Hydrol. Earth Syst. Sci. Discuss., 4, 91–123, 2007
www.hydrol-earth-syst-sci-discuss.net/4/91/2007/

Hydrology and Earth System Sciences Discussions

Papers published in *Hydrology and Earth System Sciences Discussions* are under open-access review for the journal *Hydrology and Earth System Sciences*

Soft combination of local models in a multi-objective framework

F. Fenicia[1,2], D. P. Solomatine[3], H. H. G. Savenije[2], and P. Matgen[1]

[1]Public Research Center – Gabriel Lippmann, Luxembourg
[2]Water Resources Section, Faculty of Civil Engineering and Geosciences, Delft University of Technology, The Netherlands
[3]UNESCO-IHE Institute for Water Education, Delft, The Netherlands

Received: 9 January 2007 – Accepted: 15 January 2007 – Published: 19 January 2007

Correspondence to: F. Fenicia (fenicia@lippmann.lu)

Abstract

Conceptual hydrologic models are useful tools as they provide an interpretable representation of the hydrologic behaviour of a catchment. Their representation of catchment's hydrological processes and physical characteristics, however, implies simplification of the complexity and heterogeneity of reality. As a result, those models often show a lack of flexibility in reproducing the vast spectrum of catchment responses. Hence, the accuracy in reproducing certain aspects of the system behaviour is often paid in terms of a lack of accuracy in the representation of other aspects.

By acknowledging the structural limitations of those models, a modular approach to hydrological simulation is proposed. Instead of using a single model to reproduce the full range of catchment responses, multiple models are used, each of them assigned to a specific task.

The approach is here demonstrated in the case where the different models are associated with different parameter realizations within a fixed model structure. We show that using a composite "global" model, obtained by a combination of individual "local" models, the accuracy of the simulation is improved. We argue that this approach can be useful because it partially overcomes the structural limitations that a conceptual model might exhibit. The approach is shown in application to the discharge simulation of the experimental Alzette River basin in Luxembourg, with a conceptual model that follows the structure of the HBV model.

1 Introduction

Conceptual hydrological models consist of an ensemble of fluxes and storages representing relevant processes and key zones of catchment response. In the field of hydrological research, those models are useful tools for two main reasons. First, they are based on a reasonable representation of the major hydrological processes, which enables an interpretation of the real behaviour of the catchment. Second, their data requirement and computational demand is limited, which makes them easy to apply and to operate.

Conceptual models, however, represent certain abstraction of reality, which results in a simplification of the complexity and heterogeneity of the real world. As a result, those simple models often display a lack of flexibility in capturing the dynamic and time varying nature of hydrological responses.

One solution to this problem can be to develop the model further, in such a way that more processes are included. This approach, which has the advantage of enabling a better understanding of the system through a process of testing the effects of additional modelling assumptions, is time consuming and might limit the benefits of a conceptual model as an effect of increased complexity.

A second possibility consists of using several models instead of one to better characterize the various conditions that influence the catchment hydrological behaviour. This approach, which is here investigated, is based on the idea that an integration of the results obtained by different models provides a more comprehensive and accurate representation of catchment response than what can be obtained using a single model.

Multi-model approaches have been widely used in hydrological modelling in different frameworks and for different purposes. One approach lies in the context of equifinality. In this context, an ensemble of models is generated, either by multiple realizations from a single model structure, or by single realizations from multiple model structures. Model simulations are eventually weighted or averaged or used to derive statistics of model outputs. In this approach, the models of the ensemble are assumed to be characterized by the property that they all fit the observed data equally well. This premise is at the heart of the GLUE framework (Beven and Freer, 2001), but also touches other areas of application, such as model and multi-model ensembles (Georgakakos, 2004; McIntyre, 2005).

The second approach involves the generation of several models in the way as previously mentioned, but tries to exploit possible differences in the generated models. The various models, however, are not considered “equifinal”, and it is admitted that each model could be more accurate than the others in describing a specific aspect of the system response. The models are then combined by appropriate weighing procedures that attempt to retrieve the individual strengths of each model in simulating the system response. In this context, Shamseldin et al. (1997) and Xiong et al. (2001) propose different combination methods to integrate the optimal results of different conceptual models. They show that in general the discharge estimates obtained by combining different models are more accurate than those obtained from any single model used in the combination.

In the third approach, the models are directly built or calibrated on different event types or data sequences and subsequently combined together. In this approach, the distinctive role of different models in reproducing the system response is explicitly recognized from the beginning of the model development. See and Openshaw (2000) show the application of different neural networks that were built on different event types. Wang et al. (2006) used a combination of ANNs for forecasting flow: different networks were trained on the data subsets determined by applying either a threshold discharge value, or clustering in the space of inputs (lagged discharges only but no rainfall data, however). Jain and Srinivasulu (2006) apply a mixture of neural networks and conceptual techniques to model the different segments of a decomposed flow hydrograph. Solomatine and Xue (2004) show an application of data-driven models M5 model trees and neural networks in a flood-forecasting problem, consisting of a combination of models locally valid for particular hydrologic conditions represented by specific regions of the input-output space. Corzo and Solomatine (2007) use several methods of baseflow separation, build different models for base and excess flow and combine these models ensuring optimal overall model performance.

The first two approaches can be classified as “ensemble strategies”, in a sense that different models are developed to perform similar modelling operations. The last approach corresponds to a “modular strategy”, as different models are developed to perform different tasks.

The approach introduced here can be attributed to the last approach. We in fact adopt a modular strategy to characterize different aspects of a stream hydrograph. The present work, however, focuses on conceptual model structures and it is set in a multi-objective framework.

The approach consists in calibrating a conceptual model with respect to different objectives, representing model performance towards different aspects of the simulation, and in combining the best performing models associated to each objective in such a way that the strength of each individual model used in the combination is exploited. Such an approach attempts to improve the global accuracy of the simulation overcoming possible limitations in the model structures, Moreover, it provides some insights in the model deficiencies, therefore contributing to an understanding of the real behaviour of the system.

The approach is demonstrated using a conceptual model that follows the structure of the well-known HBV model (Lindström et al., 1997). The model is analysed with respect to its ability of reproducing the rainfall-discharge behaviour of a catchment in Luxembourg, with particular reference to accurate reproduction of the high and low flows behaviour.

