(Proceedings ModelCARE2011 held at Leipzig, Germany, in September 2011) (IAHS Publ. 355, 2012).

A two-stage strategy for efficient and effective calibration of distributed hydrological models

C. FISCHER, S. KRALISCH, P. KRAUSE & W.-A. FLÜGEL

Department of Geoinformatics, Hydrology and Modelling, School of Chemical and Earth Sciences, Friedrich-Schiller-University, Jena, Germany
christian.fischer.2@uni-jena.de

Abstract Calibration of distributed environmental models requires a high amount of either computing power or time, which can be a significant issue. As the computational effort for distributed models mainly depends on the number of spatial modelling entities, a promising approach for their optimization is to decrease these entities while sustaining the major characteristics of the model. In this article, a two-stage approach is presented to reduce the computational effort. In the first stage the number of spatial units is decreased, thus simplifying the original representation of the catchment. Spatial units are eliminated or merged with other units according to different merging rules (e.g. merging of similar units). The simplified model is used to carry out an initial calibration. In the second stage the process of simplification is reversed, i.e. the spatial representation of the catchment is restored stepwise. The obtained parameter sets are recalibrated. This is reiterated until the original distribution is recovered. To test and analyse this strategy, the distributed model J2000 is applied on the two meso-scale catchments of the Wilde Gera and Ilm, both located in central Germany. Furthermore, variations of the approach to simplify the spatial representation are analysed.

Key words calibration; hydrology; spatial representation

INTRODUCTION

Analysis of the spatial distribution of hydrological states in a catchment usually requires utilization of semi-distributed or distributed hydrological modelling concepts. Distributed hydrological models subdivide the catchment into a number of spatially discrete modelling units (i.e. sub-areas). These units may either be hydrological response units, sub-catchments, square grid elements or even arbitrary triangles. Hydrological processes are simulated separately for each unit but interactions between different units are possible in most models (Reed *et al*., 2004).

Therefore, computational effort scales at least linearly with the number of spatial modelling units. This is rarely a limitation for a sole model execution, but to apply distributed hydrological models successfully it is essential to determine model parameters very carefully. The estimation of parameters is usually done by fitting model response to observed data in a trial-and-error process (Gupta *et al*., 2005).

Obtaining a solution of this nonlinear optimization problem could be difficult and time consuming. It is known that objective functions could show inconsistencies of second and sometimes even first grade, and are often superimposed with noise (Duan *et al*., 1993).

Thus, computational effort limits either the size of the representable catchment or its level of detail. This can be an issue if the scale of the spatial modelling units is above the scale of the process-scale, i.e. the scale which is exhibited by the natural processes. If this is the case, a direct physical-based simulation of these processes is hardly possible such that only a reflection of their statistical properties can be made (Blöschl, 1995).

To speed up the optimization procedure a common approach is to replace the initial problem by some relaxed sub-problems that are easier to solve (e.g. Horst & Tuy, 1996). The approach presented in this paper is to carry out the major part of the calibration routine on a simple model using only a reduced set of spatial modelling units and to transfer the attained parameters to the original model. There are two issues about this strategy. Firstly, it is not clear whether or not the substitution model is able to represent the hydrological processes well enough such that it is possible to apply this strategy efficiently. Secondly, the need for parameter scaling arises due to transferring model parameters. Several authors pointed out that this is usually a complicated task (e.g. Beven, 1989, Refsgaard, 1997). A promising approach is introduced by Samaniego *et al*. (2010). This article shows that the first issue is not a problem at all, at least with the hydrological

model J2000 which has been used for testing. The second issue is solved by stepwise refining of the coarse model until the initial model is restored. At each step of refinement a short recalibration is carried out. Although this approach is demonstrated with J2000, it is applicable to other distributed environmental models.

METHODOLOGY

Hydrological model J2000

The distributed, process oriented model J2000 was developed for hydrological simulation of the upper meso- and macro scale (Krause, 2001). It is implemented in the Jena Adaptable Modelling System (JAMS, Kralisch *et al.*, 2006), which is a software framework for component based development and application of environmental models. The model describes single hydrological processes as encapsulated process modules. Beside modules for simulation of the runoff generation and runoff concentration processes, J2000 offers routines for regionalization and correction of climate and precipitation input data, model calibration and visualization (Kralisch *et al.*, 2007).

