Evaluation of different parameterizations of the spatial heterogeneity of subsurface storage capacity for hourly runoff simulation in boreal mountainous watershed

Teklu T. Hailegeorgis*[1], Knut Alfredsen[1], Yisak S. Abdella[2] and Sjur Kolberg[2]

[1]Department of Hydraulic and Environmental Engineering, Norwegian University of Science and Technology (NTNU), NO-7491 Trondheim.

[2]SINTEF Energi AS, Sem Sælands vei 11, NO-7465 Trondheim.

***Corresponding author information:**

Name: Teklu Tesfaye Hailegeorgis

Address: S. P. Andersens veg 5, Department of Hydraulic and Environmental Engineering, NTNU, NO-7491 Trondheim, Norway

Tel. (office): [+47] 73592411

Cell phone: [+47] 45069384

Fax: [+47] 735 91298

E-mail address: tekhi09@gmail.com

Affiliations:

Researcher at Norwegian University of Science and Technology (NTNU), Trondheim (Norway) and funded by CEDREN (Centre for Environmental Design of Renewable Energy).

Keywords:

Parameterization; Spatial heterogeneity; Subsurface storage capacity; Semi-distributed and distributed; Calibration and evaluation; boreal mountainous watershed.

Abstract

Identification of proper parameterizations of spatial heterogeneity is required for precipitation-runoff models. However, relevant studies with a specific aim at hourly runoff simulation in boreal mountainous catchments are not common.

We conducted calibration and evaluation of hourly runoff simulation in a boreal mountainous watershed based on six different parameterizations of the spatial heterogeneity of subsurface storage capacity for a semi-distributed (subcatchments hereafter called elements) and distributed (1x1 km^2 grid) setup. We evaluated representation of element-to-element, grid-to-grid, and probabilistic subcatchment/subbasin, subelement and subgrid heterogeneities.

The parameterization cases satisfactorily reproduced the streamflow hydrographs with Nash-Sutcliffe efficiency values for the calibration and validation periods up to 0.84 and 0.86 respectively, and similarly for the log-transformed streamflow up to 0.85 and 0.90. The parameterizations reproduced the flow duration curves, but predictive reliability in terms of quantile-quantile (Q-Q) plots indicated marked over and under predictions. The simple and parsimonious parameterizations with no subelement or no subgrid heterogeneities provided equivalent simulation performance compared to the more complex cases. The results indicated that (i) identification of parameterizations require measurements from denser precipitation stations than what is required for acceptable calibration of the precipitation-streamflow relationships, (ii) there is challenges in the identification of parameterizations based on only calibration to catchment integrated streamflow observations and (iii) a potential preference for the simple and parsimonious parameterizations for operational forecast contingent on their equivalent simulation performance for the available input data. In addition, the effects of non-identifiability of parameters (interactions and equifinality) can contribute to the non-identifiability of the parameterizations.

Introduction

Heterogeneities across spatial scales require either explicit resolving or proper parameterization procedures, which are prevailing challenges in catchment scale precipitation-runoff modelling. Previous studies such as Myrabø (1986; 1997), Gottschalk et al. (2001), Singh et al. (2002), Smith et al. (2004) and Bogaard et al. (2005) noted growing opportunities for prediction purposes of distributed precipitation-runoff modelling, which allow for better representation of the spatial heterogeneity in climate forcing, catchment characteristics, runoff responses and state variables. These opportunities include advances in measurement techniques of input variables such as precipitation from weather radar and remotely sensed snow accumulation, and development of parameter calibration algorithms for parameter identification in distributed hydrological models. However, a thorough diagnostic evaluation of the behavior of the prediction models is indispensable since the quality of real-time forecast is dependent on the process simulation (e.g. Bell and Moore, 1998; Refsgaard, 1997).

One of the main challenges related to predictions based on distributed precipitation-runoff models is the sensitivity of the results to the degree of the spatial resolution of inputs and the computational units used to address the spatial heterogeneity. The heterogeneities to be modelled may include those of model parameters, climate forcing, land surface characteristics, storage capacity of the soils and runoff delay (travel lag). Various discretization techniques are employed in precipitation-runoff models for the representation of the spatial heterogeneities. Catchments are usually discretized into a number of units based on various catchment characteristics governing the hydrological processes namely hydrological response units (HRUs) (e.g. Leavesley and Stannard, 1990), topographic wetness index (Beven and Kirby, 1979), topographic drainage divide based subcatchments (e.g. Sivapalan and Viney, 1994) hereafter called elements, hillslopes (e.g. Goodrich, 1990) and grid squares (e.g. Abott et al., 1986) depending on the objectives of the study and the availability of data. Internal heterogeneities within the catchments or within units (e.g. elements, hillslopes, HRUs or grids) can further be parameterized by probability distribution functions (e.g. Moore, 1985). Aggregation of inputs and state variables (e.g. based on simple averaging) are also common in catchment modelling (see a review by Blöschl and Sivapalan, 1995).

There are challenges related to parameterizations and scales for boreal mountainous regions. Halldin et al. (1999) noted for northern (boreal) catchments with distinct topographic features that small-scale phenomena influence the exchange processes between the land surface and the atmosphere and the lateral distribution of water through subsurface and surface flows. The spatial observation scale of the input climate forcing is usually coarse (low resolution) from sparse hydrometeorological stations compared to a fine (high) resolution discretization that may be required to represent the underlying heterogeneity explicitly or probabilistically. In addition, there are scale mismatches between the spatial heterogeneities of climate forcing and topographic controls (e.g. the fine scale topographic driven spatial heterogeneity is dominating the grid-to-grid variability of the low intensity precipitation).

Therefore, for a reliable prediction augmented by sensitivity analysis and hence insights in to the dominant hydrological processes, it is indispensable to investigate the effects of heterogeneities at different spatial scales (i.e. subcatchment/subbasin, subelement, subgrid, element-to-element and grid-to-grid) on the simulation of runoff responses. The subcatchment, subelement and subgrid scale runoff parameterization may also enhance our understanding of saturation excess runoff generation and it allows for validation of models against spatial observations.

Several different probability distribution function based models (PDM) are described in literature with the aim to reduce the complexity of the ‘fully’ distributed precipitation-runoff models by parameterizing the spatial heterogeneity of for instance the subsurface storage and infiltration capacity by a probability distribution to model the dynamics of runoff contributing areas. Examples of such models include the Hydrologiska Byråns Vattenballansavdelning (HBV) model (Bergström, 1976), the Xinanjiang model (Zhao et al., 1980; Zhao, 1992), the Probability distributed model or PDM (Moore and Clarke, 1981; Moore, 1985), the ARNO model (Todini, 1988; 1996), the variable infiltration capacity or VIC (Wood et al, 1992), the Improved Arno model (Hagemann and Gates, 2003), and the TOPMODEL (Beven and Kirby, 1979). Bell and Moore (1998), and Cole and Moore (2009) further demonstrated the performances of a grid based PDM variants based on both rain gauge and radar precipitation data.

The main research question lies in whether it is possible to identify different parameterization approaches for representation of the spatial heterogeneity of the subsurface storage capacity for hourly runoff simulation in a boreal mountainous watershed. The approaches range from explicit representation of element-to-element and grid-to-grid heterogeneities to probabilistic parameterization of subcatchment, subelement and subgrid heterogeneities. The main objective of the present study is to investigate performances of six different parameterizations of the spatial heterogeneity of the subsurface storage capacity for semi-distributed (elements) and gridded (1x1 km^2) cases for prediction of hourly streamflow. We calibrated the routines for the Gaulfoss gauge in the Gaula watershed in mid-Norway and evaluated the calibrated parameters through spatial transfer to the internal catchments of Eggafoss, Hugdal bru and Lillebudal bru for model validation. To our knowledge, this study is the first attempt at evaluating the performance of different levels of parameterizations of the spatial heterogeneity of subsurface storage capacity for hourly runoff simulation in a boreal mountainous watershed. For the study region, there is a growing interest in streamflow prediction at fine temporal resolution for hydropeaking operation of reservoirs for production scheduling, flood forecasting and environmental flow assessment. In addition, the marked loss in performance when parameters calibrated for a daily time step are used for hourly simulation as illustrated by Bastola and

Murphy (2013) substantiates the need for hourly predictions based on parameters calibrated using hourly observations.