Multi-objective calibration with respect to two defined objectives representing model performances for the selected hydrograph characteristics shows that there are several solutions (the so-called “Pareto-optimal” set of solutions) that simultaneously optimize the selected criteria. These solutions represent a trade-off between the selected objectives and show that individual optimal models are better in matching different aspects of the observed hydrograph.

The two best performing models associated with the selected hydrograph characteristics (in this case high flows or low flows) are subsequently weighted together using a fuzzy combining scheme. The paper concludes with a discussion on the physical significance of the proposed approach.

2 Problem formulation

The process of developing a "global" model by means of aggregating multiple "local" models, each of which is specialized in simulating a certain aspect of the system response, can require a series of operations, summarized hereafter.

- *Events selection*. The events that should be correctly simulated have to be carefully chosen. Within a modular approach, which presumes switching between different models, those events should be different and non-overlapping. Consequently, such events should refer to different ranges or different time periods of a certain measured variable. The number of events, which also affects the number of individual models that are subsequently developed, should not be too high, in order to avoid a too fragmented description of the system response, which could also reduce the global efficiency for periods outside the calibration period.

- *Model selection*. The selected events could be represented by models of the same nature or of different nature (e.g. conceptual, physically based, data driven). As an example, Jain and Srinivasulu (2006) use conceptual and data driven models to simulate different segments of a flow hydrograph. They found that the considered case study models of conceptual type performed better than data driven ones in reproducing hydrograph recession.

- *Objective function definition*. Objective functions express the quality of the simulation in numerical form by aggregating model residuals in time. In the present case, they should represent the ability of the various models to reproduce the selected events, hence they should refer to different characteristics of a measured variable. Different kind of information from a single time series can be extracted either weighing the whole series of residuals in such a way that the error associated with the simulation of a certain characteristic is weighted more than other errors, achieved by power transformations (Gupta et al., 1998; Hogue et al., 2000), or by partitioning the time series into different classes each of which is associated to a different behaviour of the system. Abrahart and See (2000) use a data decomposition based on season, Jain and Srinivasulu (2006) and Boyle et al. (2000) use hydrograph separation to partition the data into different categories, Corzo and Solomatine (2007) employ baseflow separation algorithms.

- *Model calibration*. As most models contain parameters that cannot be directly measured, model parameters have to be inferred by calibration. Hence, the models associated with the different events have to be calibrated (or trained) to optimize the selected objective functions.
- *Model combination*. The calibrated models are finally reintegrated into one single model. This combination can be straight-forward (Jain and Srinivasulu, 2006) implying a switch between different models at different time steps, but can also involve some kind of weighing (See and Openshaw, 2000; Xiong et al., 2001; Solomatine, 2006). Model weighing can improve simulation results, and avoid unrealistic discontinuities in the simulated system behaviour.

2.1 Model description

The model used in this application is a lumped conceptual model that follows the structure of the HBV-96 model (Lindström et al., 1997). In this study the model is run with an hourly time step. The model consists of routines for soil moisture accounting, runoff response, and a routing procedure. The structure is composed of three storage components: a soil moisture reservoir, an upper reservoir, and a lower reservoir. The output from the lower and upper reservoir is combined and routed through a triangular transfer function.

The soil moisture routine represents the runoff generation and involves three parameters, β, FC and LP. The proportion of precipitation that produces direct runoff is related to the soil moisture by the following relation:

$$\frac{R}{P} = \left(\frac{SM}{FC}\right)^{\beta} \quad (1)$$

Where P is the total rainfall, R is the direct runoff, SM is the storage of the soil moisture reservoir, FC is the maximum soil moisture storage, and β is a parameter accounting for non linearity. The remaining part is added to the soil moisture storage.

Actual evaporation (E_a) is calculated from potential evaporation (E_p) according to the following formula:

$$E_a = E_p \cdot \min\left(1, \frac{SM}{FC \cdot LP}\right) \quad (2)$$

Where LP is the fraction of FC above which the evaporation reaches its potential level.

The direct runoff R enters the upper reservoir, and the lower reservoir is filled by a constant percolation rate ($PERC$) as long as storage in the upper reservoir is available.

Capillary flux from the upper reservoir to the soil moisture reservoir is calculated according to the following equation:

$$C = CFLUX \cdot \left(1 - \frac{SM}{FC}\right) \quad (3)$$

Outflow from the upper reservoir is expressed as

$$Q_1 = K_1 \cdot UZ^{1+\alpha} \quad (4)$$

Outflow from the lower reservoir is expressed as

$$Q_2 = K_2 \cdot LZ \quad (5)$$

Where UZ and LZ are the storage states of the upper and lower reservoirs respectively, K_1 and K_2 are storage coefficient, and α is a parameter accounting for non linearity.

The outlets from the two reservoirs are finally added and routed through a transfer function with base defined by the parameter $MAXBAS$. The model has a total of nine calibration parameters, which are summarized in Table 1.

2.2 Events selection and objective functions

Objective functions used in calibration typically aggregate model residuals in time and are used to evaluate the accuracy of the model in simulating the actual behaviour of the catchment. Different functions may enhance the error in simulating different aspects of the simulation while neglecting or downplaying the error in simulating other aspects. The minimization of one objective function through calibration of model parameters therefore might stress the simulation of certain features of the simulation at the expense of others. The aggregation of model residuals in time, moreover, inevitably results in a loss of information contained in the observed data.

For those reasons, the model calibration task can be set up as a multi-objective minimization problem, where alternative objective functions would represent different criteria to evaluate model performance with respect to specific aspects of the simulation.

For the present application we use two objective functions, one enhancing the model error with respect to low flow simulation, and the other enhancing model error with respect to high flows.

