Journal of Hydrology 377 (2009) 208–225

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Continuous simulation for flood estimation in ungauged mesoscale catchments of Switzerland – Part II: Parameter regionalisation and flood estimation results ☆

Daniel Viviroli [a,b,*], Heidi Mittelbach [c], Joachim Gurtz [c], Rolf Weingartner [a,b]

[a] Institute of Geography, University of Bern, Hallerstrasse 12, CH-3012 Bern, Switzerland
[b] Oeschger Centre for Climate Change Research, University of Bern, Zähringerstrasse 25, CH-3012 Bern, Switzerland
[c] Institute for Atmospheric and Climate Science, ETH Zürich, Universitätsstrasse 16, CH-8092 Zürich, Switzerland

ARTICLE INFO

Article history:
Received 10 April 2009
Received in revised form 8 July 2009
Accepted 17 August 2009

This manuscript was handled by K. Georgakakos, Editor-in-Chief, with the assistance of Emmanouil N. Anagnostou, Associate Editor

Keywords:
Flood estimation
Continuous simulation
Regionalisation
Ungauged basins
Precipitation–runoff-model
Switzerland

SUMMARY

Flood estimations for ungauged mesoscale catchments are as important as they are difficult. So far, empirical and stochastic methods have mainly been used for this purpose. Experience shows, however, that these procedures entail major errors. In order to make further progress in flood estimation, a continuous precipitation–runoff-modelling approach has been developed for practical application in Switzerland using the process-oriented hydrological modelling system PREVAH (Precipitation–Runoff–EVApotranspiration-HRU related model).

The main goal of this approach is to achieve discharge hydrographs for any Swiss mesoscale catchment without measurement of discharge. Subsequently, the relevant flood estimations are to be derived from these hydrographs. On the basis of 140 calibrated catchments (Viviroli et al., 2009b), a parameter regionalisation scheme has been developed to estimate PREVAH's tuneable parameters where calibration is not possible. The scheme is based on three individual parameter estimation approaches, namely Nearest Neighbours (parameter transfer from catchments similar in attribute space), Kriging (parameter interpolation in physical space) and Regression (parameter estimation from relations to catchment attributes). The most favourable results were achieved when the simulations using these three individual regionalisations were combined by computing their median.

It will be demonstrated that the framework introduced here yields plausible flood estimations for ungauged Swiss catchments. Comparing a flood with a return period of 100 years to the reference value derived from the observed record, the median error from 49 representative catchments is only −7%, while the error for half of these catchments ranges between −30% and +8%. Additionally, our estimate lies within the statistical 90% confidence interval of the reference value in more than half of these catchments. The average quality of these flood estimations compares well with present empirical standard procedures, while the range of deviations is noticeably smaller. Additionally, the availability of complete hydrographs and the process-oriented background bear potential for analyses that go beyond the mere estimation of peak flows.

Introduction

Reliable estimates for peak flow values with various return periods are an indispensable prerequisite for planning measures which reduce or even prevent flood damage (see e.g. Pilon, 2004). Particularly on the mesoscale (drainage area of roughly 10–1000 km^2 in the present case), there is a great need for such estimates, as was e.g. shown in the aftermath of the 2005 flood events in the European Alps (Bezzola and Hegg, 2007).

For catchments with long gauge records, floods with various recurrence intervals are estimated with relatively little effort using extreme value statistics (DVWK, 1999). However, the results for rare events are noticeably influenced by the choice of theoretical extreme value distribution function and parameter estimation method (see Vogel et al., 1993; Klemeš, 2000), and different conditions and processes governing individual flood events are usually not considered. Above all, gauge records are too short or totally absent in the majority of cases.

Far more frequently, however, flood estimates are sought for ungauged catchments. This refers to the concept of regionalisation,

☆ This is the companion paper of "Continuous simulation for flood estimation in ungauged mesoscale catchments of Switzerland – Part I: Modelling framework and calibration results" by Viviroli, Zappa, Schwanbeck, Gurtz and Weingartner (2009b).

* Corresponding author. Address: Institute of Geography, University of Bern, Hallerstrasse 12, CH-3012 Bern, Switzerland. Tel.: +41 31 631 80 17; fax: +41 31 631 85 11.
E-mail address: viviroli@giub.unibe.ch (D. Viviroli).

0022-1694/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2009.08.022

i.e. 'to make predictions about hydrological quantities at sites where data are absent or inadequate, frequently for design purposes' (Beran, 1990). In flood estimation, two families of methods are most commonly applied:

Regional transfer functions are used to apply information from gauged to ungauged basins, e.g. using regionally differentiated enveloping curves or regressions; a wide variety of such methods are available today (see e.g. Dalrymple, 1960; Cunnae, 1988; Hosking and Wallis, 1993; Bobée and Rasmussen, 1995; Burn et al., 1997). Although measures related to hydrological processes or meteorological conditions may be involved, these are usually considered in an empirical rather than causal manner.

Simple concept models like the rational formula (Chow, 1964) and more sophisticated derivates of it (e.g. Kölla, 1987) involve some considerations regarding the most relevant processes, but are criticised for containing parameters which are difficult to estimate (e.g. runoff coefficient) or for being founded upon questionable assumptions (e.g. identical return period for precipitation and resulting peak flow).

These methods have been widely applied in the past and proven successful for average conditions (Weingartner, 1999). For estimation concerning an individual basin, however, the disagreement between results from different approaches may be considerable (see Weingartner et al., 2003), particularly if unusual conditions prevail (e.g. regarding soil, geology, or climatology). This means that the above-mentioned methods show a lack of robustness.

In this paper, we present a deterministic, process-oriented alternative for estimating floods in catchments without gauge records. More precisely, we employ continuous long-term simulation at hourly resolution for ungauged mesoscale catchments in Switzerland. The simulated hydrographs are subsequently analysed using extreme value statistics, which leads to estimations of flood peak values with the desired recurrence interval. Our main intent is to provide robust and independent estimates which extend and improve today's flood estimation methods. The suitability of the modelling framework is tested extensively in application, by considering a large number ($n = 140$) of study catchments. This lays the foundation for a comprehensive nation-wide flood estimation system with practical relevance.

Comprehensive 'real-world' applications of continuous simulation were only rendered possible in the past decade when the numerous data sets required became available in digital form and in sufficient quality, and when computers started to be able to process them efficiently, although the idea of continuous simulation had already been conceived in the 1970s (Eagleson, 1972; see also Beven, 2001). A comprehensive review of continuous simulation applications for design flood estimation is found in Boughton and Droop (2003); they report operational systems for Australia, Europe, South Africa, the UK, and the USA most of which are aimed at gauged (i.e. calibrated) catchments. For ungauged catchments, the only application to date which is comparable in extent to ours is found in the UK (Calver et al., 2005; see also Lamb and Kay, 2004 and references therein); in contrast to our framework, however, it is not entirely based on hourly time steps.

This paper is structured as follows: After introducing the regionalisation methods in "Modelling framework" (including the attributes necessary for describing the catchments), results are presented for standard efficiencies ("Model efficiency") and, most importantly, for flood estimation ("Flood estimation"), which is then compared to popular empirical and stochastic methods ("Comparision with standard procedures"). The characteristics of estimation errors are discussed in "Estimation errors", while important issues concerning the regionalisation approaches are highlighted in "Regionalisation approaches" and "Is the proposed regionalisation scheme effective?". Concluding remarks and an outlook follow in "Conclusions and outlook".

Modelling framework

Our approach to obtaining estimations for ungauged catchments in Switzerland has been developed on the basis of 140 calibrated mesoscale catchments (see Viviroli et al., 2009b) and tested using 49 representative catchments with long and reliable gauge records. We use the conceptual process-based hydrological modelling system PREVAH (Precipitation–Runoff–EVApotranspiration-HRU related model; for definition of HRU see below) (Viviroli et al., 2009a), which has a respectable record of successful application in topographically complex regions, particularly in Switzerland (for a compilation see Viviroli et al., 2007, 2009a). The spatial resolution of PREVAH is currently based on hydrological response units (HRUs), which we aggregated for this study on the basis of $0.5 \times 0.5\ \mathrm{km}^2$ raster cells. The temporal resolution for inputs and outputs is hourly throughout. This is of special relevance for flood estimation since large peak values are smoothed severely when a coarser temporal resolution (e.g. daily time step) is used, especially in smaller mesoscale basins (Viviroli, 2007). Data interpolated from the Swiss standard meteorological gauging network are used to operate the model from 1984 to 2003, namely precipitation, air temperature, global radiation, relative sunshine duration, wind speed and relative humidity. For interpolation, Detrended Inverse Distance Weighting (e.g. Garen and Marks, 2001) and Ordinary Kriging (e.g. Isaaks and Srivastava, 1989) are used. Details of the modelling framework are furnished in the companion paper by Viviroli et al. (2009b).

Regionalisation methods

While the tuneable model parameters of PREVAH are calibrated for catchments with gauge records (Viviroli et al., 2009b), the model's application to ungauged catchments requires a parameter regionalisation procedure. The 12 tuneable model parameters are estimated using three independent approaches (Nearest Neighbours, Kriging and Regression), which are subsequently combined by calculating the median of the three respective simulated discharge hydrographs. For glaciated catchments, two more parameters need to be regionalised. Regionalisation is based on an extensive set of model parameters from 140 catchments which were calibrated using a cost-efficient procedure; particular focus was put on the appropriate representation of peak flows (for both calibration methods and results see Viviroli et al., 2009b).

This chapter describes the methods for regionalisation of the tuneable model parameters. First of all, attributes have to be evaluated which are able to describe any catchment within the study area. Then, the three regionalisation approaches are introduced and the respective outputs finally combined into a single simulation for ungauged catchments.

Catchment attributes

Characterising catchments for regionalisation purposes requires appropriate attributes. A large variety of such catchment descriptors have been presented in the past (e.g. Pearson, 1991; Sefton and Howarth, 1998; Seibert, 1999; Peel et al., 2000; Blöschl and Merz, 2002; Lamb and Kay, 2004; Merz and Blöschl, 2004; Bárdossy et al., 2005). For Switzerland and the Alpine region in general, the collection of Breinlinger et al. (1992) is most relevant and comprehensive, while a useful collection was also prepared by Pfaundler (2001). All of these studies agree that a set of attributes must always be tailor-made for the respective goals and study area.

Therefore, the above studies were screened with an emphasis on flood-relevant catchment attributes, but care was taken to characterise standard flow conditions as well since the underlying hydrological model, PREVAH, is required to produce a plausible overall simulation. Further attributes were added where necessary, with a focus on describing meteorology (particularly precipitation), which is represented only sparsely in the studies mentioned above. Besides catchment-wide averages, the higher statistical moments (standard deviation, skewness and kurtosis) were calculated for attributes which have a significant spatial variability on the catchment scale (e.g. net field capacity).

In total, 82 attributes were computed and evaluated. This number seems rather high in comparison to the relatively simple and small sets of attributes that are usually preferred (see Castellarin et al., 2001): The UK Flood Estimation Handbook (IH, 1999), for example, recommends three attributes for transfer of model parameters, as do Calver et al. (2004); Samaniego and Bárdossy (2005) used seven attributes in their study. The extensive evaluation of potentially useful attributes for the present study seems justified, however, in view of the high number of 12 tuneable model parameters to be regionalised (14 with the presence of glaciers). The high variability and heterogeneity of hydrological conditions in the study area further speaks in favour of a comprehensive assessment of potentially useful attributes. From the large set of attributes generated here, the relevant ones are selected for each regionalisation approach on the basis of statistical analyses and the efficiency of the resulting simulations (see "Regionalisation 1: Nearest Neighbours", "Regionalisation 2: Kriging", and "Regionalisation 3: Regression"). Following these selections, 72 attributes are still used in one approach or another; these are listed in the Appendix. The uncertainty of attributes is assumed to be low in comparison with the uncertainties involved in hydrological modelling (meteorological input, model structure, tuneable parameters).

Regionalisation 1: Nearest Neighbours

The Nearest Neighbour approach consists in finding a calibrated donor catchment which is as similar as possible to the ungauged target basin. All tuneable model parameters are then transferred from the donor to the target as a complete, unchanged set. This has the advantage that the mutual tuning of the tuneable model parameters is not disturbed (see Young, 2006). Furthermore, no assumptions have to be made concerning the relation between model parameters and catchment attributes (Kokkonen et al., 2003; see also Hundecha and Bárdossy, 2004; Parajka et al., 2005 and Bárdossy, 2007).

Since the parameter set from the most similar catchment will not necessarily produce the best results when used for simulation (see Oudin et al., 2008), the five most similar catchments are identified. The respective parameter sets are transferred from these to the target catchment and used for simulation. The five resulting independent simulations are then combined by calculating the median value for each (hourly) time step. Further details concerning the number of Nearest Neighbours to be used are given in "Nearest Neighbours: number of most similar catchment to be considered".

The main issue for the Nearest Neighbour approach is how to adequately define a similarity measure. A common method is to calculate the Euclidean Distance between two catchments in the n-dimensional space of catchment attributes; the attributes have to be normalised beforehand in order to account for their varying value range. Additionally, a user-specified weight w may be assigned to the individual attributes in order to consider their varying assumed importance. The user-weighted Euclidean Distance for two catchments i and j is then defined as

$$D_w(i,j) = \sqrt{\sum_{k=1}^{n} w_k [attrib_k^*(i) - attrib_k^*(j)]^2} \qquad (1)$$

where $attrib^*$ refers to the n normalised attributes. The smaller D_w, the more similar are the two catchments i and j.