The study watershed and data

We used hourly streamflow data from four gauges located inside the 3600 km^2 Gaula watershed located in mid Norway (Gaulfoss, Hugdal bru, Eggafoss and Lillebudal bru) (Figure 1). The last three catchments are located inside the Gaulfoss catchment, but are not nested. For the elements based simulation, we topographically delineated 33 elements within Gaulfoss. Seven of these elements (1-7) are located inside Eggafoss, another seven elements (9-15) are located inside Hugdal bru, the smallest gauged catchment (Lillebudal bru) was discretized as element 8, and the elements 16-33 are parts of the Gaulfoss catchment outside of Eggafoss, Hugdalbru and Lillebudalbru. Generally, the discretized elements are mesoscale sizes, which are less than but comparable to the size of the smallest gauged catchment of Lillebudal bru. The locations of the study catchments, hydro-climatic stations, elevation, and different discretization schemes are shown in Figure 1a.

The main land use is mountainous terrain, forests dominated by conifers and riparian areas (marshes/bogs) as shown in Figure 1b. Hypsometric curves (Strahler, 1952) indicate considerable variations in the elevations of the catchments (Figure 1c). The dominant loose material (soil) in the Gaula watershed is glacial till deposits underlain predominantly by metamorphic and igneous bedrock geology (http://www.ngu.no) (Table 1).

The watershed is characterized by humid temperate climate, snowmelt dependent high-flow regime (Figure 1e) and the flow duration curves (Figure 1d) show considerable contribution of the subsurface flow to the streamflow. Precipitation occurs mainly in the form of rainfall (April-October) and mainly snowfall (November-March). The climate input data are precipitation (P), temperature (T), wind speed (W_s), relative humidity (H_R) and global radiation (R_G) of hourly resolution, which matches to the simulation time step. The model was forced by a climate input distributed on a 1x1 km^2 grid scale based on the inverse distance weighed (IDW) spatial interpolator from the point measurement gauges. We used precipitation data from 12 gauging stations, three of which are located inside the Gaulfoss catchment. Table 1 provides a summary of some characteristics of the catchments and the hydro-climatic data.

Models and methods

Probability distributed parameterization of runoff response routines

The model structure used in the present study is based on a probability distributed model or the PDM (Moore, 1985). The PDM model is based on collections of subsurface reservoirs with different storage capacities (c) and maximum storage capacity (c_{max}). The pattern of subcatchment scale runoff was taken into account by parameterizing the heterogeneity of the subsurface storage capacity in the catchment by a probability distribution. The 1-shape parameter Pareto distribution was used. The maximum storage capacity on the catchment scale (the catchment scale S_{max}) are computed from the calibrated parameters c_{max} and c_{min} and the shape parameters according to the analytical solution in Appendix A.

The effective precipitation (TOSTORAGE) is partitioned into saturation excess runoff i.e. 'saturation from below' (Dunne and Black, 1970 a&b) and change in storage based on the probability distribution following the 'equal storage redistribution of interacting storage elements' concept of Moore (1985) as shown in Figure 2. The subsurface storage was conceptualized as a 'bucket type' single state reservoir with finite storage capacity (equal to S_{max}). The subsurface storage is depleted by the subsurface drainage and evaporation from the subsurface (soil).

The cumulative distribution function (CDF) and the probability density function (PDF) for the Pareto distribution for a random variable of c_n (Figure 2b) are defined as:

$$CDF: F(c_n) = 1 - (1 - c_n)^b;\ c_n[0,1] = \frac{c - c_{\min}}{c_{\max} - c_{\min}}$$
$$PDF: f(c) = \frac{dF(c)}{dc} = \left(\frac{b}{c_{\max} - c_{\min}}\right)(1 - c_n)^{b-1} \tag{1}$$

By inverting the cumulative distribution function, the quantile function for the local storage capacity (c) can be written as:

$$c = c_{\max} - \left\{(c_{\max} - c_{\min})[1 - F(c)]^{\frac{1}{b}}\right\}, \tag{2}$$

where the c_n [0, 1] is the normalized local storage capacity, c_{min} is a parameter which represents the minimum (threshold) local storage capacity below which there is no saturation excess runoff generation (Hegemann and Gates, 2003) and also it represents the threshold storage below which there is no drainage and water is being held under soil tension (Moore and Bell, 2002; Moore, 2007). The c_{max} is the maximum local storage capacity and 'b' is the shape parameter of the distribution.

The direct runoff generated due to infiltration excess or R_{iex} [L] (Horton, 1933) and the actual infiltration to the soil or *TOSOIL* [L] are given by:

$$\begin{aligned} R_{iex} &= \max[0, (SNOWOUT - INFCAP)] \\ TOSOIL &= SNOWOUT - R_{iex} \end{aligned}, \tag{3}$$

where the INFCAP [L] is a free parameter set by calibration and *SNOWOUT* [L] is outflow from the snow routine to the soil. The saturated excess direct runoff or *R* [L] is the amount of runoff in excess of the storage capacity. The change in storage with time is given as:

$$\begin{aligned} &\frac{dS}{dt} \approx \Delta S = S(t + \Delta t) - S(t) = TOSTORAGE - R; \\ &TOSTORAGE = TOSOIL - AET - D_{rv} \end{aligned} \tag{4}$$

The actual evapotranspiration from the soil (*AET* [L]) is computed as a linear function of potential evapotranspiration rate (*PET*) from the storage, the total storage (S_T) and the total maximum storage capacity (S_{Tmax}):

$$AET = PET \times \frac{S_T}{S_{T\max}} \quad (5)$$

We used the following conceptual relationships between storage and drainage for the subcatchment based runoff response:

$$D_{rv} = k\left[S(t)\right]^n, \quad (6)$$

where k is in $mm^{1-n}h^{-1}$, S [L] is the storage in mm, D_{rv} [L] is the drainage volume per unit area computed before saturation excess runoff and n is a dimensionless exponent. Drainage, D_r [L^3T^{-1}], is computed from the D_{rv} [L]. The following equation are derived from eqn. (4) for computation of saturation excess direct runoff over the interval t, t+Δt:

$$R(t) = \begin{cases} TOSTORAGE - \left(S_{\max} - S(t)\right) + S_{\max}\left\{\left[1 - \frac{S(t)}{S_{\max}}\right]^{\frac{1}{b+1}} - \frac{TOSTORAGE}{(b+1)S_{\max}}\right\}^{b+1} ; S(t+\Delta t) < S_{\max} \\ TOSTORAGE - \left(S_{\max} - S(t)\right); S(t+\Delta t) \geq S_{\max} \end{cases} \quad (7)$$

We computed the rate of total direct runoff (R_r [L^3T^{-1}]) as:

$$R_r = \frac{A_i}{\Delta t}\left\{R \times F\left(c^*(t)\right) + R_{iex}\right\}, \quad (8)$$

where the $F(c^*(t))$ is the fraction of the catchment saturated to generate the saturation excess runoff (see Appendix A) and A_i is the catchment area. However, the performance of the PDM based models may depend on the parameterization approaches used to represent the spatial heterogeneities, which we wanted to investigate. A summary of the six evaluated parameterization cases are given in Table 2 and further descriptions are given here. The lists of calibrated model parameters and their prior ranges are given in Table 3. Further details of the PDM model are given in Appendix A.

Case 1: Subcatchment heterogeneity by a probability distribution, catchment scale S_{max} and calibrated shape parameter 'b'

This case is similar to the probability distribution functions based parameterizations in the PDM model (Moore, 1985), which is explained above. This case does not represent the element-to-element heterogeneity of the S_{max} and the shape parameter ‘b’ is set by calibration.