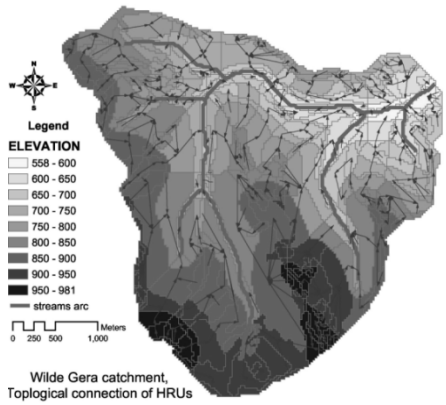


Fig. 1 the catchment is partitioned into a set of HRUs, which are laterally connected according to their flow topology.

J2000 uses Hydrological Response Units (HRUs) for the spatial representation of a catchment. HRUs are homogenous with regard to their physiography (e.g. topographic features, land use, soil properties). Therefore it is possible to assign a characteristic hydrological process response to each HRU (Flügel, 1996). They are connected by a lateral routing scheme to simulate lateral water transport processes (see Fig. 1) either with their downhill successor HRU or with a river reach they are draining into. River reaches themselves are always connected with their downstream reach. The model simulates relevant hydrological processes like evapotranspiration, snow accumulation and melt, soil water balance and groundwater processes for each HRU separately.

HRUs are delineated by combining different spatial data layers by an overlay analysis such that each HRU is spatially defined by a polygon. Advantages of this polygon based concept of HRUs is the conservation of the natural properties of the catchment, like topography, land use, soils and geology during the discretization process, which also results in a comparatively small number of modelling units (compared for example with raster based approaches).

Speeding up calibration

Let $D_{sim} \ni S = \mathcal{M}(I, \theta, \tau)$

be a hydrological model $\mathcal{M}$, mapping input data I into some output space D_{sim} and $\theta \in P$ parameters to adjust the model to the catchment characteristics. Furthermore, let $\tau \in [0,1]$ be the *relaxation parameter* controlling the *spatial complexity* of the model. If τ is set to zero the model uses a very simple spatial representation consisting only of a single unit. On the opposite the initial fully detailed model is used if τ equals one. To generate spatial representations with different complexity, four different alternatives are described in detail later.

To calibrate the model $\mathcal{M}$ some observation data O are required, which will be compared with the simulated data. In this paper the similarity between both datasets is quantified by the well-known *Nash-Sutcliffe Efficiency* E_2 (Nash & Sutcliffe, 1970). For simplicity the aim of the process of calibration is simply to find a parameter set $\theta^* \in P$ such that

$$E_2\left(\mathcal{M}\left(I, \theta^*, 1.0\right), O\right) \rightarrow max$$

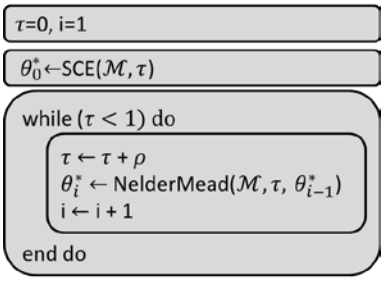
The calibration strategy is outlined in Algorithm 1. In order to apply the relaxation paradigm, the simplest possible model is calibrated first. Thus τ is set to zero, such that the model degenerates to a lumped model. Calibrating such a model can be done tremendously faster than for the case of $\tau = 1$, even if the search is very exhaustive. To perform this task the global evolutionary optimization routine *Shuffled-Complex-Evolution-UA* (SCE-UA, Duan *et al*., 1992) is used. It has shown its effectiveness, robustness and efficiency in various studies mostly in the context of hydrology.