As an important step, a suitable set of attributes (*attrib*) and corresponding weights (w) had to be compiled. A total of 13 sets were therefore defined following either hydrological or statistical reasoning; these sets are explained in more detail in Table 1. Each of these sets was used to find the five Nearest Neighbours from which the parameters were transferred. Additionally, a distinction of the catchments concerning their mean altitude (a_m) was tested since Weingartner (1999) found that this provides more homogeneous conditions for regionalisation of model parameters in regions with complex terrain. Three altitude zones were defined according to characteristics of discharge regime (Weingartner and Aschwanden, 1992) and discharge variability (Viviroli and Weingartner, 2004) in Switzerland: Swiss Plateau and Jura ($a_m < 1000$ m a.s.l.), Pre-Alps (1000 m a.s.l. $\leqslant a_m < 1550$ m a.s.l.), and Alps ($a_m \geqslant 1550$ m a.s.l.). The Nearest Neighbour search was then restricted to donor catchments which lie in the same zone of mean altitude as the target catchment; to avoid an unnaturally sharp disjunction, however, donor catchments from the adjacent altitude zone were included if the difference in altitude between target and donor catchment was less than 50 m. Using this scheme, mean altitude was omitted as a catchment attribute to avoid its multiple use.

Fig. 1 shows Nash–Sutcliffe efficiencies (*NSEs*) for simulations on basis of 26 regionalised parameter sets. These parameter sets were regionalised by using the aforementioned 13 attribute sets with (*) and without ($^\circ$) distinction of mean altitude zones. First of all, it is apparent that the results are generally better when altitude zones are taken into account. Overall, set 02* shows the best performance, with 40% of catchments attaining excellent ($NSE \geqslant 0.75$) and 83% of catchments attaining fine ($NSE \geqslant 0.5$) results. For only 6% of the sample, regionalisation has failed ($NSE < 0.2$). The highest-ranking set 02* was composed on the basis of a statistical analysis: For each model parameter, the two catchment attributes with highest correlation were identified and added to the set of attributes. To achieve a distinction of mean altitude zones, this analysis was conducted for each zone separately, thus resulting in three sets with a maximum number of 28 attributes

Table 1
Sets of attributes and corresponding weights tested for Nearest Neighbour regionalisation with Euclidean Distance (Mittelbach, 2006). See Appendix for explanation of the thematic attribute groups.

Set	Attribute selection and weighting criteria
01	All attributes with significant correlation ($\alpha \leqslant 0.05$) to model parameters, equally weighted
02	From each model parameter the two attributes with highest correlation, equally weighted
03	26 parameters selected on basis of hydrological expert judgement, equally weighted
04	Like set 03, but with two additional parameters for extreme precipitation
05	Like set 03, but considering skewness of soil-related attributes instead of their mean values
06	Like set 03, but with double weight on hydrogeology attributes
07	Like set 03, but with double weight on soil physics attributes
08	Like set 04, but with double weight on precipitation attributes
09	Like set 04, but with double weight on precipitation, hydrogeology and soil physics attributes
10	Like set 03, but with double weight on land use attributes
11	Like set 03, but with higher weight on hydrogeology ($\times 5$) and land use ($\times 2$) attributes
12	Like set 03, but with double weight on hydrogeology, soil physics and land use attributes
13	Like set 03, but with tenfold weight on land use attributes

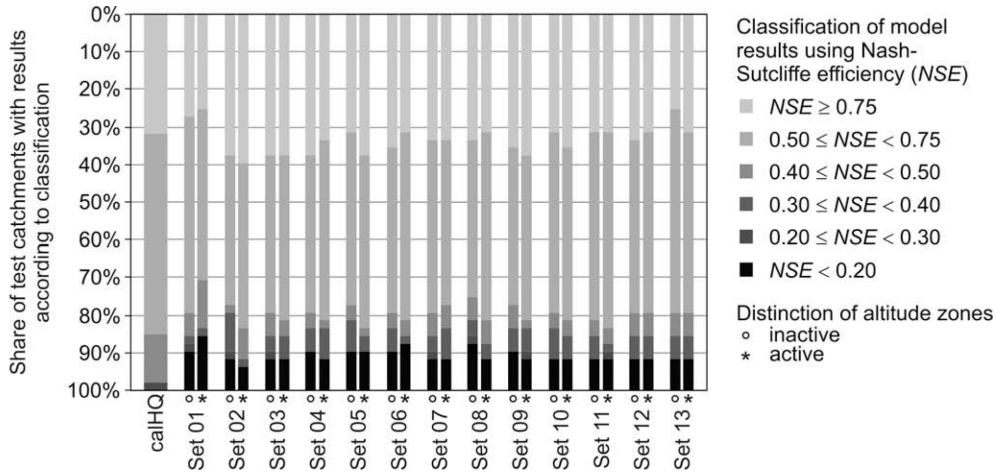

Fig. 1. Results from 13 attribute sets and corresponding weights for determination of catchment similarity (Nearest Neighbours), based on 49 representative test catchments with hourly simulation, 1984–2003.

per zone (14 parameters $\times$ 2 attributes); the actual sets are, however, smaller since some attributes appear in the top two ranks of correlation coefficient for more than one parameter. All of these attributes are given an equal weight of $w = 1$. It may be somewhat disappointing from a hydrologist's point of view that a statistical analysis with equal weights yields best results. It has to be noted, however, that at least for the analysis without distinction of altitude zones, hydrological expert judgement (set 10°) performs best; the respective set comprises 26 selected attributes and has increased weights for hydrogeology and land use attributes.

Originally, two further similarity measures based on distance in the *n*-dimensional attribute space were tested: Correlation-adjusted Euclidean Distance (Fischer, 1982) and Mahalanobis Distance (Steinhausen and Langer, 1977). Both measures account for the underlying attribute space not being orthogonal (i.e. distorted by correlations) and assign higher weight to uncorrelated attributes which are likely to contain extra information which is not present in more closely correlated attributes. Analysis has shown, however, that no noticeable improvement of regionalisation results is achieved with either of these measures (Viviroli, 2007).

Regionalisation 2: Kriging

In this approach, the calibrated model parameters are interpolated in space independently from each other. More precisely, the parameters for each catchment are associated with the respective catchment's centroid and then interpolated using Ordinary Kriging (see e.g. Goovaerts, 1997). For an ungauged catchment, the tuneable model parameters are then read from the 12 (or 14) resulting parameter maps at the respective catchment's centroid. The only catchment attributes required for this procedure are the co-ordinates in space.

In Ordinary Kriging, the estimated value P for a location $\mathbf{u}$ is a weighted linear combination of the values from n reference points. For our purposes, $\mathbf{u}$ refers to the co-ordinates of the respective catchment centroid:

$$\hat{P}(\mathbf{u}) = \sum_{i=1}^{n} \lambda_i P(\mathbf{u_i}) \tag{2}$$

Using so-called semivariograms, the weights λ_i are assigned so that the estimation variance is minimal. This involves solving a system of linear equations, the so-called Kriging system, which was handled using routines by Deutsch and Journel (1997).

Care was taken to produce parameter maps with an appropriate balance of smoothing and point accuracy (Viviroli, 2007). This was achieved with Kriging parameters close to the defaults recommended by Deutsch and Journel (1997), i.e. a nugget constant of two, a range of 25 km, a maximum search radius of 100 km, and a minimum/maximum n of 3/15 (Flach, 2007). It should be borne in mind that especially for this regionalisation scheme, the decisive restriction is the information content available from gauged catchments and not so much the interpolation scheme.

Successful regionalisations using Kriging have been reported in the past (e.g. Vandewiele and Elias, 1995), including applications for the European Alps (Merz and Blöschl, 2005; Parajka et al., 2005). From a strictly theoretical point of view, however, the use of spatial proximity as an indicator of hydrological similarity is disputed (Mosley, 1981; Nathan and McMahon, 1990a; Burn and Boorman, 1992; Reed et al., 1999; Shu and Burn, 2003). The heterogeneity of physical and climatic conditions in a mountainous region such as Switzerland may pose additional difficulties. To overcome this problem requires an extensive set of information from gauged and calibrated catchments, which is given for the present study. The actual plausibility of interpolated parameter maps will be discussed in "Kriging: interpretation of resulting parameter maps".

Regionalisation 3: Regression

In the Regression approach, model parameters are related directly to selected catchment attributes. For each of the m model parameters ($param_i$), a linear regression model containing n attributes ($attrib_j$) is set up:

$$param_i = a_i + \sum_{j=1}^{n} (b_{i,j} \cdot attrib_j) \tag{3}$$

A specific regression model is built for each parameter, containing the five attributes which show highest correlation with the respective parameter and up to 15 further attributes whose correlation with the respective parameter is significant on a 95% confidence level. These attribute–parameter relationships are generalised for all basins and are therefore based on $n = 140$ data points. It was attempted to include further regression parameters, but this did not lead to improvements. This corroborates the findings from Seibert (1999), who reported best results for regression equations with two parameters as well. Non-linear relationships were also tested, but this yielded more unstable results in some cases (Viviroli, 2007). Although the linear form of the regression model hardly represents hydrological reality (see Bárdossy, 2007), it achieves good results and is most convenient to handle (Parajka et al., 2005).

For each parameter of the hydrological model, a specific regression model (Eq. (3)) was tuned with the help of PEST (Doherty, 2002). PEST uses the Gauss–Marquardt–Levenberg method to minimise the deviation between calibrated ('true') and regionalised (regression-derived) parameter values from 140 catchments. Since catchments with better calibration results are expected to contain more reliable information, the corresponding Nash–Sutcliffe efficiencies (*NSEs*) were used as weights in calibration with PEST. Optimisation was done with PEST because it is also suitable for processing non-linear problems, which was necessary for testing non-linear regression models (see above).

Finally, the resulting parameter set for the ungauged basin is checked for plausibility: Values which exceed or fall short of the range of parameter values realised in calibration (see Viviroli et al., 2009b) are set to the respective threshold. This avoids unreasonable parameter values in catchments with exceptional conditions (see Blöschl and Grayson, 2000). Exceedance of the parameter threshold was observed in only 2.5% of the regression-regionalised parameter values. This leads us to the conclusion that our regression regionalisation scheme is stable and produces plausible parameter estimates.

A major uncertainty in this approach is the relationship between model parameters and catchment attributes. Mainly due to uncertainties in model calibration, this relationship is not always as pronounced as it would be desirable to be (Parajka et al., 2005; Beven, 2006; Wagener and Wheater, 2006). However, Merz and Blöschl (2004) have shown that this problem can be mitigated to a certain degree when a uniform and objective calibration method is used, as was done in the present study (Viviroli et al., 2009b). Furthermore, some theoretical requirements for regression are usually not fully met, such as complete independence of the catchment attributes and unequivocal causal relations between model parameters and catchment attributes (Johansson, 1994). But again, as noted for the Kriging approach, a number of successful applications (Seibert, 1999; Peel et al., 2000; Lamb and Kay, 2004; Young, 2006) have proven that regionalisation based on regressions is possible and that the theoretical reservations are of secondary relevance.

As was the case for Nearest Neighbours, better results were achieved with the Regression approach when restricting the analysis to zones of similar average basin altitude (Viviroli, 2007). For Regression regionalisation, this means that only information of the relevant catchments is used to compose and calibrate the respective zone-specific regression models. Again, mean altitude is not used as an independent catchment attribute.

Combining approaches 1–3

Since all three regionalisation methods yield good results in evaluation (Viviroli, 2007), a combination was envisaged in order to benefit from the particular advantages of each approach. Furthermore, such a procedure is suitable to avoid failures due to the shortcomings, and hence individual bad results, of a single method. Although this carries the risk of erroneously eliminating 'outliers' which actually represent the most appropriate simulation, application in practice suggests that a combination of approaches is favourable. This was shown recently by Oudin et al. (2008), who used, similarly to our study, a combination of physical similarity, spatial proximity and regression regionalisation approaches.

In our study, the combination is achieved as follows: First, complete hydrographs are simulated on basis of the Nearest Neighbour,

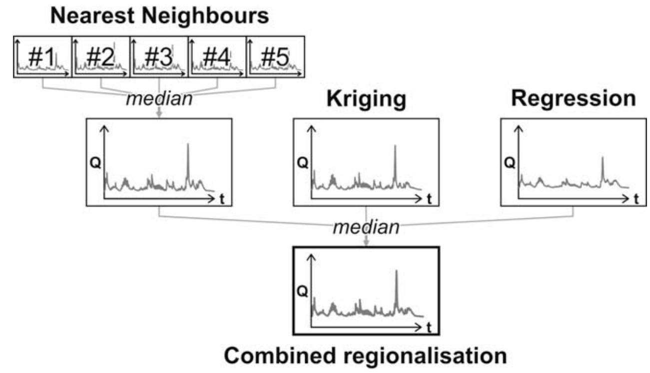

Fig. 2. Combination of the individual regionalisations into a single regionalisation.

Kriging and Regression regionalisation approaches. For each time step, the median value from the three regionalised simulations is then extracted to obtain the Combined regionalisation (Fig. 2). Using the median rather than the average reduces the undesirable smoothing effects mentioned above. The result of this procedure is a complete hydrograph for the regionalised catchment, ready to be analysed not only for peak flow, but also for discharge volumes, among other things (Viviroli, 2007). Furthermore, the three individual regionalised simulations may be used to estimate regionalisation uncertainty (see also "Nearest Neighbours: number of most similar catchment to be considered").

Results

The jack-knife technique was employed to compute and subsequently evaluate the regionalisation results. In this cross-validation approach, the parameters for each successive catchment are retrieved through the regionalisation procedure with the other calibrated catchments. Therefore, no parameter information is derived from its own calibration. This technique was applied to 49 catchments which are representative in terms of runoff regimes and climatic conditions and have a long runoff record (see Viviroli et al., 2009b). The basis for each jack-knife regionalisation was formed by 139 calibrated catchments, i.e. the full sample of 140 calibrated catchments excluding the catchment to be assessed. The regionalisation results were then compared to the values from observation and from calibrated simulation, e.g. using efficiency scores.

