Predicting hydrology in ungauged basins using the DDD model

Martin Heltberg

Thesis submitted for the degree of
Master's in Hydrology and Glaciology
60 credits

Department of Geosciences
Faculty of Mathematics and Natural Sciences

UNIVERSITY OF OSLO

Spring 2021

Predicting Hydrology in Ungauged Basins Using the DDD Model

https://www.usgs.gov/media/images/rain-falling-a-watershed-drains-rivers-valley

Martin Heltberg

Predicting hydrology in ungauged basins using the DDD model

http://www.duo.uio.no/

Printed: Reprosentralen, Universitetet i Oslo

Abstract

Streamflow data from gauged catchments plays an important role for water resources management applications such as water resource planning, flood risk management and assessment of the impact of environmental- and climate change. A hydrological model successful at predicting in ungauged basins is needed for hydrological estimation for the million basins around the globe that are ungauged and has a great potential for better predicting the hydrological consequences of climate change.

This study aimed at evaluating the DDD model and its performance on predictions in ungauged basins, as well as comparing and evaluating different regionalization methods on catchments in Norway. Regionalization is defined as methods that allow for the transfer of hydrological information from gauged catchments to ungauged catchments. The comparison of methods was done at two levels, as a whole over Norway, as well as regionally for catchments in specific regions over Norway. This study shows that the new and improved DDD model is good at predicting hydrology in ungauged basins, with average Kling-Gupta efficiency values ranging from 0.7 up to 0.77 for the different regionalization methods. The different regionalization methods perform satisfactorily, with good KGE scores.

The best regionalization method to use was the multiple regression method, in which the average KGE value were 0.77, compared to 0.72 and 0.7 for the output average and parameter average, respectively.

Acknowledgements

First and foremost, I want to thank my supervisors, Thomas Skaugen and Chong-Yu Xu, for all their expert guidance and encouragements through this thesis. Thomas for your exceptional knowledge of the DDD model and its components, as well as great discussions. Chong-Yu for your outstanding competence in the field of hydrology. You have both been huge assets for the development of this thesis.

I also want to thank Norges vassdrags- og energidirektorat, especially you Thomas, for providing the data needed to complete this thesis, along with letting me present my thesis to them.

Lastly, I want to thank my friends and family for keeping me sane under these difficult circumstances.

Table of Contents

1 Introduction

Streamflow data from gauged catchments plays an important role for water resources management applications such as water resource planning, flood risk management and assessment of the impact of environmental change (Blöschl & Montanari, 2010). However, there is also needs for streamflow data in ungauged catchments, where no measurement equipment has been set up. To predict runoff at such ungauged sites has long been recognized as one of the major challenges in scientific and operational hydrology (Skaugen, 2018). A hydrological model successful at predicting runoff in ungauged basins is needed for hydrological estimations for the millions of basins around the globe that are ungauged. These basins need a good estimation on runoff, to not only check if the area is suited for things like a hydropower plant, but also to check the consequences of climate change. For catchments with streamflow data available, the runoff is commonly predicted by a hydrological model calibrated using the observed input and streamflow data. Having said that, many hydrological models does not directly work in catchments where the observed runoff data is unavailable for model calibration (He et al., 2011). Since most of the catchments around the world lack proper discharge measurements, the International Association of Hydrological Sciences (IAHS) launched a “Decade on Predictions in Ungauged Basins (PUB): 2003-2012” with the goal of improving the PUB (Sivapalan et al., 2003).

Throughout the decade, a wide range of new methods were developed to predict discharge in catchments lacking streamflow data (e.g., Merz & Blöschl, 2004 & Parajka et al., 2007). Meeting the challenge of predictions in ungauged basins (PUB) largely depends on the ability to extrapolate hydrologically relevant information from gauged catchments to the ungauged catchments, a process called regionalization. The gauged catchments are often called donor catchments whilst the ungauged catchments are called target catchments. Regionalization methods are considered the most standard approach to predicting runoff in ungauged basins (Hrachowitz et al., 2013). In this study two regionalization methods were used, physical similarity and regression. Physical similarity assumes that catchments with similar physical characteristics have similar hydrological response (Kokkonen et al., 2003). Therefore, hydrological prediction in ungauged basins using physical similarity borrows information from physically similar catchments. In this method, the model parameter set from the most physically similar donor catchments is transferred to the target catchment, using a similarity index (McIntyre et al., 2005). In the regression-based method functions are established between the model parameters and the catchment characteristics for the donor catchments. The assumption behind the regression is that there is a linear relationship between the

hydrological model parameters and the catchment characteristics, and that this relationship can be transferred from gauged to ungauged catchments (Poissant et al., 2017).

A review of the decade by Hrachowitz (2013) have been published and points out that even though not all goals for the decade has been reached, several insights on hydrological processes, data quality and use, principles of hydrological modeling and assessments of uncertainty were found. One of the important insights learned was the advantages of parameter parsimonious models for the PUB. A common problem in hydrological modeling is overparameterization, which makes parameter identification very difficult (Kirchner, 2006). This is obviously a huge problem in predictions for ungauged basins as the model parameters are often determined from catchment characteristics or hydrological information from other gauged basins (Yadav et al., 2007). Several studies, such as Seibert (1999) and Young (2006), discussed the advantages of having few, clearly identifiable model parameters.

Several approaches have been suggested and tried to make progress in predicting ungauged basins. Seibert (1999) used the Swedish Hydrologiska Byråns Vattenbalans model (HBV; Sælthun 1996), to investigate the regionalization of model parameters. The HBV model was calibrated for 11 catchments in Sweden, and the task was to relate the model parameters to catchment characteristics. He found out that only six of the 13 model parameters could be related to the catchment characteristics. It was further pointed out that the parameter uncertainty complicated the regionalization, and the suggestion was to include additional observed data into the calibration process. Including additional observed data into the calibration process was seen as a way to constrain the model parameters further. Other studies have also attempted such a procedure using the HBV model. Bergström et al. (2002) found a decrease in precision of the run-off simulation when water quality data is included in the calibration, while Parajka & Blöschl (2008) found a small increase in precision when optical satellite scenes of snow cover were included in the calibration. Merz & Blöschl (2004) did a large-scale study where they wanted to regionalize the model parameters of the HBV model. When the HBV-model was calibrated for 308 catchments in Austria, only weak correlations were found between the model parameters and the catchment characteristics, and their conclusion was that it is extremely difficult, if at all possible to find universal relationships between the catchment characteristics and the model parameters. Young (2006) used a large dataset of 260 catchments in the United Kingdom to look at the regionalization in the Probability-Distributed Moisture (PDM) model (Moore, 1985). This model only had six model parameters to regionalize compared to the 13 in the HBV model. The study looked at two regionalization methods, relating model parameters to the

catchment characteristics using regression and nearest neighbor method. The regression of the model parameters had the best results, and the fact that there were fewer model parameters in the PDM model could have played a role in this. Yadav et al. (2007) recognized that the structural errors in conceptual models and the non-identifiability of model parameters could pose a serious threat against the ability of these models to give good PUB with regionalized parameters.

The Distance Distribution Dynamics (DDD) model (Skaugen & Onof, 2014) is a parameter parsimonious model, which hopefully can relate model parameters to catchment characteristics. A parameter parsimonious model is a model that attains a desired level of goodness of fit using as few parameters as possible (Vandekerckhove et al., 2015). That means that a parsimonious model explain data with a minimum of parameters. Skaugen et al. (2015) used the DDD model to make predictions in ungauged basins in Norway. Model parameters were estimated using the multiple regression method. The study showed that the DDD model works well with regionalized parameters and achieved good results for the predictions. Tsegaw (2019) also used the DDD model in predictions of ungauged basins, however that was done on an hourly time resolution compared to the daily resolution which is being done in this study. That study did not only use the multiple regression method, but also included the regionalization method called physical similarity. The DDD model worked well using both regionalization methods. Yang et al. (2018 & 2020) looked at different regionalization methods for different hydrological models and showed that having fewer model parameters could benefit the prediction in ungauged basins.
In this study, DDD model parameters have been regionalized in order to predict runoff at ungauged catchments all over Norway for a 24-hourly resolution. This study continuous the work done in Skaugen et al. (2015). Since that study in 2015 the DDD model has been further developed and more processes in the model, such as snowmelt and evapotranspiration, has an improved calibration-free physical founding. Both snowmelt and evapotranspiration are now calculated using an energy balance approach, with proxy models for the energy balance elements driven by temperature and precipitation (Skaugen et al., 2020). Comparatively, evapotranspiration used to be calculated using a degree-day factor.
The objectives for the study are the following:

- Evaluating the DDD models performance on predictions in ungauged basins.
- Evaluating whether the multiple regression method or the physical similarity method is better for predicting ungauged basins in the DDD model.
- Analyzing the difference of the two different sub-methods in the physical similarity method, output average method and the parameter average method.

2 Study Area and Data

2.1 Study Area

Figure 1: Map of all runoff stations used in study, the donor stations are in blue and the target stations are in yellow.

All the 107 catchments in this study are located in Norway. Norway is a country situated in northern Europe, along the northwestern part of the Scandinavian peninsula covering an area of about 385 000 km^2 (excluding Svalbard and Jan Mayen). Norway has a long, rugged coastline, with altitudes ranging from 58^o to 71^{o}N. Additionally, the elevation spans from sea level to 2469 m.a.sl. The climatic conditions vary considerably within the country, from a wet maritime climate along the coastline to a drier climate in the interior. The mean temperatures range from around 7^{o}C in the southern part to about -2^{o}C in the north, as well as the high-altitude areas in the central parts of the country. The average annual precipitation also has large spatial variations, with a countrywide average of about 1000 mm/year (Yang et al., 2018). In the southwestern part of Norway, the average annual precipitation exceeding 3500 mm/year, while the inland region to the east only receives about 700 mm/year. Due to this the hydrographs in Norway have quite distinct spatial patterns.

2.1.1 Hydrological Regions

To divide Norway into regions, Gottschalk et al. (1979) was used as a roadmap. The basis for the division is when and how the high flow and the low flow occurs. The high flow was divided into three categories, H1 to H3, depending on when the high flow occurred. H1 is dominated by snowmelt, with the high flow in the spring and early summer, H2 has a secondary high flow in the autumn due to a high amount of rainwater, and H3 is totally dominated by rain, high flow in the late autumn and early winter. The low water was divided the same way with L1 being low water in the winter due to snow accumulation, L2 occurs when the two months with the lowest flow not belonging to the same time of the year, and L3 happens when the low flow is in the summer due to high evaporation and low precipitation (Gottschalk, 1979).

