Multi-objective calibration and fuzzy preference selection of a distributed hydrological model

Rajesh Raj Shrestha\textsuperscript{a,}* , Michael Rode \textsuperscript{b}

\textsuperscript{a}Department of Civil and Environmental Engineering, The University of Western Ontario, London, Ontario, Canada N6A 5B8
\textsuperscript{b}Department of Aquatic Ecosystem Analysis and Management, UFZ – Helmholtz Centre for Environmental Research, Brueckstrasse 3a, 39114 Magdeburg, Germany

\textbf{Article info}

Article history:
Received 7 June 2007
Accepted 4 April 2008
Available online 22 May 2008

Keywords:
Distributed hydrological model
Fuzzy set
Genetic algorithm
Multi-objective calibration
Preference selection
Streamflow uncertainty
Water balance

\textbf{Abstract}

Multi-objective evaluation of distributed hydrological models enables an analysis of prediction behaviour of individual sub-systems within a catchment. The aim of this paper is to demonstrate an application of multi-response, multisite calibration strategy for a distributed hydrological model, so that model limitations can be identified and subsequently improved. The study was carried out for calibration of flows from two gauging stations in a 152 km\textsuperscript{2} catchment in Elbe Basin in Germany. The multi-objective optimisation tool NSGA-II was used for the calibration of distributed hydrological modelling code WaSiM-ETH. A fuzzy set theory based methodology was formulated for selection of preferred solution from numerous Pareto solutions in four-dimensional space. The methodology consistently led to selection of the solution which is able to reasonably represent the magnitude and dynamics of streamflow hydrograph. For a reasonable simulation of water balance in the downstream gauge, overprediction of water balance in the upstream gauge was necessary. The analysis of precipitation–discharge data and geological conditions in the river channel support the possibility of flow reduction in the upstream gauge and increase in the downstream gauge. Due to this limitation in observation data, additional optimisation runs were carried out by explicitly considering the effect. This led to a significant improvement in the performance of the model. Therefore, the study provides an effective implementation of the multi-objective calibration strategy for a distributed hydrological model, which can be used for the analysis of different catchments using a combination of different objective functions.

\textcopyright 2008 Elsevier Ltd. All rights reserved.

\textbf{1. Introduction}

The distributed hydrological models enable the representation of spatial variability of flow characteristics within a catchment. In contrast to lumped models, distributed models are characterised by complex model structures and a large number of model parameters which need to be defined. In most of the cases, only the flow measurement at the catchment outlet is available for the calibration of these models. This leads to a problem of equifinality (Beven and Freer, 2001), as many different parameter sets within a chosen model structure can give similar model performances. In view of this problem, several authors have stressed a need for an effective methodology for parameter identification of the distributed models. Grayson et al. (1995) emphasized the importance of evaluating distributed model behaviour rather than integrated value such as runoff when assessing the performance of distributed models. Refsgaard (1997) pointed out a need for a multisite calibration/validation for spatially distributed predictions and multi-variable checks for prediction behaviour of individual sub-systems within a catchment. Several studies on multi-variable and multi-criteria evaluation of distributed models (e.g. Bergström et al., 2002; Anderton et al., 2002) suggest that such a strategy allows rejection of those parameterisations for which correct outlet discharge is being erroneously simulated as a result of internal compensating errors and leads to an increased confidence of simulation.

Multi-objective calibration based on global evolutionary optimisation methods such as genetic algorithm (Goldberg, 1989) and shuffled complex evolution and Metropolis algorithm (Duan et al., 1992; Vrugt et al., 2003) provides a framework for such a multi-variable, multi-criteria evaluation of distributed models. A wide range of applications of the multi-objective optimisation have been reported in recent literature, such as, groundwater monitoring and decision making (Reed et al., 2007; Kollat and Reed, 2007), and optimal control of urban wastewater systems (Fu et al., 2008). The method recognizes that a single objective function is not adequate to measure the important characteristics of the observed data (Yapo et al., 1998) and identifies a set of optimal Pareto solutions based on a trade-off between different objective functions. Gupta et al. (1998) demonstrated that the multi-objective approach is
The calibration of rainfall–runoff model will depend upon the quality of data (Boughton, 2006) and the quality of discharge data at individual sub-catchment will determine the success of distributed model calibration. The discharge data at the sub-catchment outlets may be affected by a number of sources of uncertainties such as reduction of flow due to hyporheic flow and/or river–aquifer exchange (Fernald et al., 2001; Konrad, 2006), and limitations of the stage–discharge relationship (Shrestha et al., 2007a). These uncertainties will propagate downstream and affect the overall calibration of the model. Andersen et al. (2001) showed that good or poor model performance in the downstream catchment may be affected by either good or poor performance upstream, which suggests propagation of uncertainties and/or compensation of errors. Due to these reasons, a simultaneous calibration of multiple sub-catchments can be a useful strategy for the systematic evaluation of the trade-off between objective functions from different sub-catchments. The study by Khu et al. (2006) on multi-site calibration of an urban drainage model showed the strength of multi-objective calibration in reducing the variability of parameters in search space. The multi-objective optimisation will also enable an evaluation of the adequacy or inadequacy of the model and provide insight into the manner in which the model may be improved (Gupta et al., 1998). Therefore, the strategy can also provide a basis for the evaluation of data and model limitations and subsequent improvement of the model.