The two functions are defined as follows:

$$N_{HF} = \sqrt{\frac{1}{n}\left(\sum_{i=1}^{n}\left(Q_{s,i} - Q_{o,i}\right)^2 \cdot w_{HF,i}\right)} \quad (6)$$

$$N_{LF} = \sqrt{\frac{1}{n}\left(\sum_{i=1}^{n}\left(Q_{s,i} - Q_{o,i}\right)^2 \cdot w_{LF,i}\right)} \quad (7)$$

Where

$$w_{HF,i} = \left(\frac{Q_{o,i}}{Q_{o,\max}}\right)^2 \quad (8)$$

$$w_{LF,i} = \left(\frac{Q_{o,\max} - Q_{o,i}}{Q_{o,\max}}\right)^2 \tag{9}$$

where

n: total number of time steps

$Q_{s,i}$: simulated flow for the time step i

$Q_{o,i}$: observed flow for the time step i

$Q_{o,\max}$: maximum observed flow

The two weighing functions w_{HF} and w_{LF} allow placing a stronger weight on the low or on the high portions of the hydrograph (Fig. 1). As a result, N_{LF} places a stronger weight on low flows and a weaker weight on high flows than N_{HF}.

By computing both objective functions over the whole range of discharges, both functions constrain the model to fit the entire hydrograph.

2.3 Model calibration

We follow a standard framework of multi-objective analysis which, for hydrological models has been introduced by Gupta et al. (1998). This framework adopts the notions of Domination and Pareto-optimality, which are hereafter recalled.

We will use the term solution to mean a parameter set $\boldsymbol{x}_i$. Each solution $\boldsymbol{x}_i$ is associated to a number of objective functions' values $N_j(\boldsymbol{x}_i)$ (j=1…m, m = number of objectives), expressing the performance of the model. Lower values of $N_j(\boldsymbol{x}_i)$ indicate better performance.

- A solution $\boldsymbol{x}_1$ is said to dominate another solution $\boldsymbol{x}_2$ when $\boldsymbol{x}_1$ is better than $\boldsymbol{x}_2$ in at least one objective (meaning $N_j(\boldsymbol{x}_i)<N_j(\boldsymbol{x}_2)$ for at least one value of j), and not worse than $\boldsymbol{x}_2$ in any of the others (meaning $N_j(\boldsymbol{x}_i)\leq N_j(\boldsymbol{x}_2)$ for all values of j).

- The Pareto-optimal set of solutions is composed of those solutions that are not dominated by any solution of the feasible search space.

The outcome of a multi-objective problem in such a framework consists in the Pareto-optimal set of solutions. This set in general consists of more than one solution.

The existence of more than one solution indicates that the objective functions are conflicting to each other, meaning that an optimal performance in one objective is "paid" in terms of sub-optimal performances in the others. This has been demonstrated adopting different models and various types of objective functions (e.g. Yapo et al., 1998; Boyle et al., 2000 Vrugt et al., 2003).

When applied to hydrological models, the existence of multiple optimal solutions can be related to a systematic component of the modelling error (Gupta, 1998), which is normally attributed to model structural inadequacies (Gupta et al., 1998, Vrugt et al., 2003).

While it is plausible to think that other sources of error might contribute to this component, such as data distortion caused by incorrect rating curves or boundary conditions, it is also reasonable, when no other information is available, to put more confidence in the data than in the model, and therefore try to build models that represent the observations as correctly as possible. In this sense, the existence of multiple Pareto-optimal solutions can be regarded as a failure of the model structure. It in fact indicates that the model is not able to simultaneously represent the full variability of catchment responses.

All Pareto-optimal solutions are "equally important", in a sense that it is difficult to prefer one solution over the other without any further information about the problem.

The different solutions, however, are not "equifinal", in the sense given to this term by Beven (1993). Every solution has its strengths and limitations in describing the different aspects of the observed signal, as expressed by the selected objective functions. This observation can be exploited by trying to combine different optimal solutions in such a way that the individual strengths of each solution are exploited.

2.4 Combining scheme

Calibration of model parameters with respect to the two selected objective functions results in a set of Pareto-optimal solutions, which is delimited by the two best models that minimize each of the individual objectives N_{LF} and N_{HF}. Those models are defined as HBV_{LF} and HBV_{HF}.

The two models are then combined with an appropriate weighting procedure to generate a global model HBV_G that aims at reproducing the whole range of discharges exploiting the best parts of each individual "local" model.

The combining scheme that is used to weight the contributions of each individual model makes use of a fuzzy attribution of weights. The output of the HBV_{LF} model is assumed to be accurate in simulating low flow events, but might be not accurate in simulating high flow events. Vice versa, the output of the HBV_{HF} model is assumed to be more accurate during high flows than during low flows. In order to express this difference in the degree of believability of the outputs of the two models, each model output is associated with a certain membership function (we follow here an approach termed by Solomatine (2006) a "fuzzy committee").

The degree of membership associated with the low flow model is 1 when the simulated flow is below the threshold γ, it decreases linearly when the flow is between the thresholds γ and δ, and it is 0 when the flow is above the threshold δ (Fig. 2). The degree of membership of the high flow model follows a symmetric behaviour. Membership functions for the two local models are described in equations 10 and 11; γ and δ are named threshold for high flows and for low flows respectively and are expressed as a fraction of the maximum observed discharge $Q_{o,\max}$.

$$m_{LF}(Q) = \begin{cases} 1, if Q/Q_{o,\max} < \gamma \\ 1 - \frac{Q/Q_{o,\max} - \gamma}{\delta - \gamma}, \text{ if } \gamma \le Q/Q_{o,\max} < \delta \\ 0, if Q/Q_{o,\max} \ge \delta \end{cases} \tag{10}$$

$$m_{HF}(Q) = \begin{cases} 0, if Q/Q_{o,\max} < \gamma \\ \frac{Q/Q_{o,\max}-\gamma}{\delta-\gamma}, \text{ if } \gamma \le Q/Q_{o,\max} < \delta \\ 1, if Q/Q_{o,\max} \ge \delta \end{cases} \tag{11}$$

The outputs of the two models are finally combined according to Eq. (12). This weighing approach allows a smooth combination of the two models, and avoids discontinuities in the reproduction of the system response.

$$Q_G(t) = \frac{m_{LF}(Q_{LF}) \cdot Q_{LF}(t) + m_{HF}(Q_{HF}) \cdot Q_{HF}(t)}{m_{LF} + m_{HF}} \tag{12}$$

Note that the weighting schemes shown on Figs. 1 and 2, at first sight similar, serve different purpose: the first one is used to stress low/high flows in the objective function used to calibrate two separate models, and the second one is responsible for ensuring smooth compatibility between the models.