The search with the SCE method is carried out in an iterative fashion and finishes as soon as the improvement in the last two iterations is smaller than a given amount (in our case 5%). In the next step the model is refined stepwise, i.e. τ is increased by a *step size* ρ carrying out a recalibration every time. The method starts with the optimal parameter-set of the last calibration and finishes at the new optimal parameter. This procedure is reiterated and the optimal parameter set for the initial model is obtained. Assuming some degree of continuity between τ and the corresponding optimal set of parameters, i.e. slight changes in the spatial representation do not change the optimal set of parameters significantly, allows for the usage of the local search method Nelder-Mead (Nelder & Mead, 1969) speeding up this approach dramatically.

Approaches to simplify the spatial representation

To decrease the complexity of the spatial representation several alternatives have been implemented, which are based on pairwise merging HRUs. Every merge requires recalculation of the HRU–attributes (e.g. area, topographical parameters, soil type, etc.). While for the quantifiable properties at least its exact mean is representable, aggregation of categorical properties, like land use type, eliminates a considerable amount of information, because only one class can be preserved. As stated before, a HRU is defined to be homogenous, which is violated by this approach. Therefore it is improper to model a catchment with such an oversimplified model. However, this is not a restriction here, because it is only used as an auxiliary construct for calibration.

HRUs are always merged either with a HRU directly connected or a HRU draining into the same river reach. If there are no such HRUs, reaches are merged until two HRUs drain into the same river segment and can be merged. This procedure can be reiterated until there is only one HRU representing the whole catchment.



Algorithm 1 Outline of the calibration algorithm.

The order in which HRUs are merged is considered to be important to achieve continuity between τ and the optimal parameter set. Four different approaches are presented here.

M1 Merge smallest HRUs first. The smallest HRU in the current spatial representation is selected and merged with its downstream connected neighbour.

M2 Merge most similar HRUs. The most similar pair of HRUs is identified, where similarity *S* is quantified by a weighted mean

$$S(src, dst) = \sum_{p} w_p \left| p_{src} - p_{dst} \right| + \sum_{q_{src} \neq q_{dst}} w_q$$

where p_{src} and p_{dst} are the properties of the two HRUs for merging (see also Table 1).

M3 Merge HRUs with lowest runoff generation. The HRU with the least runoff generation is identified and merged with its downstream connected neighbour. This information is obtained by executing the model with a default parameter-set.

M4 Merge most similar HRUs with respect to size.
The method *M2* is applied, but the area of the HRUs is taken into account by multiplying the similarity measure *S* with the size of the HRU:

$$S'(src, dst) = S(src, dst) \cdot area_{src}$$

Table 1 Weights used to compute similarity between HRUs (based on the author's experiences).

p	Slope (%)	Aspect (°)	Elevation (m)	*q*	Soil type	Land use	Geology
w_p	1	0.01	0.01	w_q	5	3.5	1

Table 2 Spatial representations obtained for different values of τ. The simplification was carried out with four different merging rules. For each representation the Nash–Sutcliffe Efficiency E_2 between observed and simulated runoff was computed.

