 10 m and 1 × 1 km. The geology is taken from the Danish National Water Resources model. For each grid element a soil type is assigned with a given code and hydrogeological parameters. Six general soil types are defined (Table 1). A 2-D model is applied and hence one computational layer is defined for the saturated zone.

Table 1 Soil types.

Soil code	Soil name	Description
1	Melt water sand	Quaternary and Post-Glacial sand and gravel
2	Clay	Glacial, Inter-Glacial and Post-Glacial clay and silt
3	Quartz sand	Miocene, medium to coarse grained sand and gravel
4	Mica sand	Miocene, fine to medium grained sand
5	Mica clay/silt	Pre-Quaternary clay and silt
6	Limestone	Limestone

Preliminary sensitivity analysis shows that only the hydrogeological parameters of two of the soil types are important (melt water sand and quartz sand). The horizontal conductivities of these are subject to calibration. The vertical conductivities have been set to one tenth of the respective horizontal hydraulic conductivities.

Overland flow is only generated when the groundwater level rises and reaches the surface level. Surface runoff is routed down-gradient towards the river system using the diffusive wave approximation of the Saint Venant equations. The exact route and quantity is determined by the topography and flow resistance as well as the losses due to evaporation and infiltration along the flow path.

A drainage system is defined that includes both natural and artificial drainages in the catchment. Drainage flow is simulated using an empirical formula, which for each cell requires a drainage level and a time constant (drainage coefficient) that regulate how much and how fast water is drained. Both of these model parameters are assumed to be uniformly distributed in the catchment and are subject to calibration.

The river system collects overland and saturated zone flow. The stream–aquifer interaction is accounted for by a leakage coefficient. This coefficient is assumed to be constant for all river branches and is subject to calibration.

An empirical root zone model is applied to calculate the recharge to MIKE SHE (Henriksen, 2002) based on observed precipitation and potential evapotranspiration, land use information and estimates of the field capacity.

The five above-mentioned model parameters have been selected for automatic calibration, while the others were fixed to their previously manually-calibrated values.

The transient model has the same parameterization and recharge conceptualization as the steady-state model, but the processes are modelled in transient mode.

OPTIMIZATION PROCEDURES

The global search methodology applied is the Shuffled Complex Evolution (SCE) algorithm (Duan *et al.*, 1992) implemented in the AUTOCAL software (DHI, 2004). The SCE algorithm is an evolutionary-based procedure that simultaneously evolves a population of solutions (parameter sets) towards better solutions in the search space, trying to converge to the global optimum of the objective function.

The local method employed is the Gauss-Marquardt-Levenberg nonlinear scheme, as implemented in the PEST software (Doherty, 2004). The Levenberg-Marquardt algorithm is a gradient-based optimization strategy that combines the Gauss-Newton algorithm and the method of gradient descent, and it provides a numerical solution to the mathematical problem of minimizing a sum of squared deviations between model outcomes and corresponding field data.

The Latin Hypercube Sampling (LHS) approach (McKay *et al.*, 1979) is employed to generate an initial population of 33 parameter sets which guarantees a good coverage of the parameter space. The size of this initial population is fixed by the choice made for the SCE algorithm parameters (three complexes with 11 points in each complex). These points are then used as initial population for the SCE algorithm, while 33 independent PEST runs are conducted starting from each of these points separately.

MULTI-OBJECTIVE CALIBRATION

The calibration is conducted in a multi-objective context, i.e. different optimization criteria are used to perform the search in the feasible parameter space. In this study the sum of the mean squared errors related to groundwater levels (m) at 17 locations within the Karup catchment, MSE^{wells}, and the streamflow (m^3 s^{-1}) at the river outlet, MSE^{runoff}, are aggregated into a single objective function, F_{aggr}, as given in equation (1). The aggregated measure is a function of the parameters, θ, and of the weights assigned to the objective functions of the river discharge, w^{runoff}, and of the aggregated wells, w^{wells}.

$$F_{aggr}(\theta) = w^{runoff} MSE^{runoff}(\theta) + w^{wells} \sum_{i=1}^{17} MSE_i^{wells}(\theta) \qquad (1)$$

The use of different combinations of weights allows accounting for the different scales of magnitude of the quantities in the aggregated measure and broadens the exploration of the solution space. Therefore, several optimization runs have been conducted assigning different weights to the objective functions of the river discharge and of the aggregated wells. The weights applied are shown in Table 2.