3 Case study

3.1 Study area and data description

The study area is within the experimental Alzette river basin, which is located for the large part in the Grand-Duchy of Luxembourg. For model calibration, three years of hourly data from Hesperange, a gauging station placed along the course of the Alzette River upstream of Luxembourg-city, were used.

Catchment size is 288 km^2, and land cover is estimated as being composed of cultivated land (27%) grassland (26%), forest land (29%) and urbanized land (18%). Lithology is mainly represented by Marls and Marly-Sandstones on the left bank tributaries and Limestones on the right bank tributaries of the Alzette River.

The rainfall-runoff behaviour of these units is quite different. Marl areas are characterized by impermeable bedrock, therefore rainfall water, after losses for evaporation or transpiration, reaches the stream mostly as saturated subsurface flow that develops at the interface between the weathered zone and the underlying bedrock areas. When the weathered zone becomes saturated, or during heavy rainfall events, surface runoff occurs.

In limestone areas a large part of rainfall water infiltrates and after subtraction of losses percolates to the groundwater aquifer, which is capable of storing and releasing large quantities of water.

The response to rainfall of Marl areas is faster and characterized by larger volumes of water than that of limestone areas. Moreover, the large part of the baseflow during prolonged dry periods is mostly sustained by the limestone aquifer.

The use of an hourly time step is justified considering that the average concentration time of the basin is of the order of a few hours.

The basin is instrumented by several rain gauges including tipping-buckets and automatic samplers measuring at a time step which does not exceed 20 min. Hourly rainfall series are calculated by averaging the series at the individual stations with the Thiessen polygon method. Daily potential evaporation is estimated through the Penman-Monteith equation (Monteith, 1965). The meteorological variables needed to compute the evaporative loss are measured at the Luxembourg airport meteorological station. Hourly estimates are then calculated distributing the daily amounts through a sine function.

3.2 Multi-objective calibration

The HBV model is calibrated according to the two objectives previously defined as N_{LF} and N_{LF}. The problem is posed in a multi-objective framework and solved by determining the Pareto-optimal set of solutions. In order to efficiently sample the parameter space, the Multi-Objective Shuffled Evolution Metropolis University of Arizona (MOSCEM-UA) algorithm is used (Vrugt et al., 2003).

The MOSCEM-UA algorithm begins with a random sequence of s points sampled throughout the feasible parameter space. For each point the set of objective functions is evaluated. The sequence is partitioned into q complexes, and in each complex a parallel sequence is launched. New candidate points from each complex are generated from a multivariate normal distribution with mean at the current draw of the sequence and covariance matrix calculated based on the history of each sequence. The sequences evolve based on a Metropolis-type acceptance criterion. The algorithm proceeds until a specified maximum number of iterations m is reached.

The MOSCEM-UA has three algorithmic parameters that have to be specified by the user: s, q, and m. No theoretic guidelines exist in determining those parameters; however, a good criterion is to use a number of complexes that is at least equal to the number of parameters.

Parameter bounds were determined by analysing the results of an initial run of the algorithm on a wide parameter space.

The selected parameter bounds are reported in Table 2. The algorithm was run with the following parameters: s=2000, q=10, m=50.000.

The outcome of the optimization algorithm is presented in Fig. 3 and Fig. 4. Figure 3 shows the objective function values corresponding to the evaluated parameter sets together with the set of Pareto-optimal solutions and the optima corresponding to each individual objective. This plot clearly illustrates a trade-off in the selected objectives, and reveals the inability of the model to match the selected aspects of the hydrograph simultaneously.

The variation of the Pareto-optimal parameter sets is shown in Fig. 4. The parameter values are normalised with respect to the upper and lower bounds given in Table 2, so that the feasible range of all parameters is between 0 and 1. Each line on the plot represents one parameter set. The figure gives a visual indication on the relation between the initial feasible parameter range, and the parameter range that corresponds to the optimal solutions.

Conclusions about large or small variability of parameter values would not be meaningful, as the extent of the optimal range displayed in the figure clearly depends on the initial lower and upper limits that are selected.

Wile moving from one solution to another on the Pareto-optimal front (Fig. 3), the corresponding parameter values may show a trend from one extreme to another (Fig. 4). The existence of such a trend reveals potential deficiencies in the model structure. This behaviour is significant for the parameters β, which accounts for non linearity of the rainfall-direct runoff relation, and K_2, representing the storage coefficient of the lower reservoir.

As an example, the parameter K_2 shows higher values when calibrated towards the high flows, and lower values when calibrated towards the low flows. If we assume that the correct value for this parameter is what corresponds to the low flow calibration (as this parameter determines the behaviour of the lower reservoir which mostly affects low flow simulation), we can conclude that the calibration of the model with respect to high-flows will result in overestimating this parameter. As a result, the lower reservoir in the optimal high flow model will empty faster than it should in order to compensate for errors in the simulation of other processes. This behaviour is also evident on the hydrographs presented in Figs. 5a and b. Figure 5b shows that the best performing model with respect to N_{HF} is characterized by steeper recessions than observed, while Fig. 5a, representing the best model with respect to N_{LF}, shows a better agreement with the observations during recession periods.

3.3 Local models combination

The combining scheme aims at integrating the strengths of each individual local model in reproducing some characteristics of the simulation. As noted before, the combining approach as here interpreted, requires the selection of two discharge thresholds: γ and δ (Fig. 2). Those two thresholds can be selected based on knowledge of the system behaviour, or can be selected automatically to minimize the error of the global model.

Manual selection of thresholds could be based on the ground of a physical understanding of the behaviour of the catchment. In this case, the thresholds could represent switches in catchment behaviour that correspond, for example, to changes in rating curves or in contributing areas related to the water level in the stream. This evidence is not the case of this study, therefore the thresholds have been initially selected by a visual inspection of model performances across the range of observed discharges. A procedure to perform automatic tuning of the thresholds is described in the next paragraph.

Analysis of the observed rainfall and flow lead us to a conclusion that it would be reasonable to choose the following thresholds for flow: Q=0.12 mm/h for high flows and Q=0.07 mm/h for low flows. As the maximum discharge in the calibration period is 0.64, this results in γ=0.11 and δ=0.17. Performances of the global model with respect to the hydrograph simulations are represented in Fig. 5. Figure 5a shows the performances of the low flow local model, Fig. 5b shows the performance of the high flow local model, and Fig. 5c shows the performance of the global model developed from the combination of the two local models. It is possible to observe visually that the global model incorporates the best features of both local models, considerably improving the overall accuracy.