The Pareto set of optimal solutions from a multi-objective calibration increases as more objective functions are included in calibration. However, if the model is to be used for prediction or coupled with another model such as water quality model, usually, a preferred solution with a single set of parameters is required. This leads to a decision making problem as the users of automatic calibration methods have to face a task of selecting a preferred solution from numerous Pareto optimal sets (Khu and Madsen, 2005). Khu et al. (2006) presented a preference ordering approach for the generation of limited number of optimal solutions, which is Pareto efficient in the individual subsets of objectives. However, in a multi-objective problem with conflicting objectives, it may not always be possible to find a solution, which is Pareto efficient in all individual subsets of objectives.

The aggregation of the objective functions provides an alternative methodology for the selection of preferred solution. In a multi-objective optimisation problem, which requires a treatment of a number of conflicting objectives with different objectives having different range of values, simple aggregation methods are not suitable. It is more appropriate to transform the objective values to a comparable range of values and consider partial fulfilment of each of the objectives. The fuzzy set theory provides a framework for considering the partial fulfilment of the objectives and aggregation into a composite fulfilment function. The partial fulfilment of the objectives can be associated in terms of normalised values of membership functions of fuzzy numbers between [0, 1], such that the higher the value of the membership level, higher is the degree of satisfaction of the considered objective. Fuzzy numbers also allow the incorporation of expert judgment such as acceptable degree of fulfilment according to his/her experience and intuition (Yang and Yu, 2006), so that acceptable threshold values such as maximum or minimum can be defined. A number of authors have used fuzzy sets for the aggregation of the multiple objective functions for single objective optimisations (e.g. Yu and Yang, 2000; Cheng et al., 2002). However, when the aggregate function is used for the single objective optimisation, the method only gives an approximate Pareto solution set and makes it necessary to run a large number of independent single objective optimisation runs to obtain the full Pareto solution, making this approach inefficient and time-consuming (Vrugt et al., 2003). The application of fuzzy set theory based aggregation approach after full multi-objective optimisations will preserve the diversity of the Pareto solutions. The presence of optimal points in a multi-objective space allows an examination of the solutions for each objective, which will ensure that a particular solution does not get rejected because of poor performance for one of the objectives. Similarly, this can also avoid the selection of a particular solution because of very good performance for one of the objectives. These factors are of special importance when dealing with a number of conflicting objectives, where a good performance for one objective can lead to a poor performance for another objective. In such cases, the application of preference selection methods based on the search of Pareto efficient solutions in all individual subsets of objectives (Khu et al., 2006) will not be appropriate. The aggregation of the objectives based on the fuzzy membership levels of each objective will allow the selection of preferred solution, which may not be Pareto efficient in individual sets of objective, but gives a good overall performance for all objectives.

This paper presents a multi-response, multi-site calibration of a distributed hydrological model in the framework of multi-objective optimisation. The aim of this paper is to demonstrate an application of multi-response, multi-site calibration strategy for a distributed hydrological model, so that model limitations can be identified and subsequently improved. A methodology based on the fuzzy set theory is developed for the selection of a preferred solution from the Pareto optimal sets. The study is undertaken using data from Goeltzsch catchment in Germany. Freely available distributed hydrological modelling code WaSiM-ETH (Schulla and Jasper, 2001) and multi-objective optimisation tool NSGA-II (Deb et al., 2002) are used for this study.

2. Distributed hydrological model

The Water balance Simulation Model (WaSiM-ETH; Schulla and Jasper, 2001; Schulla, 1997) was chosen for this study, which is a process based distributed modelling system capable of simulating different hydrological processes at spatial and temporal scales. The WaSiM-ETH uses a modular, object-oriented architecture for the simulation of different hydrological components such as evapotranspiration, snow accumulation, snowmelt, infiltration and generation of surface and subsurface flow components. Spatial variability is represented in terms of orthogonal grids of digital
elevation model, land use and soil type. The model uses precipitation, temperature, global radiation, relative sunshine duration, wind velocity, relative humidity and vapour pressure as temporal data. The WaSiM-ETH model can be used in either TOPMODEL or Richards equation based versions for the generation of subsurface flow. The application of the Richards equation based method is restricted to cases where detailed soil data with vertical soil profiles are available. This is not the case in the study catchment, therefore this version is not considered in this study. The TOPMODEL version of the model is also chosen for this catchment as it has been successfully applied in similar mountainous catchments in the region (Rode and Lindenschmidt, 2001; Lindenschmidt et al., 2004; Shrestha et al., 2007b).

The TOPMODEL version of the WaSiM-ETH performs soil water balance and generates runoff based on variable saturated areas in the catchment. The model uses Green and Ampt equation for the calculation of infiltration based on soil moisture conditions, and generates surface runoff when soil infiltration capacity is exceeded. The TOPMODEL (Beven and Kirkby, 1979) is a conceptual variable contributing area approach based on the distribution of saturation deficit. In this WaSiM-ETH implementation of TOPMODEL, soil water balance and runoff generation are simulated separately for each grid cell based on the spatial distribution of soil topographic index. The model generates surface runoff when the unsaturated zone is filled, which is routed in the streams using a kinematic wave approach (Schulla and Jasper, 2001).