As seen in Figure 2, Gottschalk divided Norway into 5 different runoff regions. The different regions each got their own name, a mountainous region (H1L1 in Figure 2), an inland region (H2L1) , a transition region (H2L2), a Trøndelag region (H2L3), and a coastal region (H3L3). The mountain region is the inner and northeast parts of Norway, the inland region is between the mountain region and the west coast, the transition region follows a narrow band outside the inland region and covers parts of the southern inlands. The Trøndelag region covers parts of Møre and Trøndelag, and the coastal region follows the frontier coast of Norway. At the Norwegian west coast there is a rapid, but a gradual change when moving inland. This means that regions such as the transition region and the Trøndelag region are very narrow.

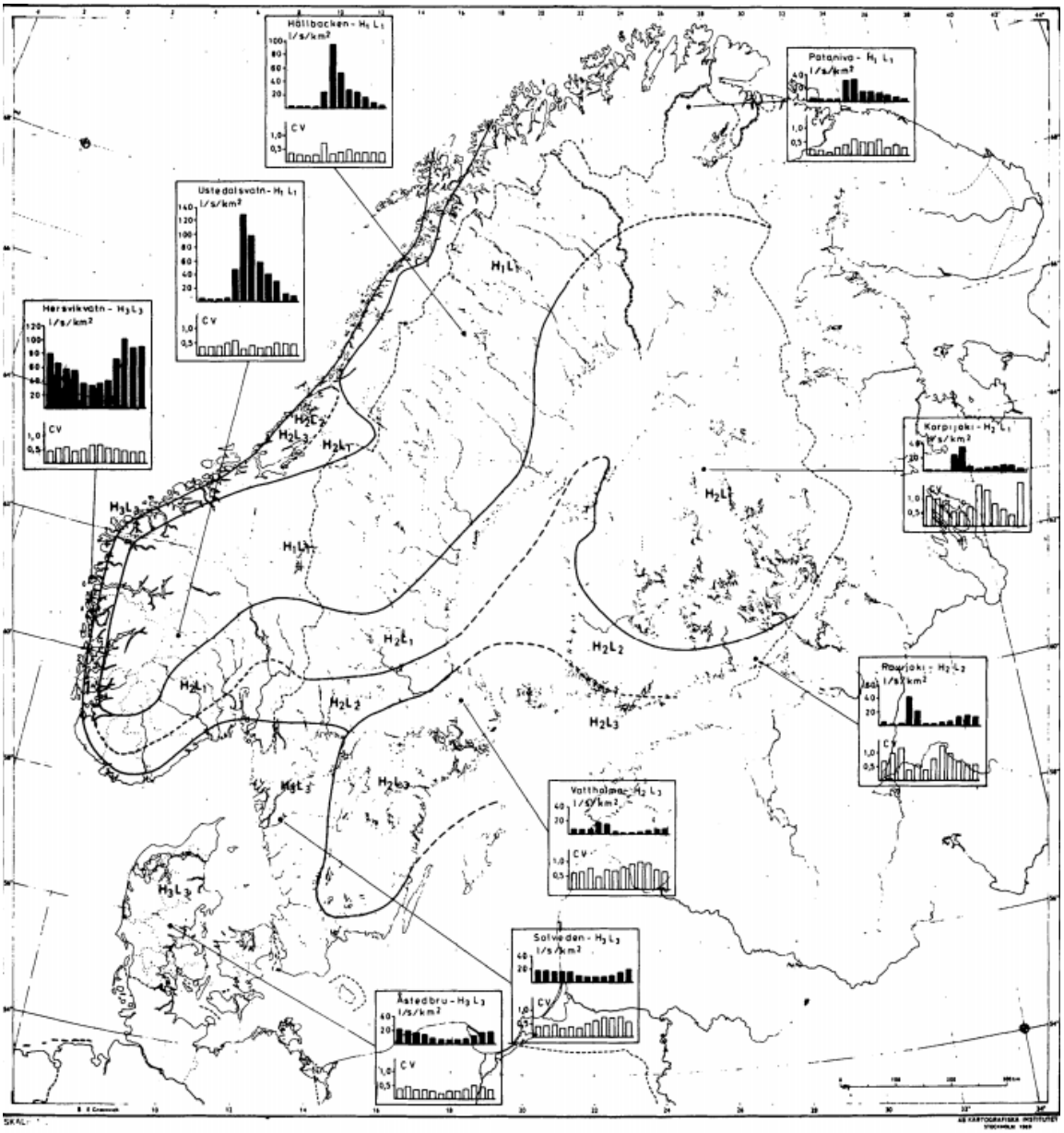


Figure 2: Map of runoff regions in the Nordic countries (Taken from Gottschalk et al. (1979)).

For each of the five regions, one donor and one target catchment were selected. This was done to get a better overview of how the regionalization methods work for the different runoff regimes in Norway.

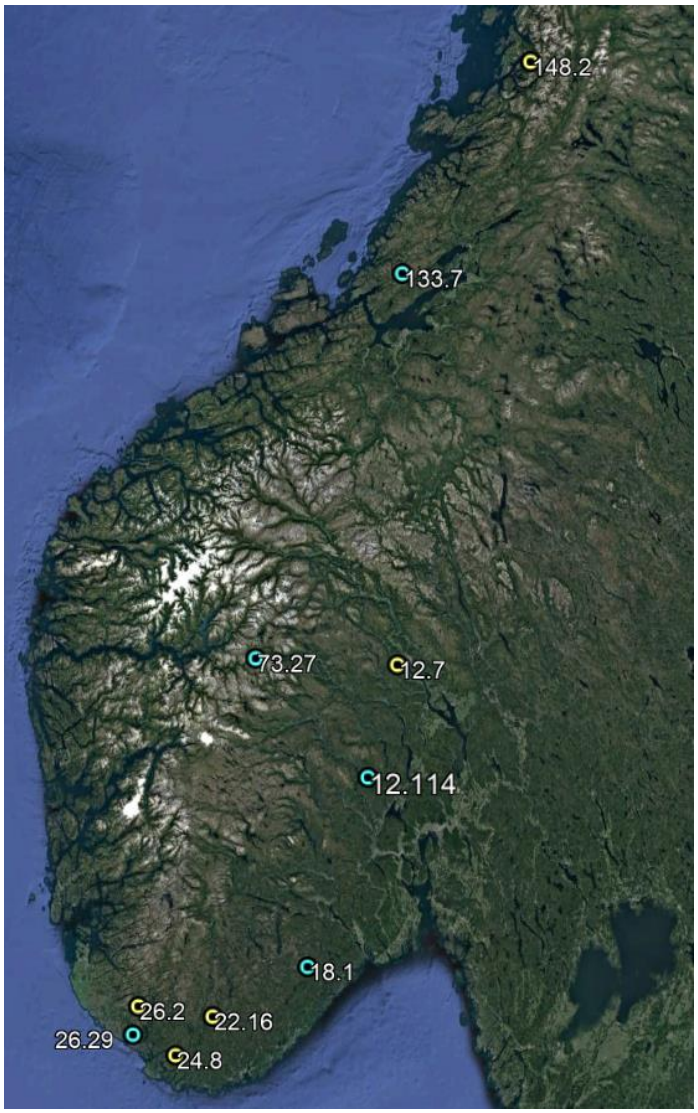


Figure 3: Location of the 5 donor and 5 target catchments chosen for regional checks. The donor catchments are in blue and the target catchments are in yellow.

The catchments chosen for each of the regions are shown in Figure 3. Due to the possible inaccuracies of choosing simply based on location, all hydrographs of the catchments were visually inspected. The catchments were then placed into the region who best fulfilled the definitions in Gottschalk (1979). Additionally, every target catchment was placed into a region, this was done to get an overview of the spatial distribution of the results. The regions and how many catchments they contain can be seen in Table 1.

Table 1: Overview of the catchments used for regional validation.

Region	Donor Catchment	Target Catchment	# of targets
Coastal (H3L3)	26.29	24.8	4
Transition (H2L2)	18.1	26.20	4
Trøndelag (H2L3)	133.7	148.2	1
Inland (H2L1)	12.114	22.16	8
Mountain (H1L1)	73.27	12.70	8

The number of catchments in each region varies greatly. From eight in the mountain and inland region to only one in the Trøndelag region. This is explained by the fact that much of Norway's geography, and thus its catchments, is dominated by the mountainous regions and the inland region. As seen in Figure 1, the majority of the target catchments used in this study covers the southern part of Norway, the part with the most mountains and inland. Some catchments fall under the other regions. The Trøndelag region was originally defined, in the Gottschalk (1979) paper as the Baltic region, for regions with similar runoff regimes as the Baltic states, but its definition also covers part of the coast in the central parts of Norway, more specifically Trøndelag, and is only covered by one catchment in this study.