Model efficiency

Prior to proceeding with flood estimation, the regionalisations are assessed concerning the hydrological plausibility of the resulting hydrographs. Although flood frequency distributions are ultimately derived from the simulations, these distributions are more trustworthy when other aspects of the simulated hydrograph are realistic as well (Lamb, 1999). For this assessment, the Nash–Sutcliffe efficiencies (*NSEs*) of the 49 representative catchments were calculated for the entire 1984–2003 simulation period (Fig. 3). It is important to note that all regionalisations were computed on the basis of the flood-calibrated parameter sets (see Viviroli et al., 2009b). With median *NSE* scores of 0.70 (Nearest Neighbours), 0.67 (Kriging), 0.65 (Regression) and 0.69 (Combined), the regionalisations perform in the range of the calibrated simulations (0.73 for standard and 0.69 for flood calibration); the few outliers at the low end seem acceptable. In order to evaluate the regionalisation methods introduced above, two very simplistic and uniform 'regionalisations' were also employed. The first one consists in simply using the default initial parameters which were

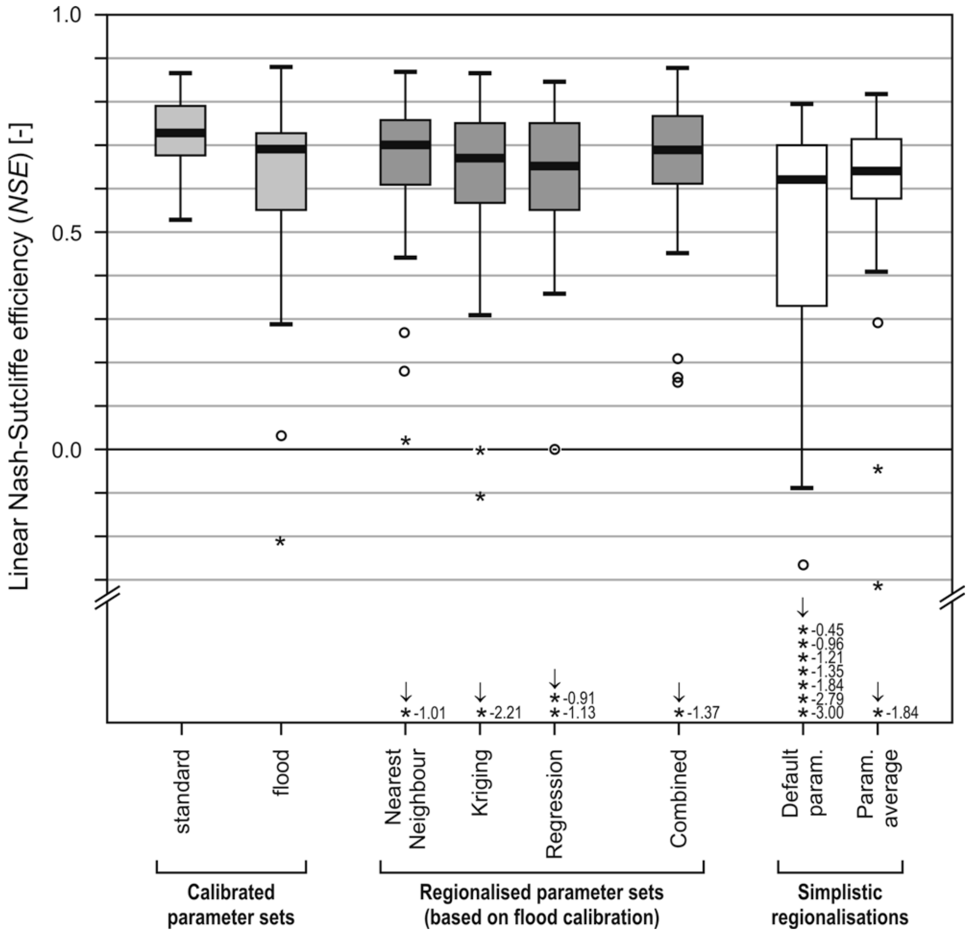

Fig. 3. Nash–Sutcliffe efficiencies (*NSEs*) from model runs with calibrated and regionalised parameter sets, based on hourly simulation, 1984–2003. Simplistic regionalisations are indicated for comparison purposes: (1) default (initial) parameter set, (2) average from calibrated catchments in same zone of average altitude for each individual parameter (flood-calibrated parameter sets). Circles denote outliers (distance from upper or lower quartile is between 1.5 and 3 times the quartile range), stars extreme values (distance from upper or lower quartile is greater than three times the quartile range).

chosen for this study (see Viviroli et al., 2009b). While the median *NSE* is still high (0.62), the results at the lower end drop off to an extent which suggests that the default parameter set is unusable. For the second simplistic approach, each parameter was determined as the respective average of all calibrated catchments in the same zone of average altitude, again on the basis of the flood-calibrated parameter sets. Although the median *NSE* score of 0.64 is only slightly inferior to that of intelligent regionalisations and the quartile range is narrow, one should be aware of the fact that an averaging of this kind, already for methodological reasons, is always less reliable since it does not take into account the actual properties of the individual catchments (see Kokkonen et al., 2003). Particularly for extreme events or unusual geophysical conditions, such local information is decisive, though.

Fig. 4 shows a similar analysis for the average of annual volumetric deviations (SVD_a; see Eq. (4) in Viviroli et al., 2009b). The agreement of simulated and observed annual water balances is clearly worse for the flood-calibrated parameter set than for the standard calibrated set. This shortcoming has been discussed by Viviroli et al. (2009b) and is clearly visible also in the regionalisations which, as noted earlier, are based on flood-calibrated parameter sets. But apart from an increase in the number of outliers, the same results are obtained as from calibration; the deviations are still acceptable. The comparatively favourable performance of the default parameter set may be explained as being due to its generic suitability for the study area; its clear deficits have already been identified above, namely the low *NSE* scores. Conversely, the distinct bias of the averaged parameter set concerning SVD_a puts into perspective the relatively high *NSE* scores.

In summary, all regionalised parameter sets based on the flood-improved calibration yield reasonable results as to hydrological standard scores. This is an essential prerequisite for the subsequent interpretation of the simulations concerning flood peaks.

Flood estimation

To assess the flood estimation results, the 100-year flood (HQ_{100}) is used as an exemplary value. This enables us to compare our results directly to earlier standard methods which frequently consider only this recurrence interval (e.g. some of the procedures relevant for Switzerland, see Barben, 2003). HQ_{100} is estimated on the basis of the 20 annual peak discharge values (HQ_a) from the 1984–2003 hourly simulation. It should be noted that by estimating HQ_{100} from $n = 20$ HQ_a values, the recommended extrapolation range of $3 \times n$ years (DVWK, 1999) is exceeded. For the present data set, however, Viviroli (2007) has shown that the relative errors for HQ_{20}, HQ_{50} and HQ_{100} are correlated, which justifies going beyond the extrapolation range of 60 (3×20) years for the benefit of comparability. In order to keep the extrapolation results consistent, the first extremal distribution (see DVWK, 1999) was used as the basis for flood estimation in all catchments, and the distribution parameters were always estimated with the help of probability-weighed moments (see Greenwood et al., 1979). This combination of extreme value distribution and parameter estima-

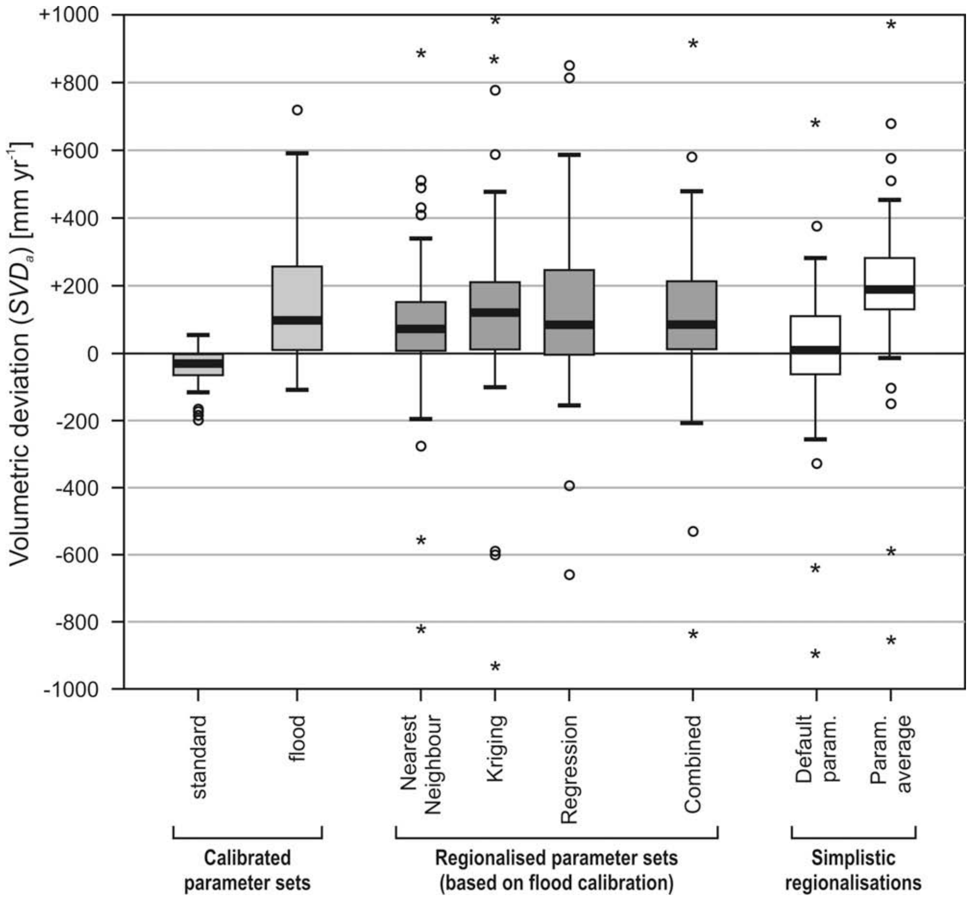

Fig. 4. Average annual volumetric deviations (SVD_a) from calibrated and regionalised model parameter runs, based on hourly simulation, 1984–2003. See also Fig. 3.

tion is very common and suitable for achieving good results in a large sample of catchments with varying properties (see Greis and Wood, 1981, 1983). The HQ_{100} estimated from the observation record (by analogy: hourly resolution, 20 annual peaks from 1984 to 2003) is used as a reference.

The upper section of Fig. 5 summarises the results from the 49 representative catchments in a box-plot. As to the calibration results, the parameter set with additional peak flow calibration ('flood') performs noticeably better than the standard calibrated parameter set; the median underestimation for HQ_{100} is reduced from −29.8% to −15.8%, with half of the catchments showing an error between −28.7% and +12.0% (interquartile range). The regionalisations perform relatively well, with underestimations of −13.6% (Nearest Neighbours), −9.9% (Kriging) and −9.7% (Regression). Combining these three regionalisations yields an even smaller underestimation of −7.2%, and the interquartile range of −30.8% to +7.5% is also smaller than for the individual approaches.

The lower part of Fig. 5 shows a more practice-relevant and straightforward assessment, the estimation results being classed in 'hits' (estimation lies within 90% confidence interval of reference, considered as excellent) and 'near misses' (90% confidence intervals of estimation and reference overlap, considered as still useful). The confidence intervals were calculated according to DVWK (1999) and are related to the statistical estimation uncertainty in flood estimation which arises from the limited sample size of the peak flow record (Maniak, 2005). In agreement with the above, the Combined regionalisation approach shows the largest number of excellent HQ_{100} estimates, with 51.0% hits. This is even a little higher than the flood-calibrated parameter set, which achieves hits in 46.9% of the test catchments.

Both for relative deviations and for 'hits' and 'near misses', regionalisation performs slightly better than calibration. This is surprising at first sight, although it has already been observed by other authors (e.g. McIntyre et al., 2005). A more detailed discussion of this will be provided in "Is the proposed regionalisation scheme effective?".

Comparison with standard procedures

Of particular interest is a comparison of the flood estimation results presented above with today's popular stochastic and empirical procedures for ungauged basins. For Switzerland, a number of such standard methods are available in HQx_meso_CH (Spreafico et al., 2003; see also Barben, 2003 and Weingartner et al., 2003). This software package enables easy peak flow estimation on the basis of regionally differentiated enveloping curves, similarity searches and the rational formula, principally for a recurrence period of 100 years. Only catchment boundaries and catchment outlet must be specified, all other data required for the individual methods are included in the program.

