Regionalization of a conceptual rainfall-runoff model based on similarity of the flow duration curve: A case study from the semi-arid Karkheh basin, Iran

I. Masih, S. Uhlenbrook, S. Maskey, M.D. Ahmad

PII:  S0022-1694(10)00443-9
DOI:  10.1016/j.jhydrol.2010.07.018
Reference:  HYDROL 17225

To appear in:  Journal of Hydrology

Received Date:  15 December 2009
Revised Date:  17 May 2010
Accepted Date:  17 July 2010

Please cite this article as: Masih, I., Uhlenbrook, S., Maskey, S., Ahmad, M.D., Regionalization of a conceptual rainfall-runoff model based on similarity of the flow duration curve: A case study from the semi-arid Karkheh basin, Iran, Journal of Hydrology (2010), doi: 10.1016/j.jhydrol.2010.07.018

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Regionalization of a conceptual rainfall-runoff model based on similarity of the flow duration curve: A case study from the semi-arid Karkheh basin, Iran

I. Masih\textsuperscript{a,b,*}, S. Uhlenbrook\textsuperscript{b,c}, S. Maskey\textsuperscript{b}, M.D. Ahmad\textsuperscript{d}

\textsuperscript{a}International Water Management Institute (IWMI), 12 km Multan Road, Choak Thokar Niaz Baig, Lahore, Pakistan
\textsuperscript{b}UNESCO-IHE Institute for Water Education, PO Box 3015, 2601 DA Delft, The Netherlands
\textsuperscript{c}Delft University of Technology, Section of Water Resources, PO Box 5048, 2600 GA Delft, The Netherlands
\textsuperscript{d}CSIRO Land and Water, GPO Box 1666, Canberra, ACT2601, Australia

*Corresponding author: Tel: +92-42 35410050-3; Fax: +92-42 35410054; E-mail: i.masih@cgiar.org

Abstract

The study examines the possibility of simulating time series of streamflows for poorly gauged catchments based on hydrological similarity. The data of 11 gauged catchments (475 to 2522 km\textsuperscript{2}), located in the mountainous semi-arid Karkheh river basin of Iran, is used to develop the procedure. The well-known HBV model is applied to simulate daily streamflow with parameters transferred from gauged catchment counterparts. Hydrological similarity is defined based on four similarity measures: drainage area, spatial proximity, catchment characteristics and flow duration curve (FDC). The study shows that transferring HBV model parameters based on the FDC similarity criterion produces better runoff simulation compared to the other three methods. Furthermore, it is demonstrated that the FDC based regionalization of HBV model parameters works reasonably well for streamflow simulations in the data limited catchments in the mountainous parts of the Karkheh river basin. In addition, it could be demonstrated that the parameter uncertainty of the model has little impact on the FDC-based regionalisation approach. The methodology presented in this paper is easy to replicate in other river basins of the world, particularly those facing decline in streamflow monitoring networks and with a limited number of gauged catchments.
Keywords: poorly gauged catchment; regionalization; catchment similarity; flow duration curve (FDC), HBV model, predictions in ungauged basins

1 Introduction

1.1 Problem statement

Streamflow data are a prerequisite for planning and management of water resources such as design of dams and hydropower plants, assessment of water availability for irrigation and other water uses, assessment of flood and drought risks, and examining the ecological health of a river system. However, in many cases, observed streamflow data are not available or are insufficient in terms of quality and quantity. This undermines the informed planning and management of water resources at a specific site as well as at the river basin scale.

Hydrologists have responded to this challenge by developing various predictive tools, which are commonly referred to as regionalization methods (e.g. Sivapalan et al., 2003; Blöschl and Sivapalan 1995; Yadav et al., 2007). These methods can be broadly classified into two groups based on their temporal dimension. The first group deals with the estimation of continuous time series of streamflows (e.g. Magette et al., 1976; Merz and Blöschl, 2004). The second group estimates selected hydrological indices, such as the mean annual flow and base flow index (e.g. Nathan and McMahon, 1990), or various percentiles of the flow instead of continuous time series (e.g. regionalization of the flow duration curve- FDC) (Castellarin et al., 2004). Further classification can be done within each group. For example, Castellarin et al. (2004) classified regionalization methods for FDC into statistical, parametric and graphical approaches. The methods used for estimating the time series of streamflows can be further categorized into three sub-groups: (i) model parameter estimation by developing regression relationships between model parameters and catchment characteristics (e.g. Magette et al., 1976); (ii) transfer of model parameters, whereby a catchment similarity analysis is conducted and parameters of gauged catchment are used in simulations for
similar ungauged or poorly gauged catchment (e.g. Kokkonen et al., 2003; Wagener et al., 2007); and (iii) other regionalization techniques such as spatial interpolation of parameters (e.g. Merz and Blöschl 2004) or regional pooling of data for parameter estimation for ungauged catchments (e.g. Goswami et al., 2007).

Despite considerable progress in hydrology, the prediction of streamflow for ungauged or poorly gauged catchments still remains a major challenge (Sivapalan et al., 2003; Wagener and Wheater 2006). A brief review of some key studies involving commonly used regionalization methods applying conceptual rainfall-runoff models for streamflow estimations in ungauged or poorly gauged catchments is presented in the following section. In this paper, we defined a catchment as ungauged when no streamflow records exist, whereas a data limited or poorly gauged catchment is defined as a catchment where some measured streamflow records are available that are usually short, have many gaps and of poor quality. These records are not enough to achieve a satisfactory level of model calibration for streamflow simulation.

1.2 Review of regionalization methods using conceptual rainfall-runoff models

An overview of some applications of the rainfall-runoff models for regionalization in different parts of the world is given in Table 1 and briefly discussed below. The selected studies estimated continuous time series of streamflows using a rainfall-runoff model and reported the performance measures in terms of at least one of the three evaluation criterion, namely, Nash-Sutcliffe Efficiency (NSE), Coefficient of determination ($R^2$) and the mean annual volume balance (VB). These points were considered in the selection in order to keep consistency in comparison of this study and the presented literature in Table 1, as we also used the above mentioned three performance measures. Moreover, on the whole, the presented studies attempted to represent wide range of hydro-climatic environments and provide reasonably good coverage of most of the regionalization methods.

Magette et al. (1976) used 21 catchments (0.04-13 km$^2$) in USA for regionalization of six
selected parameters of the Kentucky Watershed Model (KWB). They used 15 catchment characteristics in developing regression equations and found that a multiple regression technique used in stepwise manner was successful in developing equations to estimate model parameters from catchment characteristics, but simple linear regression models were totally unsuccessful. They randomly selected 5 out of 21 catchments for validation. Although the validation results showed significant variations, they concluded that the approach was useful and should be further developed. Vandewiele et al. (1991) used 24 catchments (16-3160 km²) in Belgium for developing regression equations to estimate three parameters of a monthly conceptual rainfall-runoff model using the basin lithological characteristics. They concluded that their regionalization approach was capable of generating reliable monthly time series for ungauged sites within the region.

Servat and Dezetter (1993) evaluated the performance of two conceptual rainfall-runoff models (GR3 and CREC models) for possible applications to ungauged catchments in the north-western part of the Ivory Coast. They were able to relate all model parameters to catchment characteristics (rainfall and land cover) with varying degree of success. The regionalization results in terms of $R^2$ and NSE were variable, particularly for the NSE which was quite low (i.e. close to zero) in some cases.

Ibrahim and Cordery (1995) applied a monthly water balance model for predicting stream flows in New South Wales, Australia. The used model had four parameters, of which three were estimated from rainfall data. Abdulla and Lettenmaier (1997) regionalized 7 of the 9 parameters of a large scale model (VIC-2L) for Red and White river basins in USA. They estimated two of the model parameters from STATSGO soil data. For other parameters, they used 28 catchment variables, related to soil and climate, for developing multiple regression equations between model parameters and catchment variables. Their regionalization results were generally good in most cases, although they noticed better performance in humid and sub-humid catchments and poorer in semi-arid to arid catchments.

Seibert (1999) used the HBV model for a regionalization study using 11 catchments in
Sweden and found that six of the thirteen model parameters could be estimated from the land cover features (i.e., forest and lake areas). However, the application to ungauged catchments was achieved with varying degree of success, with daily NSE ranging from 0.23 to 0.72. Merz and Blöschl (2004) compared eight regionalization methods using the HBV model with data sets from 308 catchments in Austria. Parajka et al. (2005) conducted a follow up study of the Merz and Blöschl (2004) by improving the model structure (i.e., by dividing catchments into elevation bands of 200 m interval), adding snow cover data and conducting similarity analysis on the basis of catchment attributes. They concluded that the methods based on similarity approaches produce reasonably good regionalization results. This finding is also consistent with Kokkonen et al. (2003) who concluded that “When there is reason to believe that, in the sense of hydrological behaviour, a gauged catchment resembles the ungauged catchment, then it may be worthwhile to adopt the entire set of calibrated parameters from gauged catchment instead of deriving quantitative relationships between catchment descriptors and model parameters”.