Convergence conditions in the objective function space and in the parameter space have been adopted as stopping criteria for the two search procedures.

The equifinality problem and the trade-offs between objective functions are tackled by evaluating the solutions according to the Pareto dominance criterion (Madsen, 2003). While the solution of the local search algorithm provides one result for each optimization run, thus producing 33 solutions, the global procedure is expected to provide an approximation of the Pareto front in the vicinity of the global optimum. The performance of the local and global search procedure is compared in terms of the estimated Pareto front in the objective function space.

RESULTS

The optimization results are summarized in Table 2 where the number of Pareto optimum solutions for the two optimization procedures are shown.

Table 2 shows that the performance of the local procedure is highly affected by the weights assigned to the objective functions. None of the calibration runs of the steady-state model converges to a Pareto optimum if both objective functions are simultaneously minimized.

Table 2 Quantity of Pareto optimal solutions found within the different calibration experiments.

	No. of runs	No. of optima	$w^w = 1$ $w^r = 0$	$w^w = 1$ $w^r = 1$	$w^w = 0.24$ $w^r = 1$	$w^w = 0.06$ $w^r = 1$	$w^w = 0.03$ $w^r = 1$	$w^w = 0$ $w^r = 1$
Steady-state model								
PEST	15258	14	8	0	0	0	0	6
SCE	4279	396	125	170	3	70	17	11
Transient-state model								
PEST	16332	104	19	27	24	20	8	6
SCE	4097	181	8	15	49	5	65	39

No. of runs is total number of model runs conducted for each type of model to find optimal points.
No. of optima is number of optimal points (Pareto solutions) found within the total no. of model runs.
w^w, w^r are the weights w^{wells} and w^{runoff} of equation (1). The quantity in each column represents the amount of optimal points found within a calibration experiment applying the objective functions weights w^w and w^r.

The inability of the local procedure to estimate the Pareto front is emphasized in Fig. 1, where the solutions of the 198 independent optimization runs (six combinations of weights × 33 starting points) for the steady-state model are plotted. In particular, the local optimization has a tendency to be trapped in sub-optimal regions of attraction, as seen in Fig. 1 by the two clouds with high concentration of solutions for a runoff MSE error greater than 7 $(m^3 s^{-1})^2$.

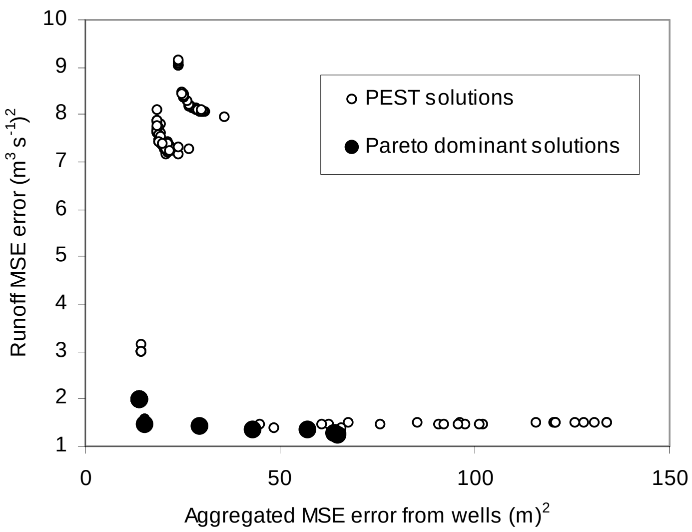

Fig. 1 PEST results for the calibration of the steady-state model.