Five relevant procedures for estimation of HQ_{100} were applied using HQx_meso_CH. In the top right part of Fig. 5, the respective estimation errors are included with the median, the minimum and the maximum of the five resulting estimates. In general, a slight tendency to overestimate HQ_{100} is observed, with a median error of +15.1% for the 49 representative test catchments. With 47% of hits, the portion of excellent estimates is almost as high as for the Combined regionalisation. However, marked overestimations are noted for the respective highest of the five estimates, with the range extended to +233% and the highest value overestimating HQ_{100} by +470%. Overall, the HQ_{100} estimates from HQx_meso_CH

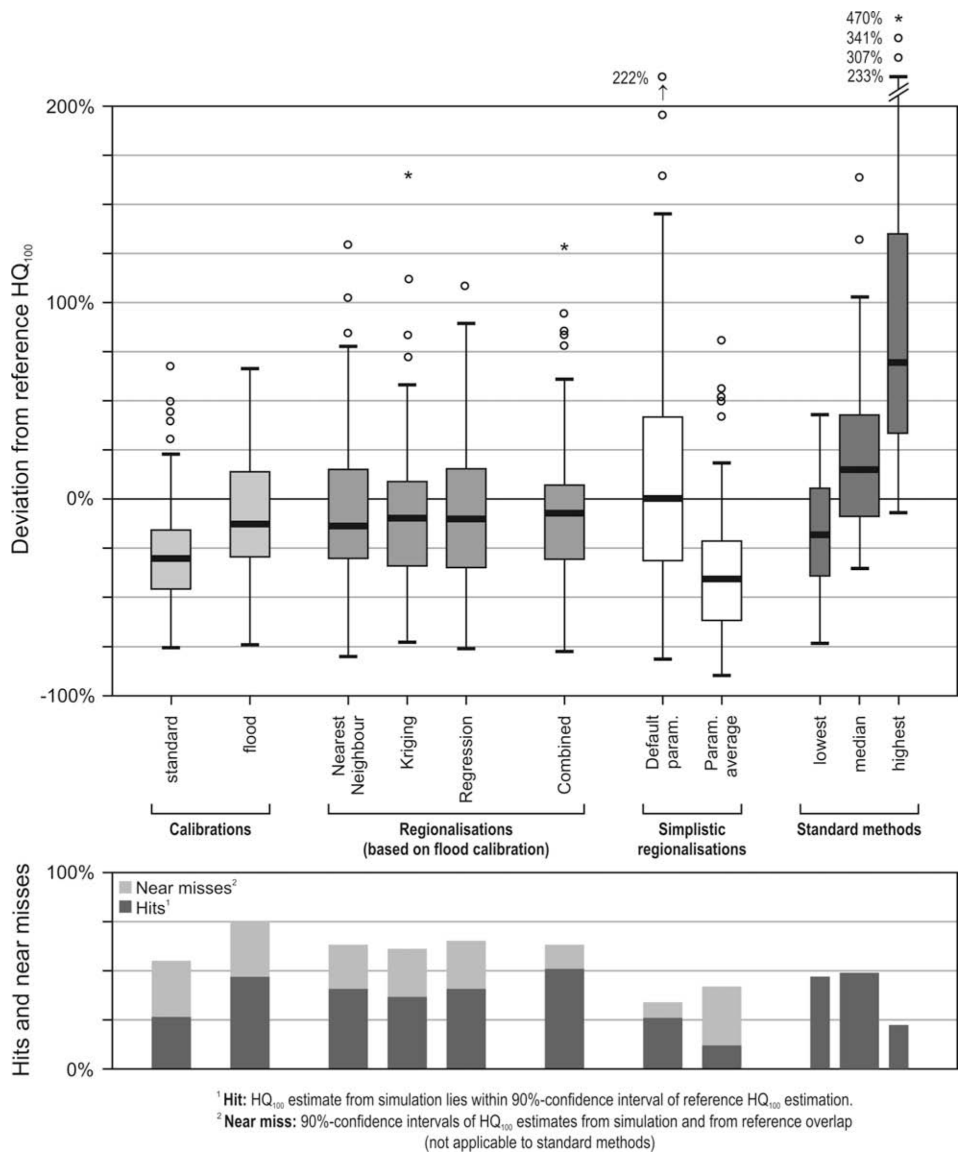

Fig. 5. Error in HQ_{100} estimation from calibrated and regionalised model parameter runs, with HQ_{100} estimated from observation serving as the reference (all on basis of hourly values, 1984–2003, 49 representative test catchments). The lower part of the figure shows the percentage of hits and near misses (see legend). Respective values are given for five empirical standard methods (Spreafico et al., 2003), see discussion in "Comparison with standard procedures". Since the standard methods have no confidence intervals, the portion of near misses cannot be determined.

perform well, but there is also noticeable uncertainty arising from the large number of estimates included in the various procedures.

A major difference between HQx_meso_CH and our procedure should be pointed out. Like most estimation procedures, HQx_meso_CH gives estimates for peak flows only. Our procedure, however, provides continuous hydrographs, which makes it possible to estimate, for example, flood volumes or the time during which critical stage levels are exceeded.

Discussion

Estimation errors

Owing to scale issues, it is assumed that the error characteristics of HQ_{100} show clear dependence on drainage area, i.e. that the accuracy of flood estimation will decrease for small catchments. This is due to deficits in process descriptions in the model as well as to issues concerning data accuracy and availability. Ultimately, the catchment's size determines the extent to which these errors are averaged in the integral runoff response (see e.g. Blöschl and Grayson, 2000 and Grayson and Blöschl, 2000). In Fig. 6, the following patterns are identified for our sample:

- In small catchments (area of 10–40 km^2), HQ_{100} is underestimated for the most part. Error is observed to increase with decreasing area (r_{xy} = 0.85), reaching −75% for the smallest catchment. Two main reasons seem to be responsible for this. First, the intense precipitation peaks triggering a large flood peak will typically be missed by the standard meteorological network, while interpolation in space further smoothes out local variations in intensity (see Wilson et al., 1979; Andréassian et al., 2001). Second, the processing time step of 1 h may be insufficient to represent fast response to short and intense precipitation (see Naef et al., 1999). In summary, capturing fast runoff processes with a conceptual process-oriented model such as PREVAH is indeed challenging; on the other hand, using a very detailed physically-based model – such as WaSiM (Schulla and

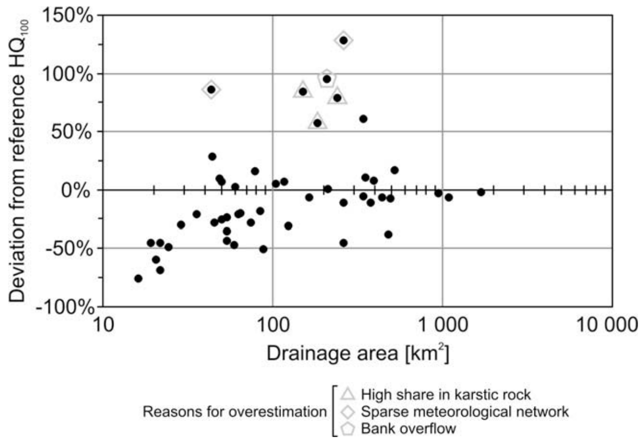

Fig. 6. Error in HQ_{100} estimation from Combined regionalisation compared to drainage area. Data from 49 representative test catchments, hourly values, 1984–2003.

Jasper, 2000), which was developed in Switzerland as well – would introduce new difficulties in determining the requisite parameter values. Advancements in this issue could potentially be achieved by integrating maps of dominant runoff processes (e.g. Schmocker-Fackel et al., 2007). An adjustment of the above estimation errors based on drainage area alone was not feasible since the sample of small catchments was considered too small to justify a scale-dependent correction factor.

- For catchments of intermediate size (40–750 km^2), the median error is only −6%, and half of the catchments show errors between −25% and + 10% (interquartile range). In spite of a slight tendency towards underestimation, a few marked overestimations occur. Out of the seven catchments with an error of more than +50%, three are located in the Swiss Jura region and show a high share in permeable karstic rock (HGKR > 50%, see Appendix). Although PREVAH is optionally able to model water losses due to karst, such was not attempted in the present study because the sample of uninfluenced karstic catchments was too small to devise a regionalisation for the karst loss parameter. Two more catchments are located in a comparatively dry part of the south-eastern Swiss Alps where representative precipitation stations at high altitudes are scarce. In addition, most of the neighbouring catchments are influenced by hydropower generation and therefore not available for calibration, which renders the regionalisation more susceptible to local particularities of precipitation characteristics and catchment dynamics. A further catchment that shows marked overestimation is known for occasional bank overflow (Haider, 1994). To represent this process, PREVAH would have to be coupled with a hydraulic model (as presented in Schwanbeck et al., 2007). Underestimations at the intermediate scale become less frequent with increasing basin size (bear in mind that the abscissa is logarithmic).
- Large catchments (750–2000 km^2) show noticeably low errors since fast response to short but intense precipitation is attenuated in the overall catchment response and becomes less dominant. Furthermore, smaller deviations (concerning both precipitation input and model response) are averaged. For catchments larger than 2000 km^2 (the largest one in this study has an area of 1696 km^2), however, it is recommended to simulate subcatchments of intermediate size separately and then to apply a routing scheme (see e.g. Schwanbeck et al., 2008).

Further attempts to adjust the above estimation errors based on the spatial distribution of deviations or on catchment attributes (see "Catchment attributes") were not successful. We also attempted to modify the estimation procedure by estimating the respective HQ_{100} from the three individual regionalisations first and then using the median HQ_{100} as the final estimate, and by estimating the parameters of the extremal distribution rather than the extrapolated HQ_{100} itself. This did not change the results significantly (see Viviroli, 2007). It must therefore be assumed that the limitations of estimation accuracy mainly stem from limitations in model structure and precipitation input.

Regionalisation approaches

Nearest Neighbours: Number of most similar catchments to be considered

Another important issue to be discussed for the Nearest Neighbour regionalisation concerns the number of most similar catchments to be included. As McIntyre et al. (2005) have noted, the parameter set from the most similar catchment does not necessarily produce the best results. This is corroborated by analysing the present data set and methodology; Fig. 7 summarises the results from the individual Nearest Neighbours used for simulation

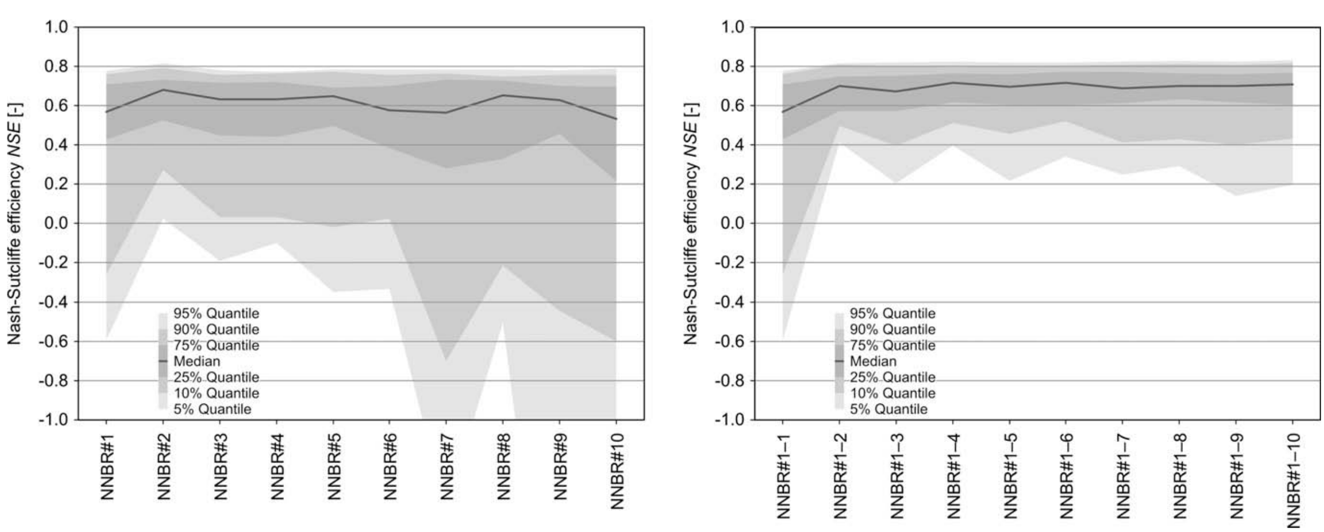

Fig. 7. Summarised Nash–Sutcliffe efficiency (*NSE*) results from Nearest Neighbour regionalisation. Left: individual results from simulation with 1st (NNBR#1) to 10th (NNBR#10) Nearest Neighbour parameter set; right: median of *n* Nearest Neighbour simulations (i.e. NNBR#1−*n*: median of Nearest Neighbour simulations 1 to *n*). Data from 49 representative test catchments, hourly simulated values, 1984–2003.

(configuration as presented in "Regionalisation 1: Nearest Neighbours"). The results obtained by employing single Nearest Neighbour parameter sets (Fig. 7, left) shows that in our case, using the second most similar catchment (NNBR#2) would produce highest efficiencies. After the 5th Nearest Neighbour (NNBR#5), there is a noticeable drop in efficiency, especially as regards the 25% quantile; this drop is even more pronounced after the 6th Nearest Neighbour (NNBR#6). When multiple Nearest Neighbour simulations are combined by computing the median simulation value for each time step (Fig. 7, right), efficiencies are always high for 2–10 Nearest Neighbours. There is, however, a slight but steady drop in the lower quantiles. This supports our findings presented in "Model efficiency", that the averaged parameter set performs well in general but carries an increased risk of failure under unusual conditions.

Using five Nearest Neighbours as proposed in "Regionalisation 1: Nearest Neighbours" yields a median *NSE* of 0.70 and is a suitable trade-off between computational cost and model efficiency. In addition, including more Nearest Neighbour parameter sets increases the risk of smoothing out specific catchment conditions.

Kriging: Interpretation of resulting parameter maps

Regarding the Kriging approach, plausibility and validity of the interpolated model parameter maps are important issues. Kriging regionalisation implies that spatial similarity, to a certain degree, is equivalent to hydrological similarity. The flaws of this theoretical restriction have already been discussed in "Regionalisation 2: Kriging", while the results from this as well as from other studies (e.g. Merz and Blöschl, 2004) have demonstrated practical success. In order to gain further insight into the plausibility of this regionalisation scheme, the underlying parameter maps interpolated from the gauged catchments will be examined below.