Case 2: Subelement heterogeneity by a probability distribution, element-to-element heterogeneities of the S_{max} and the shape parameter ‘b’

In case 2, we investigated the case when the maximum storage capacity (S_{max}) and the shape parameter were computed for each element i.e. the element-to-element heterogeneity of the S_{max} and the shape parameter (‘b’) were modelled. The influence of topography on the storage capacity and hence on the dynamics of runoff generation is considered in this case by directly utilizing the topographic information. It is useful to represent the spatial heterogeneity of hydrological variables based on readily available high-resolution spatial information such as topographic features, which can be derived from the Digital Elevation Model (DEM), both to reduce the number of model parameters and to allow transfer of parameters to ungauged catchments and in parameterization for climate models (e.g. Ducharne et al., 1998).

The role of topography in runoff response dynamics has been widely studied (e.g. Beven and Kirby, 1979; Wood et al., 1990; Wood et al., 1992; Blöschl and Sivapalan, 1995; Bell and Moore, 1998). In a study of Norwegian catchments, Beldring et al. (2003) noted a relationship between the maximum soil moisture storage and altitude with larger soil moisture storage for lowland areas than for mountains as the average thickness of surface deposits tends to decrease with altitude. Therefore, depending on the distribution of topographic and soil characteristics in the catchment, the maximum storage capacity (S_{max}) may vary throughout the catchment and hence the effects of the lumped representation of the maximum storage capacity on the runoff simulation need to be investigated.

Dumenil and Todini (1992) computed the shape parameter ‘b’ of the distribution based on the standard deviation of the subgrid elevation. Bell and Moore (1998), Hagemann and Gates

(2003), Manfreda and Fiorentino (2008), Manfreda (2008) and Liu et al. (2012) also noted the topographic influence on ´b’. Eq (1) shows that as the value of ‘b’ increases, the fraction of catchment saturated increases and hence the likelihood of more saturation excess runoff.

The maximum storage capacity and shape parameter ‘b’ for each element are computed from a functional relationship between the equations for maximum storage capacity (S_{max}) from the analytical solution in Appendix A and based on the topographic gradient (eq. 9). This approach is similar to the linkage function in the grid-to-grid (G2G) model of Bell and Moore (1998). We related the parameter ‘b’ to the maximum storage capacity of the Pareto distribution:

$$S_{\max} = \frac{c_{\max} - c_{\min}}{b+1};\ S_{\max} = \left(\frac{MaxMFDslope_{\max} - MaxMFDslope_{avg}}{MaxMFDslope_{\max}} \right) \left(c_{\max} - c_{\min} \right) \tag{9}$$

Equating the above two equations for S_{max}, the following relationship for ‘b’ and the topographic gradient can be derived:

$$b = \frac{MaxMFDslope_{avg}}{MaxMFDslope_{\max} - MaxMFDslope_{avg}} \tag{10}$$

The above relations provide

$$\begin{aligned} &S_{\max} = c_{\max} - c_{\min} \left\{ if\ \ MaxMFDslope_{avg} = 0,\ b = 0 \right. \\ &S_{\max} = \frac{c_{\max} - c_{\min}}{2} \left\{ if\ \ MaxMFDslope_{avg} = 0.5 MaxMFDslope_{\max},\ b = 1 \right. , \\ &S_{\max} = 0 \left\{ if\ \ MaxMFDslope_{avg} = MaxMFDslope_{\max},\ b\ is\ undefined \right. \end{aligned} \tag{11}$$

where the $MaxMFDslope_{avg}$ represent the average of the gradients of the 1x1 km^2 grid cells within the element while $MaxMFDslope_{max}$ is a regional parameter representing the maximum of gradients for the 1x1 km^2 grid cells in the whole catchment.

The MaxMFDslope for the grid cell is the topographic gradient in the steepest downslope flow direction among its eight neighbors in a 3x3 window. It was computed from the DEM as MFDslope = (Elevation $_{\text{upstream cell}}$ - Elevation $_{\text{downstream cell}}$) / Flow travel length. Flow travel length = grid cell size for the cardinal flow direction and (grid cell size)*√2 for diagonal flow directions. For the element based simulation (cases 1 to 3), the case of $MaxMFDslope_{avg}$ = $MaxMFDslope_{max}$ is not an issue, but for the grid based simulation (cases 1G to 3G) a $storage_{min}$ calibrated parameter was introduced to avoid S_{max} and 'b' to become zero in flat areas. Besides allowing study on the sensitivity of runoff generation to the spatial distribution of S_{max} and b, case 2 also reduces the number of calibrated parameters. The equations for actual evapotranspiration, infiltration excess runoff, saturation excess runoff and subsurface drainage were same as that of the case 1.

Case 3: No subcatchment and subelement heterogeneity of the storage capacity and no element-to-element heterogeneity of S_{max}

This case is based on the Basic-Grid-Model (Bell and Moore, 1998), but in here it is applied to an element scale rather than a grid scale. In case 3, there is no parameterization of the spatial heterogeneity by a probability distribution and we did not consider the element-to-element heterogeneity of the S_{max}, rather S_{max} was a calibrated parameter. Therefore, case 3 is a simple semi-distributed model. We update the storage for the elements and computed the direct runoff as below while the equations for actual evapotranspiration, infiltration excess runoff and subsurface drainage were same as that of cases 1 and 2:

$$R = \max\left[0, \left(S(t) + TOSTORAGE - S_{\max}\right)\right]; S(t+\Delta t) = \max\left\{0, \left[S(t) + TOSTORAGE - R\right]\right\}$$
$$R_r = \left\{R + R_{iex}\right\} \times \frac{A_i}{\Delta t} \qquad (12)$$

Cases 1G, 2G and 3G: Grid based runoff simulation

In case 1G, the subgrid heterogeneity was parameterized by a probability distribution but the parameters are calibrated for the catchment scale similar to that of case 1. In case 2G, the subgrid heterogeneity was parameterized by the probability distribution, and grid-to-grid heterogeneity of S_{max} was accounted for based on the linkage function between S_{max} and the topographic gradient. We derived the following equations for case 2G from the linkage function between topographic gradients and the S_{max}:

$$S_{\max} = Storage_{\min} + \left\{ \left(c_{\max} - Storage_{\min} \right) \times \left(\frac{MaxMFDslope_{\max} - MaxMFDslope}{MaxMFDslope_{\max}} \right) \right\} \tag{13}$$

$$b = \frac{\left(MaxMFDslope \times \left(c_{\max} - Storage_{\min} \right) \right) - \left(c_{\min} \times MaxMFDslope_{\max} \right)}{c_{\max} \left(MaxMFDslope_{\max} - MaxMFDslope \right) + \left(Storage_{\min} \times MaxMFDslope \right)} \tag{14}$$

In case 3G, we did not consider both the subgrid heterogeneity of storage capacity by a probability distribution and the grid-to-grid heterogeneity of the S_{max}. We set the S_{max} by calibration. Therefore, case 3G is a simple distributed model, which is similar to the Basic-Grid-Model (Bell and Moore, 1998).