McIntyre et al. (2005) proposed a regionalization method of ensemble modelling and model averaging and tested it using a five parameter version of the probability distributed model (PDM) on 127 catchments (1 to 1700 km²) in the United Kingdom. They selected donor catchments based on catchment similarity analysis for which three catchment characteristics, i.e., catchment area, permeability and rainfall were used. In this approach more than one donor catchment is selected, which is different from the usual approaches of using a single donor catchment for streamflow simulations at an ungauged site. Then the full parameter set of each of the donor catchment is used to predict streamflows at the ungauged catchment, thereby, generating an ensemble of flow values. Then the average streamflow could be taken from the weighted average with weights defined based upon the relative similarity. They found that the proposed method performs reasonably well as compared to the established procedure of regressing parameter values from the catchment descriptors. However, they also noted that the new method estimated the low flows better as compared to high flows for their study area. They recommended further testing of the model,
especially to test different model types and improved definition of similarity.

Goswami et al. (2007) developed a methodology that uses a regionalization and multi-model approach for simulating streamflows in ungauged catchments. Like other methods, their methodology did not involve transfer of model parameters from gauged catchment to ungauged catchment, and model parameters need not to be related to physical catchment descriptors. They used seven different models for regionalization and for each tested the three methods that involve the use of the discharge series by taking regional averages, regional pooling of data and transposition of discharge data of the nearest neighbour. They used 12 gauged catchments in France to illustrate their methodology and each time considered one of them as ungauged for the application of the method and then compared the results with observed time series of daily discharge using the NSE criterion. The results indicated a mix of success and failure for the individual catchments and tested methods. However, they concluded that the pooling method of regionalization coupled with the conceptual soil moisture accounting and routing model (SMAR) was the best approach for simulating flows in ungauged catchments in that region. The second best method was the transposition of data from the nearest neighbour provided the catchments are similar in the hydro-meteorological, physiographic characteristics and drainage area.

Oudin et al. (2008) compared three widely used regionalization approaches for (a large number of) 913 French catchments (10 to 9,390 km²) by using two conceptual rainfall-runoff models (GR4J and TOPMO models). They showed that the spatial proximity based regionalization performed the best for their sample of catchments. They also noted that the dense network of tested catchments used in their study might have resulted in favour of spatial proximity approach and recommended that this approach should also be tested in other regions, particularly where less gauged catchments are available.

The presented studies reveal that a considerable progress has been made to estimate streamflows at ungauged catchments and quite a number of promising methods have been developed over the past few decades. However, the studies also depict a mix of success and failure
of the available methods within a study region or while comparing outcomes from the different
regions. Moreover, the tested regionalization approaches indicate large variability in the achieved
performance statistics, which shows considerable scope of further improvement. Therefore, there is
every motivation to make further progress on this important subject of regionalization in hydrology.

Table 1. An overview of some studies related to regionalization of conceptual rainfall-runoff
models

1.3 Scope and objective of this paper

The main research question examined in this paper is whether or not the parameters of a conceptual
hydrological model applied to a gauged catchment can be successfully transferred for simulating
streamflows in hydrologically similar but data limited or poorly gauged catchment. In this study,
the HBV model (Bergström 1992) is used for streamflow simulations in the Karkheh river basin,
Iran. The hydrologic similarity is defined based on four measures, i.e., drainage area, spatial
proximity, catchment characteristics and flow duration curve (FDC). FDCs are frequently used for
comparing the response of gauged catchments, but their potential use for the regionalization of
conceptual rainfall-runoff models for flow estimation for the poorly gauged catchments needs to be
explored and is a main objective of this paper. It should be noted that the streamflow data is
required for the construction of a FDC. However, it should be recognized that a FDC could be
established from the catchment characteristics for ungauged catchments using available FDC
regionalization methods (e.g. Castellarin et al., 2004). For poorly gauged catchments, the available
records, though short, could be used for the FDC construction. These insufficient records may not
be used directly for rainfall-runoff modelling as indicated in the previous section. Another
limitation in their direct use for modeling purpose is the unavailability of other corresponding data
sets required for modeling, e.g. climatic data for the same period as of runoff data may not be
available. These typical limitations were faced for the poorly gauged catchments in the Karkheh
basin, Iran, providing main motivation for this regionalization study.

It is pertinent to note that the above mentioned methods evaluated in this study require very
limited data resources and were most suitable in the context of the data limited region under study.
The other commonly used methods, such as regionalization of the model parameters, generally
require data sets from a large number of gauged catchments for developing statistically sound
relationships between model parameters and catchment characteristics. Due to limited availability
of gauged catchments and necessary data sets, these data intensive methods were not tested for the
study area. Nevertheless, the results of this study were compared with those published in the
literature from some widely recommended methods tested in other regions of the world.

2 Materials and methods

2.1 Study catchments and available data

The Karkheh river basin is located in the western part of Iran (Figure 1). It drains an area of 50,764
km², of which 80% is part of the Zagros mountain ranges of Iran, from where almost all of the
basin’s runoff is generated. About 60% of the basin area is in the elevation range of 1000-2000
meters above sea level (m asl), and about 20% is below 1000 m asl. Agriculture and human
settlements are mainly found in the valleys of the upper basin and in the arid plains in the lower
parts, where the river eventually terminates in the Hoor-Al-Azim Swamp, a large transboundary
wetland shared with Iraq. The climate is semi-arid in the uplands (north) and arid in the lowlands
(south). The precipitation (P) exhibits large spatial and temporal variability. The mean annual
precipitation is about 450 mm/a, ranging from 150 mm/a in the lower arid plains to 750 mm/a in the
upper mountainous parts (JAMAB, 1999). Most of the precipitation (about 65%) falls during December to March (winter) and almost no precipitation can be observed during June to September (summer). In the mountainous parts during winter, due to temperatures often falling below zero degrees Celsius, the winter precipitation falls as snow and rain. A recent remote sensing based study on snow cover in the study area have shown that the snow water equivalent for the mountainous parts of the Karkhe basin is about 75 mm/a (Saghafian and Davtalab, 2007), which is about 17% of the long term annual precipitation in the basin. The seasonality of precipitation and distribution into rain and snow influences the streamflows, indicated by the high flows during March to May resulting from the combined effect of snowmelt and rainfall (Masih et al., 2009 and 2010). Further details on study basin can be found at Sutcliffe and Carpenter (1968); JAMAB (1999); Ahmad et al. (2009); Masih et al. (2009); Masih et al. (2010); and Muthuutta et al. (2010).

Figure 1. Salient features of the study area and location of the study catchments and used climatic stations.

In the Karkhe basin streamflow data are not available for many catchments and the existing records have gaps. There were about 50 streamflow gauging stations installed after 1950 out of which only 24 measured continuously. Filling these data gaps by estimating missing streamflow time series for the poorly gauged catchments was required for a solid understanding of the hydrology and its spatial and temporal variability, which in turn should guide informed water management decisions.

Eleven gauged catchments, draining tertiary level streams (475 to 2522 km²), located in the upper mountainous parts of the Karkhe basin were selected for this study (Figure 1 and Table 2). The study period of January 1, 1987 to September 30, 2001 was selected considering the data availability/quality and representation of dry, wet and average climatic conditions. Time series of daily precipitation data for the study period was available for 41 climatic stations, well scattered
across the study domain (Figure 1). The areal precipitation estimates were used in the model
simulations, which were obtained by interpolation of the available station data by using inverse
distance and elevation weighting (IDEW) technique (see next section). Temperature data from 8
climatic stations (Figure 1) were available and the station nearest to the catchment was used in the
simulations for that respective catchment. The missing values in the data sets were estimated based
on the values from neighbouring stations. The missing values in the temperature data sets were few
(less than 1% in most cases), with the exception of one station where records were available only
for 1996 to 2001. Generally, temperature data of a station showed very good correlation with
the corresponding data from the neighbouring stations ($R^2 > 0.90$) used for infilling of the missing
records. In case of precipitation data, 7 out of 41 stations had no missing records. On average, there
were 5% in filled precipitation events, ranging from 0 to 16%. Hargreaves’ equation (Hargreaves et
al., 1985) was used to estimate the reference evapotranspiration using daily data of maximum,
minimum and mean temperature.

Table 2. Salient features of the selected streamflow gauges.

2.2 Preparation of the precipitation input

The earlier studies for the Karkheh basin have demonstrated that topography has a strong influence
on the spatial distribution of precipitation in this mountainous region (Sutcliffe and Carpenter,
1968; JAMAB, 1999; Muthuwatta et al., 2010). Elevation is known to be an important factor
governing the spatial variability. These findings are in general agreement with other mountainous
regions of the world (e.g. Daly et al., 2002). Moreover, the rain gauge data may not adequately
represent the precipitation over an entire catchment. This issue is further exacerbated for
catchments where rain gauge density is lower, such as for the region under study. Under such
conditions, areal precipitation was considered better representative of a catchment compared to the station data.