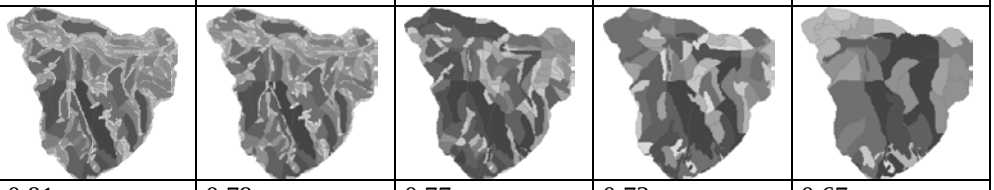
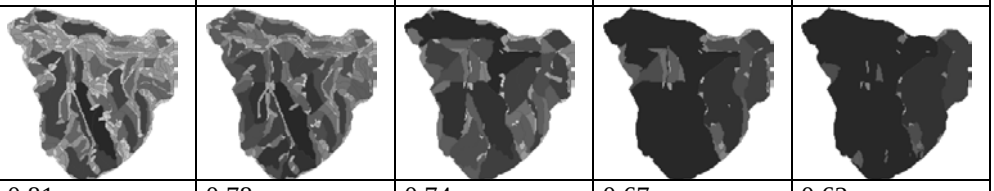
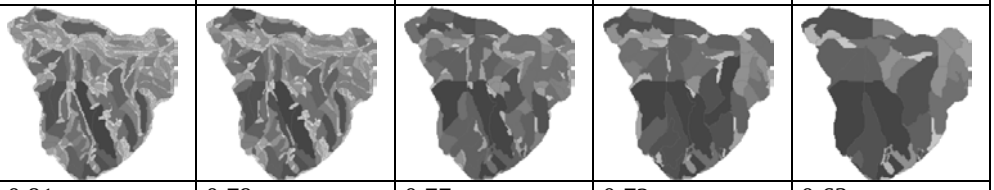
τ		1.0	0.6	0.2	0.1	0.05
M1	Spatial Modelling Units					
	E_2	0.81	0.79	0.77	0.73	0.67
M2	Spatial Modelling Units					
	E_2	0.81	0.78	0.74	0.67	0.62
M3	Spatial Modelling Units					
	E_2	0.81	0.79	0.77	0.72	0.63
M4	Spatial Modelling Units					
	E_2	0.81	0.80	0.78	0.76	0.70

Table 2 compares the spatial representation for some values of τ. For each representation a single model run with an optimal parameter-set for the initial model was carried out and the achieved Nash–Sutcliffe Efficiency is noted. As shown in Table 2 the distribution of spatial modelling units obtained by methods M1 and M3 is quite similar; this is not surprising as there is a strong correlation between size of area and runoff generation. For both methods the modelling units are similarly sized and uniformly distributed over the catchment at the end. The quality of the model decreases a bit faster with M3 than with M1. M2 produces a spatial representation with a very inhomogeneous distribution of modelling units. There are many more units around the river than in other regions of the catchment, which is explained by homogeneity differences in the original representation. At the end only three large modelling units and some very small areas are left. This is reflected by a fast decrease of model efficiency. Method M4 compensates the disadvantage of M2 by taking into account the size of the spatial units, while preserving as much of the properties as possible, which results in a more process oriented distribution of modelling units. Even in the coarse representation it is still possible to visually locate the streams inside the catchment. As M4 is capable of preserving high efficiencies even for low values of τ it seems so be best suited for simplifying the spatial representation.

Determination of step size ρ

Crucial for the success of the relaxation approach is the step size ρ (see Algorithm 1). If ρ is chosen too small, calibration becomes slow, but in contrast the difference between consecutive optimal parameter-sets becomes too large to allow the use of local search methods. Two options for choosing the step size are analysed. At first a fixed step size of 0.1 is used so that the relaxation is carried out in 10 steps until the initial problem is restored. Secondly, an adaptive method was developed, which is using a nested interval search to determine a step size such that the model response does change by approximately 5%.

APPLICATION

Study area

Two well-investigated catchments in Thuringia, Germany were selected to demonstrate the described methods with the model J2000. Both catchments border each other and are similar in their physio-graphical and climatological properties. The average annual temperatures are around 6–7°C. In the higher parts of the catchments annual precipitation is larger than 1400 mm.

The Wilde Gera is located at the northern part of the sub mountainous area of the Thüringer Wald (Fig. 2). The catchment covers 13 km^2. Despite the small size, there is an elevation range of 420m. The catchment is covered mostly with coniferous forest (94%). In the lower part is some agriculture (2%) and deciduous forest (4%).

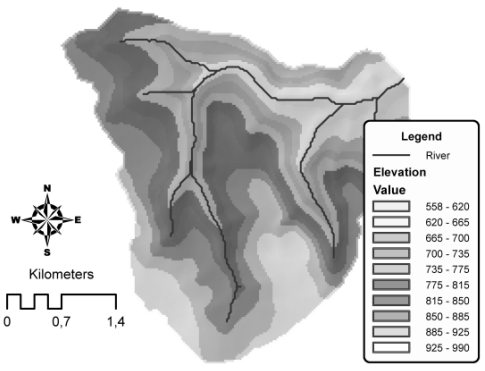


Fig. 2 Wilde Gera to gauge Gehlberg.