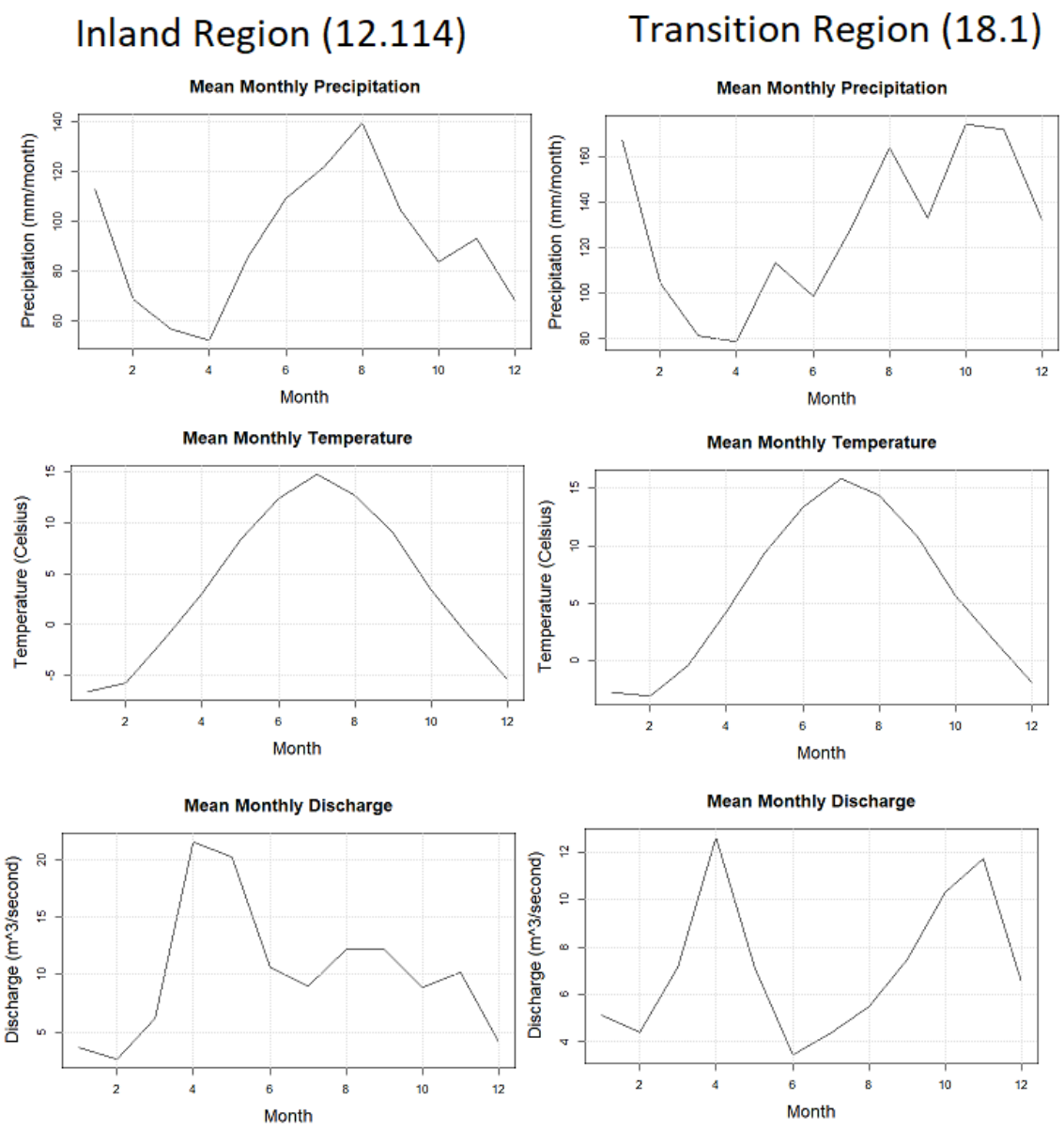
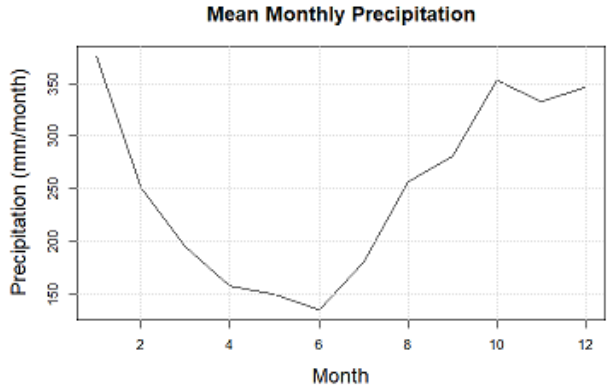


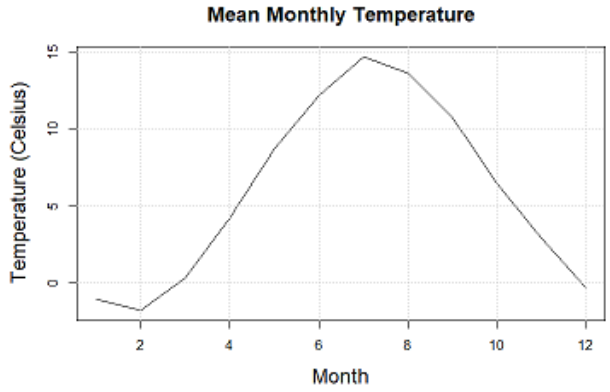
Figure 4: Mean monthly precipitation, temperature, and discharge for the inland and the transition regions.

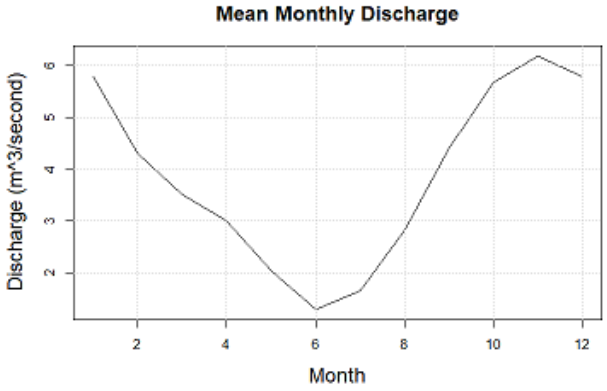
Figure 4 shows the mean monthly precipitation, temperature, and discharge for the two stations in the inland and the transition regions. The inland regions high flow is in the spring, during the snowmelt season, but is defined by the second or third highest monthly runoff taking place in the autumn. This can be seen as the September discharge is the largest high flow after the snowmelt. During the winter, the flow is lower, due to precipitation being stored as snow. This is seen as the lowest flows occurs in December, January, and February. The transition region is also defined as having the second or third highest runoff in the autumn, which can be seen in Figure 4 as the mean monthly discharge in November is the second highest discharge. The transition region is also defined as having the two months

with the lowest discharge not belonging to the same time of the year. This can be seen in Figure 4 as June and February have the lowest flows.

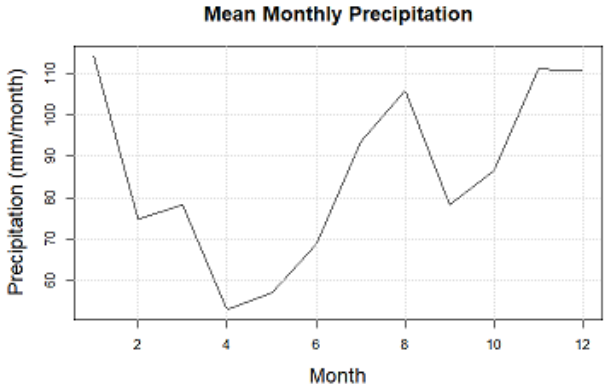
Coastal Region (26.29)

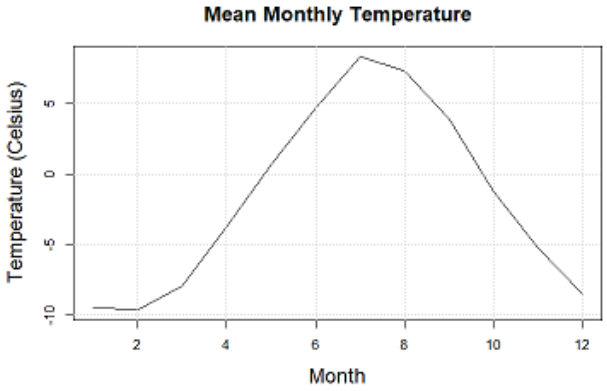


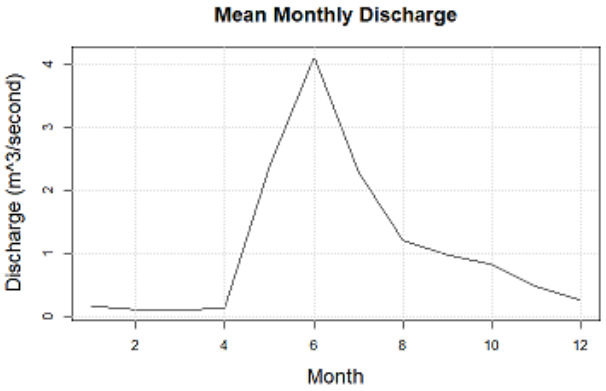




Mountian Region (73.27)