The daily station data were interpolated and aggregated at the catchment scale using the IDEW technique. The hydrological data processing software HyKit was used (Maskey, 2007). The distance weighting method has already proven to perform better compared to some other standard methods of regionalization for the Karkheh and its neighbouring basins in the Zagros mountains, Iran (Saghafian and Davtalab, 2007). HyKit is a grid-based interpolation technique and offers also the possibility of defining elevation weighting along with the distance weighting, making it more suitable for mountainous regions like the Karkheh basin where topographic impacts on precipitation are important. The mathematical form of the equation used for interpolation is as follows:

\[
\hat{p}_k = W_D \sum_{i=1}^{N} \frac{1}{D} w(d), p_i + W_Z \sum_{i=1}^{N} \frac{1}{Z} w(z), p_i
\]  

where, \( \hat{p} \) in mm per time step is the interpolated precipitation for a grid cell, \( W_D (\cdot) \) and \( W_Z (\cdot) \) are the total weighting factors for distance and elevations, respectively, \( p_i \) is the precipitation value in mm per time step of the \( i \)-th gauge station and \( N \) is the number of gauges that are used in the interpolation for the current grid cell. Similarly, \( w(d)_i(\cdot) \) and \( w(z)_i(\cdot) \) are the individual gauge weighting factors for distance and elevation, respectively, and \( D(\cdot) \) and \( Z(\cdot) \) are the normalization quantities given by the sum of individual weighting factors \( w(d)_i \) and \( w(z)_i \), respectively, for all the gauges used in the interpolation. The weighting factors \( w(d)_i \) and \( w(z)_i \) based on inverse of distance and elevation are given by

\[
w(d) = 1/d^a \quad \text{for} \quad d > 0
\]

\[
w(z) = \begin{cases} 
1/z^b_{\min} & \text{for} \quad z \leq z_{\min} \\
1/z^b & \text{for} \quad z_{\min} < z < z_{\max} \\
0 & \text{for} \quad z \geq z_{\max}
\end{cases}
\]

where, \( d \) is the distance in km between the current grid and the gauge station used for interpolation,
\( z \) is the absolute elevation difference in m between the current grid cell and the gauge station used for interpolation, \( a (\cdot) \) and \( b (\cdot) \) are constants for distance and elevation weightings, respectively, and \( z_{\text{min}} \) (m) and \( z_{\text{max}} \) (m) are the minimum and maximum limiting values for computing elevation weightings. The reasons for using limits in elevation are discussed in Daly et al. (2002). Note that in this interpolation technique, no grid cell can contain more than one gauging station and that the grid cell which contains a station will retain the same precipitation as of the gauge station.

Daily time series of precipitation from all available gauges were used for interpolation in 5\times5 \text{ km}^2 grids, which are then aggregated at the catchment level. The parameters of interpolation, i.e. the exponents \( a \) and \( b \), the importance factors \( W_D \) and \( W_Z \) and the radius of influence, were determined by cross validating the interpolated rainfall using Jack-Knife method (e.g. Varljen et al., 1999). The radius of influence determines the number of gauges to be included in interpolation, which may vary for each grid cell. The cross validation was done for the 10 selected rain gauge locations/grid cells scattered across the whole study area. The interpolated values were in good agreement with the observed ones. The mean and standard deviation of monthly \( R^2 \) were 0.91 and 0.04, respectively. As expected, the daily \( R^2 \) values were comparatively lower than the monthly ones (with mean \( R^2 \) of 0.62 and standard deviation of \( R^2 \) of 0.13). However, considering high spatial variability of precipitation in this mountainous terrain, the achieved \( R^2 \) values were considered satisfactory. It is noteworthy that a detailed comparison of model efficiency under station and areal precipitation data was beyond the scope of this paper. However, we have conducted a detailed comparison of streamflow simulations under both cases using a semi-distributed hydrological model, Soil Water Assessment Tool (SWAT). The results have shown better streamflow simulations using areal precipitation input compared to those simulated with the gauge rainfall without areal interpolation (Masih et al. in review).

The final parameters used for the interpolation were: radius of influence = 70 km, \( a = 2, b = 1 \), \( W_D = 0.8 \) and \( W_Z = 0.2 \). Analysis of rainfall correlations among the gauging stations was additionally carried out to determine the radius of influence. Most of the stations within a distance
of 70 km exhibited good correlation with each other (i.e. greater than 0.8 at monthly time scale). This indicates the dominance of frontal rainfall in the study area, especially during the winter times. Generally, the interpolation method used more stations in the catchments located in the upper part of the study area because of the higher station density compared to the middle and lower parts. The limiting values for elevation weighting $z_{min}$ and $z_{max}$ are selected as 100 m and 1500 m, respectively, which are within the range prescribed by Daly et al. (2002).

2.3 Naturalization of the streamflows

The abstraction of river water for irrigation purposes influenced the river flows in some of the study catchments. Therefore, naturalization of streamflows was carried out by adding abstraction rates, if any, to the observed streamflows. The direct pumping from the streams is the main mode of irrigation diversions by the farmers. However, no pumping records or data for other means of surface water diversions were available. Therefore, abstractions were estimated using the available information on crop evapotranspiration, cropping patterns and cropped area, estimates of irrigation efficiencies and total annual abstractions. The procedure used is summarized below.

Calculation of crop water demand. The daily potential crop evapotranspiration ($ET_c$) was calculated using the following equation:

$$ET_c = \sum_{j=1}^{n} A_j Kc_j ET_o$$  \hspace{1cm} (4)

where $ET_c$ is the total potential crop evapotranspiration in m$^3$/d, $A_j$ is the area under the $j^{th}$ crop in m$^2$, $ET_o$ is the reference evapotranspiration expressed in m/d, $Kc_j$ is the crop coefficient for the $j^{th}$ crop (according to Allen et al., 1998), and $n$ is the number of crop types, which are mainly wheat, barley, alfa alfa, sugarbeat, maize and orchards. The data on cropping patterns and cropped area
were obtained from JAMAB (1999) whereas sowing and harvesting dates were based on field surveys. The total ETc was obtained by the summation of the values for the individual crops.

*Calculation of irrigation demand and streamflow abstractions.* The irrigation demand was estimated using the following equation:

\[ I_d = ET_c \left(1 - \frac{e_p P}{ET_0}\right) \]  

(5)

where \( I_d \) is irrigation demand, m\(^3\)/d; the ratio of effective precipitation and reference evapotranspiration was computed using monthly data of precipitation (\( P \)), mm/month, and \( ET_0 \) mm/month; \( e_p \) (-) is fraction of the precipitation effectively used as evapotranspiration. For the whole Karkheh basin, JAMAB (1999) estimated that 66% of the annual precipitation is consumed as evapotranspiration and 34% forms the renewable water resources. For this study conducted in the upper catchments of Karkheh basin, the value of \( e_p \) was assumed as 0.5, since the evaporation rates are lower in upper mountainous part of the basin compared to the lower arid plains.

The abstractions from the streams were estimated using the following equation:

\[ I_{sw} = f_{sw} \frac{I_d}{\eta} \]  

(6)

where \( I_{sw} \) is the surface water withdrawals, m\(^3\)/d, \( f_{sw} \) is the fraction of surface supplies in the total irrigation withdrawals and \( \eta \) (-) is the irrigation efficiency. The used values of \( \eta \) were in the range of 0.3 to 0.7 (JAMAB, 1999). The lower values of \( \eta \) correspond to catchments with higher surface water withdrawals and vice versa. The annual values of \( f_{sw} \) were also available from the study of JAMAB (1999) who estimated total irrigation withdrawals from surface water and groundwater sources in the study catchments for the period of 1993-94. The catchments where surface water was
the main source of irrigation (i.e., $f_{sw} > 0.9$), the same value of $f_{sw}$ was used for each day of the year. For catchments where conjunctive use of surface water and groundwater was present, the annual value of $f_{sw}$ was distributed into monthly values following the supply-demand principle whereby higher values were assigned to the months having higher streamflows (i.e., March to June) and lower values to the months having lower streamflows (i.e., August to October). In this way, $f_{sw}$ was varied for each month but was kept constant for each day of a month. The estimated values of $I_{sw}$ were compared with the available estimates at annual scale for the year 1993-94. If the difference was more than 15%, the procedure was repeated by modifying the values of $\eta$ and monthly distribution of $f_{sw}$. Finally, $I_{sw}$ values were added to the observed streamflow in order to get the naturalized streamflows. The observed and naturalized streamflows are given in Table 2, which indicates the extent of the influence of naturalization on each of the study catchment. As an example, Figure 2 shows the observed and naturalized streamflows of one catchment (Aran). This illustrates the streamflow differences in particular during the late spring and summer, when the crop water requirements are the largest. Discussions with local experts concluded that these corrections are reasonable and reflect the impact of local practices.

Figure 2. Naturalized and observed daily time series of streamflows of Aran catchment.