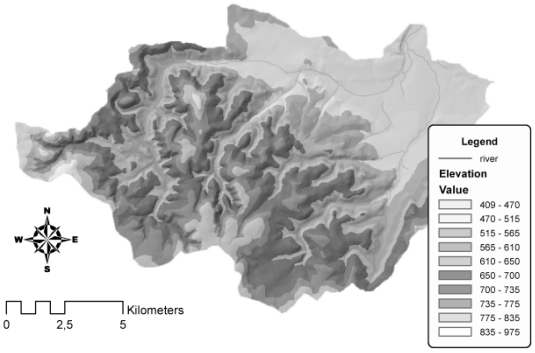


Fig. 3 Ilm to gauge Gräfinau-Angstedt.

The Ilm has its source in the central part of the Thüringer Wald (Fig. 3). The catchment has a size of 155 km^2 and an elevation range of 500 m. Dominating land use of the catchment is 60% coniferous forest, 12% grassland and 10% sealed areas. The hydrological regime is influenced mainly by lateral flow processes and snow melt.

RESULTS

The relaxation strategy is applied to both models of the Wilde Gera and Ilm catchments. The 10 most dominant parameters are chosen for calibration, mainly controlling the simulation of snow processes, lateral flow processes and groundwater processes. The results of this study are summarized in Table 3. For the purpose of comparison with a traditional calibration, the SCE method is used first. That calibration was reiterated 10 times, because SCE contains some randomness. To compare the efficiency of the calibration routines, two measures are taken into account. Firstly the quality is measured by the Nash-Sutcliffe Efficiency and secondly the computational effort is measured by *unit iterations*. For a single model execution this is defined to be the number of spatial units the model uses. For a whole calibration it is the sum of the unit iterations required for each individual simulation of the calibration process.

In the case of the Wilde Gera an average Nash-Sutcliffe Efficiency of 0.816 is achieved, with 1.1 million unit iterations. The catchment of the Ilm is much larger, containing about four times more spatial units. The average effort scaled to nearly linearly to 4.3 million unit iterations achieving an efficiency of 0.769.

Table 3 Comparison of the efficiency of the relaxation strategy for two catchments with different merging rules.

Catchment	Method	Nelder Mead (A)		SCE (B)		Adaptive (C)	
		Time in 10^3 Unit Iterations	Quality in E_2	Time in 10^3 Unit Iterations	Quality in E_2	Time in 10^3 Unit Iterations	Quality in E_2
Wilde Gera	Reference	–		1102	0.816	–	
	M1	480	0.777	990	0.803	627	0.808
	M2	1323	0.819	1050	0.795	540	0.788
	M3	593	0.813	948	0.820	692	0.808
	M4	297	0.814	827	0.780	535	0.752
Ilm	Reference	–		4273	0.769	–	
	M1	2078	0.731	4228	0.754	1530	0.737
	M2	2007	0.768	3800	0.744	1906	0.767
	M3	2027	0.772	3940	0.752	2993	0.784
	M4	2481	0.743	2811	0.757	1811	0.750

The relaxation strategy is used to test the four merging rules. For each of them the calibration is repeated once with the fixed step size and once with the adaptive step size. To achieve a considerable speed-up this strategy should be applied with a local search method, like the Nelder-Mead method. To see whether or not it is capable of finding the global optimum for each sub-calibration the SCE method is used additionally (column B of Table 3). Therefore 44 calibrations were performed on a *Linux* cluster computer with 64 computing cores. This is already a huge amount of computing effort, so it was not possible to reiterate the individual calibrations to get more significant results.

However, it becomes apparent that the relaxation is beneficial. Using the Nelder-Mead method with a fixed step size (column A), the required computing effort is reduced by 45% on average compared to the traditional approach, while the Nash-Sutcliffe Efficiency is altered by <5%, which is not significant. The only exception occurred during the application of method *M2* to the Wilde Gera, but this outlier probably results from unfavorable starting conditions.