3. Multi-objective optimisation

The solution of the multi-objective optimisation problem is based on all feasible trade-offs between the multiple objectives for a set of optimal solutions (Goldberg, 1989). A general multi-objective problem can be formulated as a vector function \( f_i \) that maps \( m \) number of decision variables (parameters) to \( n \) number of objectives (Zitzler and Thiele, 1998):

\[
\begin{align*}
\text{minimise/maximise} & \quad y = \{f_1(x), f_2(x), \ldots, f_n(x)\} \\
\text{subject to} & \quad x = \{x_1, x_2, \ldots, x_m\} \in X \\
& \quad y = \{y_1, y_2, \ldots, y_n\} \in Y
\end{align*}
\]

where \( x \) is the decision vector, \( X \) is the parameter space, \( y \) is the objective vector, and \( Y \) is the objective space.

The solution of the optimisation problem given by Eq. (1) is not unique, but consists of a set of Pareto optimal solution, which can be described as a vector of non-dominated solutions in the objective space. For example, in a minimisation problem with two decision vectors \( a, b \), \( a \) is said to dominate \( b \) (denoted by \( a \prec b \)), if and only if:

\[
y \in \{1, 2, \ldots, n\} : f_i(a) \leq f_i(b) \forall i \in \{1, 2, \ldots, n\} : f_i(a) < f_i(b)
\]

All decision vectors of a given set which are not dominated by any other decision vector using the criteria specified by Eq. (2) are called non-dominated regarding this set. A set of decision variables that are non-dominated within the entire search space constitute the Pareto optimal set (Zitzler and Thiele, 1998).

3.1. Nondominated sorting genetic algorithm-II (NSGA-II)

The NSGA-II is a fast and elitist multi-objective genetic algorithm (Deb et al., 2002; Deb, 2006), which is capable of finding multiple Pareto solutions in a single optimisation run. Key features of the NSGA-II are efficient sorting algorithm and maintaining of a diverse set of elite population. The NSGA-II does not require any additional implicit or explicit parameters other than the standard genetic algorithm parameters, such as population size, operator probabilities, etc., which is probably the reason for its popularity (Deb, 2006).

In general, the NSGA-II initializes by generating a random set of initial population of size \( p \). Once the population is initialized, it is sorted based on nondomination into Pareto optimal fronts. In order to reduce the computational complexity, the NSGA-II uses a special book-keeping strategy for a faster and efficient comparison of the solutions. Each solution is assigned a fitness (rank) equal to its nondomination level using the criteria by Goldberg (1989). The algorithm uses binary tournament selection, recombination and mutation operators to create offspring population of size \( p \). Since elitism is important for the effectiveness of the search (Zitzler and Thiele, 1998), the NSGA-II introduces elitism by including all population from current and previous generations and a combined population of size \( 2p \) is formed. The population is sorted again based on nondomination and a new population of size \( p \) is selected. The diversity amongst the non-dominated solution is introduced by crowding distance comparison procedure, which is used during tournament selection and population reduction phase. The crowding distance measures how close an individual is to its neighbours and a larger crowding distance will result in better diversity in the population. From the selected population, a new offspring is created and the procedure is continued for subsequent generations until the stopping criteria specified by number of generations are met.

3.2. Fuzzy preference selection

In this study, a fuzzy preference selection scheme is proposed, which is based on the aggregation of partial fulfilment of each of the objective functions obtained from the multi-objective optimisation. Although the approach is similar to the aggregation of multiple objective functions for a single objective optimisation, the fundamental difference is that the aggregation is performed after the full multi-objective optimisation run. Hence, diversity of the non-dominated solution is maintained.

A basic tool for the evaluation of partial fulfilment of objectives is the fuzzy number, which can be defined on a fuzzy set, with numerical value of the objective functions in a domain assigned to a specific grade of membership between and including 0 and 1. Since the non-dominated solutions consist of a diverse set of populations with wide range of values of objective functions, it may be possible that a number of solutions may have unacceptable range of values for one or more objective functions. Therefore, maximum threshold value \( y_{\text{max}} \) (in a minimisation problem) for each of the objective is defined, which corresponds to a membership grade of 0. The membership grade of 1 corresponds to the minimum objective values \( y_{\text{min}} \). Hence, the membership function \( \mu_\alpha(y) \) of a fuzzy set \( A \) can be taken as continuous strictly decreasing function \( g(\cdot) \), defined on \([0, 1]\) and satisfying the conditions (Fig. 1):

\[
\mu_\alpha(y) = \begin{cases} 
1 & \text{if } y_1 \leq y_{\text{min}} \\
1 - \alpha (y_{\text{max}} - y) & \text{if } y_{\text{min}} < y < y_{\text{max}} \\
0 & \text{if } y \geq y_{\text{max}}
\end{cases}
\]

where \( \alpha \) is the spread of the membership function given by \( y_{\text{max}} - y_{\text{min}} \).
The definition of fuzzy numbers for each of the objective allows the calculation of the membership levels of the objective vectors. For the evaluation of the overall performance of an objective vector, a combined response can be derived from the individual membership levels of all the objective functions. This is similar to the calculation of degree of fulfilment (ν) in a fuzzy rule based system as defined in Bárdozy and Duckstein (1995). The ν value can be determined using several logical operators such as product inference and min–max inference. The product inference accounts for the fulfillment of all elements of a vector in contrast to the min–max inference and min–max inference. The product inference is calculated using several logical operators such as product inference and min–max inference. The product inference allows for the calculation of the membership levels of all the objective vectors. This is similar to the calculation of degree of fulfilment (ν) in a fuzzy rule based system as defined in Bárdozy and Duckstein (1995). The ν value can be determined using several logical operators such as product inference and min–max inference. The product inference accounts for the fulfillment of all elements of a vector in contrast to the min–max inference where only the limiting or extreme element is considered (Bárdozy and Duckstein, 1995). As the aim of the combination in this study is the evaluation of degree of fulfilment of all elements of the objective vector, only the product inference is considered, which can be expressed as

\[ r = \mu_{A_1}(y_1) \ast \mu_{A_2}(y_2) \ast \ldots \ast \mu_{A_n}(y_n) \]  