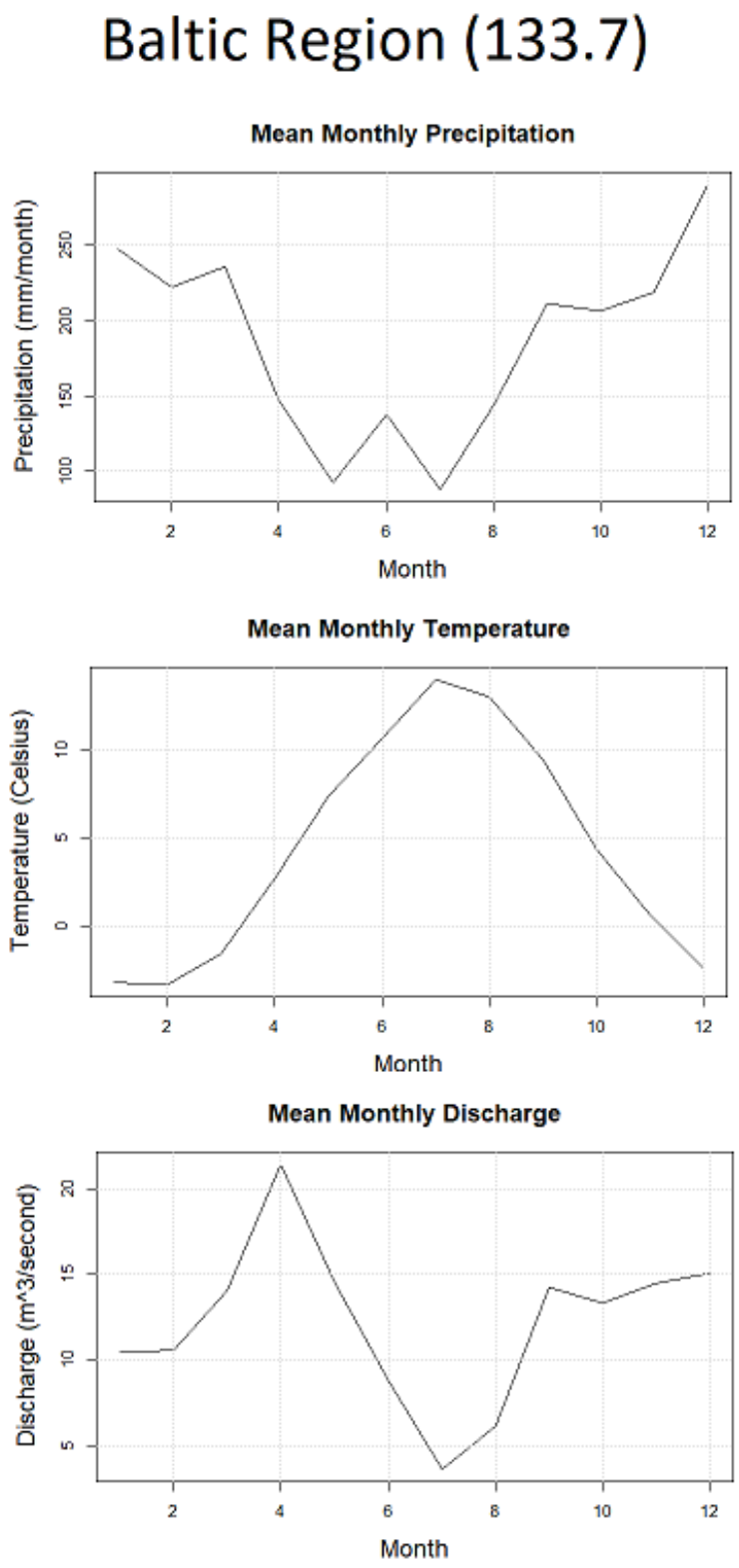


Figure 5: Mean monthly precipitation, temperature, and discharge for the coastal, mountain and the Trøndelag regions.

Figure 5 shows the mean monthly precipitation, temperature, and discharge for the coastal region, mountain region and the Trøndelag region. The coastal regions high flows are dominated by rain, the highest discharge takes place in the autumn and early winter, in this

case November and December. The low flow on the other hand happens in the summer and is caused by a combination of low precipitation and high evaporation. In Figure 5 the low flows occur in June and July.
The mountain region has its high flow dominated by snowmelt, the three months with the highest discharge belong to the late spring and early summer. The low flow occurs in the winter and is caused by snow accumulation, with virtually no runoff in the winter months.
The Trøndelag region is defined by having its second or third highest flow in the autumn due to rain, with the highest coming in the snowmelting season. The low flow occurs in the summer, due to low precipitation and high evaporation, which here can be seen with low flows in July and August.

2.2 Data

In this study 107 gauged donor catchments, used for regionalization, and 25 ungauged target catchments have been used. The stations are located all over Norway, with most of the target catchments in the southern half of Norway. All of these stations have continuous meteorological and discharge data during the period 1999-2018, and this period is thus used in this study. The size of the catchments varies from 2,84 km^2 to 5543,98 km^2, and 85 out of the 107 catchments are under 500 km^2. Figure 1 shows the location of the center for each catchment.

Time series of precipitation, temperature and discharge are the main input data for running and calibrating the model. Precipitation and temperature data are extracted from a 1x1 km gridded product of the Norwegian Meteorological Institute, with 24-hourly temporal resolution the seNorge_2018v20.05 (Lussana, 2020).
24-hourly discharge data has been obtained from the Norwegian Water Resources and Energy Directorates (NVE) Hydra II database.

3 Methods

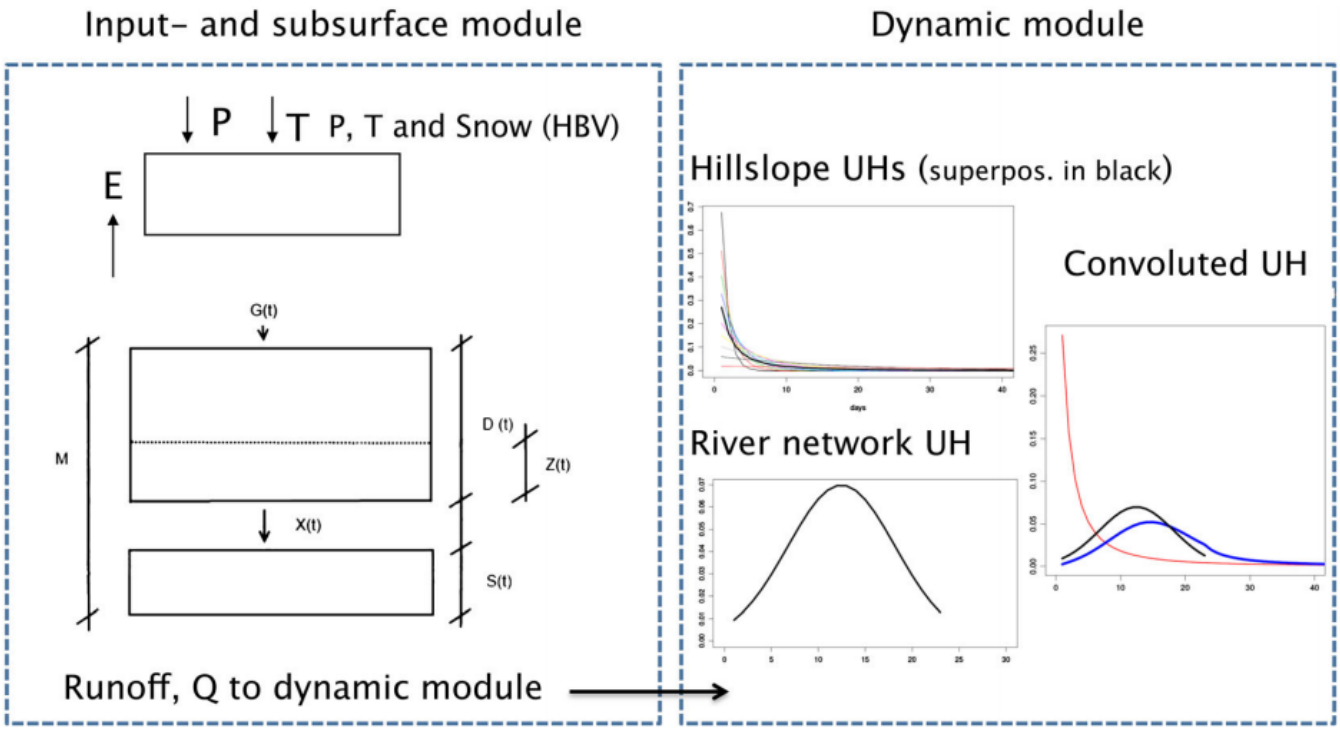


Figure 6: The model structure of the Distance Distribution Dynamics model. In the representation of the subsurface water reservoir M (bottom left), the dotted horizontal line is the actual level Z of soil moisture in D. The ratio of (G(t) + Z(t))/D controls the release of excess water to S and hence to discharge. Note that D, S and Z are functions of time, while M is fixed. In the dynamic module (right), the superpositioned hillslope unit hydrograph is convoluted with the river network unit hydrograph to give runoff. Sketch borrowed from Skaugen & Onof (2014).

Figure 6 shows a schematic overview of the model structure in the DDD model, which is further explained below.

3.1 The Model Structure.

The DDD model (Skaugen & Onof, 2014; Skaugen et al. 2015) is a physically based, rainfall-runoff model written in the Julia programming language (Shah et al., 2017) and currently runs operationally with daily and 3-hourly timesteps at the Norwegian flood forecasting service at NVE. It is also an example of a parameter parsimonious model. The inputs to the model are precipitation and temperature data. The DDD model is an extension of the unit hydrograph method (Skaugen & Onof, 2014). The unit hydrograph of a catchment is defined as the direct runoff hydrograph resulting from one unit volume of excess rainfall at a constant intensity and uniformly distributed over the catchment area for a duration of time (Ramirez, 2000). The duration of the unit hydrograph, or the response time, is the time it takes for the water that is the furthest away from the outlet to reach the outlet.

The two main modules of the DDD model are the hydrological module and the runoff dynamics module (Skaugen & Onof, 2014). The hydrological module estimates the amount

of water that enters the runoff dynamics module, and the runoff dynamics module uses distance distribution analysis to describe the transport of water in the catchment, from hillslope to river network and from the river network to the outlet (Stavang, 2019).

3.2 Hydrologic Module

The volume capacity of the subsurface water reservoir, M [mm], is shared between an unsaturated zone with volume D [mm], called the soil moisture zone and a saturated zone with volume S_S [mm], called the groundwater zone (Skaugen & Onof, 2014). The actual volume of water present in the unsaturated zone D, is called Z [mm]. The subsurface state variables are updated after evaluating whether the current soil moisture Z(t), together with the input of rain, snowmelt, and discharge.

Firstly, in the DDD model, the precipitation is distributed to the 10 elevation zones of equal area defined for each catchment. To distinguish between precipitation as snow or rain the following is computed,

$$If\ T > TX, P_{rain} = P * pkorr, P_{Snow} = 0 \quad (1)$$

$$If\ T < TX, P_{rain} = 0, P_{Snow} = P * skorr \quad (2)$$

In which, T [ºC] is the observed temperature, TX [ºC] is the temperature threshold for determining is precipitation is snow or rain, P [mm] is the amount of precipitation and the pkorr [-] and skorr [-] model parameters are the correction factors for precipitation as rain and snow, respectively.

Input of water can, additionally to precipitation, come from snowmelt. The estimation of meltwater is done using an energy balance equation.

$$SWE = \frac{K+H+L_{net}+LE+GH+R-CC}{\lambda_F \rho_W} \quad (3)$$

In which SWE [m] is the change in the snowpack's water equivalent, K [kJm^{-2}] is the net shortwave radiation, H [kJm^{-2}] is the sensible heat exchange, L_{net} [kJm^{-2}] is the net longwave radiation, LE [kJm^{-2}] is the latent heat exchange, GH [kJm^{-2}] is the heat from the ground, R [kJm^{-2}] is the heat from the precipitation and CC [kJm^{-2}] is the snowpack heat storage. λ_F [$kJkg^{-1}$] is the latent heat fusion and ρ_W [1000 kgm^{-3}] is the density of water (Dunne, 1976).