Model performances in term of the selected criteria are presented in Fig. 6. The solution corresponding to the global model lies beyond the Pareto-optimal set, showing that the global model improves the accuracy of the simulations.

3.4 Automatic tuning of the thresholds

When no evidence exists in determining the thresholds corresponding to changes in the system behaviour, those thresholds can be calculated by trying to maximise the performance of the model. With this purpose, a sequence of thresholds was generated on a grid in the (γ, δ) space, and the Pareto-optimal set of solutions corresponding to different values of the thresholds has been calculated. Results are represented in Fig. 7. It is possible to observe that even the employed simple type of search improves the global model accuracy.

The thresholds values corresponding to the Pareto solutions are represented in Fig. 8. With respect to the manually selected values, the Pareto values are smaller for the thresholds for low flows, and larger for the high flow threshold, which enlarge the area where both models are evaluated and weighted.

It is also noteworthy that the performance of some Pareto-optimal models is higher than that of both single best solutions corresponding to each individual criterion. This is due to the fact that the selected objectives calculate the error with respect to the whole range of flows, even with different weights. An improvement in low flow description, for example, given the same performances in high flows, reduces the total error as calculated by N_{HF}.

4 Discussion and conclusion

The synthetic view of reality that is incorporated in conceptual hydrological models does not allow a simultaneous optimal representation of different aspects of the system behaviour.

To overcome this problem, a modular approach to hydrologic simulation has been presented. This approach allows for different models to operate simultaneously, each of them developed to reproduce a specific aspect or phase of the system behaviour. The various models are then combined through an appropriate weighing procedure, to produce a global representation of the catchment behaviour. The combining scheme exploits the strengths of each individual model in a synergistic manner.

The presented method allows for different parameter sets of a fixed model structure, but, in principle, could be applied allowing for different model structures too (e.g. conceptual, physically based, data driven). Specifically, we build separate models for high flow and low flow simulation, which are subsequently combined through a soft combination approach. Results show that the combined "global" model reaches a higher overall accuracy than what can be obtained using any individual parameter set.

A drawback of the proposed approach is that the switching between different models causes a loss of continuity between model internal states. This might limit its application in cases where water balance of different model compartments is an issue. A possible way forward is to incorporate the switching of model parameters directly within the model structure.

While the practical utility of such an approach relies on an improvement of the simulation accuracy, the physical implications involved require interpretation and justification. The switching between different models, in fact, implies an alternation between different views or descriptions of reality. If the natural system is not modified by natural phenomena or artificial activities, it might seem physically inconsistent to represent it by means of separate descriptions.

The hydrological processes involved in the rainfall-runoff transformation, however, are extremely complex and characterized by a high degree of spatial and temporal variability. As a result, the catchment behaviour can be characterized by visibly different hydrological responses. The different "personalities" that a catchment might display are determined by a series of phenomena and processes that can be in general identified, but that is difficult to separate and quantify.

The different interacting causes of variability in hydrological behaviour include:

– *Seasonal dependencies*. As an effect of vegetation or biologic activities, aspects like land cover or macropore distribution in the top-soil vary, affecting processes such as interception, infiltration, pathways of water in on the soil surface and in the weathered zone

– *Environmental forcing conditions*. Forcing conditions influence the amount and distribution of water in the soil, determining the condition of the catchment hydrological "state". With changes in hydrological states, such as from low to high flow or from dry to wet conditions, the compartments of the catchment that contribute to discharge (e.g. saturated and unsaturated zone, near stream saturated areas) vary dynamically, leading to different domains of formation and integration of the hydrological processes.

– *Threshold behaviour*. The occurrence of several hydrological processes is characterized by clear threshold behaviour. Groundwater levels, rainfall intensities, soil moisture conditions control the occurrence of processes such as surface runoff or rapid subsurface flow, and can trigger the contribution of different areas of the catchment.

– *Variability of hydrological relations*. The relation between hydrological processes is extremely variable, depending on time, location and antecedent conditions. As a result of the interaction between many different physical phenomena, the overall catchment response to rainfall is a highly non linear and dynamic process, which is difficult to capture with simple conceptualizations.

A multi-model approach can implicitly take into account the variability in hydrologic behaviour that is not explicitly considered in the realization of a single model.

By allowing different models to operate for the simulation of different aspects of the system response it is implicitly recognized that a single model cannot explain by itself the full variability of catchment responses. In the specific case, where the different models are represented by individual parameter sets, it can be assumed that model parameters, depending on the particular stage of the simulation, describe different behaviours of the catchment by expressing different processes.

In conceptual modelling it is typically assumed that model parameters, if not physically based or clearly related to catchment attributes, are representative of inherent properties of the catchment, and therefore not supposed to vary (Wagener et al., 2003). The fact that model parameter may have different values depending on aspects such as the length of the calibration period or the performance measure used for calibration is an indication of potential inadequacies of the model structure, which, ideally, should be refined and corrected.

Understanding where the model fails, and where the catchment shows a certain "personality" that is different than what is estimated a-priori, can guide towards a better understanding of the system behaviour. When building a model, in fact, we use a possible representation of the most relevant processes and their interrelation. The analysis of the performance of the model represents a possibility to test the hypotheses made. In this sense, identifying a switching between different states can clarify triggers or thresholds in the catchment behaviour, helping to guide model refinement and providing new understanding that could be a base for field research.

It can be concluded that the presented approach can be seen as an effective way to improve model accuracy by representing different aspects of the system behaviour by differently parameterized models. The multi-objective framework makes it possible to perform the detailed analysis of the models' performance and to construct an optimal model structure. The use of a "fuzzy committee" allows for soft combination of local models and prevents discontinuities between the model predictions. The approach is quite universal and can be used to combine different types of models, from conceptual to data-driven ones. A first challenge is to complement the presented method by the algorithms aimed at discovering various regimes in the time series representing the modelled system; this would allow for optimal combination of domain (hydrologic) knowledge incorporated in models with the automatic machine learning or time series analysis routines. A second challenge is to implement different states of catchment behaviour directly within the model structure, in order to obtain a comprehensive description of the overall catchment behaviour within a single representation of reality.