(4)

The degree of fulfilment is used for the selection of preferred solution, so that higher the value of the degree of fulfilment, higher is the preference level of that particular solution.

4. Study area and data

4.1. Study area

The Goeltzsch catchment comprises an area of 152.2 km² in Vogtland mountainous region in the state of Saxony in Germany. The Goeltzsch River is a tributary of the Weisse Elster River in the Elbe River basin. The major tributaries of the Goeltzsch River are Eulenwasser, Wernesbach and Plohnbach. There are two gauging stations in the Goeltzsch catchment with the upper gauge at Rodewisch covering an area of 71.5 km² and the lower gauge at Mylau covering a sub-catchment of 80.7 km² downstream of the Rodewisch gauging station (Fig. 2). The Rodewisch gauging station was established as a part of the flood warning network and has been in operation since 1997.

The spatial data obtained from the catchment included a digital terrain model of 100 m grid resolution, land use data based on the classification of Landsat ETM images from the year 1999 and soil data based on 1:400 000 soil map. There are substantial differences in the characteristics of the upstream and the downstream subcatchments. The elevation in the upstream sub-catchment varies between 417 and 787 m and the downstream sub-catchment between 312 and 653 m. The land use in the upstream sub-catchment is dominated by forest (53%) followed by settlement (17%), grassland (16%) and arable land (11%). In the downstream sub-catchment, the main land use types include forest (32%), grassland (28%), arable land (29%) and settlement (9%). Soil types in the two sub-catchments are quite similar and dominated by Brown soil and Podsol-Brown soil.

Geology of both the sub-catchments is dominated by metamorphic formations of phyllite and siltstone. The geological characteristics of the catchment can affect the flows in the gauging stations as flow reduction can occur from deep fissures’ aquifers and/or flow in hyporheic zone and river–aquifer exchange. The analysis by LiUG (2007) did not find evidence of water leakage from upstream sub-catchment from fissured aquifer and noted a possibility of other sources of reduction in the upstream gauging station. In the reach of about 7 km where the tributaries Eulenwasser and Wernesbach flow into the Goeltzsch River, the river flows along unconsolidated sedimentary deposits of sand and gravel (Fig. 2). The porous unconsolidated composition creates conditions conducive to hyporheic flow and river–aquifer exchange (Fernald et al., 2001; Konrad, 2006). A substantial fraction of streamflow can pass through the hyporheic zone as a part of ground water/surface water continuum without a net downstream change in streamflow (Fernald et al., 2001; Hinkle et al., 2001). River–aquifer exchanges may also affect the flow in the channels when a river loses flow to a shallow aquifer that discharges back to the river in a downstream reach (Konrad, 2006). Due to location of Rodewisch gauging station in the porous bed composition, there is a possibility that the discharge data from the gauge are affected by these phenomena, especially during low flow period.

4.2. Analysis of precipitation and discharge data

The discharge and average precipitation data from Rodewisch and Mylau gauging stations were analysed from the years 1997 to 2003. The precipitation data used for the analysis consisted of average upstream sub-catchment precipitation for the Rodewisch gauging station and average precipitation for the whole catchment for the Mylau gauging station. The cumulative average precipitation and specific discharge data presented in Fig. 3 show that although average precipitation at Rodewisch sub-catchment is about 5.5% higher than the total average catchment precipitation, the specific discharge at Rodewisch is about 13.9% lower than the total catchment specific discharge at Mylau. The effect of higher precipitation in the upstream area was also reflected in the LiUG (2007) simulation results, which produced higher runoff generation in the upstream. The statistical characteristics of the discharge data from the two gauging stations (Table 1) show that the maximum, minimum, mean and median discharges at the Rodewisch gauge are less than the Mylau gauge during both the dry period (April–September) and wet period (October–March). The substantially lower mean and median discharge values during the wet and dry periods point to the fact that the low flow discharge at Rodewisch gauging station is significantly lower than that at Mylau gauging station. These factors affect the total discharge in the Rodewisch gauge leading to lower cumulative values. This combination of lower cumulative discharge with geological conditions suggests that flow reduction in the upstream and increase in the downstream gauging station may be prevalent in the catchment.
5. Model set up and calibration

The distributed model WaSiM-ETH was set up for the Goeltzsch catchment for the calibration of flows at Rodewisch and Mylau gauging stations. Spatial data consisting of digital elevation model, soil and land use classes of 100 m grid resolution and temporal data consisting of precipitation at daily time step from 10 stations in the catchment and vicinity were used. In addition,

![Study area: Goeltzsch catchment.](image)

**Fig. 2.** Study area: Goeltzsch catchment.

![Comparison of cumulative (a) precipitation, and (b) specific discharge at Rodewisch and Mylau gauging stations.](image)

**Fig. 3.** Comparison of cumulative (a) precipitation, and (b) specific discharge at Rodewisch and Mylau gauging stations.
climatic data of relative humidity, relative sunshine hours, temperature and wind velocity from four climatic stations in the vicinity of the catchment were used. Spatial interpolation of the time series data was carried out using inverse distance weighting algorithm.