2.4 Model calibration and validation at the gauged catchments

The HBV model was selected for this study due to the following reasons: (i) its model structure is simple but flexible such that a catchment can be sub-divided into different elevation and vegetation zones, which was important to model the mountainous Karkheh basin, (ii) it is not very data intensive and most of the data needed were readily available, and (iii) it has been widely used world-wide in particular in snow-influence climates, but recent studies demonstrate also its
applicability in semi-arid environments (e.g. Lidén and Harlin 2000; Love et al., in press). A
number of studies have demonstrated its suitability in regionalization studies (e.g. Seibert 1999;
Merz and Blöschl 2004; Götzinger and Bárddossy 2007).
The HBV model (Bergström, 1992) is a conceptual rainfall-runoff model which simulates
daily discharge using as input variables daily rainfall, temperature and daily or monthly estimates of
$ET_o$. The model consists of different routines representing the snow accumulation and snowmelt by
a degree-day method, recharge and actual evapotranspiration as functions of the actual water
storage in a soil box, runoff generation by two linear reservoirs with three possible outlets (i.e.
runoff components), and channel routing by a simple triangular weighting function. Further
descriptions of the model can be found elsewhere (Bergström, 1992; Seibert 1999 and 2002;
Uhlenbrook et al., 1999). The version of the model used in this study, “HBV light” (Seibert, 2002),
corresponds to the version HBV-6 described by Bergström (1992) with only two slight changes.
Instead of starting the simulation with some user-defined initial state values, this version uses a
warming-up period during which the state variables evolve from standard initial values to more
appropriate values for the given hydro-meteorological conditions. Furthermore, the restriction that
only integer values are allowed for the routing parameter MAXBAS has been removed, which
enables a somewhat more realistic parameterisation of the runoff routing processes. In the original
version of the HBV model (Bergström, 1992) computations in both the snow and soil routine are
performed individually for each elevation zone before the groundwater recharge of all zones is
lumped in the response routine. In the model version used in this study, the upper box in the
response function is treated individually for each elevation zone additionally to the separate
computations in the snow and soil routines. This version is considered more logical than the
standard HBV versions, especially for a mountainous area like the Karkheh basin. It is important to
note that the parameters of the snow and soil routines (Table 3) are estimated in a distributed
manner for each land use category. But these parameters remain the same for each elevation zone
within a land use category. The parameters of the response and routing routines (Table 3) are
estimated in a lumped way for each catchment.

The HBV model was applied to each of the 11 gauged catchments and was calibrated using daily climatic and streamflow data from January 1, 1987 to September 30, 2001. The data was split into calibration (October 1, 1987 to September 30, 1994) and validation (October 1, 1994 to September 30, 2001) periods. Before calibration, a warming-up period of 273 days was used for initialization so that model parameters attain appropriate initial values. Each catchment was divided into a number of elevation zones at an interval of 200 meters. Each elevation zone was divided into three vegetation zones, namely forest (zone 1), cropland (zone 2) and range/bare lands (zone 3). Since the elevation is known to have major impacts on the distribution of rainfall and temperature and have already been studied in the region, the values of the two parameters for lapse rates of precipitation and temperature were based on the earlier studies of Sutcliffe and Carpenter (1968) JAMAB (1999) and Muthuwatta et al. (2010). The values of lapse rates were kept constant for all catchments and set to an increase of 5.5 % per 100 m increase in elevation for precipitation and to a decrease of 0.4 °C per 100 m increase in elevation in case of temperature. A Genetic Algorithm (GA) based automatic calibration method, which is built-in in the present version of the model by Seibert (2002), was applied during model calibration. Similar calibration methods have been widely used as a global optimization tools (e.g. Wang, 1991; Seibert, 2000; Maskey et al., 2004). The ranges of parameter values (Table 3) were selected based on our understanding of the study region, experiences of other studies (Seibert, 1999; Uhlenbrook et al., 1999; Uhlenbrook and Leibundgut, 2002) and initial model runs for the study catchments. For instance, the threshold temperature (TT) for snow was set to fall in the range of -2.5 to 2.5 °C. The optimized threshold value of this parameter defines whether the precipitation falls in the form of rain or snow. As indicated in section 2.1, during winter months, the temperature may fall below optimized snow temperature threshold causing precipitation to occur in the form of snow fall apart from the rain events during this period. The parameters of the snow and soil routines were estimated, using above mentioned GA based optimization procedure, in a distributed manner, thus having different values for each of the three
vegetation zones. The parameters of the response and routing routines could only be estimated uniformly in the current version of the HBV model and, therefore, were representative of the whole catchment. The $NSE$ estimated at the at daily time step was used as an objective function to estimate the model performance (Nash and Sutcliffe, 1970). The $NSE$ is considered as a robust approach to assess the model goodness of fit in hydrological modelling and is widely used (e.g. ASCE, 1993). However, it is also worth noting that the results based on $NSE$ optimisation could be biased towards high flows, which warrant caution in interpretations (e.g. Wagener et al., 2004).

Further, other commonly used measures also have their own merits and constraints. For instance, widely used performance measure, $R^2$, may reflect higher values (good performance) if the variability of two data sets is well synchronized despite their volumetric difference. Therefore for having better picture of the results, in addition to $NSE$, we examined $R^2$ and $VB$. The use of more performance evaluation measures and their comparison is beyond the scope of this paper and could be found elsewhere in the literature (e.g. ASCE 1993, Gupta et al., 2009).

Table 3. Model parameters and their ranges used during the Genetic Algorithm based automatic calibration procedure.

### 2.5 Regionalization of model parameters based on catchment similarity analysis

In this study, the hydrological similarity was defined based on four similarity measures: drainage area, spatial proximity, catchment characteristics and flow duration curve (FDC). Once the similarity was established among 11 gauged catchments, the best parameter set of one catchment was transferred to another catchment (temporarily considered as ungauged, termed as *pseudo ungauged*) for streamflow simulations. The whole parameter set was adopted from a donor catchment. The main advantage of adopting complete parameter set is that the parameter
interdependencies are not neglected. The results were then compared, in terms of NSE, $R^2$ and $VB$, by using the observed streamflow time series of the *pseudo ungauged* catchment.

In terms of similarity in area, each of the 11 catchments was compared with other catchments and was rendered similar to the one which had the closest drainage area. Similarly for spatial proximity, the two catchments located nearest to each other were defined as similar. In cases, where more than one catchments were available in the neighbourhood, the catchment with least distance from the centroid and/or having the longest common boundary was considered the most similar one. The similarity based on catchment characteristics was defined comparing the climate (ratio of mean annual precipitation and reference evapotranspiration), topography (average catchment slope, elevation and stream density), land use (area under forest and crop land), soil (area under rock outcrop type soils) and geology (area under limestone dominated geology). These characteristics are generally considered as the major drivers of the hydrological processes and catchment runoff response (Nathan and McMahon, 1990; Wagener et al., 2007). The similarity index ($S$) was calculated by using equation 7 and the variables given in Table 4.

$$S = 1 - \sum_{i=1}^{M} \alpha_i \frac{\Delta V_i}{\max(\Delta V_i, \bar{V})}$$

(7)

Where, $S$ is the similarity index (-) which takes a value between 0 and 1 and defines the degree to which catchment 1 is similar to catchment 2, $M$ is the number of catchment characteristics (variables) used for computing the similarity index. The $\alpha_i$ are the weights (-) between 0 and 1 for the given characteristics such that sum of the weights is equal to 1. In this study, equal weights are used for all the characteristics. The variables $V$, $\Delta V$, and $\bar{V}$ refer to the value of the respective catchment characteristics, the absolute difference between catchment 1 and 2, and the average value of catchment 1 and 2, respectively.
Table 4. Catchment characteristics used in calculating the similarity index.

In the fourth approach, similarity in the FDCs was compared both by means of visual inspection and by using a statistical criterion, Relative Root Mean Square Error (RRMSE). FDCs are very useful for comparing the hydrological response of catchments (e.g. Linsley et al., 1949; Hughes and Smakhtin, 1996; Yilmaz et al., 2008). Their shape is an indicator of catchment response to rainfall and also depicts the storage characteristics of the catchments and influence of topography, geology, vegetative cover and land use. In this study, the FDCs were plotted using daily discharge data which were normalized by the drainage area in order to facilitate comparison. The shape of the FDC for each catchment was visually compared with the FDCs of the other catchments; the catchments showing best match for both high and low flow percentiles were considered hydrologically similar. A commonly used objective criterion based on the Relative Root Mean Square Error (RRMSE), termed here as $\varepsilon (-)$, equation 8, was applied in this study to facilitate defining the similarity between the FDCs.

$$\varepsilon = \sqrt{\frac{1}{N} \sum_{i} (Q_i - \hat{Q})^2}$$

(8)

Where $Q_i$ is the $i$-th flow percentile (mm/d) of one FDC and $i$ ranges from 1 to $N$; $\hat{Q}$ is the corresponding $i$-th flow percentile (mm/d) of another FDC; and $\bar{Q}$ is the mean discharge of the first (base line) FDC. The $\varepsilon$ values were calculated for the whole FDC corresponding to the flow percentiles $Q_o$ to $Q_{100}$ using daily discharge data.