To estimate the potential evapotranspiration, E_P [mm day^{-1}], the DDD model uses the Priestley-Taylor method.

$$E_P = a_{PT} * \left(\frac{\delta}{\delta+y}\right) * (K + L_{net}) * \left(\frac{1000}{LE*\rho_W}\right) \quad (4)$$

Wherein, a_{PT} [-] is the empirical Priestley-Taylor constant, δ [kPa °C^{-1}] is the slope of saturation vapor pressure-temperature relation, y [kPa °C^{-1}]is the psychometric constant, K [kJm^{-2}] is the net shortwave radiation, L_{net} [kJm^{-2}] is the net longwave radiation, LE [kJm^{-2}] is the latent heat exchange, and ρ_W [1000 kgm^{-3}] is the density of water (Priestley & Taylor, 1972).

The actual evapotranspiration, E_A [mm day^{-1}], is the evaporation potential, scaled non-linearly by the amount of water that is in the subsurface (Skaugen et al., 2020).

$$E_a = \min\left(E_P, E_P\left(1 - \exp\left(-4 * \frac{M - D + Z + G}{M}\right)\right)\right) \quad (5)$$

Where M [mm] is the capacity of the subsurface, D [mm] is the unsaturated zone, Z [mm] is the soil moisture and G [mm] is the precipitation/snowmelt. Having a non-linear method to calculate the actual evapotranspiration as a function of the degree of soil saturation has previously been proposed by Chanzy and Bruckler (1993), Arnell (2002) and Skaugen et al. (2020).

When the input water, I, reaches the unsaturated zone, D, the water is added to the volume Z. The movement of water from D to S_S happens when the actual water content, Z, reaches the field capacity, R, of 0.3 of the capacity of D. The field capacity is fixed at 30 % (R=0.3) of D(t) (Skaugen & Onof, 2014).

$$\frac{dZ}{dt} = G(t) - X(t) \quad (6)$$

Where X is the excess water transferred to the saturated zone, when the field capacity of 0.3 is reached.

$$X(t) = Max\left\{\frac{G(t) + Z(t)}{D(t)} - R, 0\right\} D(t) \quad (7)$$

The excess water, X, is then added to the volume of the saturated zone, S_S.

$$\frac{dS_s}{dt} = X(t) - Q(t) - Ea(t) \quad (8)$$

In which Q is the water output from the saturated zone.

3.3 Runoff Dynamics Module

To assess the direct runoff, Q, at the outlet, the convolution integral of the excess water from the hillslope Q_h and the unit hydrograph of the river network, u_r, is used.

$$Q(t) = \int_0^t Q_h(\tau) u_r(t-\tau) d\tau \tag{9}$$

Where Q_h is the excess water contributed by the hillslope. It is estimated using the convolution integral of excess water, X, and the unit hydrograph of the hillslope, u_h.

$$Q_h(t) = \int_0^t X(\tau) u_h(t-\tau) d\tau \tag{10}$$

3.3.1 Deriving the Unit Hydrograph for Hillslope

The unit hydrograph is derived from recession analysis and distance distribution. The distance distribution is defined as the distribution of distances from points in the catchment to the river network (Stavang, 2019). The distribution is modelled as an exponential distribution with a cumulative distribution function.

$$U_h(d) = 1 - e^{-\Psi(d-d_0)} \tag{11}$$

In which Ψ is the parameter of the exponential function. If the velocity of water down a hillslope is constant, then Δd is the distance travelled by the water during timestep Δt (Skaugen & Onof, 2014). The distance distribution then becomes a distribution of travel times, where the response time is $t_{h,\,max}$, which is defined in equation 15.

$$U_h(t) = 1 - e^{-\Psi(t-t_0)} \tag{12}$$

The derivative of U_h is the unit hydrograph of the hillslope, u_h, where Ψ is the parameter of the exponential function. The unit hydrograph is then defined as

$$u_h(t) = \Psi e^{-\Psi(t-t_0)} \tag{13}$$

3.3.2 Deriving the Unit Hydrograph for Each Subsurface Saturation Level

The DDD model assumes different levels of saturation, i=1,2,…,I, based on the distribution of Λ. The saturation levels all have different velocities and recession characteristics. The level specific unit hydrograph is,

$$u_{h,i}(t) = \lambda_i e^{-\lambda_i(t-t_0)} \tag{14}$$

Where λ_i is the specific recession characteristics of that saturation level. The $u_{h,i}$ is then further cut into intervals, f_i = 1,2,…,F_i, to acquire the weights, $\mu_{i,j}$, that each level contributes at each time step. For each level, the number of intervals is,

$$f_i = \frac{t_{h,i,max}}{\Delta t} \tag{15}$$

Where $t_{h,\,i,\,max}$ is defined as

$$t_{h,i,max} = \frac{d_{max}}{v(i)_h} \tag{16}$$

In which d_{max} is the maximum distance in the distance distribution and $v(i)_h$ is the level specific celerity. The maximum amount of time intervals, f_i, is the time it takes for the wave furthest away to reach the river network. The weight of each time interval is defined as

$$\mu_{i,f} = \int_{(f-1)\Delta t}^{(f)\Delta t} u_{h,i}(t)dt \tag{17}$$

The sum of all weights should be equal to 1. The runoff from each interval and saturation level is then,

$$Q_h(f\Delta t) = \frac{1}{I}\sum_{i=1}^{I} X(\Delta t)\xi_{i,j}\mu_{i,j} \tag{18}$$

Where ξ are the weights distributing X to each saturation level, i, and interval, f, depending on the degree of saturation for each level.

3.3.3 Estimating Celerity of Subsurface Flow

To estimate the celerities needed for estimations of the temporal scale of the travel time distribution, and thus the unit hydrographs, associated with the different levels of saturation, the runoff series has to be investigated. Specifically, the recession data is investigated under the assumption that one can observe the superpositioned runoff response composed of contributions from several different saturation levels (Skaugen & Onof, 2014). Furthermore, one can assume that this hydrograph too can be approximated by an exponential unit hydrograph. The exponential recession curve has been a quite popular choice for runoff modelling and base-flow recession for many decades, also the assumption of considering runoff recession as superpositioned exponential unit hydrographs has a long history (Tallaksen, 1995). The procedure where the superpositioned unit hydrographs are parameterized from levels of saturation was first used by Skaugen & Onof (2014).
The recession runoff can be modelled as $Q(t) = Q_0 \Lambda e^{-\Lambda(t-t0)}$, where Q_0 is the peak discharge immediately before the recession starts. Therefore, one can determine the parameter Λ from equation 19.

$$\Lambda = \frac{\log\big(Q(t)\big) - \log\left(Q(t+\Delta t)\right)}{\Delta t} \tag{19}$$

A high value for Λ implies a large change in storage and a high celerity and vice versa. The parameter Λ is therefore the slope per Δt of the recession. Since the mean of the distance distribution is $\bar{d}$, the mean of the travel time distribution is $\frac{\bar{d}}{v_h}$. Furthermore, the exponential unit hydrograph with mean $\Delta t \frac{1}{\Lambda}$ conveys the same travel time distribution, so that one can write,

$$\Delta t \frac{1}{\Lambda} = \frac{\bar{d}}{v_h} \tag{20}$$

Where $\bar{d}$ is the mean of the distance distribution v_h [m/s] is the celerity and Λ is the slope of the recession curve. The celerity is thus associated with Λ as

$$v_h = \frac{\Lambda \bar{d}}{\Delta t} \tag{21}$$

Because one can assume that the variability of Λ is due to the different levels of saturation in the catchment, it follows that events during saturated conditions will give out the highest values of Λ, and by equation 20, also the highest values of the celerity v_h. If the highest level of saturation defines the subsurface water reservoir *M*, one can let the probability space of the distribution of Λ also represent the saturation level of the reservoir M,

$$F(\Lambda) = \frac{S}{M} \tag{22}$$

Since *M* is divided into saturation levels, i=1,2,…,I, one can let the probability F(Λ) detect which levels, *L*, yields runoff,

$$L = G \; x \; F(\Lambda) \tag{23}$$

Which has to include all levels up to I due to the fact that levels are saturated from below. If F(Λ) = 1, the subsurface soil reservoir is fully saturated and L=G.
If, on the other hand, there is a saturation shortage, F(Λ) < 1, only the levels i= 1,2,…,L produces runoff.
If the distribution of Λ is estimated by an exponential distribution, and Λ_i is estimated so that $F(\Lambda_i)=i/I$, Λ_i is then considered to be the parameter of the unit hydrograph resulting from the super-positioning of the unit hydrographs for all levels below and up to i:

$$\Lambda_i e^{-\Lambda_i (t-t_o)} = w_1 \lambda_1 e^{-\lambda_1 (t-t_o)} + w_2 \lambda_2 e^{-\lambda_2 (t-t_o)} + \cdots + w_{i+1} \lambda_{i+1} e^{-\lambda_{i+1} (t-t_o)} \tag{24}$$

The variables, *w*, are weights corresponding with each saturation level. The weights show the discharge produced by the level specific unit hydrographs, the response for each level for a given saturation. In each level the discharge is directly proportional to the celerity, which is again proportional with *λ,* so equal weights for different levels are not appropriate, but should

rather increase with increasing saturation levels. To estimate the weights, for each level *i*, one can use the empirical distribution of Λ which is derived from observed recession and can be seen in equation 25.

$$w = \frac{\Lambda_i}{\sum_{k=1}^{i} \Lambda_k} \tag{25}$$

The parameter λ for the different saturation levels can be estimated by solving equation 24 for the increasing levels under the assumption that $\Lambda_1 = \lambda_1$. The celerity for the individual levels, $v_h(i)$ can be estimated by equation 21 using λ_i instead of Λ:

$$v_h(i) = \frac{\lambda_i \vec{d}}{\Delta t} \tag{26}$$

3.3.4 Distribution of Saturation Level and Estimation of the Mean Storage

In the DDD model, it is assumed that the variability in celerity for each level is due to the variation in storage and saturation degree. Higher saturation gives higher celerity. This is shown by a greater difference in runoff values in equation 19. Λ is modelled as a gamma distribution.

$$f(\Lambda) = \frac{1}{\beta^a \Gamma(\mathrm{a})} \Lambda^{a-1} e^{-\frac{\Lambda}{\beta}}, \quad a > 0, \quad \beta > 0 \tag{27}$$

Where a is the shape parameter and β is the scale parameter of the gamma distribution. This distribution reflects the distribution of the saturation levels, so the saturation levels are also showed as a gamma distribution.