The calibration of Goeltzsch catchment model was first carried out without considering any flow reduction/exchange in the gauging stations. In the second case, the flow reduction in the upstream gauging station and increase in the downstream gauging station were explicitly considered. A set of eight parameters of the WaSiM-ETH model for each of the sub-catchment were used for the calibration. The calibration parameters are summarised in Table 2. Previous study by Schulla and Jasper (2001) has identified the most sensitive parameters as recession parameter \( m \) and correction factor for the transmissivity of the soil \( T_{\text{cor}} \). Most of the parameters are dependent on each other, for example, the scaling factor for capillary rise \( r_c \) and threshold precipitation for infiltration \( P_{\text{thres}} \) have opposite effects on soil water storage. Therefore, it is necessary to calibrate all of these parameters together. For the calibration of these parameters in a multi-objective framework, two objective functions for each of the sub-catchments were chosen. These include Nash Sutcliffe coefficient of efficiency (NSCE\(_k\)) and overall water balance (WBa\(_k\)). The 1–NSCE\(_k\) was used in this study to express the objective function for minimisation, which corresponds to the ratio of mean square error and variance of observed data. As it includes the mean square error term, it provides a suitable measure of “goodness of fit” of the model (Legates and McCabe, 1999). The WBa\(_k\) represents the ratio of the difference between cumulative observed and simulated discharges to the cumulative observations and indicates how well the model is able to reproduce the overall water cycle in the catchment. This criterion is especially important as the characteristics of average rainfall and specific discharge in the upstream and downstream sub-catchments show important as the characteristics of average rainfall and specific discharge in the upstream and downstream sub-catchments show important as the characteristics of average rainfall and specific discharge in the upstream and downstream sub-catchments show important

### Table 1

<table>
<thead>
<tr>
<th>Period</th>
<th>Station</th>
<th>Maximum</th>
<th>Minimum</th>
<th>Mean</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agr.–Sep.</td>
<td>Rodewisch</td>
<td>16.393</td>
<td>0.027</td>
<td>0.539</td>
<td>0.307</td>
</tr>
<tr>
<td></td>
<td>Mylau</td>
<td>20.141</td>
<td>0.068</td>
<td>0.717</td>
<td>0.481</td>
</tr>
<tr>
<td>Oct.–Mar.</td>
<td>Rodewisch</td>
<td>8.462</td>
<td>0.065</td>
<td>1.326</td>
<td>0.951</td>
</tr>
<tr>
<td></td>
<td>Mylau</td>
<td>9.391</td>
<td>0.143</td>
<td>1.396</td>
<td>1.161</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Name</th>
<th>Unit</th>
<th>Lower limit</th>
<th>Upper limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recession parameter</td>
<td>( m )</td>
<td>[mm]</td>
<td>0.001</td>
<td>0.1</td>
</tr>
<tr>
<td>Correction factor for transmissivity of soil</td>
<td>( T_{\text{cor}} )</td>
<td>[–]</td>
<td>0.00001</td>
<td>0.001</td>
</tr>
<tr>
<td>Correction factor for vertical percolation</td>
<td>( K_{\text{cor}} )</td>
<td>[–]</td>
<td>1.0</td>
<td>50</td>
</tr>
<tr>
<td>Recession constant for surface runoff</td>
<td>( k_D )</td>
<td>[h]</td>
<td>1.0</td>
<td>80</td>
</tr>
<tr>
<td>Recession constant for interflow</td>
<td>( k_H )</td>
<td>[h]</td>
<td>20</td>
<td>250</td>
</tr>
<tr>
<td>Precipitation threshold for infiltration</td>
<td>( P_{\text{thres}} )</td>
<td>[mm h (^{-1})]</td>
<td>0.01</td>
<td>10</td>
</tr>
<tr>
<td>Scaling factor for capillary rise</td>
<td>( r_c )</td>
<td>[–]</td>
<td>0.1</td>
<td>1.0</td>
</tr>
<tr>
<td>Fraction of snowmelt on surface runoff</td>
<td>( c_{\text{mel}} )</td>
<td>[m]</td>
<td>0.01</td>
<td>1.0</td>
</tr>
</tbody>
</table>

\[
1 - \text{NSCE}_k = \frac{\sum_{j=1}^{N} (Q_{j,\text{obs}} - Q_{j,\text{sim}})^2}{\sum_{j=1}^{N} (Q_{j,\text{obs}} - \bar{Q}_{\text{obs}})^2}
\]

\[
\text{WBa}_k = \frac{\sum_{j=1}^{N} Q_{j,\text{obs}} - \sum_{j=1}^{N} Q_{j,\text{sim}}}{\sum_{j=1}^{N} Q_{j,\text{obs}}}
\]

where \( N \) is the number of observations, \( Q_{j,\text{obs}} \) and \( Q_{j,\text{sim}} \) are the observed and simulated discharges at time step \( j \) for sub-catchment \( k \) and \( \bar{Q}_{\text{obs}} \) is the mean of the observed discharge.