Fig. 8 shows that the most reasonable pattern is observed for the water balance correction factor for rain (PKOR): while negative

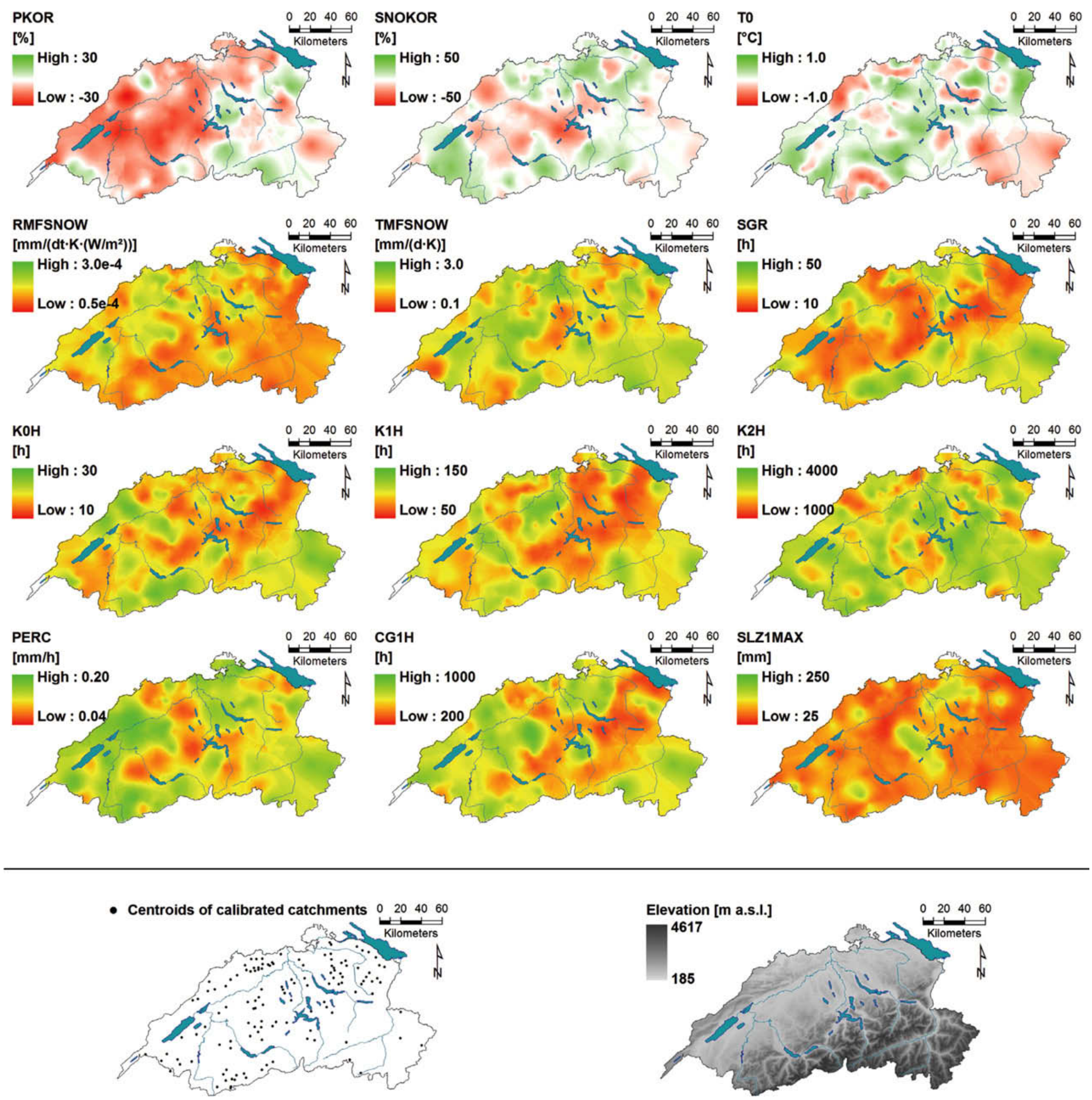

Fig. 8. Maps for 12 PREVAH tuneable model parameters, based on 140 successfully calibrated catchments. Maps for the two tuneable ice melt parameters are not shown due to poor spatial representativity. The two maps at the bottom show distribution of the 140 calibrated catchments (left, as used for creating the maps) and elevation in the study area (right).

values are generally predominant, positive values occur in the eastern Pre-Alps as well as in the central and eastern Alps. In its shape, this pattern is similar to those of annual precipitation (Kirchhofer and Sevruk, 1992) and of gauge error correction factors (Sevruk and Kirchhofer, 1992). It is therefore assumed that PREVAH's wind-dependent gauge error correction for liquid precipitation (see Viviroli et al., 2007 and Viviroli et al., 2009a) is slightly too large in general, corresponding to the average PKOR of −15% which compensates for this. On the other hand, PKOR seems to be too small in regions with high precipitation. The values for the water balance correction factor for snow (SNOKOR) are more balanced in general, while the patterns of positive and negative values are similar to the ones observed for PKOR. The distribution of the radiation melt factor for snow (RMFSNOW) follows topography to a certain degree, with decreasing values for higher altitudes. Since RMFSNOW controls the diurnal variability of snowmelt, the parameter distribution can be explained as being due to the stronger diurnal snowmelt cycle observed in lower altitudes. Difficult to interpret are the distributions of the temperature melt factor for snow (TMFSNOW, controlling the basic intensity of snowmelt) and of the threshold temperature for snowmelt (T0).

For the parameters of the runoff generation module (SGR, K0H, K1H, K2H, PERC, CG1H, and SLZ1MAX), strong correlations with soil type and underground properties would be expected. A comparison with maps of hydrogeology (Bitterli et al., 2004) and geology (SFSO, 2003), however, has revealed only moderate similarities; probably most discernible are the low percolation values (PERC) in the Pre-Alps. To a significant extent, this low correspondence must be attributed to PREVAH's runoff generation concept following the popular HBV model type (Bergström, 1976), which allows for various parameters to fulfil each other's functions to a certain degree (e.g. interflow and baseflow; see "Discussion" in Viviroli et al., 2009b). Due to this equifinality (see Beven, 2002), the calibrated parameter sets used for interpolation will not be entirely accurate in terms of physical meaning (see also Young, 2006). A further difficulty in finding reasonable connections between tuneable model parameters for runoff formation and soil properties is the large heterogeneity of soils in Switzerland and the comparatively low resolution of soil maps.

Altogether, the parameter maps created for Kriging regionalisation are only of an auxiliary nature and should not be interpreted on a pixel-by-pixel basis. Despite the restrictions concerning spatial plausibility of certain maps, results have shown that the derived regionalisation is successful and reliable.

Regression approach

Plausibility of parameter–attribute relationships. The Regression approach is founded on the assumption that there is a sufficiently strong relation between model parameters and catchment attributes (the Kriging approach is a special case, with the x and y positions in physical space being the only attributes). Due to parameter uncertainty, it must be expected that this theoretical assumption is not entirely met in reality. From a total of 984 possible parameter–attribute combinations examined (12 parameters, 82 attributes), 160 show correlation coefficients that differ significantly from zero at $\alpha \leqslant 0.05$ and out of these, 51 are significantly different from zero even at $\alpha \leqslant 0.025$. As Parajka et al. (2005) note, however, combinations with high correlation do not automatically imply physical plausibility. Therefore, the parameter–attribute relationships found for our data set are examined with the help of Table 2, which shows the 20 parameter–attribute pairs with the highest absolute correlation coefficient ($0.29 \leqslant |r_{xy}| \leqslant 0.41$); for easier interpretation, the relationships are examined here on the basis of all catchments, i.e. without separation into altitude zones. Reasonable relationships are found between water balance correction for rain (PKOR) and attributes of intense precipitation (PXXA, PXXB, PXXC, PXXD, and PXXG) or between certain runoff formation module parameters (SGR, K0H, and CG1H) and soil-related attributes (TSM2, KWM1, KWM2, and ROCK), as well as attributes for inclination (IN15) and relief (REL8). The relation between surface runoff (K0H) and wind speed (WIND) is difficult to interpret; it might be linked to the wind-dependent precipitation gauge error correction (see "Kriging: interpretation of resulting parameter maps").

Reproduction of the actual parameter values. It is expected that the Regression approach reproduces to a certain degree the parameters found in calibration, although this is neither entirely feasible (due to parameter uncertainty) nor necessarily required (due to parameter equifinality). This question of reproducibility is examined in Fig. 9, which compares, for each tuneable parameter, the actually calibrated values (abscissae) to the Regression regionalised values (ordinates) from all 140 catchments examined. The weights allocated to each individual data point for optimising the regression model are indicated by the shade of the dots. Fig. 9 depicts the

Table 2
The 20 highest correlated parameter–attribute pairs with respective Pearson correlation coefficient r_{xy}; all correlations are significant on a level of α = 0.025. The right column lists the generic meaning of the model parameters (see Viviroli et al., 2009b) and attributes (see Appendix). Data from 140 calibrated test catchments.

Param.–Attrib.	r_{xy}	Subject area
SGR–TSM2	0.41	Surface runoff and interflow – topographic index
PKOR–PXXD	0.40	Water balance correction – maximum 24 h-precipitation intensity
PKOR–PXXB	0.39	Water balance correction – maximum 1 h-precipitation intensity
PKOR–PXXC	0.38	Water balance correction – maximum 24 h-precipitation intensity
SGR–IN15	0.36	Surface runoff and interflow – inclination
SGR–ROCK	0.35	Surface runoff and interflow – rocky areas
PKOR–PXXA	0.35	Surface runoff – maximum 24 h-precipitation intensity
K0H–TSM2	0.34	Surface runoff – topographic index
K0H–IN15	0.34	Surface runoff – inclination
PKOR–PXXG	0.33	Water balance correction – maximum 15 min-precipitation intensity
PKOR–FBTR	0.33	Water balance correction – contributing areas
SGR–IAVG	0.33	Surface runoff and interflow – inclination
CG1H– KWM1	0.32	Fast response baseflow – hydraulic conductivity
K0H– REL8	0.32	Surface runoff – relief energy
K0H–WIND	0.31	Surface runoff – meteorology
K0H–KWM1	0.30	Surface runoff – hydraulic conductivity
K0H–IAVG	0.30	Surface runoff – inclination
SGR–REL8	0.30	Surface runoff and interflow – relief energy
SGR–HAVG	0.29	Surface runoff and interflow – average altitude
CG1H–KWM2	0.29	Fast response baseflow – hydraulic conductivity

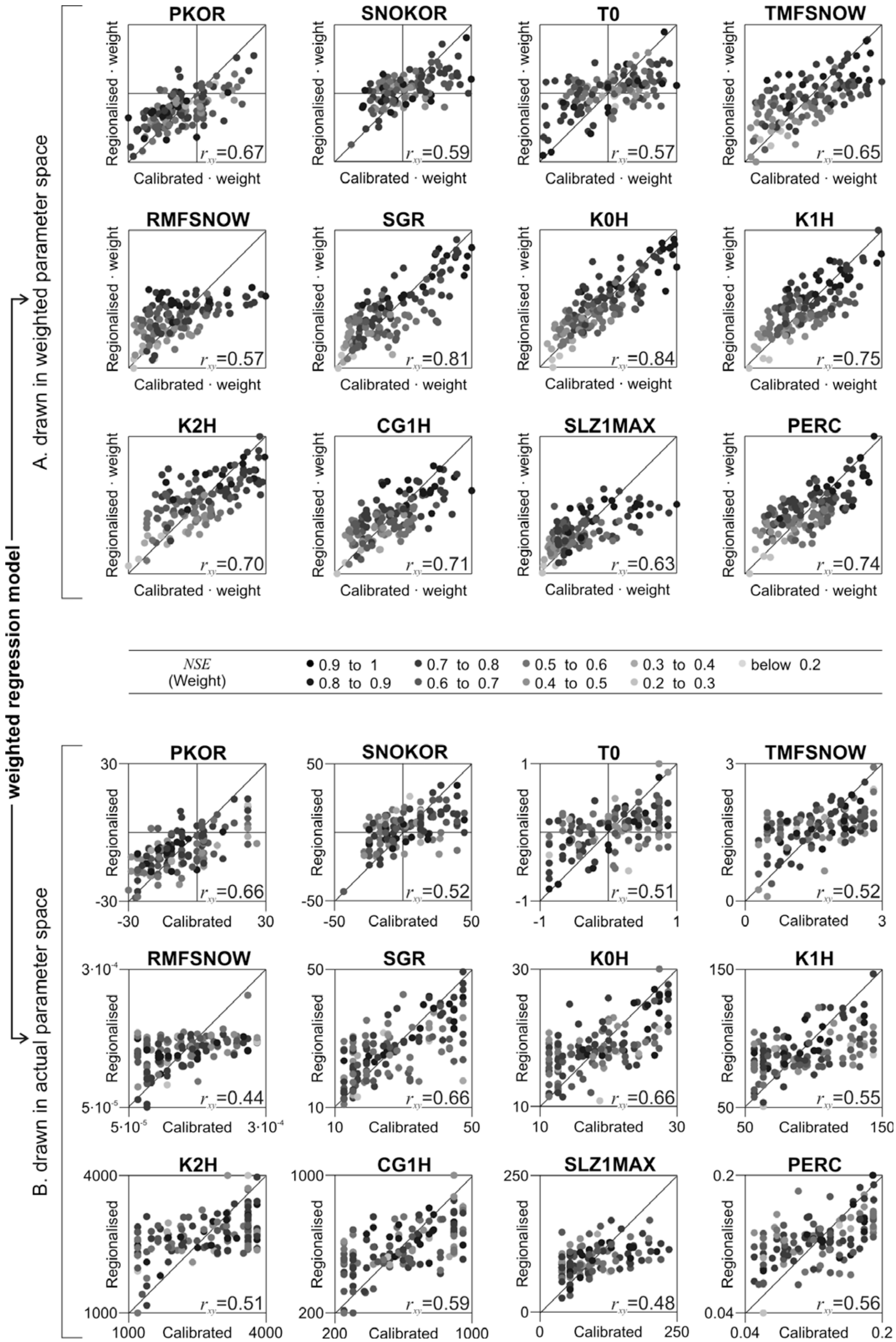

Fig. 9. Comparison of tuneable parameter values from hydrological model calibration (abscissae) with regression regionalisation (ordinates) on the basis of 140 catchments. The results of the same weighted regression model are drawn (A) in the weighted parameter space (as used to calibrate the regression model) and (B) in the actual (unweighted) parameter space. The shading of the dots indicates the weight of the data points, i.e. their Nash–Sutcliffe efficiency (*NSE*). The Pearson correlation coefficients (r_{xy}) relate to the data as plotted in the respective parameter spaces: In (A), r_{xy} is therefore the correlation coefficient of the weighted regression model; in (B), r_{xy} refers to the same weighted regression model, with the data points, however, plotted in the unweighted parameter space.

results in two versions: In the upper part (A), the data are drawn in the weighted parameter space, i.e. as 'seen' by the regression model, whereas in the lower part (B), the same data are drawn in the actual (unweighted) parameter space; both versions are based on the same weighted regression model.