2.6 Assessment of the impact of parameter uncertainty on the regionalization results

The issue of parameter uncertainty is well recognized in hydrological modelling (Uhlenbrook et al., 1999; Beven 2001; Wagener et al., 2004; McIntyre et al., 2005). Generally, parameter values are
not unique and cause large uncertainty bands in the discharge predictions. Furthermore, similar
model simulations can be achieved by using different combinations of parameter values, which is
generally termed in hydrology as equifinality or non-uniqueness of the model parameters (Beven
2001). In this study, the impact of parameter uncertainty on the regionalization results was also
investigated. First, the best parameter set of a study catchment in the regionalization procedure was
used, as indicated in section 2.5. Then to check the consistency of the results, we selected 50
different parameter sets of a catchment that yielded in the highest NSE values during the automatic
calibration process, and used them for the regionalization in a similar way as of using the single
best parameter set. As mentioned in section 2.4, the automatic calibration was based on GA based
optimization procedure. Therefore, the 50 best parameter sets are the ones resulting in highest NSE
out of the many good parameter sets that GA based optimization method generates. The
regionalization results were considered reliable given the results remain consistent in terms of
studied performance indicators (NSE, $R^2$ and VB) while using different parameter sets (e.g. both in
case of the best parameter set and the 50 other good parameter sets).

3 Results and discussion

3.1 Model results of automatic parameter estimation

The calibration results showing the comparison of observed and simulated streamflows are
provided in Table 5, summarizing the daily NSE, $R^2$ and VB estimates. The NSE values were quite
good for most of the catchments (i.e., >0.6), with the exception of two catchments indicating values
in the range of 0.41-0.46. Similar patterns were indicated by $R^2$ and VB, depicting reasonably good
model performance in most cases. Although, during validation period, NSE and $R^2$ values were
lower as compared to their corresponding values during calibration period, the values were
reasonably good in most cases (i.e., $NSE > 0.5$). Furthermore, the performance results obtained in
this study are in good agreement with those of other model regionalisation studies (e.g. Abdulla and
Lettenmaier 1997; Merz and Blöschl 2004).

The calibration and validation results suggest that the optimized parameter sets could
simulate the rainfall-runoff relationships reasonably well in most cases. However, it should be
noted that the models are not perfect and may involve uncertainties resulting from uncertainties in
the model structure, input data and parameter values (which is further discussed in section 3.5).
Therefore, the results should be interpreted cautiously. For example, in the case of the Sange
Sorakh (ID: 3) catchment the low performance was attributed to the high influence of groundwater
discharge of a spring which the model was not able to simulate well given the high uncertainties in
locating the boundaries of the karstified recharge area. The low performance of the Afarineh (ID: 8)
could be mainly attributed to possibly high uncertainty in the climatic input data for this particular
catchment due to less density of the climatic gauges in this area. In this catchment, the model
consistently overestimated the average flows resulting in a high volume error and underestimated
the high flood peaks.

Table 5. HBV model calibration and validation results, showing daily Nash-Sutcliffe
efficiency ($NSE$), daily coefficient of determination ($R^2$) and annual volume balance ($VB$)

3.2 Regionalization results based on drainage area, spatial proximity and catchment
characteristics

The summary of the catchment similarity analysis is presented in Table 6, indicating most
similar catchment whose parameters were transferred for the regionalization purpose under each of
the four tested methods. The regionalization results for the calibration period are presented in
Figure 3. The results of transferring the model parameters based on similarity in area show that in
most cases the simulations were far away from the observed values in terms of NSE, $R^2$ and VB, with the exception of Kaka Raza (ID: 11) where the results were reasonably good. The regionalization based on spatial proximity showed much better simulations compared to those based on drainage area. Promising results were obtained for four catchments, namely, Aran (ID: 1), Firoz Abad (ID: 2), Doabe Merek (ID: 4) and Sarab Seidali (ID: 10), with NSE in the range of 0.51 to 0.78. But a large number of catchments resulted in poor simulations, i.e., four catchments had negative NSE values (ranging from -3.4 to -0.10). Similar to drainage area and spatial proximity, the regionalization results based on catchment characteristics were not better in most cases (Figure 3). Four out of 11 catchments produced comparatively better results with NSE and $R^2$ values in the range of 0.24 to 0.64 and 0.69 to 0.77, respectively. Rest of the catchments yielded poor results, particularly in terms of VB and NSE. On the whole, the results suggest that the above mentioned regionalization approaches are likely to produce unacceptable results in most cases. Therefore, none of them could be recommended for the regionalization purposes in the study region.

Table 6. Results of the catchment similarity analysis for the four tested methods.

Figure 3. Regionalization results of the four tested methods. (The used catchment numbers in x-axis correspond to the names as follow: 1: Aran; 2: Firoz Abad; 3: Sange Sorakh; 4: Doabe Merek; 5: Khers Abad; 6: Noor Abad; 7: Dartoot; 8: Afarineh; 9: Cham Injeer; 10: Sarab Seidali; 11: Kaka Raza)

3.3 Regionalization results based on FDC

The FDC plots for all the study catchments are shown in Figure 4 and their similarities in terms of RRMSE ($\varepsilon$) are given in Table 6. In general, visual comparison and the used objective criteria
indicated good correspondence with each other. Both visual comparison and $\epsilon$ values indicate that 7
out of 11 studied catchments revealed good similarity with at least one catchment in the study
group. The $\epsilon$ values in these 7 cases range from 0.25 to 0.61. The FDC based regionalization results
for these catchments were reasonably good, with 5 out 7 catchments resulted in the $NSE$ values in
the range of 0.23 to 0.78 (Figure 3). The $R^2$ values were also good ranging from 0.54 to 0.87.
Similarly, most of them depicted reasonably good performance in terms of $VB$. For instance, only 2
out of these 7 catchments produced, negative $NSE$ values, but still could simulate reasonably well
annual yields, i.e., $VB$ for Sange Sorakh (ID: 3) and Noor Abad (ID: 6) was 1% and 24%,
respectively. It is important to note that the Sange Sorakh catchment yielded lower $NSE$, $R^2$ values
even during calibration. The lower performance, both during calibration, validation and
regionalization could be attributed to the significant contribution from a perennial spring, which the
model was not able to simulate well given the high uncertainties in locating the geographical
boundaries of the recharge area and the complexities in the hydrological processes in this region.

The FDCs of the remaining 4 catchments were not very similar to rest of the study
catchments. However, for the purpose of keeping consistency in the number of catchments used in
all of the tested regionalization methods, we also executed FDC based regionalization for these
catchments by transferring the parameters from the catchment having the least value of $\epsilon$. As
expected, the results were not very good when compared to those catchments where similarity was
adequately defined. Nevertheless, the outcome was comparable to the other three methods.

Furthermore, it is pertinent to note that, in most cases, the good regionalization results in
case of tested methods other than FDC based method correspond to the pair of catchments having
quite similar FDCs. For example, 3 out of 4 good performing catchments in case of spatial
proximity (e.g. Aran, Firoz Abad and Sarab Seidali) also depicted similarity in the FDC of the

619

Figure 4. Comparison of FDCs for the similarity analysis.
3.4 Impact of parameter uncertainty on the regionalization results

The summary of the regionalization results using 50 best parameter sets for the FDC based regionalization method is presented in Table 7 and Figure 5. The resulting statistics given in Table 7 are reported in terms of median, 25th and 75th percentile, minimum and maximum. The presented statistics were obtained by arranging the results in descending order and then calculating various exceeding percentile in a similar way as of well known flow duration analysis. This analysis helped to quickly view the degree of consistency when different parameter sets were used in the regionalization. For instance, if the range of different percentiles is small, then the impact of parameter uncertainty could be considered negligible. The results reveal that, despite different parameter sets, the regionalization results were reasonably consistent. This suggests that parameter uncertainty did not have considerable impact on the regionalization outcome. For example, maximum NSE values, which were achieved using the best parameter sets (as discussed in the previous sections 3.2 and 3.3) were not markedly different in most cases. This is further supported by the fact that the good performing catchments continue to perform well for all of the 50 tested parameter sets (Table 7 and Figure 5). Moreover, none of the low performing catchments showed significant improvement as result of using different parameter sets. The similar inferences were drawn regarding impact of parameter uncertainty on the regionalization results of the other three tested method (not shown here).

Table 7. Impact of parameter uncertainty on regionalization results, illustrated by the Nash-Sutcliffe efficiency (NSE) and coefficient of determination (R²) results achieved for the 50 parameter sets used for the FDC based regionalization method.

Figure 5. Impact of parameter uncertainty on regionalization results, illustrated by the
exceeding percentiles of Nash-Sutcliffe efficiency ($NSE$) obtained from the 50 parameter sets used during regionalization based on similarity in the FDC.