The main differences between the distributed simulations (cases 1G to 3G) and the semi-distributed simulations (cases 1 to 3) are related to the equations used for simulation of the subsurface drainage. In a boreal landscape dominated by till soils, hydraulic conductivity decreases with depth, the groundwater table largely follows the topography and the catchment runoff depend on soil moisture conditions and the depth to groundwater (Lind and Lundin, 1990; Hinton et al., 1993; Myrabø, 1997; Beldring, 1999; 2002). We computed the rate of subsurface drainage/flow from derived equation based on assumptions of Dupuit-Forchheimer to Darcy's law for saturated subsurface flow (Freeze and Cherry, 1979; Wigmosta and Lettenmaier, 1999) by assuming a power-law transmissivity decay with depth (Ambroise et al. 1996; Wigmosta and Lettenmaier, 1999):

$$D_r = \left\{ \frac{w}{n_T} \psi MaxMFDslope \left[S_{\max} \right]^{1-n_T} \frac{\Delta t}{A_i} \right\} \left[S\left(t\right) \right]^{n_T}, \tag{15}$$

where ψ [LT^{-1}] is diffusivity or saturated hydraulic conductivity at the surface divided by porosity, w [L] is size of the grid cell, and n_T is the transmissivity decay exponent, D_r [L^3T^{-1}] is drainage volume and D_{rv} (L) is drainage volume per unit area computed from the D_r. Eq. (15) for drainage accounts for the grid-to-grid heterogeneity of topographic gradient. Based on preliminary tests of parameter sensitivity, a hyperbolic (Duan and Miller, 1997) transmissivity decay profile (i.e. exponent $n_T = 2.0$) was used.

The evapotranspiration routine

We used the Priestley Taylor method (Priestley and Taylor. 1972) to estimate the potential evapotranspiration (mm/h):

$$PET = \alpha \frac{\Delta}{\Delta + \gamma} \left(R_n \right) \left(\frac{\Delta t}{L_v} \right), \tag{16}$$

where α is Priestley Taylor constant, the Δ is the slope of saturation vapor pressure curve corresponding to an air temperature at 2m (kPa/^{o}C), γ is the psychrometric constant (0.066 kPa/^{o}C), R_n (W/m^2) is the net radiation which is the sum of net shortwave radiation and the net longwave radiation, L_v (kJ/m^3) is volumetric latent heat of vaporization and Δt (s) is the simulation time step in seconds. We used α = 1.26 (see Priestley and Taylor. 1972; Teuling et al., 2010) in the present study rather than setting by calibration. The net short wave radiation was computed from global radiation and land albedo while the net long wave radiation was computed based on Sicart et al. (2006). We used eq (5) for the computation of the actual evapotranspiration (*AET*). The *AET* is set to zero for the proportion of the grid cells covered by snow.

The snow routine

The snow accumulation and snowmelt processes exert significant influence on the hydrological cycle of the study area. The outflow melt water release from saturated snow (i.e. *SNOWOUT*) was computed by a snow routine based on the Gamma distributed snow depletion curve (SDC) (Kolberg and Gottschalk, 2006; 2010), which was implemented in ENKI hydrological modelling platform (Kolberg and Bruland, 2012). The free parameters in this routine are snow-rain threshold temperature parameter (TX) and snowmelt sensitivity to wind speed or windscale (WS). Simulation of potential evapotranspiration, snow accumulation and snowmelt-runoff processes were based on the 1x1 km^2 grid scale. For the semi-distributed (element) simulations (cases 1, 2 and 3), we aggregated the gridded (1x1 km^2) outflow from the snow routine

(*SNOWOUT*) and potential evapotranspiration (*PET*) to the element scale based on simple averaging and provided as an input to the runoff response routines.

Runoff routing

We used the source-to-sink (STS) routing algorithm (Olivera, 1996; Olivera and Maidment, 1999) to route the generated runoff at each source to the sink (catchment outlet). The instantaneous runoff generated at the source are related to the outlet response by a flow path response function or $U_i(t)$ [T^{-1}]. The flow path response function used in the present study was based on the first passage time distribution (Hayami, 1951; Nauman, 1981). Olivera (1996) showed the relationships among the total expected travel time from the source to the outlet (T_i), its corresponding variance or $Var(T_i)$, the flow dispersion coefficient (D_i) and Peclet number or Π_i[-] based on the statistical properties of mean and variance. The gridded (1x1 km^2) flow path response function or $U_i(t)$ [T^{-1}] is given by:

$$U_i(t) = \frac{1}{2t\sqrt{\pi\left(\frac{t}{T_i}\right)/\Pi_i}} \exp\left\{-\frac{\left[1-\left(\frac{t}{T_i}\right)\right]^2}{4\left(\frac{t}{T_i}\right)/\Pi_i}\right\} \tag{17}$$

The flow path responses function for grid cell '*i*' represents the probability distribution of flow travel time (t) from the source (grid cell) to the sink (outlet) which has a mean value of T_i. The flow path Peclet number is a representative measure of the relative importance of advection with respect to dispersion whereby the flow dispersion coefficient represents the effects of storage and spreading. Fig. 3c shows typical response functions.

For the semi-distributed runoff simulations (cases 1, 2 and 3), the generated runoff at the element scale were distributed over the 1x1 km^2 grid cells within the elements. We coupled the generated runoff to the flow path response function to perform the flow routing on the grid scale, rather than aggregating to the element scale. This grid scale would enable us to account for the differences in the flow travel time and hence response functions among the grid cells especially for elongated elements. Beldring et al. (2003) noted that the permanent river network including streams and lakes lies within 1x1 km^2 of almost every point in the Norwegian landscape and all the lateral transfers of water at 1x1 km^2 grid cells take place within the river network.

The sum of direct runoff and subsurface drainage generated at the source grid cell are routed to the outlet. From the unit hydrograph model for a spatially distributed linear system subdivided into uniform non-overlapping sub-areas (Maidment et al., 1996; Olivera and Maidment, 1999), runoff routing can be performed by convolution:

$$Q(t) = \sum_{i=1}^{N} \left\{ R_{ri}\left(t\right) + D_{ri}\left(t\right) \right\} \otimes U_{i}\left(t\right) \tag{18}$$

Q [L^3T^{-1}] is routed simulated streamflow, N is the total number of grids in the watershed, and $\otimes$ is the convolution operator. The routing routine involves two free parameters namely the velocity of flow (V) and the dispersion coefficient of flow (D). We assumed the parameters to be ‘time-invariant’ and set as calibration parameters for the whole catchments i.e. not spatially distributed.

Model calibration and evaluation

The Differential Evolution Adaptive Metropolis algorithm or DREAM (Vrugt et al, 2008; 2009) with residuals based log-likelihood objective function was used for the calibration. DREAM is an adaptive random walk Metropolis algorithm to enhance the applicability of the MCMC methods to complex, non-linear and high-dimensional problems such as calibration of watershed models (Vrugt et al, 2008; 2009).

For the hourly streamflow series, the serial correlation is expected to be high and hence the actual amount of information obtained from the data is much less. Therefore, we introduced the fraction of effectively independent observations from the total observations denoted as ‘f’. We used the logarithmic likelihood function for simplicity and numerical stability (Vrugt et al., 2013). The residual based log-likelihood (*L-L*) is given as:

$$L-L\left(\delta/\sigma_{\varepsilon}^{2},\sum_{t=1}^{n_i}\left(Qsim_t^{(\theta)}-Qobs_t^{(\theta)}\right)^{2}\right)=\left(\frac{-n_i}{2}\log(2\pi)-\frac{n_i}{2}\log\left(\sigma_{\varepsilon}^{2}\right)-\frac{\sum_{t=1}^{n_i}\left(Qsim_t^{(\theta)}-Qobs_t^{(\theta)}\right)^{2}}{2\sigma_{\varepsilon}^{2}}\right)\times f, \qquad (19)$$

where $Qsim^{(\theta)}$ and $Qobs^{(\theta)}$ respectively are Box-Cox (Box and Cox, 1964) transformed observed and simulated streamflow time series (t), n_i is the length of non-missing records of streamflow for the catchment, δ represents model parameter, θ is the Box-Cox transformation parameter and σ_{ε}^{2} is variance of error.

Transformation was carried out in order to obtain an approximately Normal distributed series with homoscedastic residuals. If $\theta = 0$, the streamflow is assumed to be lognormal distributed i.e. high weightage to low flows. If $\theta = 1$, the streamflow series is assumed to be Gaussian i.e. high weightage to high flow. A value of $\theta = 0.3$ is common in literature (e.g. Vrugt et al., 2002). However, we computed the θ values from the observed streamflow data set using the ‘fminsearch’ algorithm in matlab, which calls for finding the θ value that maximizes a likelihood function (http://www.mathworks.com). We computed the fraction f based on a AutoRegressive or AR(1) model of error covariance (Zięba, 2010).