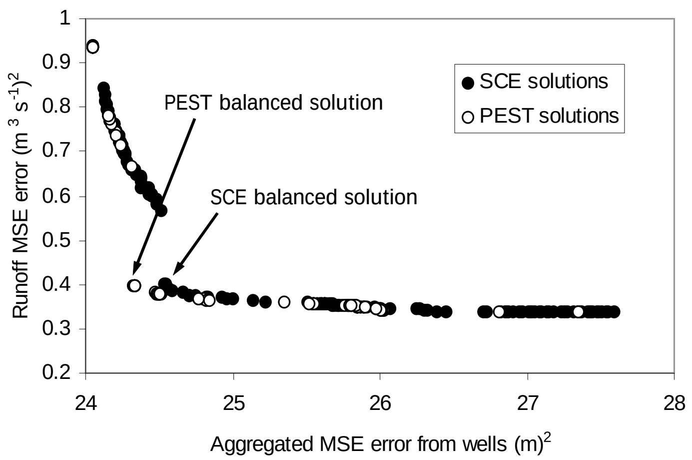

Fig. 2 Pareto front plot for the transient-state model.

In Fig. 2 the calibration results obtained for the transient-state model using the two procedures are compared in terms of estimated Pareto fronts. Despite the fact that both methods fail to estimate a particular region of the front (probably due to the inappropriate or incomplete choice of the combination of objective functions weights), the capability of SCE in exploring the objective function space and estimating the Pareto front is evident.

It must be noted that these results are achieved by a number of Pareto solutions that is about twice the number of points found by PEST and a total number of model runs that is about four times smaller than that required by the local method (results in Table 2).

Despite the better performance of SCE in exploring the solution space with higher efficiency, some of the Pareto dominant solutions obtained for the steady-state model with the local search procedure are better than those found with SCE (results not shown herein). Also, for the transient-state model some of the local search solutions are better than the SCE solutions (as shown in Fig. 2). This demonstrates that local search methods can be quite effective in parameter estimation, but at the expense of a high computational cost if a complete exploration of the solution space has to be conducted.

A "balanced optimum solution" has been chosen for the two different optimization algorithms with the transient-state model as a compromise between the two objective functions. The solution of the local method (indicated in Fig. 2 as "PEST balanced solution") appears the best compromise between the two objective functions.

SCE provides a better exploration of the solution space and a more complete estimation of the Pareto front, although it may be less effective at certain points. However, a better performance can be obtained by increasing the population size in SCE.

CONCLUSIONS

The results show that PEST can have severe problems in locating the global optimum and in being trapped in local regions of attractions, as especially demonstrated by the steady-state model calibration. The global SCE procedure is, in general, more effective and provides a better coverage of the Pareto optimal solutions at a lower computational cost. However, SCE may also be trapped in local points of attraction.

A future issue for this research is to extend the comparison of performance of the SCE algorithm and the Gauss-Marquardt-Levenberg scheme to a more complex and fully integrated MIKE SHE model of the same catchment.

Another issue to investigate is the combined application of the global and local search techniques, using the global method as an initial screening procedure with which to approach the Pareto front and subsequently employing the local method to refine the estimation of the optima. In this way the effectiveness of the local procedure can be reached at a lower computational cost.

REFERENCES

DHI (2004) *AUTOCAL, Autocalibration Tool, User Guide*. DHI Software 2004, DHI Water & Environment, Denmark.

Doherty, J. (2004) *PEST: Model-Independent Parameter Estimation, User Manual, Version 5*. Watermark Numerical Computing, Brisbane, Australia.

Duan, Q., Sorooshian, S. & Gupta, V. (1992) Effective and efficient global optimization for conceptual rainfall–runoff models. *Water Resour. Res.* **28**(4), 1015–1031.

Henriksen, H. J. (2002) Water balance on the basis of results from the DK-model. In: *Water Balance at Field and Catchment Scale* (ed. by F. Plauborg, J. C. Refsgaard, H. J. Henriksen, G. Blicher-Mathiasen & C. Kern-Hansen). DJF Rapport 70, Danish Institute of Agricultural Sciences, Tjele, Denmark.

Madsen, H. (2003) Parameter estimation in distributed hydrological modelling using automatic calibration with multiple objectives. *Adv. Water Resour.* **26**, 205–216.

McKay, M. D., Conover, W. J. & Beckman, R. J. (1979) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21**, 239–245.

Refsgaard, J. C. (1997) Parameterisation, calibration and validation of distributed hydrological models. *J. Hydrol.* **198**, 69–97.

Refsgaard, J. C. & Storm, B. (1995) MIKE SHE. In: *Computer Models of Watershed Hydrology* (ed. by V. P. Singh), 809–846. Water Resources Publications, Colorado, USA.