3.5 **Comparison of the FDC based regionalization results with other studies**

The results of this study indicate that the regionalization based on the similarity of the FDC perform superior compared to other three tested methods. Although, we could not test more methods due to limitations of the available data, we compared our findings with related studies conducted elsewhere using other methods. The comparison was made between the results of the FDC based regionalization (Figure 3) with the results of the studies presented in Table 1. The main aim of this comparison is to have an overview of the comparative position of the proposed FDC based regionalization method among other widely recommended regionalization methods. Moreover, this comparison can not replace a rigorous comparative assessment and is recommended as a future research activity. Therefore, it is acknowledged that this comparison should be interpreted cautiously because of inherent differences in the studies i.e., differences in the amount and quality of the used data sets and varying hydro-climatic environments, among others.

The comparison reveals that the FDC based regionalization approach stands very well among the most promising techniques developed elsewhere. For instance, the regionalization results based on the estimation of model parameters using catchment characteristics, indicated variable degree of success, as demonstrated by the wide range of calculated performance measures (Table 1). The reported daily $NSE$ values were in the range of 0.02 to 0.45 and 0.23 to 0.72 for the parameter regionalization studies of Servet and Desetter (1993) and Seibert (1999), respectively. Similarly, the studies of Servet and Dezetter (1993) and Abdulla and Lettenmaier (1997) reported $R^2$ values in the range of 0.62 to 0.99 and 0.05 to 0.81, respectively. A similar trend of variable performance can be seen in many methods other than parameter regionalization, e.g. Merz and
Blöschl (2004) achieved median $NSE$ values in the range of 0.32 to 0.56 for their 8 regionalization methods tested for the 308 catchments and Goswani et al. (2007) indicated $NSE$ values in the range -27.66 to 0.94 for their regional pooling method. The reported FDC based regionalization results of this study for 5 out of 7 catchments (where FDC similarity was well established) were in the range of 0.54 to 0.87 in terms of daily $R^2$ values and 0.23 to 0.78 in terms of daily $NSE$ values. These encouraging results suggest that model regionalization based on the FDC similarity is a very good addition to the available regionalization methods.

However, it should be noted that all of the tested methods, including the FDC based regionalization, resulted in some cases where the performance was not good. This suggests that the problem of achieving successful outcomes for all model applications for the poorly gauged or ungauged catchments still remains a challenging undertaking, thus needs further research in the future. This could be attributed to the nature of the problem at hand, as hydrology is a context dependent science and involves high degree of variability in the hydrological processes among different catchments. Therefore, supporting the regionalization results through other sources of data and qualitative information is extremely desirable to avoid erroneous results. Nonetheless, the chances of invalid results drawn by applying the FDC based regionalization method to poorly gauged catchments are likely to be small, because at least some estimates of the streamflows characteristics are available for comparison in such cases (e.g. mean annual and monthly flows; various exceeding percentiles).

4 Conclusions

This study examined the application of the HBV model for streamflow time series generation in data limited catchments of the Karkheh river basin, Iran, using model parameters transferred from similar gauged catchments. The similarities of the catchments for model parameter transfer were
determined based on drainage area, spatial proximity, catchment characteristics and flow duration
curve (FDC). Although the streamflow validation results based on spatial proximity and catchment
characteristics are better than those based on geographical area, the overall results remain
unsatisfactory in most cases. The study has shown that catchment similarity analysis based on
FDCs provides a sound basis for transferring model parameters from gauged catchments to data
limited catchments in the Karkheh basin. In most cases, the simulated time series of streamflows
resulted in reasonably good values of the examined performance indicators (i.e., NSE, R² and VB)
with negligible impact of the parameter uncertainty on the regionalization outcome. Furthermore,
this new method also compares well with the studies conducted elsewhere using other promising
methods. These demonstrations suggest that the new FDC based regionalization method is a
valuable addition to the available regionalization methods and could be recommended for the
practical applications for estimating time series of streamflows for the poorly gauged catchments in
the mountainous parts of the Karkheh river basin, Iran. However, it is important to recognize that
the poor performance in some cases for the promising regionalization methods indicate the
complexity of the hydrological issues and of the regionalization problem and clearly highlights the
scope of further improvements. This essentially requires more efforts on better understanding the
hydrology of ungauged or poorly gauged catchments and further developments in the
regionalization procedures, in particular with regard to widely testing and improving existing
methods, finding new regionalization approaches and exploring innovative ways use of available
(scarce) data sets.

The methodology presented in this paper is easy to replicate in other river basins of the
world. Moreover, it can work well in the river basins, like the Karkheh basin of Iran, facing a
decline in streamflow monitoring networks and/or have limited number of gauged catchments.
Further testing of the proposed FDC based regionalization method is highly recommended, i.e., by
using different rainfall-runoff models, application under different hydro-climatic conditions, and for
different extents of water resources development in the catchments (e.g. from more pristine to more
regulated catchments).

Acknowledgements

Funds were provided through the Basin Focal Project for the Karkheh basin of the Challenge Program on Water and Food (CPWF) and from the capacity building program of the International Water Management Institute (IWMI). We are thankful to the Ministry of Energy, Iran, and to the Meteorological Organization of Iran for providing necessary data sets. Thanks are also due to Agricultural Research and Extension Organization of Iran (AREO) and Soil Conservation and Watershed Management Research Institute (SCWMRI) for their support during data collection and field visits. We are thankful for the constructive discussion of initial results with the scientists from SCWMRI, Tehran, notably to the Prof. Bahram Saghatian. We are thankful for the help of colleagues at IWMI, Sri Lanka. Special thanks are due to Dr. Hugh Turrell and Dr. Vladimir Smakhtin for providing useful insights during the initial phase of this work and to Md. A. Islam for his support in the GIS analysis. The used HBV model was programmed and provided for this study by Prof. Jan Seibert (University of Zurich, Switzerland).

References


Kokkonen, T.S., Jakeman, A.J., Young, P.C., Koivusalo, H.J., 2003. Predicting daily flows in ungauged catchments: model regionalization from catchment descriptors at the Coweeta Hydrologic...


UNESCO-IHE Institute for Water Education, Delft, the Netherlands.


## Tables

### Table 1. An overview of some studies related to regionalization of conceptual rainfall-runoff models

<table>
<thead>
<tr>
<th>Country</th>
<th>Catchments</th>
<th>Drainage area, km²</th>
<th>Simulation time step</th>
<th>Evaluation catchments</th>
<th>Volume Balance, VB, mm/a</th>
<th>Coefficient of determination, ( R^2 ) (-)</th>
<th>Nash-Sutcliff Efficiency, NSE, (-)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>16 (5)</td>
<td>0.04 to 12 (0.02 to 10)</td>
<td>Hourly</td>
<td>NA (-372 to 155)</td>
<td>NA</td>
<td>NA</td>
<td>Magette et al., 1976</td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>20 (4)</td>
<td>16 to 2163 (73 to 148)</td>
<td>Monthly</td>
<td>-8 to 12 (-29 to 54)</td>
<td>NA</td>
<td>NA</td>
<td>Vandewiele et al., 1991</td>
<td></td>
</tr>
<tr>
<td>Ivory</td>
<td>11 (5)</td>
<td>100 to 4500</td>
<td>Daily</td>
<td>NA</td>
<td>0.23 to 0.99 (0.62 to 0.99)</td>
<td>0.69 to 0.94 (0.62 to 0.89)</td>
<td>Servat and Dezetter, 1993</td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td>18 (8)</td>
<td>10 to 1870 (156 to 1792)</td>
<td>Monthly</td>
<td>NA (-1 to 4)</td>
<td>(0.67 to 0.76)</td>
<td>Ibrahim and Cordery, 1995</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>34 (40)</td>
<td>168 to 5226 (442 to 6894)</td>
<td>Daily</td>
<td>NA (-11 to 134)</td>
<td>(0.05 to 0.81)</td>
<td>Abdulla and Lettenmaier, 1997</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sweden</td>
<td>11 (7)</td>
<td>7 to 950 (7 to 1284)</td>
<td>Daily</td>
<td>NA</td>
<td>0.70 to 0.88 (0.23 to 0.72)</td>
<td>NA</td>
<td>Seibert, 1999</td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>308 (308)</td>
<td>3 to 5000 (3 to 5000)</td>
<td>Daily</td>
<td>NA</td>
<td>0.67 (0.32 to 0.56)*</td>
<td>NA (-27.66 to 0.94)</td>
<td>Merz and Blöschl, 2004</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>12 (11)</td>
<td>32 to 371 (32 to 371)</td>
<td>Daily</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>Goswami et al., 2007</td>
<td></td>
</tr>
</tbody>
</table>

*Figures in parenthesis correspond to the test catchments.

*NA refers to information not available.

* Efficiency values refer to median of all 308 catchments during calibration phase and (in parenthesis) minimum and maximum median values of tested regionalization methods.