$$f(S_s) = \frac{1}{\eta^a \Gamma(\mathrm{a})} S_s^{a-1} e^{-\frac{S_s}{\eta}}, \quad a > 0, \quad \eta > 0 \tag{28}$$

Where η = β/c and c=$\bar{\Lambda}$/m_s. a is the shape parameter, η is the scale parameter and m_s is the mean storage. The mean storage is calculated through the daily excess water transferred to the saturated zone, X, which depends on the mean annual discharge (MAD) [m^3s^{-1}] and the area of the catchment [m^2].

$$X = \frac{1000 * MAD * 86\,400\ sec}{area} \tag{29}$$

The total sum of X after F days is if a steady state is reached,

$$F * X = S_{ss} + Q_{ss} \tag{30}$$

Wherein Q_{ss} is the total runoff after F days and S_{ss} is the water still in the soil and therefore the mean storage, m_s.

$$Q_{ss} = \sum_{k=1}^{F} \sum_{f=1}^{k} X * \mu(\bar{\Lambda})_j \tag{31}$$

$$S_{ss} = \sum_{k=1}^{F-1} \sum_{f=k+1}^{F} X * \mu(\bar{\Lambda})_j \tag{32}$$

The distribution of each level, S_i, is estimated as quantiles of $f(S_s)$ where the subsurface capacity, M, is the 99 % quantile of the distribution of S_s.

$$\frac{S_i}{M} = \int_0^{S_i} \frac{1}{\eta^a \Gamma(\mathrm{a})} S_s^{a-1} e^{-\frac{S_s}{\eta}} dS_S \tag{33}$$

3.3.5 Deriving the Unit Hydrograph of the River Network

The same principles as presented above can be implemented for the derivation of the river network hydrograph, U_r. The distance distribution of points in the river network to the outlet of the catchment is determined as the distance from point in the river network to the outlet. One can then use the mean celerity of the river network, v_r, to transform the function to a distribution of travel times. The unit hydrograph of the river, u_r, is the derivative of the cumulative distribution function of travel with

$$t_{r,max} = \frac{R_{d\,max}}{v_r R_{dmax}} \tag{34}$$

Where v_r is the mean celerity of the river network, $R_{d\,max}$ is the maximum distance measured from the outlet to the river network.

As the max distance from the points in the network to the outlet.

$$U_r(t) = \int_0^{S_i} u_r(t)dt\,, \qquad U_r\left(t_{r,max}\right) = 1 \tag{35}$$

Where U_r is the river network hydrograph and u_r is the unit hydrograph of the river.

3.4 Model Parameters and Calibration

Model parameters in the DDD model include estimations from GIS analysis, fixed values, and regionalized parameters. All the parameters and how they are determined are shown in Table 2.

Table 2: Overview of the parameters in the DDD model

Parameter	Description	Method of Estimation
Hypsographic Curve	11 values describing the quantiles, 0, 10,…, 100	GIS
Mean Elevation	Mean elevation of the catchment	GIS
Pkorr	Correction factor for precipitation	Regionalization

Skorr	Correction factor for precipitation as snow	Regionalization
U	Mean windspeed	Regionalization
Pro	Maximum liquid water content of snow	Regionalization
TX	Threshold temperature rain/snow	Regionalization
CGLAC	Degree-day factor for glacial melt	Regionalization
a0	Parameter for new spatial distribution of SWE, shape parameter	Estimated from observed spatial variability of precipitation
d	Parameter for new spatial distribution of SWE, decorrelation length	Estimated from observed spatial variability of precipitation
MAD	Long term mean annual discharge	GIS
Area	Catchment area	GIS
maxLbog	Max distance of marshland portion of hillslope	GIS
midLbog	Mean distance of marshland portion of hillslope	GIS
BogFrac	Areal fraction of marshland	GIS
zsoil	Areal fraction for soils (area with distance zero to the river)	GIS
zbog	Areal fraction for marsh land (area with distance zero to the river)	GIS
R	Parameter defining field capacity	GIS
GshInt	Shape parameter	Regionalization
GscInt	Scale parameter	Regionalization
rv	Celerity for riverflow	Fixed/Regionalization
midFL	Mean distance for river network	GIS
stdFL	Standard deviation of distance for river network	GIS
maxFL	Max distance for river network	GIS
maxDL	Max distance of non-marsh land of hillslope	GIS
midDL	Mean distance of non-marsh land of hillslope	GIS
Glacfrac	Fraction of glaciers in catchment	GIS
midGL	Mean distance for glacier	GIS
stdGL	Standard deviation of distance for glacier	GIS
maxGL	Max distance for glacier	GIS
Glacier	Areal fraction of glaciers in 10 elevation zones	GIS

As seen in Table 2, most of the model parameters are estimated using GIS, but eight parameters need to be estimated from the relationship between model parameters and catchment characteristics. The parameter *rv* could also be estimated using regionalization but is a fixed value due to the celerity of the riverflow having little to no effect on the results on a 24-hourly scale.

The calibration of the model is performed using the global optimization package BlackBoxoptim. The Kling-Gupta efficiency criteria (KGE) has been used as an objective function for the calibration (Gupta et al., 2009). The KGE formula used in this study is:

$$KGE = 1 - ((sqrt(r-1)^2 + (y-1)^2 + (\beta-1)^2) \quad (36)$$

With r being the linear correlation coefficient between Q[sim] and Q[obs], $y = \frac{CV[sim]}{CV[obs]}$ and $\beta = \frac{\mu[sim]}{\mu[obs]}$. KGE ranges from a perfect score of 1 to negative infinity and assesses the error in the mean bias and bias in the variance, in addition to the correlation between observed and simulated values. The widely used Nash-Sutcliffe efficiency criterion only assesses the correlation, while factors such as the bias must be assessed separately.

$$Bias = \frac{Q_{Sim} - Q_{Obs}}{Q_{Obs}} \quad (37)$$

Both KGE and bias, which is the ratio of the mean simulated to observed discharge, has been used to evaluate the calibration results and the PUB results.

3.5 Sensitivity of Calibrated Parameters

Sensitivity analysis is the study of how output response of a model is affected by input uncertainty (Zhang et al., 2015). As most of the parameters are determined by GIS, there are only eight parameters needed to be analyzed for sensitivity.

Parameter sensitivity plots was also produced for several of the catchments in the study. This is essentially a plot with KGE values on the y-axis and relative error of parameter values on the x-axis. The relative error is found by this equation:

$$Relative\ Error = \frac{a_t - a_{opt}}{a_{opt}} \quad (38)$$

Where a_{opt} is the optimal parameter value and a_t is the parameter value that is being evaluated. In this study this was done by varying the optimal parameter by 10 %, both increasing and decreasing, up until 50 %.

Dottyplots were also created for all catchments. A dottyplot illustrates the variation in KGE as a function of the variation of the model parameters for each catchment (Tegegne et al., 2017). A sharp curve (big difference in KGE with changing model parameter value) shows that a parameter is strongly influencing the model performance, while a flatter curve (small difference in KGE with changing model parameter value) shows weak influence of the model parameter on the model performance.

3.6 Correlation Analysis

An exploratory correlation analysis between the catchment characteristics and the model parameters were carried out for the dataset.

In this study, the spearman rank correlation was used. The spearman rank correlation is a nonparametric measure of rank correlation, and one of the oldest and most well-known nonparametric procedures (Artusi et al., 2002). The rank correlation coefficient, r_s, is expressed as

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} \tag{39}$$

In which *n* is the number of measurements and d_i is the ranked difference between the ith measurements for the variates (Zar, 1972). The Spearman rank correlation is being used here because it is a robust correlation method and presupposes no fixed shape of possible functional relationship (Seibert, 1999). In addition to the correlation, the p-value of each correlation was also evaluated. A very low p-value means that such an extreme observed outcome would be very unlikely under the null hypothesis. Low p-values means that there is a small chance of the results being random.

3.7 Regionalization Methods

Predictions of streamflow by hydrological models in ungauged basins are essentially data driven. The models need to be calibrated with observed flow data. When applying these models to ungauged catchments, the challenge is that no flow data is available and hence no calibration is possible. Therefore, hydrologists have been attempting to develop strategies to estimate the model parameters without using calibration (Oudin et al., 2008). Regionalization originated in the process of regime classification and catchment grouping (Gottschalk et al., 1979) and was later extended, in the rainfall-runoff modeling context, to the transfer of parameters from gauged catchments to ungauged catchments. The concept of regionalization has also evolved into a term which covers all methods aimed at estimating model parameter values on any ungauged catchment in a definable region of consistent hydrological response (Oudin et al., 2008).

3.7.1 Multiple Regression

Arguably the most popular approach when it comes to regionalization is the multiple regression (MR) method (Wilkerson & Merwade, 2010). In this method, functions are established between the model parameters and the catchment characteristics for the donor catchments. These functions coupled with the catchment characteristics from the target catchments are used for the prediction of the ungauged catchments. To use regression methods, it is assumed that there is a well-behaved relationship between the catchment characteristics and the model parameters. In addition, it is assumed that the catchment characteristics used in the regression provide information that is relevant to the hydrological behavior at the ungauged sites (Merz et al., 2006). This study assumes that the catchment characteristics and the model parameters in the DDD model are related. The relationships (multiple regression functions), which is built for the gauged basins are then transferred to the ungauged catchments (Yang et al., 2020).

The multiple regression equations, which is used to relate catchment characteristics to model parameters, are trained using the calibrated model parameters of the 107 gauged catchments. To do this a stepwise regression procedure is used for building the regression model. The model was built by adding and removing catchment characteristics in a stepwise manner until there are no compelling reason to add or remove catchment characteristics anymore.

Both linear and non-linear (logarithmic) relationships between the model parameters and the catchment characteristics are tested in the regression model.

3.7.2 Physical Similarity

The physical similarity method is based on the idea that catchments far apart, with similar attributes could show similar hydrologic behavior (Acreman & Sinclair, 1968; Gottschalk, 1985; Nathan & McMahon, 1990; Parajka et al, 2005). The method consists of transferring hydrological information from gauged catchments, known as donors, to the ungauged target catchments as catchment descriptors. In this study, the physical similarity index from Burn & Boorman (1993) is used (Yang et al., 2018).

$$SI_{td} = \sum_{i=1}^{k} \frac{\left|CD_{d,i} - CD_{t,i}\right|}{\Delta CD_i} \tag{40}$$

Where CD is the catchment characteristics, k is the total number of catchment descriptors, d is the donor catchment, t is the target catchment, and ΔCD_i is the range of the ith catchment descriptor. Multiple donor catchments works better than one for this method (Kjeldsen et al., 2014). There are two options to combine the information from the donor catchments to the ungauged catchment (McIntyre et al., 2005).