A comparison of the Nelder-Mead method with SCE (columns A and B) shows that the local method is able to find the global optimum, or at least a very good approximation of it, because it is <5% worse than SCE. Sometimes it is even better, but requires less than half of the computational effort. It seems that the relaxation is also favorable when using SCE, because the computing effort is slightly (~13%) below the reference, which is somewhat surprising.

Concerning the use of an adaptive step size, there is still some potential. Although a notable increase of efficiency is achieved (~30–60%), it is often more time-consuming than in the case of a fixed step size. Therefore, better strategies for adaptation of the step size could lead to more improvement.

Table 4 Short description of some important parameters.

Parameter	Description
FCAdaptation	Factor to adjust maximum middle pore storage capacity of the soil
flowRouteTA	Stream flow routing parameter
gwRG2Fact	Groundwater recession parameter
soilConcRD1	Surface runoff recession parameter

The merging rules were not very sensitive, as seen especially for the Ilm catchment (Table 3, Nelder-Mead column). This is also supported by Fig. 4, which shows the influence of the spatial model complexity τ on the nearly optimal parameter values. Therefore, the range of the best 10% of all parameter sets occurred during a SCE optimization are plotted against each value of E_2. Due to space restrictions, Fig. 4 shows only the parameters *soilConcRD1*, *flowRouteTA*, *aRain*, *gwRG2Fact* (explained in Table 4) and some of the methods. But the analysis was carried out for each parameter, simplification method and catchment. It can be concluded that the different methods do not show significant differences, e.g. as seen in Fig. 4 (Wilde Gera M1 *vs* M4).

Figure 4 shows the distribution of the best parameter sets around the parameter space for different values of τ. In the case of the lumped model (i.e. $\tau=0$), the range of the best parameter sets is very small. Therefore, the optimal set of parameters is well identifiable for this simple model. With increasing spatial model complexity the parameter identifiability begins to decrease rapidly. With τ between 0.1 and 0.2, the nearly optimal parameters are scattered around a broad range of the parameter space. But this process is reversed with further increase of τ. At 40% of the initial spatial complexity, the range of the best parameters is getting smaller and is stagnating at a low level. Therefore it should be possible to calibrate the models on a spatial representation containing only 40% of all HRUs.

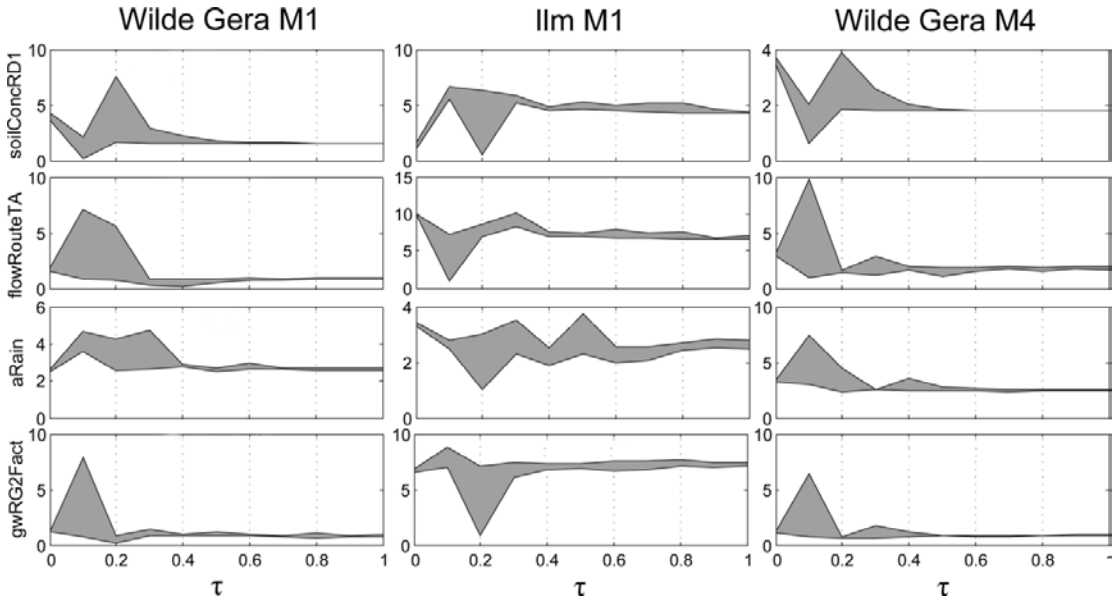


Fig. 4 Dependency between the best sets of parameters and the spatial complexity τ.