The maximum Nash-Sutcliffe efficiency or R^2 (Nash and Sutcliffe, 1970), which emphasizes high flows, and the maximum R^2ln for log-transformed series, which emphasizes low flows), were used for further comparisons and evaluations. We also evaluated the performances of the routines based on their predictive reliability (Kavetski and Fenicia, 2011) using quantile-quantile (Q-Q) plots. The Q-Q plots were in the form of the probability of non-exceedance or empirical cumulative distribution functions (CDF) of the observed and simulated streamflow time series. The departures of the plots from the theoretical uniform distribution indicate the discrepancy between the predictive distribution and the observed data. In addition, we evaluated the routines based on their prediction performances of temporal variability of the streamflow or the flow duration curves. The ‘split-sample’ test (Klemeś, 1986) and ‘proxy basin’ test were used for temporal, spatial and spatio-temporal validation of the models against independent data to test the reliability of model prediction outside the calibration conditions (Seibert, 2003). We performed the spatial and spatio-temporal validation of the model through direct transfer of calibrated parameter sets, which correspond to the maximum R^2 and maximum R^2ln of the Gaulfoss catchment, to the internal catchments of Eggafoss, Hugdalbru and Lillebudal bru for the calibration and validation periods.

Results and discussion

Model calibration

Hydrographs, quantile-quantile (Q-Q) plots and flow duration curves of observed versus simulated streamflow for Gaulfoss are given in Figure 3a-b, Figure 4 and Figure 5 respectively. We presented the hydrographs only for the R^2 performance measure for a part of calibration period for clear presentation, the Q-Q plots for the R^2 for the calibration and validation periods and the flow duration curves for both the R^2 and R^2ln for the calibration and validation periods. We presented the performance measures for calibration, temporal, spatial and spatio-temporal validation of the calibrated parameters in Table 4.

We obtained the goodness-of-fits of R^2/R^2ln respectively up to 0.84/0.86 for the calibration and up to 0.85/0.90 for the temporal and spatial validation, which indicate satisfactory fits between the observed and simulated hydrographs for the six different parameterization cases for semi-distributed and distributed runoff simulation. Therefore, the simulations based on different parameterizations of the single state and single drainage outlet subsurface storage provided satisfactory runoff simulation in terms of the goodness-of-fit tests.

However, a more stringent test for reliability of prediction based on the Q-Q plots of the observed and simulated streamflow indicated that there is a considerable prediction uncertainty for all the parameterization cases (see Figure 4). Nearly symmetrical deviations from the perfect fit uniform distribution (diagonal line) show both under and over predictions. The results of the flow duration curves indicate better simulation of the temporal variability of the high flow compared to the low flow as shown in Figure 5. The performance of the calibration of the parameterizations in reproducing the hydrographs based on R^2 and R^2ln performance measures found to be better than reproducing the Q-Q plots and the FDC. Calibration of hydrological models for specific objectives of reproducing the flow-duration curves (e.g. Westerberg et al., 2011) and the Q-Q values may improve their respective performances.

It was impossible to consistently distinguish the best performing parameterization since different parameterization cases provided only marginally different performance for different seasons (snowmelt versus rainfall) and ranges of flow (low, medium and high).

Model validation

The investigation of performances of distributed models calibrated to streamflow at basin outlet for the simulation at internal catchments was one of the science question tested by the Distributed Model Intercomparison Project, DMIP (Smith et al., 2004). Spatial transferability of calibrated parameters from the Gaulfoss catchment to the internal catchments of Eggafoss and Hugdal bru (Table 4) provided satisfactory performances for all parameterization cases. However, parameter transfer to Lillebudal catchment showed poor performance especially for R^2ln. For Lillebudal bru catchment, Hailegeorgis and Alfredsen (2014) found poor performance of parameter transfer from the Gaulfoss watershed also for the HBV conceptual model especially for low flow simulation. The Lillebudal bru catchment is characterized by high elevation mountainous terrain with a mean altitude above the altitude of all climate stations used for the calibration (Table 1). Moreover, there are no climate stations inside or nearby to the Lillebudal bru catchment and hence less representativeness in precipitation data may cause poor streamflow simulation. Effects of the dominantly mountainous terrain are expected to cause significant temporal and spatial variability of precipitation fields. Performance in simulation of low flow, which is mainly contributed by the baseflow, was also poor for Lillebudal bru that shows the importance of examination of the quality of observed streamflow data for the Lillebudal bru.

Parametrical parsimony

The effects of interactions or correlations among the parameters during calibration may cause poor identifiability of the parameterizations. Improving the parsimony of the routines can be suggested as a possible solution to reduce parameter interactions. Parsimony can be achieved by reducing the number of free parameters for instance by fixing the insensitive parameters. For instance, the calibrated values of c_{min} were less than 7.5 mm against a prior range of 0.0 - 100 mm and hence c_{min} can be set to zero and excluded from the free parameters to improve the parsimony and to avoid its interaction with c_{max} and other parameters. In addition, the calibrated values for the exponent parameter of the conceptual subsurface drainage-storage relationship (n) of the parameterization cases 1 to 3 ranges from 1.5 to 3.0 against a prior range of 0.20-5.0. Hence, there is a possibility to fix this parameter to some representative value within this range to improve the parsimony and to avoid its interaction with k and other parameters. For instance, Wittenberg and Sivapalan (1999) and Moore and Bell (2002) respectively assumed quadratic ($n = 2.0$) and cubic ($n = 3.0$) relationships between ground water storage and baseflow. However, parsimony alone may not guarantee improvement in the identifiability and predictive uncertainty of the parameterizations since there are also other sources of uncertainty related to the input data and scale issues.

The effects of input data for parameter calibration

We conducted the semi-distributed and distributed runoff simulations for the boreal mountainous catchments based on precipitation data from 12 gauging stations, which were spatially interpolated by inverse distance weighing (IDW) on 1x1 km^2 grids. However, Goodrich et al. (1995) reported an inadequacy of meteorological gauging networks in the higher altitudes. In addition, for high latitude mountainous regions, Moine et al. (2003) noted the complexity of hydrological modelling due to the complexity of local processes and the difficulty of estimating spatially-distributed inputs such as rainfall and temperature due to sparse networks. Beldring et al. (2003) noted that the spatial interpolation procedure with correction for altitude differences is unable to describe all effects caused by the various precipitation formation mechanisms and wind directions in Norwegian catchments. Das et al. (2008) found that a distributed HBV model structure do not outperform the simpler model structures, which they attributed to the interpolated climate inputs that cannot reflect the true spatial variability. Wrede et al. (2013) compared a distributed HBV model complemented by the subgrid scale parameterization for distinct land use classes to a less parameterized lumped HBV model for a Swedish lowland catchment. The authors found the results to be indistinguishable, which they attributed to the deficiency of calibration against only the observed streamflow at the catchment outlet. In the present study, we performed the calibration based on only the catchment integrated observed streamflow. Calibration based on climate data from dense gauging stations and spatial distributed observations, which were not available for the present study, may provide more insights out of the simulations.

Parameterization and scale issues

Both discretization and aggregation techniques in precipitation-runoff models are dependent on the scales and hence the results of simulation from parameterization across a range of scales may be sensitive to the spatial scales used (e.g. Wood et al., 1990; Becker and Braun, 1999; Koren et al., 1999; Haddeland et al., 2002; Merz et al., 2009). Beldring et al. (1999; 2000) suggested elements at scales of approximately 1x1 km^2 sufficient to parameterize the hydrological processes in till soils. Gottschalk et al. (2001) also identified a hillslope scale of 1-2 km^2 for the NOPEX region. In addition, Wood et al. (1988; 1990) identified a

'Representative Elementary Area (REA)' of subcatchments of about 1x1 km^2. Scale issues in hydrological modelling (Blöschl and Sivapalan, 1995) are one of the major challenges in parameterization of precipitation-runoff models.