The first option is the parameter averaging method, in which the model parameters from the donor catchments are averaged and then used in the model for the target catchment. While for the second option, the output averaging method, the model is run with the parameters from the donor catchment on the target catchment, and the outputs of the model are then averaged.

The maximum number of donor catchments used in the physical similarity method, is chosen to be 5. This is less than the number of donor catchments used by many previous studies (Oudin et al., 2008; Arsenault et al., 2015), but according to Bao et al. (2012) and Yang et al. (2018) more than 5 donor catchments do not increase the accuracy significantly.

4 Results

4.1 Relations between Catchment Characteristics and Model Parameters

The correlations between the model parameters and catchment characteristics are not very high, as seen in Table 3, but several are found to be significant. The most striking correlations are found between *Pkorr* and forest cover (skog), as well as bare rock % (sn_fj). Between *u* and specific discharge (SpQ) and between *GscInt* and precipitation (P), all includes correlations over 0.6.

Table 3: Correlation between model parameters and catchment characteristics, only significant correlations at p value < 0.01 and 0.05 are shown.

Model Param	d_mean	MAD	L.	Myr.	Area	EL.	1085.00	Feltl	Bre	Skog	Sn_fjell	P	T	Elevation	SpQ
Pkorr				-0.33					0.32	-0.67	0.68		-0.60	0.45	0.36
Skorr									0.25	-0.38	0.48		-0.42		
u						0.31			0.26	-0.33		0.53			0.61
pro					0.25		-0.26	0.26		0.40	-0.41				-0.40
TX					-0.26					0.41	-0.46	0.34	0.45	-0.34	
CGlac									0.29						
CritFlux															
GshInt												0.55	0.36		0.50
GscInt				-0.28	-0.40		0.44	-0.42				0.61	0.38		0.46
rv															

P value < 0.01

Model Param	d_mean	MAD	L.	Myr.	Area	EL.	1085.00	Feltl	Bre	Skog	Sn_fjell	P	T	Elevation	SpQ
Pkorr			0.20	-0.33					0.32	-0.67	0.68		-0.60	0.45	0.36
Skorr		0.20							0.25	-0.38	0.48		-0.42	0.23	
u			0.24			0.31			0.26	-0.33		0.53		0.19	0.61
pro					0.25		-0.26	0.26		0.40	-0.41	-0.24			-0.40
TX	-0.20	-0.22			-0.26					0.41	-0.46	0.34	0.45	-0.34	
CGlac									0.29						
CritFlux														-0.20	
GshInt	-0.23				-0.19		0.24					0.55	0.36		0.50
GscInt	-0.22	-0.19	-0.19	-0.28	-0.40		0.44	-0.42				0.61	0.38		0.46
rv			-0.24									-0.19			-0.25

P value < 0.05

Figure 7a shows that the correlation of the catchment characteristics, while Figure 7b shows the significant correlations. The correlations are generally not very high but includes quite a few significant correlations. The most striking correlations is between area and catchment

length (L.), mean annual discharge (MAD) and area, and mean annual discharge and catchment length.

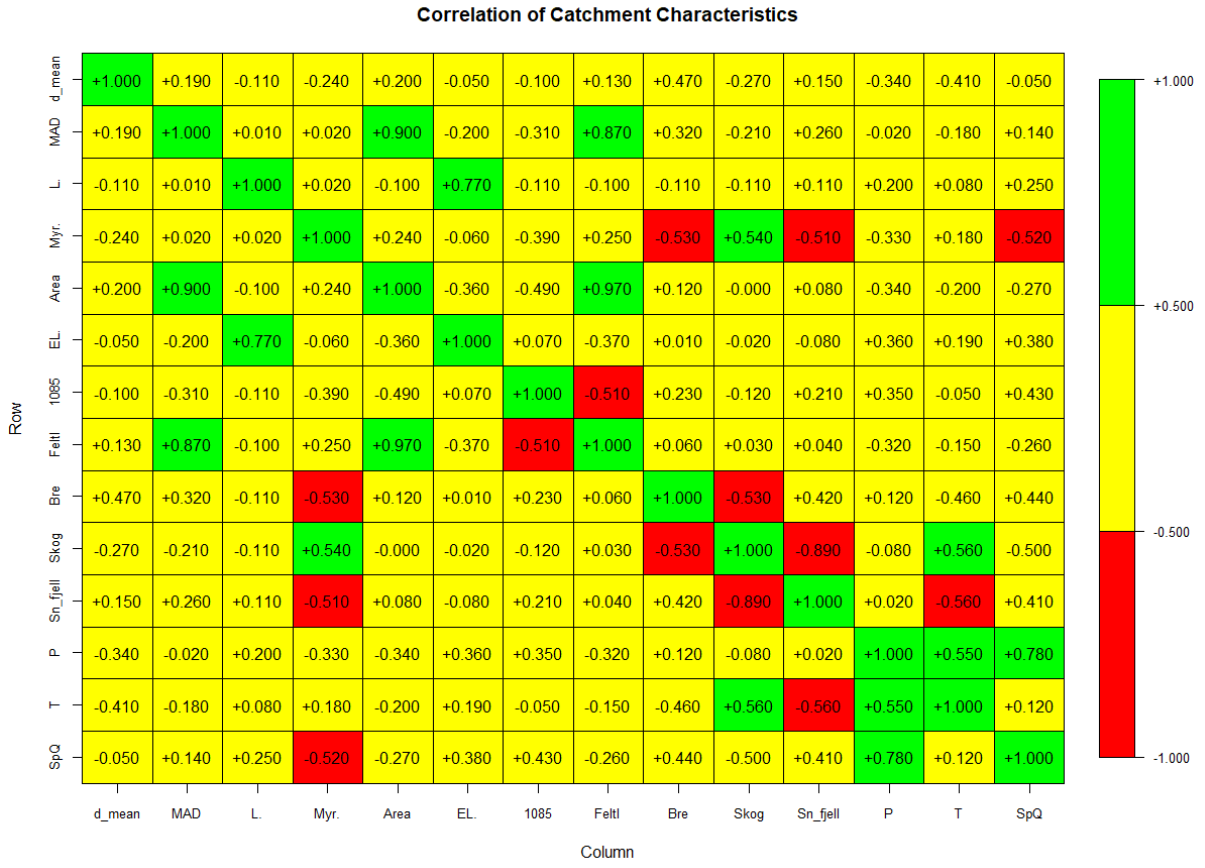


Figure 7a: Correlations between the Catchment Characteristics.

Catchment Characteristics	d_mean	MAD	L.	Myr.	Area	EL.	1085	FeltI	Bre	Skog	Sn_fjell	P	T	SpQ
d_mean	1								0.47	-0.27		-0.34	-0.41	
MAD		1			0.9		-0.31	0.87	0.32					
L.			1			0.77								
Myr.				1			-0.39		-0.53	0.54	-0.51	-0.33		-0.52
Area		0.9			1	-0.36	-0.49	0.97				-0.34		-0.27
EL.			0.77		-0.36	1		-0.37				0.36		0.38
1085		-0.31		-0.39	-0.49		1	-0.51				0.35		0.43
FeltI		0.87			0.97	-0.37	-0.51	1				-0.32		
Bre	0.47	0.32		-0.53					1	-0.53	0.42		-0.46	0.44
Skog	-0.27			0.54					-0.53	1	-0.89		0.56	-0.5
Sn_fjell				-0.51					0.42	-0.89	1		-0.56	0.41
P	-0.34			-0.33	-0.34	0.36	0.35	-0.32				1	0.55	0.78
T	-0.41				-0.2				-0.46	0.56	-0.56	0.55	1	
SpQ				-0.52	-0.27	0.38	0.43		0.44	-0.5	0.41	0.78		1

P value < 0.01

Figure 7b: Correlation of the catchment characteristics, only significant correlations with p-value < 0.01

Figure 8a shows the correlation between the model parameters, while Figure 8b shows the significant correlations. The correlations between the model parameters are not very high but includes significant correlations. Skorr is strongly correlated with both Pkorr and TX, positively and negatively respectively.