References

Abrahart, R. J. and See, L.: Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments, Hydrol. Process., 14, 2157–2172, 2000.

Beven, K. and Freer, J.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, J. Hydrol., 249, 11–29. 2001.

Beven, K. J.: Prophesy, reality and uncertainty in distributed hydrological modelling, Adv. Water Resour. 16, 41–51, 1993.

Boyle D. P., Gupta H. V., and Sorooshian S.: Towards improved calibration of hydrologic models: combining the strengths of manual and automatic methods, Water Resour. Res., 36, 3663–3674, 2000.

Corzo, G. and Solomatine, D. P.: Baseflow separation techniques for modular artificial neural network modelling in flow forecasting, Hydrol. Sci. J. accepted, 2007.
Georgakakos, K. P., Seo, D. J., Gupta, H., Schaake, J., and Butts, M. B.: Towards the characterization of streamflow simulation uncertainty through multimodel ensembles, J. Hydrol, 298, 222–241, 2004.
Gupta H. V., Sorooshian S., and Yapo P. O.: Toward improved calibration of hydrologic models: multiple and noncommesurable measures of information, Water Resour. Res., 34, 751–763, 1998.
Hogue, T. S., Sorooshian, S., Gupta, H., Holz, A., and Braatz, D.: A Multi-step Automatic Calibration Scheme for River Forecasting Models, J. Hydrometeorol., 1, 524–542, 2000.
Jain, A. and Srinivasulu, S.: Integrated approach to model decomposed flow hydrograph using artificial neural network and conceptual techniques, J. Hydrol., 317, 291–306, 2006.
Lindström, G., Johansson, B., Persson, M., Gardelin, M., and Bergström, S.: Development and test of the distributed HBV-96 hydrological model, J. Hydrol., 201, 272–288, 1997.
McIntyre, N., Lee, H., Wheater, H., Young, A., and Wagener, T., Ensemble predictions of runoff in ungauged catchments, Water Resour. Res., 41, W12434, doi:10.1029/2005WR004289, 2005.
Monteith, J. L.: Evaporation and the environment. Symp. Soc. Exp. Biol., 19: 205–234, 1965.
See, L. and Openshaw, S.: A hybrid multi-model approach to river level forecasting, Hydrolog. Sci. J., 45, 523–536, 2000.
Shamseldin, A. Y., O'Connor, K. M., and Liang, G. C.: Methods for Combining the Output of Different Rainfall-Runoff Models, J. Hydrol., 197, 203–229, 1997.
Solomatine D. P. and Xue, Y.: M5 model trees compared to neural networks: application to flood forecasting in the upper reach of the Huai River in China, J. Hydrol. Eng., 6(5), 491–501, 2004.
Solomatine D. P.: Optimal modularization of learning models in forecasting environmental variables. Proc. of the iEMSs 3rd Biennial Meeting: "Summit on Environmental Modelling and Software", edited by: Voinov, A., Jakeman, A., and Rizzoli, A., Burlington, USA, July 2006. CD ROM. Internet: http://www.iemss.org/iemss2006/sessions/all.html, 2006.
Vrugt, J. A., Gupta, H. V., Bastidas, L. A., Bouten, W., and Sorooshian, S.: Effective and efficient algorithm for multiobjective optimization of hydrologic models, Water Resour. Res., 39(7), 1214, doi:10.1029/2002WR001746, 2003.
Wagener, T., Boyle, D. P., Lees, M. J., Wheater, H. S., and Gupta, H. V., and Sorooshian, S.:

A framework for the development and application of hydrological models, Hydrol. Earth Sys. Sci., 5, 13–26, 2001.
Wagener, T., McIntyre, N., Lees, M. J., Wheater, H. S., and Gupta, H.V.: Towards reduced uncertainty in conceptual rainfall-runoff modeling: dynamic identifiability analysis, Hydrol. Process., 17(1), 455–476, 2003.
Wang, W., van Gelder, P. H. A. J. M., Vrijling, J. K., and Ma, J.: Forecasting daily streamflow using hybrid ANN models, J. Hydrol., 324(1–4), 383–399, 2006.
Xiong, L., Shamseldin, A. Y., and O'Connor, K. M.: A nonlinear combination of the forecasts of rainfall-runoff models by the first-order Takagi-Sugeno fuzzy system, J. Hydrol., 245(1), 196–217, 2001.
Yapo, P. O., Gupta, H. V., and Sorooshian, S.: Multi-Objective Global Optimization for Hydrologic Models, J. Hydrol., 204, 83–97, 1998.

Table 1. Model parameters and corresponding units.

Parameter name	Description	Units
FC	Maximum soil moisture storage	mm
LP	Limit for potential evaporation	–
β	Non linear runoff parameter	–
$PERC$	Percolation rate	mm/h
$CFLUX$	Maximum capillary rate	mm/h
α	Non linear response parameter	–
K_1	Upper storage coefficient	mm/h
K_2	Lower storage coefficient	mm/h
$MAXBAS$	Transfer function length	h

HESSD

4, 91–123, 2007

Local models

F. Fenicia et al.

EGU

Table 2. Parameter ranges.

Parameter name	Units	Lower Bound	Upper Bound
FC	mm	200	450
LP	–	0.01	1
β	–	0.01	2
$PERC$	mm/h	0.01	1
$CFLUX$	mm/h	0	0.1
α	–	0	0.5
K_1	mm/h	0.001	0.1
K_2	mm/h	0.001	0.1
$MAXBAS$	h	7	15

HESSD
4, 91–123, 2007
Local models
F. Fenicia et al.

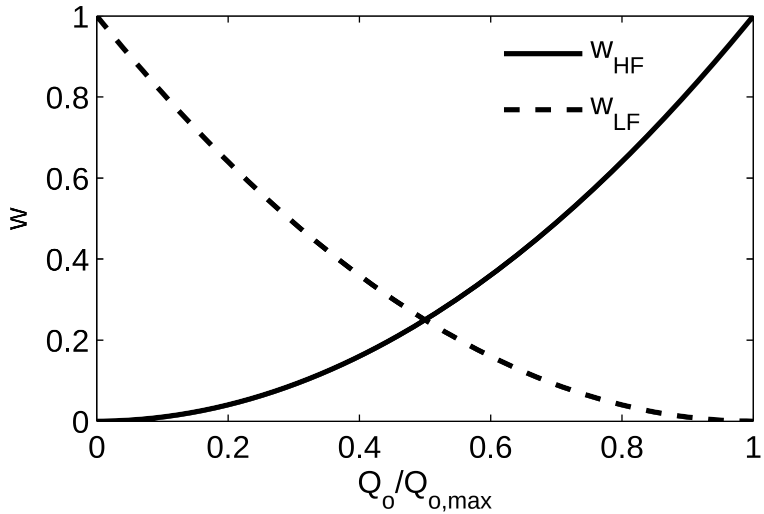

Fig. 1. Weight distribution.