Due to the sparse hydro-meteorological stations, i.e. only 12 precipitation stations distributed over the study region, it is clear that the resolution of the forcing field is low. Even though the resolution of climate forcing was much lower than the resolution at which the model was parameterized for the case 2G, the performances of the case 2G and case 2 were found to be indistinguishable. There is a scale mismatch between spatial heterogeneities of climate control and topographic control due to the prevailing terrain heterogeneity at a finer hillslope scale (e.g. 25mx25m). For the boreal watershed, the topographic driven influence on the spatial heterogeneity of soil moisture, subsurface storage and hence lateral movement of subsurface flow is expected to dominate the grid-to-grid variability of the low intensity precipitation. However, the only advantage of distributed (gridded) simulations (cases 1G to 3G) over the semi-distributed (cases 1 to 3) was found to be the simplicity in preparing gridded input maps for the distributed model than preparing topographically delineated elements for the semi-distributed model, rather than marked improvement in the runoff simulation.

For the boreal catchments, topographic control heterogeneities at finer spatial scales is expected to dominate the runoff generation processes and hence parameterizations for the finer scale hillslope processes may be required (see Halldin et al., 1999). Therefore, the grid cell-to-grid cell routing in the hillslopes (e.g. 25mx25m grids) towards the stream networks by considering the hillslope topographic gradients within the 1x1 km^2 grid like in the distributed hydrology-soil-vegetation model or DHSVM (Wigmosta et al., 1994) may further allow more representativeness and utility of the terrain features.

Conclusions

We evaluated the performances of six different parameterizations of the spatial heterogeneity of subsurface storage capacity based on the probability distributed model for semi-distributed and distributed (1x1 km^2 grids) hourly runoff simulation in boreal mountainous watershed in mid Norway.

Calibration of all parameterization cases for the study watershed provided satisfactory but indistinguishable simulation of hourly runoff hydrographs and reproduced the temporal variability of streamflow in terms of duration curves. Transferability of the calibrated parameters to two internal catchments of relatively representative climate data indicated validation of the models. The marginal differences in the hourly runoff simulation performance indicated that case 3 (a simple semi-distributed model) and case 3G (a simple distributed model) are preferable due to their simplicity and parsimony. This study showed that the subelement and subgrid scale parameterizations of the subsurface storage capacity did not provide better results for the hourly runoff simulation than the coarser parameterizations, which indicate:

i. Identification of parameterizations require climate records from denser precipitation gauging stations than what is sufficient to provide acceptable calibration of precipitation-streamflow relationships;
ii. Challenges towards identification of parameterizations based on model calibration only to the catchment integrated streamflow observations;
iii. Equivalent simulation performance for the available data set showed a potential preference for the simple and parsimonious parameterizations in operational forecast mode related to model updating.

Previous studies are lacking pertinent to comparisons of different parameterizations of the subsurface storage capacity for hourly runoff simulation in boreal catchments. Further insights would be expected from studies based on better quality climate data (e.g. dense climate gauging stations). Both the precipitation control (e.g. the density of the climate stations) and topographic control (at further finer spatial scales) driven heterogeneities need to be thoroughly explored. The effects of input uncertainties related to precipitation and streamflow, and parameter non-identifiability on identification of the parameterizations require further investigations, which were not the scope of the present study. In addition, we did not consider the preferential flow, which may be apparent in the glacial till soils (e.g. Jansson et al., 2004).

Acknowledgements

The Center for Environmental Design of Renewable Energy (CEDREN) supported the research financially under HydroPEAK hydrology (Project number: 50043420). We obtained the climate data from the Norwegian Meteorological Institute, TrønderEnergi, Statkraft and bioforsk, and the streamflow data from the Norwegian Water Resources and Energy Directorate. We wish to thank the anonymous reviewers who contributed to improving the manuscript.

References

Abbott, M. B., Bathurst, J. C., Cunge, J. A., O'Connell, P. E., and Rasmussen, J.,1986. 'An Introduction to the European Hydrological System - Systéme Hydrologique Européen, "SHE", 2: structure of a physically-based, distributed modelling system', J. Hydrol. 87, 61-77.

Ambroise, B., Beven, K., Freer J., 1996. Towards a generalization of the TOPMODEL concepts: topographic indices of hydrological similarity. Water Resour. Res. 32(7), 2135–2145.

Bastola, S. & Murphy, C., 2013. Sensitivity of the performance of a conceptual rainfall-runoff model to the temporal sampling of calibration data. Hydrology Research 44 (3), 484-494.

Becker, A., and P. Braun, 1999. Disaggregation, aggregation, and spatial scaling in hydrological modelling. J. Hydrol. 217, 239–252.

Beldring, S., Gottschalk, L., Seibert, J., Tallaksen, L.M., 1999. Distribution of soil moisture and groundwater levels at patch and catchment scales. Agr. Forest Meteorol. 98-99, 305-324.

Beldring, S., Gottschalk, L., Rodhe, A., Tallaksen, L.M., 2000. Kinematic wave approximations to hillslope hydrological processes in tills. Hydrol. Process. 14, 727-745.

Beldring, S., 2002. Runoff Generating Processes in Boreal Forest Environments with Glacial Tills. Nord. Hydrol., 33 (5), 347-372.

Beldring, S., Engeland, K., Roald, L. A., Sælthun, N. R., Vøkso, A., 2003. Estimation of parameters in a distributed precipitation-runoff model for Norway. Hydrol. Earth. Syst. Sc. 7(3), 304-316.

Bell, V.A., Moore, R.J., 1998. A grid-based distributed flood forecasting model for use with weather radar data: Part 1. Formulation. Hydrol. Earth. Syst. Sc. 2, (2-3), 265-281.

Bergström, S.,1976. Development and application of a conceptual runoff model for Scandinavian catchments, SMHI Reports RHO, No. 7, Norrköping.

Beven, K. J., Kirby, M. J.A, 1979. Physically-based variable contributing area model of basin hydrology. Hydrolog. Sci. Bull. 24(1), 43–69.

Blöschl, G., Sivapalan, M., 1995. Scale issues in hydrological modeling: a review. Hydrol. Process. 9, 251-290.

Bogaard, T.A., Luxemburg, W.M.J., Wit, M. de., Douben, N., Savenije, H.H.G., 2005. Some hydrological challenges in understanding discharge generation processes in the Rhine and Meuse basins. Phys. Chem. Earth, Pt. A/B/C 30 (4-5), 262-266.

Box, G. E. P., Cox, D. R., 1964. An analysis of transformations. Journal of the Royal Statistical Society, Series B *26*, 211-252.

Cole, S.J., Moore, R.J., 2009. Distributed hydrological modelling using weather radar in gauged and ungauged basins. Advances in Water Resources, 32, 1107-1120.

Das, T., Bárdossy, A., Zehe, E., He, Y., 2008. Comparison of conceptual model performance using different representations of spatial variability. J. Hydrol. 356, 106-118.

Duan, J., Miller, N.L., 1997. A generalized power function for the subsurface transmissivity profile in TOPMODEL. Water Resour. Res., 33(11), 2559–2562.

Ducharne, A., Laval, K., Polcher, J., 1998. Sensitivity of the hydrological cycle to the parameterization of soil hydrology in GCM. Clim. Dynam.14, 307-327.

Dumenil, L. and Todini, E., 1992. A rainfall-runoff scheme for use in the Hambrurg climate model, in: O'Kane, J.P. (Ed), Advances in Theoretical Hydrology, a Tribute to James Dooge, Amsterdam, Netherlands: Elsevier,129-157.

Dunne, T., Black, R. D., 1970a. An experimental investigation of runoff production in permeable soils. Water Resour. Res. 6(2), 478–490.