Fig. 9 shows that the comparatively reliable parameters (i.e. those which achieve high *NSEs* in calibration) control the regression while those considered unreliable (low *NSEs*) are reproduced with lower accuracy. This is quantified with the Pearson correlation coefficients (r_{xy}) between calibrated and Regression regionalised values: In weighted notation, i.e. with the values drawn as they were used to calibrate the regression model (Fig. 9A), r_{xy} ranges between 0.57 and 0.84 ($\tilde{r}_{xy} = 0.69$). The data from the same weighted regression model drawn in unweighted notation (Fig. 9B) achieve r_{xy} values between 0.44 and 0.66 ($\tilde{r}_{xy} = 0.54$). The general decrease in r_{xy} from weighted (A) to unweighted (B) notation (Δr_{xy}) indicates the shortcomings mentioned above. Differences occur particularly for K1H ($\Delta r_{xy} = -0.20$), K2H ($\Delta r_{xy} = -0.19$), K0H ($\Delta r_{xy} = -0.18$), PERC ($\Delta r_{xy} = -0.18$), SLZ1MAX ($\Delta r_{xy} = -0.15$) and SGR ($\Delta r_{xy} = -0.15$). This means that for most of the runoff formation module parameters, the use of weights has a noticeable effect on the regression model. Only small decreases occur for SNOKOR ($\Delta r_{xy} = -0.07$), T0 ($\Delta r_{xy} = -0.06$) and PKOR ($\Delta r_{xy} = -0.01$), which suggests that regionalisation of water balance adjustment factors and snowmelt threshold temperature is relatively robust.

For RMFSNOW and SLZ1MAX, the ranges of calibrated and regionalised values differ visibly, meaning that either the regression model formulation or the information content of the catchment attributes is insufficient to accurately reproduce the parameter range.

Is the proposed regionalisation scheme effective?

A superficial look at the results presented in this paper might give the impression that the overall calibration and regionalisation framework is not as effective as would be desirable. First, regionalisation outperforms calibration in some cases, and second, simplistic regionalisations do not seem to perform distinctly worse than intelligent regionalisations. These issues are addressed below, drawing on additional analyses and providing further explanations.

Prior to this discussion, a general comment must be made about the performance of the regionalisations. Choosing the hydrological model for an extensive regionalisation task entails a compromise: On the one hand, the model should be able to capture hydrological processes in adequate detail. On the other hand, sufficiently detailed and reliable catchment descriptors should be available for regionalising the model parameters. Since such local information is generally sparse and subject to considerable uncertainty, regionalisation will be far more difficult for more sophisticated models. PREVAH is a model of intermediate complexity, pursuing a conceptual yet process-oriented and semi-distributed approach. This ensures that the model's tuneable parameters can actually be estimated for ungauged basins. In turn, the model's sensitivity is limited by its conceptual orientation. Furthermore, PREVAH uses topography, soil maps and land use information to parameterise a considerable number of model components a priori (see Viviroli et al., 2009b). This facilitates regionalisations and is one explanation why PREVAH simulations yield acceptable results even if the tuneable parameters are estimated only roughly.

Calibration outperforms regionalisation

Regionalisation outperforming calibration to some extent has been observed e.g. by McIntyre et al. (2005), who, however, provide no explanation for this phenomenon. For the present study, the uncertainty in model forcing conditions must be assumed to be an important factor. First of all, considerable uncertainties must be assumed in a mountainous region such as Switzerland due to precipitation measurement errors (Sevruk and Kirchhofer, 1992; Sevruk, 1997). Furthermore, meteorological gauging networks are usually inadequate in higher altitudes (see e.g. Goodrich et al., 1995; Briggs and Cogley, 1996). In Switzerland, merely one fifth of the gauging stations are located above the Swiss average altitude of 1312 m a.s.l. Given the sensitivity of hydrological models towards precipitation input, on-site calibration will be inherently affected by the bias caused by the aforementioned uncertainties. It can therefore be assumed that combining calibration information from many sites through regionalisation will reduce on-site errors in calibration. A further, if less important effect might be caused by the limited information available for calibration: As in most hydrological modelling studies, discharge is the only target value for calibration, and this might not be sufficient for identifying all model parameters. It is therefore plausible that combining calibration information from many sites through regionalisation is sometimes more effective than local calibration.

The aforementioned errors and uncertainties are further reduced by using an ensemble of three regionalisation approaches. In support of our explanation, it can be assumed that the uncertainty in catchment attributes used for parameter estimation is small as compared to the uncertainty in forcing conditions. Furthermore, Perrin et al. (2008) have recently proposed a successful method for selecting model parameters from an extensive library of calibrated catchments instead of relying on on-site calibrations.

It should, however, be noted that in our study, the number of outliers and extreme values is always higher in the regionalisations than in calibration. This applies to all of the three scores we have discussed in detail, namely the Nash–Sutcliffe efficiency (*NSE*) (Fig. 3), the volumetric score SVD_a (Fig. 4) and the estimation error for HQ_{100} (Fig. 5). This proves that calibration is still more robust than regionalisation. Further evidence of the model's sensitivity towards calibrated and regionalised parameters is given in "Box-plots tend to mask deficits in individual basins".

Simplistic regionalisations seem surprisingly effective

The results in "Results" were presented in the form of box-plots which emphasise the average performance of a parameter set and do not account for the behaviour of different parameter sets in the individual basins. It was nevertheless noted that the simplistic regionalisations (default parameters and parameter average) are inferior to the intelligent regionalisations as to their *NSE* scores. Although the difference is not pronounced, it is still noticeable in median, quartiles and maximum value (Fig. 3), particularly for the default parameter set. The high number of extremely bad *NSE* scores for the default set (10 out of 49 catchments have a *NSE* below 0) proves that this 'regionalisation' is not robust.

To explain the relatively high *NSE* performance of the averaged parameter set, we have to recall how this set is composed: The parameters are not averaged universally across the entire set of catchments, but individually for the three zones of mean altitude (Swiss Plateau, Pre-Alps, Alps), i.e. each altitude zone has its own average parameter set which is computed from the respective calibrated catchments. Since altitude determines many key factors in hydrological behaviour, the fair *NSE* scores of this regionalisation are less surprising. But as already noted in "Results", averaged parameter regionalisation has clear deficits concerning the volumetric score SVD_a (Fig. 4). The median of SVD_a is +191 mm yr^{-1}, which means that for a simulation of 20 years, the model has a runoff bias of +3820 mm or more in 50% of the catchments (the average runoff in Switzerland is 991 mm yr^{-1} according to Weingartner et al., 2007).

Since regionalisation, in this study, was devised for the ultimate goal of flood estimation for ungauged sites, the most compelling argument not to use simplistic regionalisations is furnished by an analysis of estimation errors for HQ_{100} (Fig. 5). Although the default parameter set shows a median error of only +1% (25 out of 49 catchments show positive, 24 negative errors), its range of errors is extended clearly. Therefore, these regionalised flood peak estimates are not robust at all. The averaged parameter set has a more narrow range of errors, but shows a clear tendency towards underestimation of HQ_{100} (36 out of 49 catchments have errors of more than −25%, 18 of even more than −50%,). It should therefore not be used for flood estimation, either.

Box-plots tend to mask deficits in individual basins

It has already been mentioned that box-plots tend to emphasise the overall average behaviour and mask the differences of the various regionalisation approaches in individual basins. Therefore, a slightly different analysis is proposed to complete the picture: Instead of characterising the performance of a parameter set across all catchments, a hierarchy of the parameter sets is determined for each individual catchment. After that, the hierarchies from all catchments are summarised for each parameter set, which makes it possible to conclude, for example, which parameter set performs best, second-best, etc. in the majority of catchments.

Fig. 10 shows the results from this analysis. In terms of best or second-best performance in individual catchments (brightest shade in Fig. 10), the parameter sets are ranked as follows (in descending order): standard calibration (#1 or #2 in 63% of catchments), flood calibration (33%), Nearest Neighbour and Combined regionalisations (27% each), Kriging and Regression regionalisations (20% each), parameter average (8%), default parameters (4%). This ranking is compatible with what would be expected if the scope and preconditions of the individual parameter sets were considered: Flood calibration performs worse than standard calibration due to the compromises described in Viviroli et al. (2009b). The intelligent regionalisations perform worse than the flood calibration on which they are based since some calibration information is lost in the process of transferring it to ungauged sites. Finally, the simplistic regionalisations perform worse than the intelligent ones because they do not or not sufficiently capture the spatial variability of hydrological conditions, which is a prerequisite for successful information transfer.

To corroborate these findings, a similar analysis was conducted using the volumetric efficiency score *VE*, which has recently been proposed by Criss and Winston (2008). *VE* circumvents some of the deficits of *NSE* such as sensitivity to high flows, runoff variance and meteorological model input and is defined as

$$VE = 1 - \frac{\sum_{t=1}^{n}|q_t - Q_t|}{\sum_{t=1}^{n}Q_t}, \quad VE \in [0,1] \tag{4}$$

where Q_t is the observed runoff at time step t and q_t the simulated runoff at time step t. *VE* represents the fraction of water delivered at the proper time, the best value being 1. Since it was not used in our study at all, it seems well suited for an independent assessment of our results. The hierarchy of parameter sets performing best and second-best is comparable to the result from the above analysis based on *NSE*: standard calibration (#1 or #2 in 84% of catchments), flood calibration (31%), Nearest Neighbour regionalisation (27%), Regression and Combined regionalisations (18% each), Kriging regionalisation (16%), parameter average (6%), default parameters (0%).

In summary, it is concluded that simplistic regionalisations, although they may be surprisingly effective on average, have a clearly higher probability of failing or at least delivering inferior results than intelligent regionalisations. This particularly applies to flood estimation, which is the main purpose of the modelling framework presented in this article and its companion paper (Viviroli et al., 2009b). It is therefore not recommended to spare the effort of a complex regionalisation. At the same time, the above analyses provide evidence of the sensitivity of our model towards calibrated, intelligently regionalised and simplistically regionalised parameters.

Conclusions and outlook

The results of estimating a 100-year flood (HQ_{100}) have shown that the continuous modelling framework introduced here is suitable for application in ungauged catchments of Switzerland. The estimation errors are in the same order of magnitude as those of today's standard empirical and stochastic approaches employed for ungauged catchments. While current procedures tend to overestimate flood peaks with a large recurrence interval, our estimation approach shows a slight tendency towards underestimation. However, the range of deviations is noticeably smaller for the continuous simulation.

Reviewing the entire flood estimation task, it becomes clear that our framework leads to a significant gain in information for estimating large floods in ungauged catchments. Today's approaches and the one presented here differ strongly in their methodological set-up, the former being empirical or statistical, the latter process-oriented. Particularly when these approaches

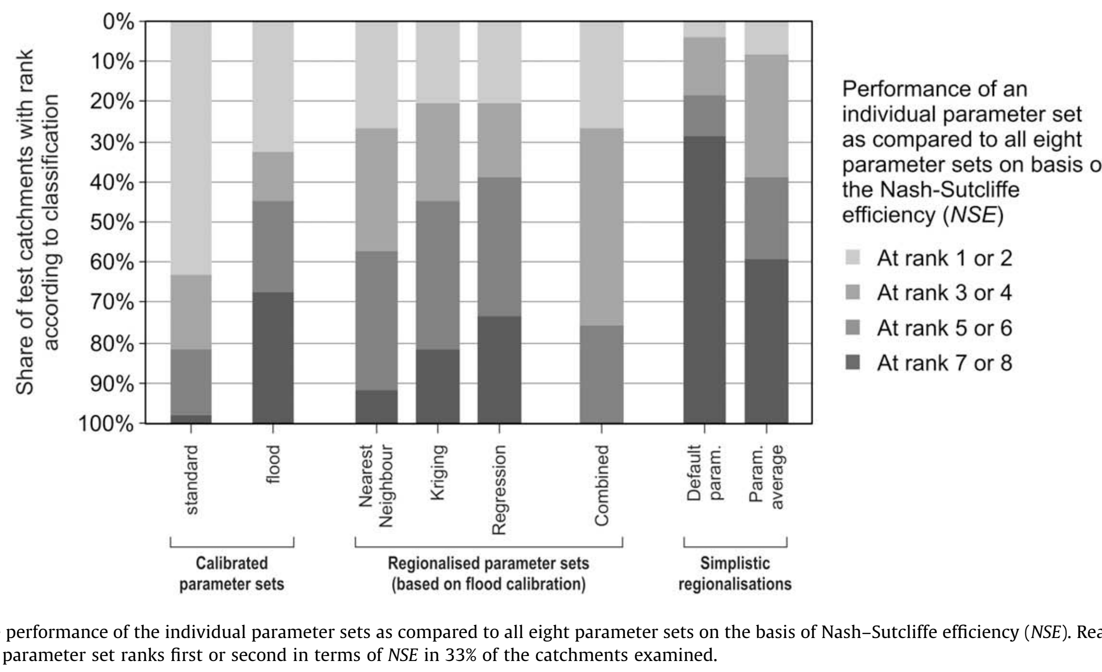

Fig. 10. Relative performance of the individual parameter sets as compared to all eight parameter sets on the basis of Nash–Sutcliffe efficiency (*NSE*). Reading example: The flood-calibrated parameter set ranks first or second in terms of *NSE* in 33% of the catchments examined.

are used in parallel, more robust estimates can be expected from our approach since we use discharge information from gauged sites in a completely different manner. So far, hydrological records and further data describing meteorology and physiography have been used to design empirical and statistical procedures; such approaches are always generalised to some extent. As to our solution, on the other hand, a complete modelling system has been designed which considers local conditions and processes to a clearly greater extent. We use entire discharge hydrographs for calibration and furthermore rely on physiographical and detailed meteorological records.