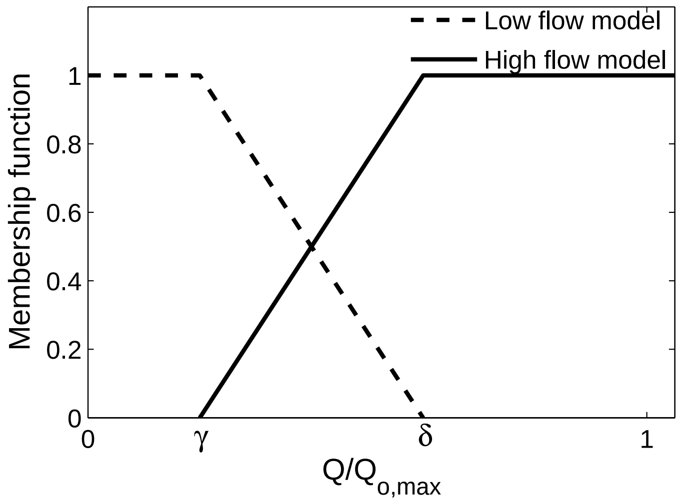

Fig. 2. Combining scheme.

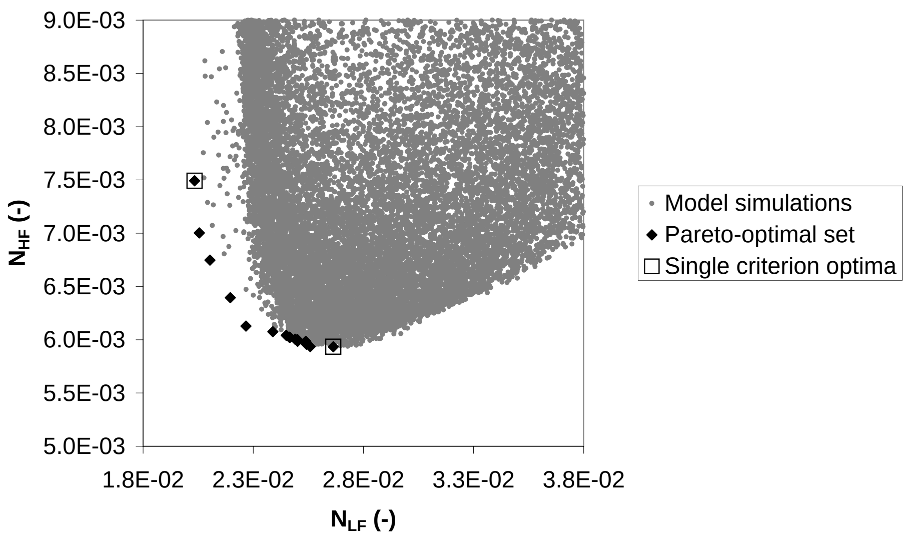

Fig. 3. Scatter plot of model simulations in the objective space.

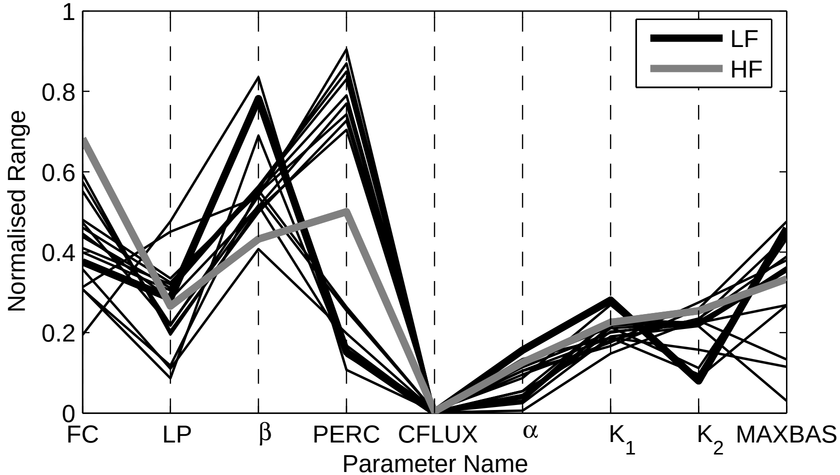

Fig. 4. Normalized parameter plot. Pareto-optimal solutions from Fig. 3 are shown. The two thicker lines represent optimal solutions with respect to one single criterion (LF or HF).

HESSD

4, 91–123, 2007

Local models

F. Fenicia et al.

EGU

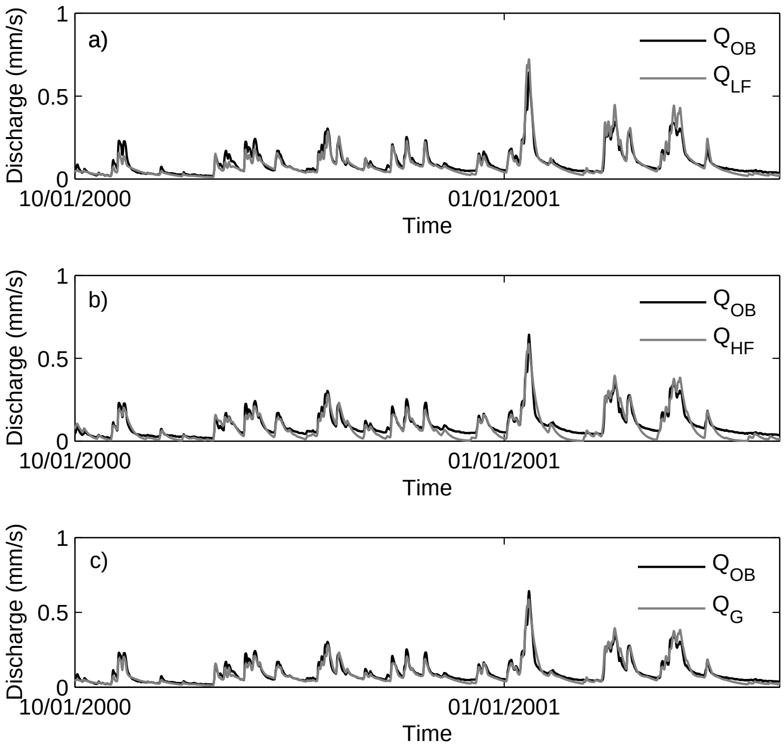

Fig. 5. Hydrograph comparison **(a)** low-flow model; **(b)** high-flow model; **(c)** model combination (Eq. 12).