In addition, the mean absolute error (MAE\(_k\)) was used for the evaluation of results of model calibration and validation:

\[
\text{MAE}_k = \frac{1}{N} \sum_{j=1}^{N} |Q_{j,\text{obs}} - Q_{j,\text{cal}}|
\]

For the NSGA-II multi-objective optimisation runs, minimisation of the objective functions was taken as the optimisation problem. Therefore, NSCE\(_k\) subtracted from unity was used as the objective. Ten independent optimisation runs of NSGA-II were carried out with population size between 100 and 150 and number of generation between 20 and 40. The range of parameter values was taken from the calibrated WaSiM-ETH model from the studies of similar catchments in the region (Rode and Lindenschmidt, 2001; Lindenschmidt et al., 2004; Shrestha et al., 2007b), which is given in Table 2. Three years and nine months data from April 2000 to December 2003 were used for model calibration and further three years data from April 1997 to March 2003 were used for independent evaluation (validation) of model results. A warm-up period of 1 year was used so that initial conditions do not affect the model calibration.

For the selection of preferred optimal solution using the fuzzy degree of fulfilment (Eq. (4)), each of the objective values was mapped with their predefined fuzzy numbers as given in Eq. (3). For the selection of the threshold values (\( \mu_{\text{max}} \)), the acceptable ranges were considered with regard to the physical meaning of both 1 – NSCE\(_k\) and WBa\(_k\). A 1 – NSCE\(_k\) value of 0.5 can be regarded as poor model performance as the mean square error is 50% of the variance. Therefore 0.5 was taken as the maximum acceptable threshold value (\( \mu_{\text{max}} \)) for the 1 – NSCE\(_k\), which corresponds to membership level 0. The minimum value (\( \mu_{\text{min}} \)) of 1 – NSCE\(_k\) for membership level 1 was taken as 0.2, which corresponds to minimum values 1 – NSCE\(_k\) obtained (0.2–0.3). Similarly, for the WBa\(_k\), a value of 0.25 means that the deviation between observed and simulated discharge is 25% of the observed discharge, which can be regarded as poor model performance. Therefore value of 0.25 was taken as membership level 0. A WBa\(_k\) of 0 means that the difference between observed and simulated discharge is zero, therefore can be represented by membership level 1. The membership function of the fuzzy numbers of both 1 – NSCE\(_k\) and WBa\(_k\) were chosen to vary linearly between 0 and 1. From the values of the membership levels of each of the objective functions, the degree of fulfilment is obtained by multiplication for each of the NSGA-II simulation.

### 6. Results and discussion

#### 6.1. Model calibration without considering flow reduction/increase in the gauging stations

From 10 independent NSGA-II multi-objective optimisation runs, a total of 1150 Pareto points in the four-dimensional space were obtained. The trade-off between the various objective functions in the two-dimensional space is shown in Fig. 4. The results
illustrate that the use of four objective functions from the upstream and downstream gauging stations leads to a large trade-off between the objective functions. The solutions show that there exist only a limited number of Pareto optimal points in the two-dimensional space. Further, none of the Pareto optimal points in the two-dimensional space belongs to the Pareto optimal points in all four trade-offs. Due to this reason, the preference ordering methodology as proposed by Khu and Madsen (2005) could not be used for the selection of the preferred solution. The trade-off between the WBal from Rodewisch and Mylau gauging stations shows a peculiar pattern consisting of three straight lines (Fig. 4d), which indicate linear relationships in water balance between the gauging stations.

The correlation analysis between the objective functions taken two at a time resulted in the correlation coefficient of less than 0.1 between the most of the objectives, which indicate a lack of any significant relationships. The only exception is the WBal1 and WBal2 (where the indices 1 and 2 correspond to Rodewisch and Mylau gauging stations), which has a correlation coefficient of −0.38. This suggests that a good water balance in the downstream gauge can be obtained with a poor water balance in the upstream gauge.

A total of 196 solutions with degree of fulfilment greater than zero were obtained from the total of 1150 solutions. Fig. 5 shows the membership levels of each of the solutions with degree of fulfilment greater than zero. The preferred solution obtained from the fuzzy preference selection methodology results in reasonably high membership levels of the objectives except the WBal1. The preferred solution lies only in the 1 – NSCE1 – WBal1 2D Pareto fronts (Fig. 4). This implies that the obtained band of solutions has a high variability and the best solution in the trade-off between two parameters does not necessarily mean that the solution also lies in the Pareto trade-off between other objective functions.

The visualisation of partial degree of fulfilment of each of the objectives provides information about the performance of the

![Fig. 4. Trade-off between various objective functions and 2D Pareto optimal solutions without consideration of flow reduction/increase in the gauging stations (a) 1 – NSCE2 in the upstream and downstream gauging stations, (b) 1 – NSCE1 and WBal in the upstream gauging station, (c) 1 – NSCE2 and WBal2 in the downstream gauging station and (d) WBal in the upstream and downstream gauging stations.](image1)

![Fig. 5. Membership levels of four objective functions with degree of fulfilment greater than zero (without consideration of flow reduction/increase in the gauging stations). The thick line represents the preferred solution.](image2)
model with regard to each of the objective function. There are some apparent patterns such as higher membership value of the \(1/\text{NSCE}_1\) in the upstream gauge leads to lower membership value of the \(1/\text{NSCE}_2\) in the downstream gauge and vice versa (Fig. 5). The membership levels of the \(\text{WBal}_k\) in the upstream and downstream gauges show a band of solutions with lower membership levels in the upstream and higher membership level in the downstream. This indicates not only a poor water balance in the upstream gauge, but also the fact that better water balance in downstream can be obtained by poor water balance upstream.