Thanks to the availability of complete simulated hydrographs, the approach presented in this paper is also suitable for estimating rare volumes of direct discharge (Viviroli, 2007). These figures are of equal importance to hydraulic engineering as peak values, but are neglected by most common estimation procedures. In contrast to standard methods, which are usually limited to predefined recurrence intervals (e.g. HQ_{100}), our approach is also very flexible concerning recurrence intervals provided that the extrapolation limits imposed by the length of the simulated hydrograph are respected. This flexibility is particularly important for modern flood mitigation schemes which seek to adjust the level of protection (i.e. the recurrence period on which the designed flood is based) to the importance and value of the object in question (Loat and Petraschek, 1997). Furthermore, the underlying model is process-oriented, which allows sensitivity and scenario analyses to be conducted, e.g. concerning the reaction of catchments to various scenarios of precipitation fields with high intensity (e.g. Schwanbeck et al., 2008) or to global change scenarios. Another wide field of applications opens up when weather generators are employed to simulate long series of model forcing; this could be used for a more thorough uncertainty quantification and for estimating floods with even longer recurrence intervals (see e.g. Cameron et al., 1999; Eberle et al., 2002; Leander et al., 2005).

The flood estimation framework presented in this article and its companion paper (Viviroli et al., 2009b) is currently being expanded into a nation-wide set of process-oriented flood estimations for Switzerland on behalf of the Swiss Federal Office for the Environment. For this purpose, long-term hydrographs for most of the relevant (and mainly ungauged) mesoscale catchments in Switzerland will be simulated and disseminated. This is an important step in transferring academic research findings into more practice-oriented applications.

Acknowledgements

This work was mainly funded through a PhD grant for Daniel Viviroli from the Swiss Federal Office for the Environment (FOEN). We gratefully acknowledge Manfred Spreafico's support in accompanying the project on the part of FOEN. We also thank Philipp Flach for his contributions to examining selected regionalisation schemes. The constructive comments of an anonymous reviewer were very helpful in improving the manuscript.

Appendix A

See Table A1.

Table A1
Catchment attributes used in the present study. More detailed descriptions of the individual attributes are found in Viviroli (2007).

Name	Description	Unit
Physiography[a]		
AREA	Area	(km^2)
ASPE	East-exposed surfaces	(%)
ASPN	North-exposed surfaces	(%)
ASPS	South-exposed surfaces	(%)
ASPW	West-exposed surfaces	(%)
CRCL	Circularity index	(–)
HMAX	Maximum altitude	(m a.s.l.)
HMIN	Minimum altitude	(m a.s.l.)
IAVG	Average inclination	(°)
IN03	Surfaces with inclination < 3°	(%)
IN15	Surfaces with inclination > 15°	(%)
PERI	Catchment perimeter	(m)
REL8	Relief energy of intermediate 80% altitude range	(m)
SHP1	Shape parameter as to Hundecha and Bárdossy (2004)	(–)
Land use[b]		
AGRC	Pastures and arable land	(%)
BAGR	Pasture and arable land in contributing areas	(%)
BBLT	Urban area in contributing areas	(%)
BFST	Forest area in contributing areas	(%)
FBTR	Contributing areas (average distance to channel 250 m)[c]	(%)
FRST	Forest areas	(%)
GLCC	Glaciated areas, accumulation zone	(%)
ROCK	Hard-rock areas	(%)
SOLS	Soil-covered areas	(%)
Soil[b]		
KWM1	Hydraulic conductivity, average	($mm\ h^{-1}$)
KWM2	Hydraulic conductivity, standard deviation	($mm\ h^{-1}$)
KWM3	Hydraulic conductivity, skewness	(–)
KWM4	Hydraulic conductivity, kurtosis	(–)
NFM1	Net field capacity, average	(%)
NFM2	Net field capacity, standard deviation	(%)
NFM3	Net field capacity, skewness	(–)
NFM4	Net field capacity, kurtosis	(–)
TSM1	Soil-topographic index[d], average	(–)
TSM2	Soil-topographic index[d], standard deviation	(–)
TSM3	Soil-topographic index[d], skewness	(–)
TSM4	Soil-topographic index[d], kurtosis	(–)

Table A1 (*continued*)

Name	Description	Unit
Hydrogeology[e]		
HG_1	Unconsolidated rock, high permeability	(%)
HG_2	Unconsolidated rock, intermediate permeability	(%)
HG_3	Unconsolidated rock, low permeability	(%)
HG0A	Unconsolidated rock, impermeable	(%)
HG0B	Hard rock, impermeable	(%)
HGFG	Hard rock, generic	(%)
HGKR	Karstic rock	(%)
Geology[b]		
G_F1	Hard rock – pores, fissures or karst	(%)
G_F2	Hard rock – variable permeability	(%)
G_F3	Hard rock – impermeable	(%)
G_L1	Unconsolidated rock, low permeability	(%)
G_L2	Unconsolidated rock, variable permeability	(%)
G_L3	Unconsolidated rock, high permeability	(%)
Precipitation[f]		
PSUM	Average of annual precipitation sum	(mm)
PAVG	Hourly precipitation ($\geqslant$0.02 mm), average	(mm)
PSDV	Hourly precipitation ($\geqslant$0.02 mm), standard deviation	(mm)
PSKW	Hourly precipitation ($\geqslant$0.02 mm), skewness	(–)
PKRT	Hourly precipitation ($\geqslant$0.02 mm), kurtosis	(–)
PCVA	Hourly precipitation ($\geqslant$0.02 mm), coefficient of variation	(–)
P_MD	Hourly precipitation ($\geqslant$0.02 mm), average Julian Date[g]	(–)
P_RR	Hourly precipitation ($\geqslant$0.02 mm), variability of Julian Date[g]	(–)
PXXG	Maximum 15 min-precipitation intensity, return period 2.33 a	(mm)
PXXA	Maximum 1 h-precipitation intensity, return period 2.33 a	(mm)
PXXB	Maximum 1 h-precipitation intensity, return period 100 a	(mm)
PXXC	Maximum 24 h-precipitation intensity, return period 2.33 a	(mm)
PXXD	Maximum 24 h-precipitation intensity, return period 100 a	(mm)
PDMD	Maximum 24 h-precipitation, average Julian date[g]	(–)
PDRR	Maximum 24 h-precipitation, variability of Julian Date[g]	(–)
PMRL	Relation of PMXX to PMXA	(–)
PMXX	Maximum precipitation intensity	(mm h^{-1})
PMXA	Average of maximum annual precipitation intensities	(mm h^{-1})
Climate[h]		
SSDR	Average annual sunshine duration	(%)
T_SP	Range of monthly average temperatures	(°C)
VAPO	Average vapour pressure	(hPa)
WIND	Average wind speed	(m s^{-1})
Position in space		
CTRX	East–west co-ordinate of catchment centroid	(km)
CTRY	North–south co-ordinate of catchment centroid	(km)

[a] Based on $100 \times 100\ m^2$ digital elevation model from SFSO (2003).
[b] See SFSO (2003) for details.
[c] See also Kölla (1987).
[d] See e.g. Ambroise et al. (1996).
[e] See Bitterli et al., 2004 for details.
[f] Based on Geiger et al. (1992), Jensen et al. (1997) and MeteoSwiss (2008).
[g] See Burn (1997).
[h] Based on data from MeteoSwiss (2008).

References

Ambroise, B., Freer, J., Beven, K.J., 1996. Toward a generalization of the TOPMODEL concepts: topographic indices of hydrological similarity. Water Resources Research 32, 2135–2145.

Andréassian, V., Perrin, C., Michel, C., Usart-Sanchez, I., Lavabre, J., 2001. Impact of imperfect rainfall knowledge on the efficiency and parameters of watershed models. Journal of Hydrology 250, 206–223.

Barben, M., 2003. Beurteilung von Verfahren zur Abschätzung seltener Hochwasserabflüsse in der Schweiz, Geographica Bernensia G71. Department of Geography, University of Bern, CH.

Bárdossy, A., 2007. Calibration of hydrological model parameters for ungauged catchments. Hydrology and Earth System Sciences 11, 703–710.

Bárdossy, A., Pegram, G.G.S., Samaniego, L., 2005. Modeling data relationships with a local variance reducing technique: applications in hydrology. Water Resources Research 41, W08404.

Beran, M.A., 1990. Preface. In: Beran, M.A., Brilly, M., Becker, A., Bonacci, O. (Eds.), Regionalization in Hydrology. Proceedings of an International Symposium held at Ljubljana, April 1990. IAHS Publication No. 191. IAHS Press, Wallingford, UK, pp. v–vi.

Bergström, S., 1976. Development and Application of a Conceptual Runoff Model for 421 Scandinavian Catchments. Bulletin Series A, No. 52. University of Lund, SE.

Beven, K.J., 2001. Rainfall–Runoff Modelling: The Primer. Wiley, Chichester, UK.

Beven, K.J., 2002. Towards a coherent philosophy for modelling the environment. Proceedings of the Royal Society of London, Series A – Mathematical, Physical and Engineering Sciences 458 (2026), 2465–2484.

Beven, K.J., 2006. A manifesto for the equifinality thesis. Journal of Hydrology 320, 18–36.

Bezzola, G.R., Hegg, C., (Eds.), 2007. Ereignisanalyse Hochwasser 2005. Teil 1 – Prozesse, Schäden und erste Einordnung. Umwelt-Wissen 0707. Federal Office for the Environment (FOEN), Bern, CH.

Bitterli, T., Aviolat, P., Brändli, R., Christe, R., Fracheboud, S., Frey, D., George, M., Matousek, F., Tripet, J.-P., 2004. Groundwater Resources. Hydrological Atlas of Switzerland, Plate 8.6. Federal Office for the Environment (FOEN), Bern, CH.

Blöschl, G., Grayson, R.B., 2000. Spatial observations and interpolation. In: Grayson, R.B., Blöschl, G. (Eds.), Spatial Patterns in Catchment Hydrology. Observations and Modelling. Cambridge University Press, Cambridge, pp. 17–50.

Blöschl, G., Merz, R., 2002. Methoden der hydrologischen Regionalisierung im Zusammenhang mit der Niederschlag-Abflussmodellierung. In: Gutknecht, D., (Ed.), Niederschlag-Abfluss Modellierung. Simulation und Prognose. Wiener