Dunne, T., and Black, R. D., 1970b. Partial area contributions to storm runoff in a small New England watershed. Water Resour. Res. 6(5), 1296–311.

Freeze, R.A., Cherry, J.A., 1979. Groundwater, Prentice-Hall, Englewood Cliffs, N.J.

Goodrich, D. C., 1990. 'Geometric simplification of a distributed rainfall-runoff model over a large range of basin scales', PhD Thesis, Univ. Arizona, Tucson, 361 pp.

Goodrich, D.C., Faures, J.-M., Woolhiser, D.A., Lane, L.J., Sorooshian, S., 1995. Measurment and analysis of small-scale convective storm rainfall variability. J. Hydrol. 173, 283-308.

Gottschalk, L., Beldring, S., Engeland, K., Tallaksen, L.M., Sælthun, N. R., Kolberg, S.,Yury, M., 2001. Regional/macroscale hydrological modelling: a Scandinavian experience. Hydrolog. Sci. J. 46 (6), 963-982.

Haddeland, I., Matheussen, B.V., Lettenmaier, D.P., 2002. Influence of spatial resolution on simulated streamflow in a macroscale hydrological model. Water Resour. Res. 38 (7), 1124.

Hagemann, S., Gates, L.D., 2003. Improving a subgrid runoff parameterization scheme for climate models by the use of high resolution data derived from satellite observations. Clim. Dynam. 21, 349-359.

Hailegeorgis, T., Alfredsen, K., 2014. Comparative evaluation of performances of different conceptualizations of distributed HBV runoff response routines for prediction of hourly streamflow in boreal mountainous catchments. Hydrology Research, doi:10.2166/nh.2014.051.

Halldin, S., Gottschalk, L., Gryning, S.E., Jochum, A., Lundin, L.C., Van de Griend, A.A., 1999. Energy, water and carbon exchange in a boreal forest-NOPEX experiences. Agric. For. Meteorol. 98-99, 5-29.

Hinton, M.J., Schiff, S.L., English, M.C., 1993. Physical properties governing groundwater flow in a glacial till catchment. J. Hydrol. 142, 229-249.

Horton, R.E., 1933. The role of infiltration in the hydrologic cycle. Trans. Am. Geophys. Union 14, 446-460.

Jansson, C. Espeby, B., and Jannson, P-E., 2004. Preferential flow in glacial till soil. Nord. Hydrol. 36 (1), 1-11.

Kavetski, D., Fenicia, F., 2011. Elements of a flexible approach for conceptual hydrological modeling: 2. Application and experimental insights. DOI: 10.1029/2011WR010748.

Klemeś, V, 1986. Operational testing of hydrological simulation models. Hydrolog. Sci. J. 31, 13-24.

Kolberg, S. A., Bruland, O., 2012. ENKI - An Open Source environmental modelling platform. Geophysical Research Abstracts, 14, EGU2012-13630,EGU General Assembly.

Kolberg, S. A., Gottschalk, L., 2006. Updating of snow depletion curve with remote sensing data. Hydrol. Process. 20(11), 2363-2380.

Kolberg, S. A., Gottschalk, L., 2010. Interannual stability of grid cell snow depletion curves as estimated from MODIS images. Water Resour. Res., 46 (11),doi:10.1029/2008WR007617.

Koren, V., Schaake, J., Mitchell, K., Duan, Q.-Y., Chen, F., and Baker, J.M., 1999, A parameterization of snowpack and frozen ground intended for NCEP weather and climate models. Journal of Geophysical Research, 104, (D16), 19,569-19, 585.

Leavesley, G. H. and Stannard, L. G. 1990. Application of remotely sensed data in a distributed-parameter watershed model in Kite, G. W. and Wankiewicz, A. (Eds), Proc. Workshop on Applications of Remote Sensing in Hydrology, Saskatoon, February 1990. pp. 47-64.

Lind, B.B., Lundin, L., 1990. Saturated hydraulic conductivity of Scandinavian tills. Nord. Hydrol. 2(1), 107- 118.

Liu, J., Chen, X., Wu, J., Zhang, X., Feng, D., Xu, C.-Y, 2012. Grid parameterization of a conceptual distributed hydrological model through integration of a sub-grid topographic index: necessity and practicability. Hydrological Sciences Journal, 57(2).

Manfreda, S., Fiorentino, M., 2008. A stochastic approach for the description of the water balance dynamics in a river basin. Hydrol. Earth Syst. Sci. 12, 1189–1200.

Manfreda, S., 2008. Runoff generation dynamics within a humid river basin. Nat. Hazards Earth Syst. Sci. 8, 1349–1357.

Maidment, D.R., Olivera, J.F., Calver, A., Eatherral, A., Fraczek, W., 1996. A unit hydrograph derived from a spatially distributed velocity field. Hydrol. Process.10 (6), 831–844.

Merz, R., Parajka, J, Blöschl, G., 2009. Scale effects in conceptual hydrological modelling. Water Resour. Res. 45, W09405, doi:10.1029/2009WR007872.

Moine, N., Hendrickx, F., Gailhard, J., 2013. Rainfall–runoff modelling as a tool for constraining the reanalysis of daily precipitation and temperature fields in mountainous regions Cold and Mountain Region Hydrological Systems Under Climate Change: Towards Improved Projections. Proceedings of H02, IAHS-IAPSO-IASPEI Assembly, Gothenburg, Sweden, July 2013 (IAHS Publ. 360, 2013).

Moore, R. J., Clarke, R. T., 1981. A distribution function approach to rainfall runoff modeling. Water Resour. Res. 17, 1367–1382.

Moore, R.J., 1985. The probability-distributed principle and runoff production at point and basin scales. Hydrolog. Sci. J. 30 (2), 273-297.

Moore, R.J., Bell, V.A., 2002. Incorporation of groundwater losses and well level data in rainfall-runoff models illustrated using the PDM. Hydrol. Earth Syst. Sci. 6 (1), 25-38.

Moore, R.J., 2007. The PDM rainfall-runoff model. Hydrol. Earth Syst. Sci.11 (1), 483-499.

Myrabo, S., 1986. Runoff studies in a small catchment. Nord. Hydrol. 17, 335–346.

Myrabo, S., 1997. Temporal and spatial scale of response area and groundwater variation in Till. Hydrol. Process.11,1861–1880.

Nash, J. E., Sutcliffe, J. V., 1970. River flow forecasting through conceptual models, I. A discussion of principles. J. Hydrol. 10, 228–290.

Olivera, F., 1996. Spatially distributed modeling of storm runoff and nonpoint source pollution using geographic information systems, PhD Thesis, Department of Civil Engineering, University of Texas at Austin, USA.

Olivera, F., Maidment, D.R.,1999. Geographic information system (GIS)-based spatially distributed model for runoff routing. Water Resour. Res. 35 (4), 1155–1164.

Priestley, C.H.B.,Taylor, R.J., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Mon. Weather Rev. 100, 81-82.

Refsgaard, J.C., 1997. Validation and Intercomparison of different updating procedures for real-time forecasting. Nord. Hydrol. 28, 65-84.

Seibert, J., 2003. Reliability of model predictions outside calibration conditions. Nordic Hydrology 34(5), 477-492.

Sicart, J. E. Pomeroy, J. W., Essery, R. L. H. and Bewley, D., 2006. Incoming longwave radiation to melting snow: observations, sensitivity and estimation in northern environments. Hydrol. Process. 20, 3697-3708.

Singh, V.P., Woolhiser, D.A., 2002. Mathematical modeling of watershed hydrology. J. Hydrol. Eng. 7 (4), 270-292.

Sivapalan, M. and Viney, N. R., 1994. 'Application of a nested catchment model for predicting the effects of changes in forest cover'. In: Ohta, T. (Ed.) Proc. Int. Symp. Forest Hydrology, Univ. Tokyo, Tokyo. IUFRO, pp. 315-322.