### Table 2. Salient features of the selected streamflow gauges.

<table>
<thead>
<tr>
<th>River Name</th>
<th>Station Name</th>
<th>Station ID</th>
<th>Long Lat</th>
<th>Elevation at the gauge, m asl</th>
<th>Drainage area, km²</th>
<th>Observed flow, mm/a</th>
<th>Naturalized flow, mm/a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Khorram Rod</td>
<td>Aran</td>
<td>1</td>
<td>47.92 34.42</td>
<td>1440</td>
<td>2320</td>
<td>59</td>
<td>87</td>
</tr>
<tr>
<td>Toyserkan</td>
<td>Firoz Abad</td>
<td>2</td>
<td>48.12 34.35</td>
<td>1450</td>
<td>844</td>
<td>55</td>
<td>102</td>
</tr>
<tr>
<td>Gamasaiib</td>
<td>Sange Sorakh</td>
<td>3</td>
<td>48.23 34.03</td>
<td>1800</td>
<td>475</td>
<td>254</td>
<td>294</td>
</tr>
<tr>
<td>Qarsou</td>
<td>Doab Merek</td>
<td>4</td>
<td>46.78 34.55</td>
<td>1310</td>
<td>1260</td>
<td>148</td>
<td>148</td>
</tr>
<tr>
<td>Ahe Mareg</td>
<td>Khers Abad</td>
<td>5</td>
<td>46.73 34.52</td>
<td>1320</td>
<td>1460</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>Bad Avar</td>
<td>Noor Abad</td>
<td>6</td>
<td>47.97 34.08</td>
<td>1780</td>
<td>590</td>
<td>202</td>
<td>315</td>
</tr>
<tr>
<td>Ace Chamar</td>
<td>Dartoot</td>
<td>7</td>
<td>46.40 35.45</td>
<td>1110</td>
<td>2522</td>
<td>71</td>
<td>95</td>
</tr>
<tr>
<td>Chalood</td>
<td>Afarineh</td>
<td>8</td>
<td>47.38 33.30</td>
<td>800</td>
<td>800</td>
<td>160</td>
<td>170</td>
</tr>
<tr>
<td>Khorramabad</td>
<td>Cham Injeer</td>
<td>9</td>
<td>48.23 33.45</td>
<td>1140</td>
<td>1590</td>
<td>223</td>
<td>341</td>
</tr>
<tr>
<td>Doab Aleshtar</td>
<td>Sarab Seidali</td>
<td>10</td>
<td>48.22 33.80</td>
<td>1520</td>
<td>776</td>
<td>345</td>
<td>516</td>
</tr>
<tr>
<td>Har Rod</td>
<td>Kaka Raza</td>
<td>11</td>
<td>48.27 33.72</td>
<td>1530</td>
<td>1130</td>
<td>355</td>
<td>428</td>
</tr>
</tbody>
</table>

* Data source: Ministry of Energy, Iran, with the exception of station ID and naturalized flow.

*m asl refers to meters above sea level.
879 Table 3. Model parameters and their ranges used during the Genetic Algorithm based
880 automatic calibration procedure.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Unit</th>
<th>Explanation</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Snow routine</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>°C</td>
<td>Threshold temperature</td>
<td>-2.5 to 2.5</td>
</tr>
<tr>
<td>CFMAX</td>
<td>mm °C⁻¹d⁻¹</td>
<td>Degree-day factor</td>
<td>1 to 6</td>
</tr>
<tr>
<td>SFCF</td>
<td>-</td>
<td>Snowfall correction factor</td>
<td>0.8 to 1.25</td>
</tr>
<tr>
<td>CFR</td>
<td>-</td>
<td>Refreezing coefficient</td>
<td>0.05 to 0.05</td>
</tr>
<tr>
<td>CWH</td>
<td>-</td>
<td>Water holding capacity</td>
<td>0.1 to 0.1</td>
</tr>
<tr>
<td><strong>Soil routine</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FC</td>
<td>mm</td>
<td>Maximum of SM (storage in soil box)</td>
<td>50 to 500</td>
</tr>
<tr>
<td>LP</td>
<td>-</td>
<td>Threshold for reduction of evaporation (SM/FC)</td>
<td>0.5 to 0.7</td>
</tr>
<tr>
<td>BETA</td>
<td>-</td>
<td>Shape coefficient for soil storage/percolation</td>
<td>1 to 6</td>
</tr>
<tr>
<td><strong>Response routine</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERC</td>
<td>mm d⁻¹</td>
<td>Maximal flow from upper to lower box</td>
<td>0.1 to 6</td>
</tr>
<tr>
<td>UZL</td>
<td>mm</td>
<td>Threshold for Q₀ outflow in upper box</td>
<td>10 to 100</td>
</tr>
<tr>
<td>K₀</td>
<td>d⁻¹</td>
<td>Recession coefficient (upper in upper box)</td>
<td>0.05 to 0.5</td>
</tr>
<tr>
<td>K₁</td>
<td>d⁻¹</td>
<td>Recession coefficient (lower in upper box)</td>
<td>0.01 to 0.15</td>
</tr>
<tr>
<td>K₂</td>
<td>d⁻¹</td>
<td>Recession coefficient (lower box)</td>
<td>0.001 to 0.05</td>
</tr>
<tr>
<td><strong>Routing routine</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAXBAS</td>
<td>d</td>
<td>Routing, length of weighting function</td>
<td>1 to 5</td>
</tr>
</tbody>
</table>

881

882 Table 4. Catchment characteristics used in calculating the similarity index.

<table>
<thead>
<tr>
<th>Catchment</th>
<th>Catchment characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>Name</td>
</tr>
<tr>
<td>-----------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td>Aran</td>
</tr>
<tr>
<td>2</td>
<td>Firoz Abad</td>
</tr>
<tr>
<td>3</td>
<td>Sange Sorakh</td>
</tr>
<tr>
<td>4</td>
<td>Doabe Merek</td>
</tr>
<tr>
<td>5</td>
<td>Khers Abad</td>
</tr>
<tr>
<td>6</td>
<td>Noor Abad</td>
</tr>
<tr>
<td>7</td>
<td>Dartoot</td>
</tr>
<tr>
<td>8</td>
<td>Afarineh</td>
</tr>
<tr>
<td>9</td>
<td>Cham Injerd</td>
</tr>
<tr>
<td>10</td>
<td>Sarab Seidal</td>
</tr>
<tr>
<td>11</td>
<td>Kaka Raza</td>
</tr>
</tbody>
</table>
Table 5. HBV model calibration and validation results, showing daily Nash-Sutcliffe efficiency ($NSE$), daily coefficient of determination ($R^2$) and annual volume balance ($VB$)

<table>
<thead>
<tr>
<th>Catchment ID</th>
<th>Name</th>
<th>Nash-Sutcliffe efficiency ($NSE, -$)</th>
<th>Coefficient of determination, $R^2$ (-)</th>
<th>Volume Balance ($VB$)</th>
<th>Difference, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Observed, mm/a</td>
<td>Simulated, mm/a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calibration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Aran</td>
<td>0.91</td>
<td>0.91</td>
<td>95</td>
<td>90</td>
</tr>
<tr>
<td>2</td>
<td>Firoz Abad</td>
<td>0.76</td>
<td>0.78</td>
<td>118</td>
<td>104</td>
</tr>
<tr>
<td>3</td>
<td>Sange Sorakh</td>
<td>0.46</td>
<td>0.46</td>
<td>332</td>
<td>332</td>
</tr>
<tr>
<td>4</td>
<td>Doabe Merek</td>
<td>0.88</td>
<td>0.89</td>
<td>171</td>
<td>148</td>
</tr>
<tr>
<td>5</td>
<td>Kkers Abad</td>
<td>0.66</td>
<td>0.67</td>
<td>39</td>
<td>39</td>
</tr>
<tr>
<td>6</td>
<td>Noor Abad</td>
<td>0.64</td>
<td>0.70</td>
<td>349</td>
<td>326</td>
</tr>
<tr>
<td>7</td>
<td>Dartoot</td>
<td>0.80</td>
<td>0.81</td>
<td>95</td>
<td>111</td>
</tr>
<tr>
<td>8</td>
<td>Afarineh</td>
<td>0.41</td>
<td>0.48</td>
<td>196</td>
<td>294</td>
</tr>
<tr>
<td>9</td>
<td>Cham Injeer</td>
<td>0.80</td>
<td>0.80</td>
<td>367</td>
<td>349</td>
</tr>
<tr>
<td>10</td>
<td>Sarab Seidali</td>
<td>0.73</td>
<td>0.76</td>
<td>560</td>
<td>498</td>
</tr>
<tr>
<td>11</td>
<td>Kaka Raza</td>
<td>0.83</td>
<td>0.84</td>
<td>483</td>
<td>405</td>
</tr>
<tr>
<td>Validation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Aran</td>
<td>0.67</td>
<td>0.81</td>
<td>79</td>
<td>95</td>
</tr>
<tr>
<td>2</td>
<td>Firoz Abad</td>
<td>0.45</td>
<td>0.64</td>
<td>85</td>
<td>94</td>
</tr>
<tr>
<td>3</td>
<td>Sange Sorakh</td>
<td>0.56</td>
<td>0.71</td>
<td>271</td>
<td>238</td>
</tr>
<tr>
<td>4</td>
<td>Doabe Merek</td>
<td>0.66</td>
<td>0.69</td>
<td>129</td>
<td>96</td>
</tr>
<tr>
<td>5</td>
<td>Kkers Abad</td>
<td>0.68</td>
<td>0.69</td>
<td>30</td>
<td>37</td>
</tr>
<tr>
<td>6</td>
<td>Noor Abad</td>
<td>0.44</td>
<td>0.57</td>
<td>279</td>
<td>309</td>
</tr>
<tr>
<td>7</td>
<td>Dartoot</td>
<td>0.25</td>
<td>0.46</td>
<td>94</td>
<td>106</td>
</tr>
<tr>
<td>8</td>
<td>Afarineh</td>
<td>0.11</td>
<td>0.58</td>
<td>144</td>
<td>298</td>
</tr>
<tr>
<td>9</td>
<td>Cham Injeer</td>
<td>0.56</td>
<td>0.66</td>
<td>315</td>
<td>331</td>
</tr>
<tr>
<td>10</td>
<td>Sarab Seidali</td>
<td>0.59</td>
<td>0.68</td>
<td>471</td>
<td>450</td>
</tr>
<tr>
<td>11</td>
<td>Kaka Raza</td>
<td>0.75</td>
<td>0.77</td>
<td>371</td>
<td>370</td>
</tr>
</tbody>
</table>