Mitteilungen Wasser Abwasser Gewässer 164. Technical University of Vienna, pp. 149–178.
Bobée, B., Rasmussen, P.F., 1995. Recent advances in flood frequency analysis. Reviews of Geophysics 33 (S1), 1111–1116.
Boughton, W., Droop, O., 2003. Continuous simulation for design flood estimation – a review. Environmental Modelling & Software 18 (4), 309–318.
Breinlinger, R., Gamma, P., Weingartner, R., 1992. Characteristics of small basins. In: Hydrological Atlas of Switzerland, Plate 1.2, Federal Office for the Environment (FOEN), Bern, CH.
Briggs, P.R., Cogley, J.G., 1996. Topographic bias in mesoscale precipitation networks. Journal of Climate 9, 205–218.
Burn, D.H., 1997. Catchment similarity for regional flood frequency analysis using seasonality measures. Journal of Hydrology 202, 212–230.
Burn, D.H., Boorman, D.B., 1992. Estimation of hydrological parameters at ungauged catchments. Journal of Hydrology 143, 429–454.
Burn, D.H., Zrinji, Z., Kowalchuck, M., 1997. Regionalisation of catchments for regional flood frequency analysis. Journal of Hydrologic Engineering 2 (2), 76–82.
Calver, A., Crooks, S., Jones, D., Kay, A.L., Kjeldsen, T.R., Reynard, N.S., 2005. National river catchment flood frequency method using continuous simulation. R&D Technical Report FD2106/TR. Department for Environment, Food and Rural Affairs (DEFRA), London.
Calver, A., Kay, A.L., Jones, D.A., Kjeldsen, T.R., Reynard, N.S., Crooks, S., 2004. Flood frequency quantification for ungauged sites using continuous simulation: A UK approach. In: Pahl-Wostl, C., Schmidt, S., Rizzoli, A.E., Jakeman, A.J., (Eds.), Complexity and Integrated Resources Management. Transactions of the 2nd Biennial Meeting of the International Environmental Modelling and Software Society (iEMSs). iEMSs, Manno.
Cameron, D., Beven, K.J., Tawn, J., Blazkova, S., Naden, P., 1999. Flood frequency estimation by continuous simulation for a gauged upland catchment (with uncertainty). Journal of Hydrology 219, 169–187.
Castellarin, A., Burn, D.H., Brath, A., 2001. Assessing the effectiveness of hydrological similarity measures for flood frequency analysis. Journal of Hydrology 241, 270–285.
Chow, V.T., 1964. Handbook of Applied Hydrology. McGraw-Hill, New York.
Criss, R.E., Winston, W.E., 2008. Do Nash values have value? Discussion and alternate proposals. Hydrological Processes 22, 2723–2735.
Cunnae, C., 1988. Methods and merits of regional flood frequency analysis. Journal of Hydrology 100, 269–290.
Dalrymple, T., 1960. Flood Frequency Analysis. Water Supply Paper 1543-A. US Geological Survey, Reston, VA, USA.
Deutsch, C.V., Journel, A.G., 1997. GSLIB: Geostatistical Software Library and User's Guide. Oxford University Press, New York.
Doherty, J., 2002. PEST – Model-Independent Parameter Estimation. User manual, fourth ed. Watermark Numerical Computing, Brisbane, AUS.
DVWK [Deutscher Verband für Wasserwirtschaft und Kulturbau], 1999. Statistische Analyse von Hochwasserabflüssen. DVWK-Merkblätter zur Wasserwirtschaft, 238, Wirtschafts- und Verlagsgesellschaft Gas und Wasser, Bonn, D., 42 pp.
Eagleson, P.S., 1972. Dynamics of flood frequency. Water Resources Research 8, 878–989.
Eberle, M.U., Krahe, P., Buiteveld, H., Wilke, K., Beersma, J., 2002. Estimation of extreme floods in the River Rhine basin by combining precipitation–runoff modelling and a rainfall generator. In: Spreafico M., Weingartner R. (Eds.), International Conference on Flood Estimation (March 6–8, 2002 in Bern, Switzerland). CHR-Report II-17. International Commission for the Hydrology of the Rhine Basin (CHR), Lelystad, NL, pp. 459–467.
Fischer, M.M., 1982. Eine Methodologie der Regionaltaxonomie: Probleme und Verfahren der Klassifikation und Regionalisierung in der Geographie und Regionalforschung. Bremer Beiträge zur Geographie und Raumplanung 3. University of Bremen, D.
Flach, P., 2007. Grundlagen zur Regionalisierung von Modellparametern des hydrologischen Modellsystems PREVAH. Publikationen Gewässerkunde 385. Institute of Geography, University of Bern, CH.
Garen, D.C., Marks, D., 2001. Spatial fields of meteorological input data including forest canopy corrections for an energy budget snow simulation model. In: Dolman, A.J., Hall, A.J., Kavvas, M.L., Oki, T., Pomeroy, J.W., (Eds.), Soil–Vegetation–Atmosphere Transfer Schemes and Large Scale Hydrological Models. IAHS publication No. 270, pp. 349–353, IAHS.
Geiger, H., Röthlisberger, G., Stehli, A., Zeller, J., 1992. Extreme point rainfall of varying duration and return period 1901–1970. In: Hydrological Atlas of Switzerland, Plate 2.4. Federal Office for the Environment (FOEN), Bern, CH.
Goodrich, D.C., Faurès, J.-M., Woolhiser, D.A., Lane, L.J., Sorooshian, S., 1995. Measurement and analysis of small-scale convective storm rainfall variability. Journal of Hydrology 173, 283–308.
Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York, US.
Grayson, R., Blöschl, G., 2000. Spatial modelling of catchment dynamics. In: Grayson, R., Blöschl, G. (Eds.), Spatial Patterns in Catchment Hydrology. Observations and Modelling. Cambridge University Press, Cambridge, UK, pp. 51–81.
Greenwood, J.A., Landwehr, J.M., Matalas, N.C., Wallis, J.R., 1979. Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. Water Resources Research 15, 1049–1054.
Greis, N.P., Wood, E.F., 1981. Regional flood frequency estimation and network design. Water Resources Research 17, 1167–1177.
Greis, N.P., Wood, E.F., 1983. Correction Regional flood frequency estimation and network design. Water Resources Research 19, 589–590.
Haider, S., 1994. Der Beitrag von Vorlandüberflutungen zur Verformung von Hochwasserwellen. Mitteilungen der Versuchsanstalt für Wasserbau, Hydrologie und Glaziologie (VAW) 129. ETH Zürich.
Hosking, J.R.M., Wallis, J.R., 1993. Some useful statistics in regional frequency analysis. Water Resources Research 29 (2), 271–281.
Hundecha, Y., Bárdossy, A., 2004. Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model. Journal of Hydrology 292, 281–295.
IH [Institute of Hydrology], 1999. Flood Estimation Handbook, vol. 5. Wallingford, UK.
Isaaks, E.H., Srivastava, R.M., 1989. An Introduction to Applied Geostatistics. Oxford University Press, New York.
Jensen, H., Lang, H., Rinderknecht, J., 1997. Extreme point rainfall of varying duration and return period 1901–1970 (updated map). In: Hydrological Atlas of Switzerland, Plate 2.4^2. Federal Office for the Environment (FOEN), Bern, CH.
Johansson, B., 1994. The relationship between catchment characteristics and the parameters of a conceptual runoff model: a study in the south of Sweden. In: Seuna, P., Gustard, A., Arnell, N.W., Cole, G.A. (Eds.), FRIEND: Flow Regimes from International Experimental and Network Data, IAHS publication No. 221. IAHS Press, Wallingford, UK, pp. 475–482.
Kirchhofer, W., Sevruk, B., 1992. Mean annual corrected precipitation depths 1951–1980. In: Hydrological Atlas of Switzerland, Plate 2.2. Federal Office for the Environment (FOEN), Bern, CH.
Klemeš, V., 2000. Tall tales about tails of hydrological distributions. Parts I and II. Journal of Hydrologic Engineering 5 (3), 227–239.
Kokkonen, T.S., Jakeman, A.J., Koivusalo, H.J., 2003. Predicting daily flows in ungauged catchments: model regionalization from catchment descriptors at the Coweeta Hydrologic Laboratory, North Carolina. Hydrological Processes 17, 2219–2238.
Kölla, E., 1987. Estimating flood peaks from small rural catchments in Switzerland. Journal of Hydrology 95, 203–225.
Lamb, R., 1999. Calibration of a conceptual rainfall-runoff model for flood frequency estimation by continuous simulation. Water Resources Research 35 (10), 3103–3114.
Lamb, R., Kay, A.L., 2004. Confidence intervals for a spatially generalized, continuous simulation flood frequency model for Great Britain. Water Resources Research 40, W07501.
Leander, R., Buishand, T.A., Aalders, P., De Wit, M., 2005. Estimation of extreme floods of the River Meuse using a stochastic weather generator and a rainfall–runoff model. Hydrological Sciences Journal 50 (6), 1089–1103.
Loat, R., Petraschek, A., 1997. Berücksichtigung der Hochwassergefahren bei raumwirksamen Tätigkeiten. Federal Office for Water Management (BWW), Federal Office for Spatial Planning (BRP) and Federal Office for Environment, Forest and Landscape (BUWAL), Bern, CH.
Maniak, U., 2005. Hydrologie und Wasserwirtschaft. Eine Einführung für Ingenieure, fifth ed. Springer Verlag, Berlin and Heidelberg, D.
McIntyre, N., Lee, H., Wheater, H.S., Young, A.R., Wagener, T., 2005. Ensemble predictions of runoff in ungauged catchments. Water Resources Research 41, W12434.
Merz, R., Blöschl, G., 2004. Regionalisation of catchment model parameters. Journal of Hydrology 287, 95–123.
Merz, R., Blöschl, G., 2005. Flood frequency regionalisation–spatial proximity vs. Catchment attributes. Journal of Hydrology 302, 283–306.
MeteoSwiss [Federal Office for Meteorology and Climatology], 2008. Time Series of Meteorological Variables. Zürich.
Mittelbach, H., 2006. Kalibrierung und Regionalisierung des flächendifferenzierten hydrologischen Modellsystems PREVAH in ausgewählten Einzugsgebieten der Schweiz. Diploma thesis, Technical University of Dresden and ETH Zürich.
Mosley, M.P., 1981. Delimitation of New Zealand hydrologic regions. Journal of Hydrology (NZ) 49, 173–192.
Naef, F., Scherrer, S., Zurbrügg, C., 1999. Major floods – differing reactions of catchments to intense rainfall. In: Hydrological Atlas of Switzerland, Plate 5.7, Federal Office for the Environment (FOEN), Bern, CH.
Nathan, R.J., McMahon, T.A., 1990. Identification of homogeneous regions for the purposes of regionalisation. Journal of Hydrology 121, 217–238.
Oudin, L., Andréassian, V., Perrin, C., Michel, C., Le Moine, N., 2008. Spatial proximity, physical similarity and ungaged catchments: a comparison based on 913 French catchments. Water Resources Research 44, W03413.
Parajka, J., Merz, R., Blöschl, G., 2005. A comparison of regionalisation methods for catchment model parameters. Hydrology and Earth System Sciences 9, 157–171.
Pearson, C.P., 1991. New Zealand regional flood frequency analysis using L-Moments–2. Flood frequency groups. Journal of Hydrology (NZ) 30, 77–92.
Peel, M.C., Chiew, F.H.S., McMahon, T.A., 2000. Extension of Unimpaired Monthly Streamflow Data and Regionalisation of Parameter Values to Estimate Streamflow in Ungauged Catchments. Natural Heritage Trust, Canberra, AUS.
Perrin, C., Andréassian, V., Rojas Serna, C., Mathevet, T., Le Moine, N., 2008. Discrete parameterization of hydrological models: Evaluating the use of parameter sets libraries over 900 catchments. Water Resources Research 44, W08447.
Pfaundler, M., 2001. Adapting, analysing and evaluating a flexible Index Flood regionalisation approach for heterogeneous geographical environments. Dissertation N° 14253, ETH Zürich, CH.

Pilon, P.J., 2004. Guidelines for reducing flood losses. United Nations Inter-Agency Secretariat of the International Strategy for Disaster Reduction (UNISDR), Geneva, CH.

Reed, D.W., Jakob, D., Robinson, A.J., Faulkner, D.A., Stewart, E.J., 1999. Regional frequency analysis: a new vocabulary. In: Gottschalk, L., Olivry, J.-C., Reed, D.W., Rosbjerg, D., (Eds.), Hydrological Extremes: Understanding, Predicting, Mitigating. IAHS Publication No. 255, pp. 237–243, IAHS, Wallingford, UK.

Samaniego, L., Bárdossy, A., 2005. Robust parametric models of runoff characteristics at the mesoscale. Journal of Hydrology 303, 136–151.

Schmocker-Fackel, P., Naef, F., Scherrer, S., 2007. Identifying runoff processes on the plot and catchment scale. Hydrology and Earth System Sciences 11, 891–906.

Schulla, J., Jasper K., 2000. Model Description WaSiM-ETH, Institute for Atmospheric and Climate Science, ETH Zürich.

Schwanbeck, J., Viviroli, D., Weingartner, R., Röser, I., Trösch, J., 2007. Prozessbasierte Abschätzung von Extremhochwassern im Einzugsgebiet der Sihl. Projektschlussbericht zum Auftrag des Amtes für Abfall, Wasser, Energie und Luft des Kantons Zürich (AWEL). Geographisches Institut der Universität Bern und TK Consult AG, Zürich.

Schwanbeck, J., Viviroli, D., Weingartner, R., 2008. Modellbasierte Sensitivitätsanalysen für das Berner Oberland. In: Hegg, C., Bezzola, G.-R. (eds.), Ereignisanalyse Hochwasser 2005. Teil 2 – Analyse von Prozessen, Massnahmen und Gefahrengrundlagen. Federal Office for the Environment (FOEN), Bern.

Sefton, C.E.M., Howarth, S.M., 1998. Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales. Journal of Hydrology 211, 1–16.

Seibert, J., 1999. Regionalisation of parameters for a conceptual rainfall–runoff-model. Agricultural and Forest Meteorology 98–99, 279–293.

Sevruk, B., 1997. Regional dependency of precipitation–altitude relationships in the Swiss Alps. Climatic Change 36, 355–369.

Sevruk, B., Kirchhofer, W., 1992. Mean annual corrections of measured precipitation depths 1951–1980. In: Hydrological Atlas of Switzerland, Plate 2.3, Federal Office for the Environment (FOEN), Bern, CH.

SFSO [Swiss Federal Statistical Office], 2003. GEOSTAT Database Products. Licence N° G158000315, ©SFSO. Neuchâtel, CH.

Shu, C., Burn, D.H., 2003. Spatial patterns of homogeneous pooling groups for flood frequency analysis. Hydrological Sciences Journal 48, 601–618.

Spreafico, M., Weingartner, R., Barben, M., Ryser, A., 2003. Hochwasserabschätzung in schweizerischen Einzugsgebieten – Praxishilfe. Berichte des BWG, Serie Wasser, 4. Federal Office for the Environment (FOEN), Bern.

Steinhausen D., Langer K., 1977. Clusteranalyse. Einführung in Methoden und Verfahren der automatischen Klassifikation. Walter de Gruyter, Berlin and New York.

Vandewiele, G.L., Elias, A., 1995. Monthly water balance of ungauged catchments obtained by geographical regionalization. Journal of Hydrology 170, 277–291.

Viviroli, D., 2007. Ein prozessorientiertes Modellsystem zur Ermittlung seltener Hochwasserabflüsse für ungemessene Einzugsgebiete der Schweiz. PhD thesis, Faculty of Science, University of Bern. Geographica Bernensia G77. Institute of Geography, University of Bern, CH, ISBN 978-3-905835-00-7.

Viviroli, D., Weingartner, R., 2004. The hydrological significance of mountains – from regional to global scale. Hydrology and Earth System Sciences 8, 1016–1029.

Viviroli, D., Gurtz, J., Zappa, M., 2007. The Hydrological Modelling System PREVAH. Geographica Bernensia P40. Institute of Geography, University of Bern, ISBN 978-3905835-01-0.

Viviroli, D., Zappa, M., Gurtz, J., Weingartner, R., 2009a. An introduction to the hydrological modelling system PREVAH and its pre- and post-processing-tools. Environmental Modelling & Software 24 (10), 1209–1222.

Viviroli, D., Zappa, M., Schwanbeck, J., Gurtz, J., Weingartner, R., 2009b. Continuous simulation for flood estimation in ungauged mesoscale catchments of Switzerland – Part I: Modelling framework and calibration results. Journal of Hydrology 377 (1–2), 191–207.

Vogel, R.M., Thomas, W.O., McMahon, T.A., 1993. Flood-flow frequency model selection in southwestern United States. Journal of Water Resources Planning and Management 119 (3), 353–366.

Wagener, T., Wheater, H.S., 2006. Parameter estimation and regionalization for continuous rainfall–runoff models including uncertainty. Journal of Hydrology 320, 132–154.

Weingartner, R., 1999. Regionalhydrologische Analysen – Grundlagen und Anwendungen. Beiträge zur Hydrologie der Schweiz 37. Swiss Hydrological and Limnological Society, Bern, CH.

Weingartner R., Aschwanden H., 1992. Discharge regime – the basis for the estimation of average flows. Hydrological Atlas of Switzerland, Plate 5.2, Federal Office for the Environment (FOEN), Bern, CH.

Weingartner, R., Barben, M., Spreafico, M., 2003. Floods in mountain areas – an overview based on examples from Switzerland. Journal of Hydrology 282, 10–24.

Weingartner, R., Viviroli, D., Schädler, B., 2007. Water resources in mountain regions: a methodological approach to assess the water balance in a highland–lowland-system. Hydrological Processes 21, 578–585.

Wilson, C.B., Valdes, J.B., Rodríguez-Iturbe, I., 1979. On the influence of the spatial distribution of rainfall on storm runoff. Water Resources Research 15 (2), 321–328.

Young, A.R., 2006. Stream flow simulation within UK ungauged catchments using a daily rainfall–runoff model. Journal of Hydrology 320, 155–172.