CONCLUSION AND FUTURE WORK

This article describes the investigation of a two-stage strategy to speed up the calibration of spatially distributed hydrological models. The strategy is based mainly on relaxation of the model during calibration by reduction of spatial model complexity. The comparison with traditional approaches showed that this strategy is able to calibrate the model with only about half the computational effort, while maintaining the original quality.

Different strategies were tested to optimize the procedure. Firstly, four different merging rules were applied to reduce the spatial complexity of the model, but the results showed that this did not have a significant influence. In addition, an adaptive adjustment of spatial complexity was integrated but found to degrade the method's performance. There is still potential to modify the adaptation rule and to speed up the process of calibration.

The analysis of the optimal parameter values depend on the spatial complexity showed that the parameters of models with very simple spatial representations are very identifiable, where the best parameter sets of slightly more complex models are distributed more broadly. The reason for this is not obvious and must be investigated in more detail.

If the spatial representation has >40% of its initial complexity, parameters become identifiable again and the influence of the spatial representation on the optimal parameterization stagnates at a low level. This property can be used to calibrate the model even more efficiently. Furthermore, it would be very interesting to compare these results with other modelling systems and other spatial representations (e.g. grids).

REFERENCES

Beven, K. (1989) Changing ideas in hydrology – the case of physically-based models. *J. Hydrol.* 105(1–2), 157–172.

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

Duan, Q. Y., Gupta, V. K. & Sorooshian, S. (1993) Shuffled complex evolution approach for effective and efficient global minimization. *Journal of Optimization Theory and Applications* 76(3), 501–521.

Flügel, W. (1996) Hydrological response units (HRU's) as modelling entities for hydrological river basin simulation and their methodological potential for modelling complex environmental process systems. Results from the Sieg catchment. *Die Erde* 127(1), 43–62.

Gupta, H. V., Beven, K. J. & Wagener, T. (2005) Model calibration and uncertainty estimation. *Encyclopedia of Hydrological Sciences*. Wiley Online Library.Horst, R. & Tuy, H. (1996) *Global Optimization: Deterministic Approaches*, 3rd edn. Berlin: Springer-Verlag.

Irvem, A., Topaloglu, F. & Uygur, V. (2007) Estimating spatial distribution of soil loss over Seyhan River Basin in Turkey. *J. Hydrol.* 336(1–2), 30–37.

Kralisch, S. & Krause, P. (2006) JAMS-A framework for natural resource model development and application. In: *Proc. iEMSs Third Biannual Meeting Summit on Environmental Modelling and Software* (Burlington, USA).

Kralisch, S., Krause, P., Fink, M., Fischer, C. & Flügel, W.-A. (2007) Component based environmental modelling using the JAMS framework. In: *Proceedings of MODSIM07*.

Krause, P. (2002) Quantifying the impact of land use changes on the water balance of large catchments using the J2000 model. *Physics and Chemistry of the Earth* 27(9–10), 663–673.

Nash, J. E. & Sutcliffe, J. V. (1970) River flow forecasting through conceptual models – Part I: a discussion of principles. *J. Hydrol.* 10, 282–290.

Nelder, J. A. & Mead, R. (1965) A simplex method for function minimization. *The Computer J.* 7(4), 308–313.

Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D. J. & DMIP, P. (2004) Overall distributed model intercomparison project results. *J. Hydrol.* 298(1–4), 27–60.

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

Samaniego, L., Kumar R. & Attinger, S. (2010): Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, *Water Resour. Res.* 46(5), 2347.