Smith, M.B., Seo, D. -J., Koren, V. I., Reed, S., Zhang, Z., Duan, Q.-Y., Moreda, F., Cong, S., 2004. The distributed model intercomparison project (DMIP): motivation and experiment design. J. Hydrol. 298, 1-4, 4-26.

Strahler, A.N., 1952. Hypsometric (area–altitude) analysis of erosional topography. Geol Soc Amer Bull 63, 1117–1141.

Teuling, J. Lehner, I., Kirchner, J. W. & Seneviratne, S. I., 2010. Catchments as simple dynamical systems: Experience from a Swiss prealpine catchment. Water Resour. Res. 46, W10502, doi:10.1029/2009WR008777.

Todini, E., 1988. Rainfall-runoff modeling-past, present, and future. J. Hydrol 100, 341–352.

Todini, E., 1996. The Arno rainfall-runoff model. J. Hydrol. 175, 339-382.

Vrugt, J.A., Bouten, W., Gupta, H.V., Sorooshian, S., 2002. Toward improved identifiability of hydrologic model parameters: The information content of experimental data. Water Resour. Res. 38 (12), 1312, doi:10.1029/2001WR001118.

Vrugt, J.A., Ter Braak, C.J.F., Clark, M.P., Hyman, J.M., Robinson, B.A., 2008. Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation. Water Resour. Res. 44, W00B09, doi:10.1029/2007WR006720.

Vrugt J.A., Ter Braak C.J.F., Diks C.G.H., Robinson B.A., Hyman J.M., Higdon D., 2009, Accelerating Markov Chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. Journal of Nonlinear Sciences and Numerical Simulation 10(3), 273-290.

Vrugt, J.A., Ter Braak, C.J.F., Diks, C.G.H., Schoups, G., 2013. Hydrologic data assimilation using particle Markov chain Monte Carlo simulation: Theory, concepts and applications. Advances in Water Resources 51, 457–478.

Westerberg, I. K., Guerrero, J.-L., Younger, P. M., Beven, K. J., Seibert, J., Halldin, S., Freer, J. E., Xu, C.-Y., 2011. Calibration of hydrological models using flow-duration curves. Hydrol. Earth Syst. Sci., 15, 2205-2227.

Wigmosta, M.S., Vail L.W., Lettenmaier, D.P., 1994. A distributed hydrology-vegetation model for complex terrain. Water Resour. Res. 30 (6), 1665-1679.

Wigmosta, M.S., Lettenmaier, D.P., 1999. A comparison of simplified methods for routing topographically driven subsurface flow. Water Resour. Res. 35 (1), 255-264.

Wittenberg, H., Sivapalan, M., 1999. Watershed groundwater balance estimation using streamflow recession analysis and baseflow separation. J. Hydrol. 219, 20-33.

Wood, E.F., Sivapalan, M., Beven, K.J., B and, L., 1988. Effects of spatial variability and scale with implications to hydrologic modeling. J. Hydrol, 102, 29-47.

Wood, E.F., Sivapalan, M., Beven, K., 1990. Similarity and scale in catchment storm response. Rev. Geophys. 28, (1), 1-18.

Wood, E.F., Lettenmaier, D.P., Zartarian, V.G., 1992. A Land-Surface Hydrology Parameterization with subgrid variability for General Circulation Models. J. Geophys. Res. 97, 2717-2728.

Wrede, S., Seibert, J., Uhlenbrook, S., 2013. Distributed conceptual modelling in a Swedish lowland catchment: a multi-criteria model assessment. Hydrol. Res. doi:10.2166/nh.2012.056. (in press).

Zhao, R.J., Zuang, Y., Fang, L.R., Lin, X.R., Zhang, Q.S., 1980. The Xinanjiang model, in Hydrological Forecasting (Proc. Oxford Symp., April 1980), 351-356, IAHS Publ.

Zhao, R.-J., 1992. The Xinanjiang model applied in China. J. Hydrol. 135, 371-381.

Zięba, A., 2010. Effective number of observations and unbiased estimators of variance for autocorrelated data an overview. Metrol. Meas. Syst. XVII (1), 3-16.

Appendix A: Further details on the PDM

The actual storage S [L] is the sum of the unsaturated (S_{US}) and saturated (S_S) portions (Fig. 2a):

$$
\begin{aligned}
&S_{US}\left(t+\Delta t\right)=\left\{1-F\left(c^*\left(t+\Delta t\right)\right)\right\}c^*\left(t+\Delta t\right)\\
&S_{S}\left(t+\Delta t\right)=F\left(c^*\left(t+\Delta t\right)\right)c^*\left(t+\Delta t\right)-\int_{c_{\min}}^{c^*\left(t+\Delta t\right)}F\left(c\right)dc\\
&S\left(t+\Delta t\right)=c^*\left(t+\Delta t\right)-\int_{c_{\min}}^{c^*\left(t+\Delta t\right)}F\left(c\right)dc=\int_{c_{\min}}^{c^*\left(t+\Delta t\right)}\left\{1-F\left(c\right)\right\}dc=\int_{c_{\min}}^{c^*\left(t+\Delta t\right)}\left(1-c_n\right)^b dc
\end{aligned}
\tag{A.1}
$$

The total actual storage, S_T [L] for the grid cell is computed as:

$$S_T\left(t+\Delta t\right)=c_{\min}+S\left(t+\Delta t\right), \tag{A.2}$$

where $F(c^*(t))$ = probability ($c \leq c^*(t)$) indicates the fraction of grid cell with local storage capacity less than or equal to $c^*(t)$ and is saturated to generate runoff at time t (Fig. 2a and b). The c is the local storage capacity, c_n is the normalized storage capacity, $c_{\min}$ is the minimum local storage capacity, and 'b' is the shape parameter. Based on the 'equal storage redistribution of interacting storage elements' assumption, $c^*(t)$ is the critical store capacity at which all stores have water content of c^*, irrespective of their capacity, unless this is less than c^* when they will be full at time t (Moore, 1985). The maximum possible storage at saturation ($S_{\max}$ [L]) and the total maximum possible storage at saturation ($S_{T\max}$ [L]) for the grid cell are:

$$S_{\max}=\int_{c_{\min}}^{c_{\max}}\left\{1-F\left(c\right)\right\}dc=\int_{c_{\min}}^{c_{\max}}\left(1-c_n\right)^b dc \text{ and } S_{T\max}=c_{\min}+S_{\max} \tag{A.3}$$

The analytical solutions for the Pareto distribution are given as below:

$$S_{T\max}=c_{\min}+\int_{c_{\min}}^{c_{\max}}1-F\left(c\right)dc=c_{\min}+\int_{c_{\min}}^{c_{\max}}\left(\frac{c_{\max}-c}{c_{\max}-c_{\min}}\right)^b dc=\frac{bc_{\min}+c_{\max}}{b+1}$$

$$S_{\max}=S_{T\max}-c_{\min}=\frac{c_{\max}-c_{\min}}{b+1}$$

$$S_T\left(t\right)=c_{\min}+\int_{c_{\min}}^{c^*\left(t\right)}\left(\frac{c_{\max}-c}{c_{\max}-c_{\min}}\right)^b dc=S_{T\max}\left\{1-\left(\frac{c_{\max}-c^*\left(t\right)}{c_{\max}-c_{\min}}\right)^{b+1}\right\}$$

$$S(t) = \int_{c_{\min}}^{c^*(t)} \left(\frac{c_{\max} - c}{c_{\max} - c_{\min}} \right)^{b} dc = S_{\max} \left\{ 1 - \left(\frac{c_{\max} - c^*(t)}{c_{\max} - c_{\min}} \right)^{b+1} \right\}$$

$$c^*(t) = c_{\max} - \left\{ \left(c_{\max} - c_{\min} \right) \left[1 - \frac{S(t)}{S_{\max}} \right]^{\frac{1}{b+1}} \right\}$$