*a* Dartoot and Sange Sorakh had missing streamflow data. For Dartoot the calibration and validation results refer to the periods October 1, 1994 to September 30, 2001 and October 1, 1990 to September 30, 1992, respectively. For Sange Sorakh the calibration and validation results refer to the periods October 1, 1987 to September 30, 1994 and October 1, 1999 to September 30, 2001, respectively. For all other catchments the calibration and validation periods refer to October 1, 1987 to September 30, 1994 and October 1, 1994 to September 30, 2001, respectively.
Table 6. Results of the catchment similarity analysis for the four tested methods.

<table>
<thead>
<tr>
<th>Catchment</th>
<th>Catchment similarity based on the studied methods</th>
<th>drainage area</th>
<th>Spatial proximity</th>
<th>Similarity index</th>
<th>Flow duration curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>Name</td>
<td>Similar catchment</td>
<td>Similar catchment</td>
<td>Similar catchment</td>
<td>Value of S</td>
</tr>
<tr>
<td>1 Aran</td>
<td>Dartoot</td>
<td>Firoz Abad</td>
<td>Noor Abad</td>
<td>0.85</td>
<td>Firoze Abad</td>
</tr>
<tr>
<td>2 Firoz Abad</td>
<td>Afarineh</td>
<td>Aran</td>
<td>Aran</td>
<td>0.82</td>
<td>Aran</td>
</tr>
<tr>
<td>3 Sange Sorakh</td>
<td>Noor Abad</td>
<td>Sarab Seidali</td>
<td>Kaka Raza</td>
<td>0.70</td>
<td>Cham Injeer</td>
</tr>
<tr>
<td>4 Doabe Merek</td>
<td>Kaka Raza</td>
<td>Khers Abad</td>
<td>Noor Abad</td>
<td>0.81</td>
<td>Firoze Abad</td>
</tr>
<tr>
<td>5 Khers Abad</td>
<td>Cham Injeer</td>
<td>Dartoot</td>
<td>Doabe Merek</td>
<td>0.75</td>
<td>Aran</td>
</tr>
<tr>
<td>6 Noor Abad</td>
<td>Sange Sorakh</td>
<td>Sarab Seidali</td>
<td>Aran</td>
<td>0.85</td>
<td>Cham Injeer</td>
</tr>
<tr>
<td>7 Dartoot</td>
<td>Aran</td>
<td>Khers Abad</td>
<td>Cham Injeer</td>
<td>0.79</td>
<td>Aran</td>
</tr>
<tr>
<td>8 Afarineh</td>
<td>Sarab Seidali</td>
<td>Cham Injeer</td>
<td>Cham Injeer</td>
<td>0.67</td>
<td>Doabe Merek</td>
</tr>
<tr>
<td>9 Cham Injeer</td>
<td>Khers Abad</td>
<td>Kaka Raza</td>
<td>Dartoot</td>
<td>0.79</td>
<td>Noor Abad</td>
</tr>
<tr>
<td>10 Sarab Seidali</td>
<td>Afarineh</td>
<td>Noor Abad</td>
<td>Kaka Raza</td>
<td>0.83</td>
<td>Cham Injeer</td>
</tr>
<tr>
<td>11 Kaka Raza</td>
<td>Doabe Merek</td>
<td>Cham Injeer</td>
<td>Sarab Seidali</td>
<td>0.83</td>
<td>Sarab Seidali</td>
</tr>
</tbody>
</table>

Table 7. Impact of parameter uncertainty on regionalization results, illustrated by the Nash-Sutcliffe efficiency (NSE) and coefficient of determination (R²) results achieved for the 50 parameter sets used for the FDC based regionalization method.

<table>
<thead>
<tr>
<th>Catchment</th>
<th>Median</th>
<th>P25</th>
<th>P75</th>
<th>Min</th>
<th>Max</th>
<th>Median</th>
<th>P25</th>
<th>P75</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID 1 Aran</td>
<td>0.79</td>
<td>0.80</td>
<td>0.79</td>
<td>0.78</td>
<td>0.80</td>
<td>0.86</td>
<td>0.86</td>
<td>0.84</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>2 Firoz Abad</td>
<td>0.66</td>
<td>0.67</td>
<td>0.65</td>
<td>0.57</td>
<td>0.69</td>
<td>0.73</td>
<td>0.74</td>
<td>0.73</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>3 Sange Sorakh</td>
<td>-0.94</td>
<td>-0.80</td>
<td>-1.33</td>
<td>-1.77</td>
<td>-0.41</td>
<td>0.25</td>
<td>0.27</td>
<td>0.25</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>4 Doabe Merek</td>
<td>0.36</td>
<td>0.39</td>
<td>0.35</td>
<td>0.24</td>
<td>0.45</td>
<td>0.79</td>
<td>0.79</td>
<td>0.79</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>5 Khers Abad</td>
<td>0.26</td>
<td>0.44</td>
<td>0.11</td>
<td>-0.05</td>
<td>0.47</td>
<td>0.54</td>
<td>0.54</td>
<td>0.53</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>6 Noor Abad</td>
<td>-0.59</td>
<td>-0.48</td>
<td>-0.64</td>
<td>-0.88</td>
<td>-0.18</td>
<td>0.26</td>
<td>0.26</td>
<td>0.25</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>7 Dartoot</td>
<td>0.03</td>
<td>0.11</td>
<td>0.00</td>
<td>-0.11</td>
<td>0.26</td>
<td>0.27</td>
<td>0.28</td>
<td>0.26</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>8 Afarineh</td>
<td>-1.23</td>
<td>-1.22</td>
<td>-1.24</td>
<td>-1.31</td>
<td>-1.11</td>
<td>0.11</td>
<td>0.12</td>
<td>0.11</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>9 Cham Injeer</td>
<td>0.39</td>
<td>0.41</td>
<td>0.37</td>
<td>0.29</td>
<td>0.57</td>
<td>0.59</td>
<td>0.60</td>
<td>0.59</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>10 Sarab Seidali</td>
<td>0.26</td>
<td>0.28</td>
<td>0.23</td>
<td>0.16</td>
<td>0.31</td>
<td>0.61</td>
<td>0.62</td>
<td>0.59</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>11 Kaka Raza</td>
<td>0.59</td>
<td>0.60</td>
<td>0.58</td>
<td>0.56</td>
<td>0.62</td>
<td>0.73</td>
<td>0.74</td>
<td>0.73</td>
<td>0.71</td>
<td></td>
</tr>
</tbody>
</table>

37
Figures
Figure 1. Salient features of the study area and location of the study catchments and used climatic stations.

Figure 2. Naturalized and observed daily time series of streamflows of Aran catchment.
Figure 3. Regionalization results of the four tested methods. (The used catchment numbers in x-axis correspond to the names as follow: 1: Aran; 2: Firoz Abad; 3: Sange Sorakh; 4: Doabe Merek; 5: Khers Abad; 6: Noor Abad; 7: Dartoot; 8: Afarineh; 9: Cham Injeer; 10: Sarab Seidali; 11: Kaka Raza)
Figure 4. Comparison of FDCs for the similarity analysis.

Figure 5. Impact of parameter uncertainty on regionalization results, illustrated by the exceeding percentiles of Nash-Sutcliffe efficiency (NSE) obtained from the 50 parameter sets used during regionalization based on similarity in the FDC.