The comparison of the results showing observations, Pareto solution band with degree of fulfillment greater than zero and preferred solution from fuzzy preference selection (without consideration of flow reduction/increase in the gauging stations) for (a) Rodewisch gauging station (calibration), (b) Mylau gauging station (calibration), (c) Rodewisch gauging station (validation), (d) Mylau gauging station (validation) and (e) partial results from Rodewisch gauging station (validation).

![Fig. 6. Comparison of observations, Pareto solution band with degree of fulfillment greater than zero and preferred solution from fuzzy preference selection (without consideration of flow reduction/increase in the gauging stations) for (a) Rodewisch gauging station (calibration), (b) Mylau gauging station (calibration), (c) Rodewisch gauging station (validation), (d) Mylau gauging station (validation) and (e) partial results from Rodewisch gauging station (validation).](image)

Reasonably well, the observed low flows are either outside or at the bottom end of the Pareto solution band (Fig. 6e). The incorrect simulation of low flow in the upstream gauge causes water balance of the model to be overestimated. The statistical performance of the preferred solution summarised in Table 3 also indicates that poor

<table>
<thead>
<tr>
<th>Data sets</th>
<th>Gauging station</th>
<th>NSCE(_k)</th>
<th>WBal(_k)</th>
<th>MAE(_k) [mm/d]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calibration</td>
<td>Rodewisch</td>
<td>0.715</td>
<td>0.183</td>
<td>0.352</td>
</tr>
<tr>
<td></td>
<td>Mylau</td>
<td>0.660</td>
<td>0.019</td>
<td>0.318</td>
</tr>
<tr>
<td>Validation</td>
<td>Rodewisch</td>
<td>0.718</td>
<td>0.218</td>
<td>0.313</td>
</tr>
<tr>
<td></td>
<td>Mylau</td>
<td>0.729</td>
<td>0.027</td>
<td>0.328</td>
</tr>
</tbody>
</table>
water balance in the upstream gauging station at Rodewisch is the main problem of the model.

The trade-offs between objective functions from the two gauging stations using multi-objective optimisation method led to consistent results showing overestimation of flow in the upstream gauge. This provides further evidence on the limitations of the Rodewisch gauge data. It will not be possible to detect such limitations in the model without considering the trade-offs between the objective functions from the two gauging stations. The results also provide insights into how the model can be improved, i.e. flow reduction in the upstream gauge, and increase in the downstream gauge, which is assessed in the following section.

6.2. Model calibration with consideration of flow reduction/increase in the gauging stations

The geological conditions, analysis of rainfall and discharge data and the problem of water balance in the upstream gauging station all point to flow reduction in the upstream gauge and gain in the downstream. Due to these reasons the effect of flow reduction in the upstream gauge and increase in the downstream gauge was considered explicitly for further multi-objective calibration runs. For this purpose, the WaSiM-ETH model was set up with abstraction from the upstream sub-catchment and inflow to the downstream sub-catchment. This requires an estimation of three parameters for abstraction: threshold discharge for abstraction to occur, fraction of abstracted water beyond threshold and maximum rate of abstraction (Schulla and Jasper, 2001). A number of preliminary multi-objective optimisation runs were carried out to find an appropriate combination of these parameters. Based on the preliminary runs a suitable combination of these parameters was selected and fixed for further multi-objective calibration runs. It is to be noted that these parameter values could have been calibrated together with other 16 parameters. However, higher number of parameters would require higher population size and number of iterations, which would make the optimisation process more time-consuming. The use of the same number of parameters

![Diagram](image_url)

**Fig. 7.** Trade-off between various objective functions and 2D Pareto optimal solutions with consideration of flow reduction/increase in the gauging stations. (a) $1 - \text{NSCE}_1$ in the upstream and downstream gauging stations, (b) $1 - \text{NSCE}_2$ and $\text{WBal}_1$ in the upstream gauging station, (c) $1 - \text{NSCE}_2$ and $\text{WBal}_2$ in the downstream gauging station and (d) $\text{WBal}_k$ in the upstream and downstream gauging stations.

![Diagram](image_url)

**Fig. 8.** Membership levels of four objective functions with degree of fulfilment greater than zero (with consideration of flow reduction/increase in the gauging stations). The thick line represents the preferred solution.
as in the previous case without consideration of flow reduction/increase in the gauging stations also allows the comparison of results of two cases.

After setting the appropriate abstraction rates, 10 independent NSGA-II multi-objective optimisation runs were performed in this case too and a total of 1150 Pareto points in the four-dimensional space were obtained. The trade-off between various objective functions in the two-dimensional space are shown in Fig. 7. The results show a noticeable improvement in performance, especially that of the \( 1 - \text{NSCE}_1 \) (Rodewisch). It can also be seen that the variability of \( \text{WBal}_k \) in both the gauging stations have reduced from 0 to 0.6 in the previous case to 0–0.3. A linear pattern between the upstream and downstream water balance was also observed in this case.