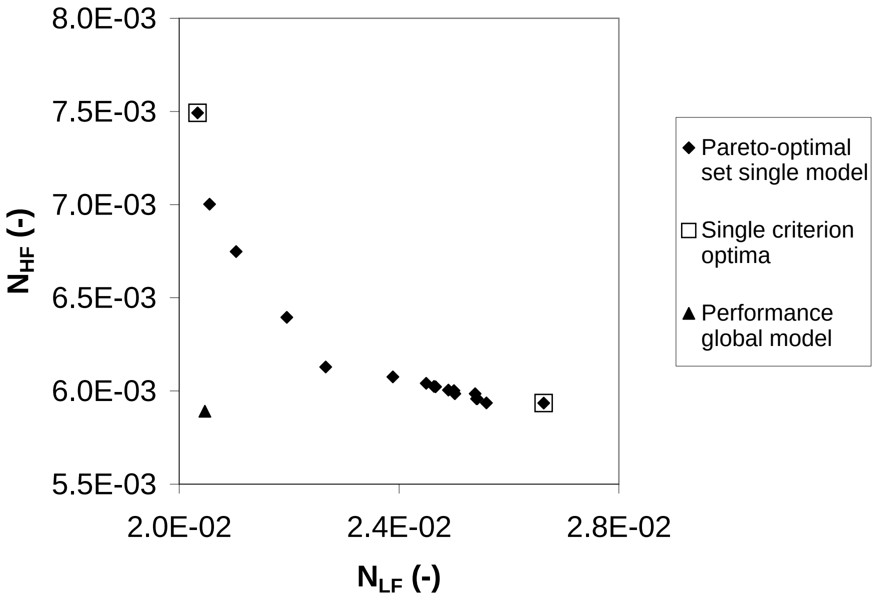

Fig. 6. Comparison of the performances of the single and the global models.

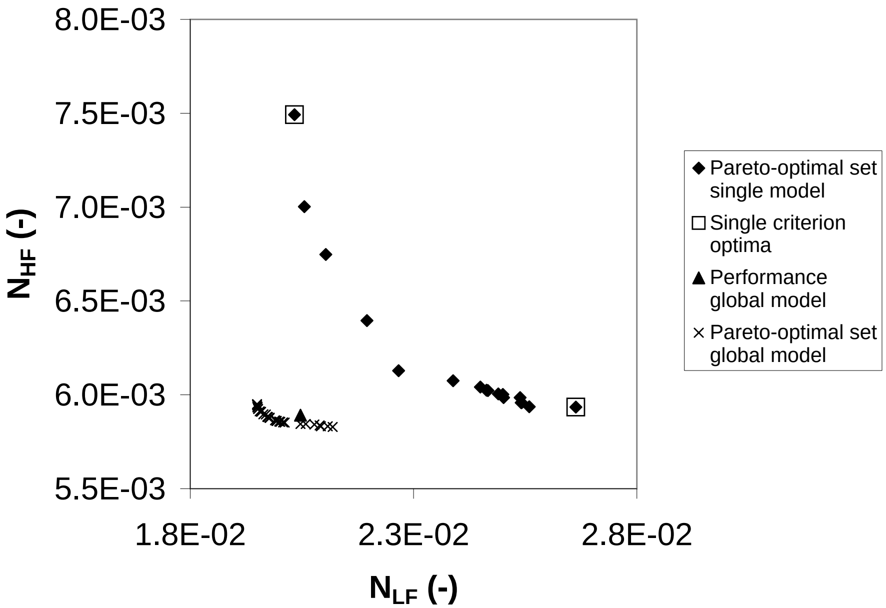

Fig. 7. Comparison of model performances between single model, global model with manually selected thresholds, and automatic tuning of the thresholds.

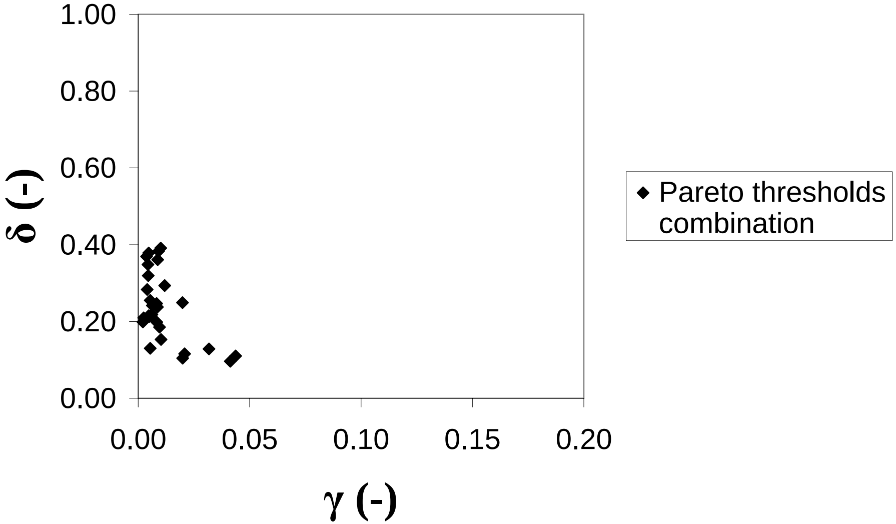

Fig. 8. Combination of thresholds corresponding to the Pareto solutions.

HESSD

4, 91–123, 2007

Local models

F. Fenicia et al.