Unlike in the previous case, without the consideration flow reduction/increase in the gauging stations, significant correlation was obtained in this case between various objective functions. The correlation coefficients of 0.40, −0.42, −0.35, 0.54 were obtained between the \( 1 - \text{NSCE}_1 \) and \( 1 - \text{NSCE}_2 \), \( 1 - \text{NSCE}_1 \) and \( \text{WBal}_1 \), \( 1 - \text{NSCE}_2 \) and \( \text{WBal}_2 \), and \( \text{WBal}_1 \) and \( \text{WBal}_2 \), respectively. This indicates relationships between these variables, such as lower value of \( 1 - \text{NSCE}_1 \) leads to lower value \( 1 - \text{NSCE}_2 \). A positive value of correlation coefficient between \( \text{WBal}_1 \) and \( \text{WBal}_2 \) indicates that a good water balance in the upstream gauge also leads to a good water balance in the downstream gauge, which is opposite to the previous case.

From the selected of preferred optimal solution using the fuzzy preference selection methodology, 569 solutions with degree of fulfilment greater than zero were obtained, which is considerably higher than 196 solutions obtained in the previous case. Fig. 8 shows the membership levels of each of the solutions with degree

---

**Fig. 9.** Comparison of observations, Pareto solution band with degree of fulfilment greater than zero and preferred solution from fuzzy preference selection for (with consideration of flow reduction/increase in the gauging stations); (a) Rodewisch gauging station (calibration), (b) Mylau gauging station (calibration), (c) Rodewisch gauging station (validation), (d) Mylau gauging station (validation) and (e) partial results from Rodewisch gauging station (validation).
of fulfilment greater than zero. The preferred solution selected with the highest value of the degree of fulfilment results in a reasonably high membership levels and therefore reasonably low 1 – NSCE and WBal values for all four objective functions. In this case also, only the trade-off between NSCE and WBal of the preferred solution lies in the Pareto optimal fronts of the 2D plots (Fig. 7). Fig. 9 presents the comparison of the results of observations, Pareto solution band with degree of fulfilment greater than zero and preferred solution from fuzzy preference selection. In this case, the Pareto solution band is able to represent the magnitude and dynamics of the high as well as low flows better than the previous case (Fig. 9e). The preferred solution obtained from the fuzzy preference selection methodology also shows an improvement over the previous case, which is also apparent in the membership level plot (Fig. 8), where higher membership levels of the 1 – NSCE and WBal were obtained. The statistical performance of the preferred solution summarised in Table 4 also shows significant improvement in the model performance for the upstream gauging station. The statistical performance in terms of the criteria NSCE and WBal has improved considerably and there is also an overall improvement of the MAE in comparison to the previous case.

The overall improvement of the performance of the model with consideration of the flow reduction/increase in the gauging stations leads to the fact that the model is better able to represent the catchment characteristics. As the proposition of flow reduction in the upstream gauging station is further supported by geological conditions in the stream channel and analysis of the rainfall–runoff data, it can be concluded that the assumption is very reasonable. The best solution obtained from the fuzzy preference selection methodology, therefore, can be taken for further analyses and in combination with other models.

7. Conclusions

This paper has presented a multi-objective optimisation and fuzzy set theory based preference selection methodologies for the calibration of a distributed hydrological model. The methodologies have been demonstrated for a multi-response, multi-site calibration using data from two gauging stations in the Goeltzsch catchment in Germany. The methods are independent of catchment characteristics and also applicable for different catchment type and a combination of different objective functions.

The results of the study show that the application of multi-objective optimisation method with the trade-off between four objective functions in the upstream and downstream sub-catchments leads to a good overall calibration and a better representation of the catchment characteristics. Using the fuzzy preference selection methodology, which is based on the composite degree of fulfilment of each of the objective functions, the best solution from numerous four-dimensional Pareto solution was chosen. The methodology considers the quality of all the objective functions and consistently led to the solution which is able to represent the magnitude and dynamics of the streamflow hydrograph. The method also allows the visualisation of partial degree of fulfilment of each of the objectives, therefore the performance of the model with regard to each of the objective function. Due to these reasons, the fuzzy preference selection can be viewed as an effective and objective methodology for the selection of the preferred solution from numerous Pareto solutions.

Using the methodologies of the multi-objective optimisation and fuzzy preference selection, the limitations of the available calibration data were identified. The poor water balance in the upstream sub-catchment, together with the geological conditions of the river channel, and analysis of the precipitation and discharge data led to the proposition that the flow reduction and increase are prevalent in the upstream and downstream gauging stations, respectively, which might be due to hyporheic flow and/or river–aquifer exchange. It will not be possible to detect such limitations in the model without considering the trade-offs between the objective functions from the two sub-catchments. Therefore it can be concluded that multi-objective optimisation method has further facilitated the identification of the limitations of the model. Based on the identification of the limitations, flow reduction/increase in the gauging stations was introduced to the model, which led to the improvement of the overall performance of the model. Therefore, the best solution obtained from the fuzzy preference selection methodology with the consideration of flow reduction/increase can be used for further analysis and in combination with other models. The results of the study can also be used for the improvement of the process description of the model such as the incorporation of hyporheic flow and river–aquifer exchange in this particular case. The methodology described in this study can also be used as the basis for the evaluation of different catchments using a combination of different objective functions. Further work in this study will be an investigation of parameter and input uncertainties and their effect on the model output.

References