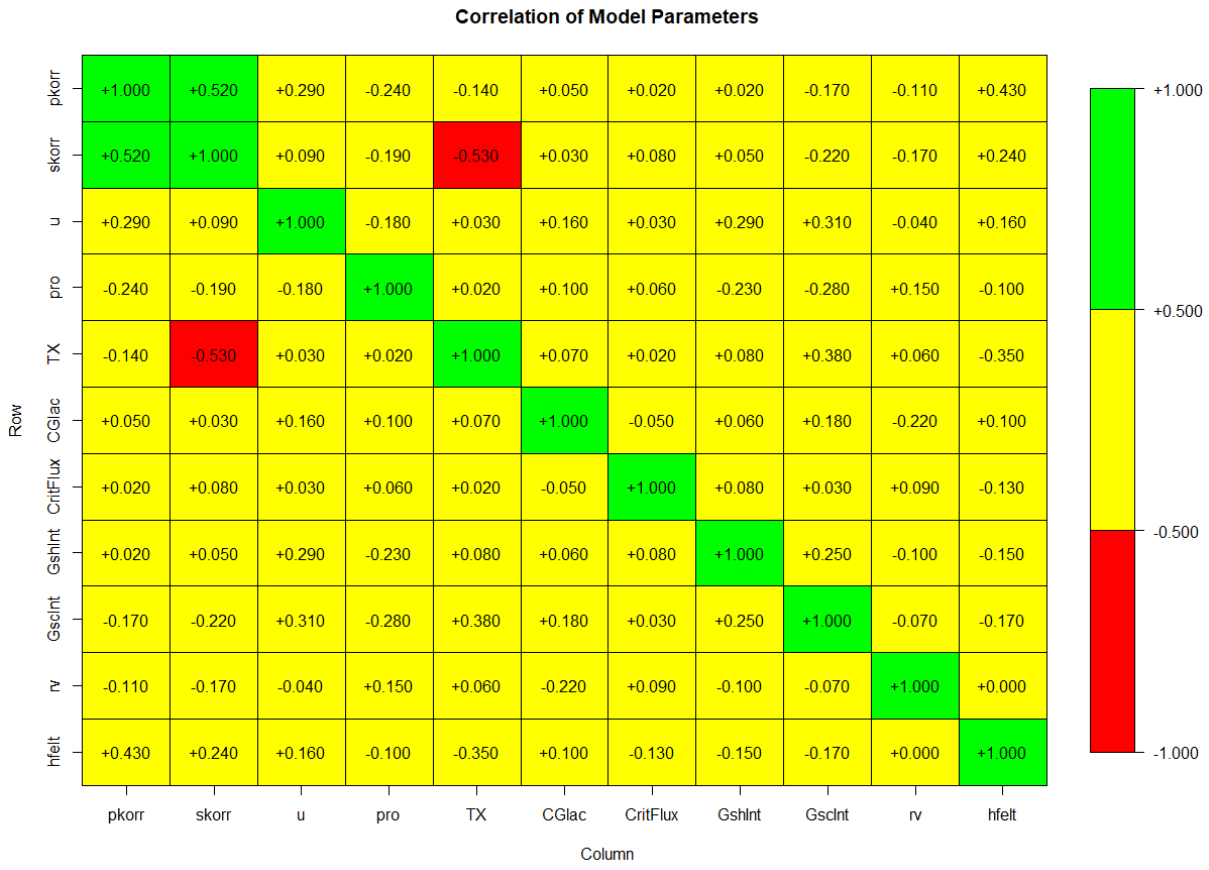


Figure 8a: Correlation of the model parameters.

Model Param	Pkorr	Skorr	U	Pro	TX	Cglac	CritFlux	GshInt	GscInt	rv	hfelt
Pkorr	1	0.52	0.29	-0.24							0.43
Skorr	0.52	1			-0.53						0.24
U	0.29		1					0.29	0.31		
Pro	-0.24			1					-0.28		
TX		-0.53			1				0.38		-0.35
Cglac						1					
CritFlux							1				
GshInt			0.29					1	0.25		
GscInt			0.31	-0.28	0.38			0.25	1		
rv				0.15						1	
hfelt	0.43	0.24			-0.35						1

P value < 0.01

Figure 8b: Correlation of the model parameters, only significant correlations with p-value < 0.01

4.2 Parameter Sensitivity

Figure 9 shows the sensitivity plot for station 12.70 (Etna). It is clear that the *Pkorr* parameter is the most sensitive, followed by *Skorr*. On the other hand, *TX*, and *pro* shows very little sensitivity. *GscInt*, *GshInt* and *u* are somewhere in the middle, closer to the least sensitive parameters.

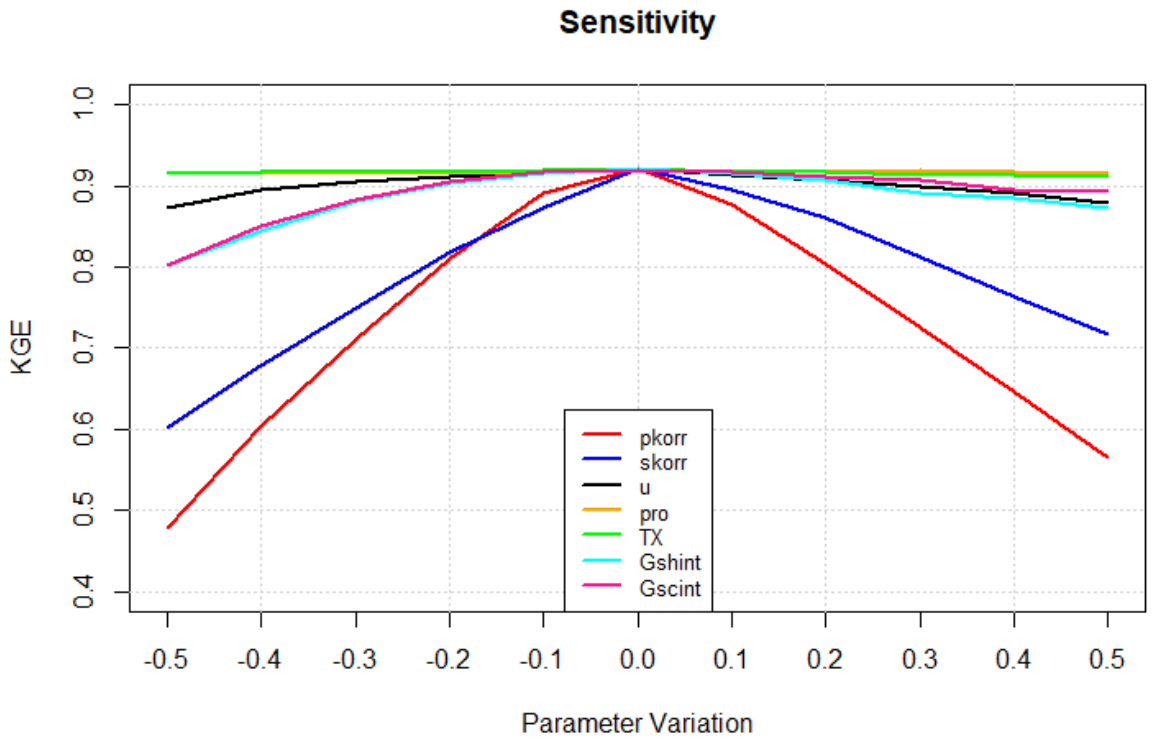


Figure 9: Sensitivity plot for station 12.70 (Etna), x-axis shows variation in parameter values by percent, from 10 % up to 50 %.

Figure 10 shows a dottyplot for the mountain region. The Figure shows much the same as Figure 9. *Pkorr* and *Skorr* being the most sensitive parameters, with well-defined curves. The dottyplot also shows a weak defined curve for the *u* parameter. The rest of the parameters shows no defined curve. The *pro* parameter has a flat curve here, but still follows the other parameters with good KGE values for several different *pro* values.

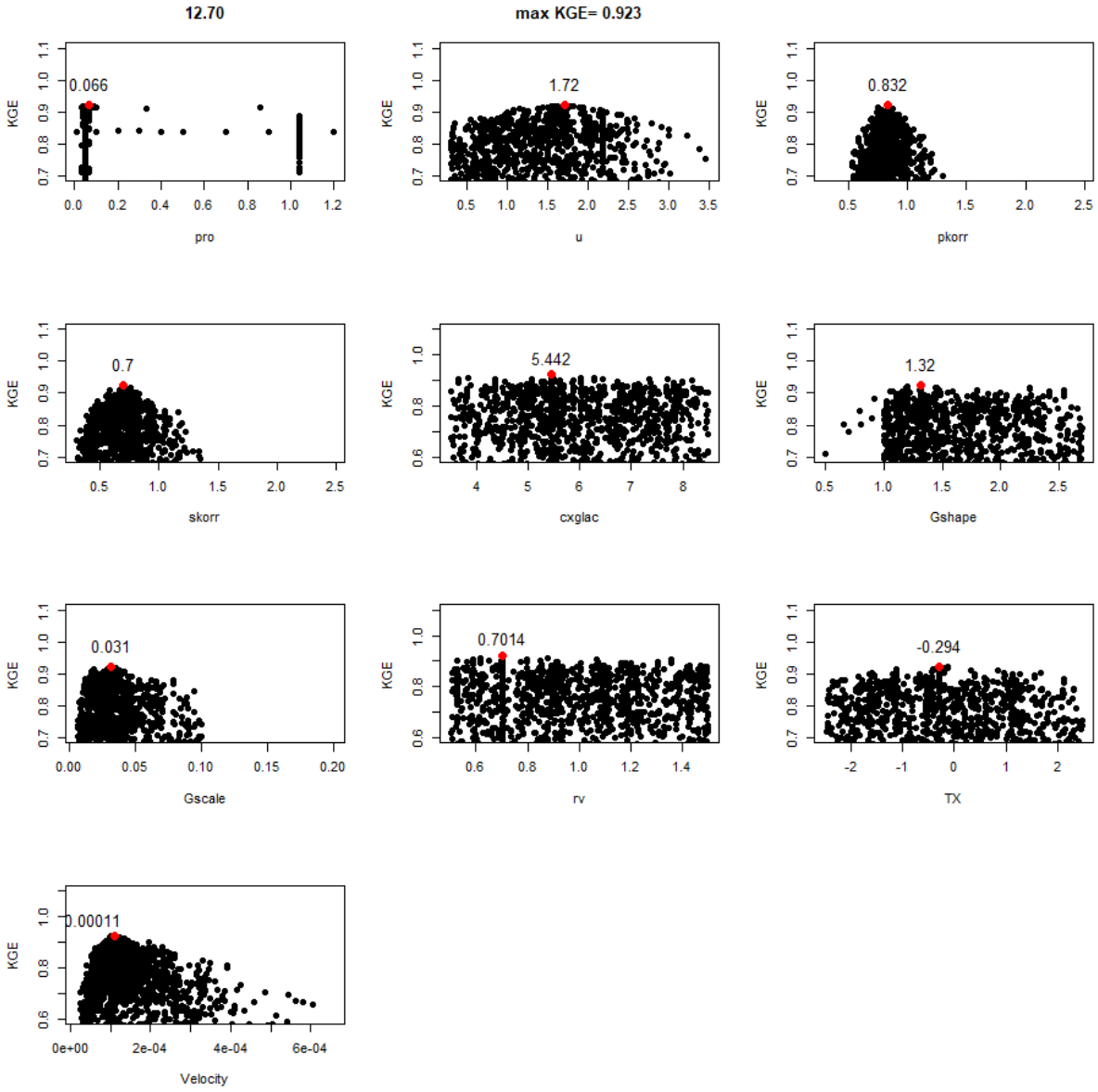


Figure 10: Dottyplot for station 12.70 (Etna).

Figure 11 shows the dottyplot for the catchment in the inland region. Much of the same can be seen as in Figure 10. A defined curve for the *skorr* and *pkorr* parameters, although not as defined for the *skorr* parameter, with less defined curves for the other parameters. The *pro* parameter is quite interesting, but show good KGE values for different values of *pro*.

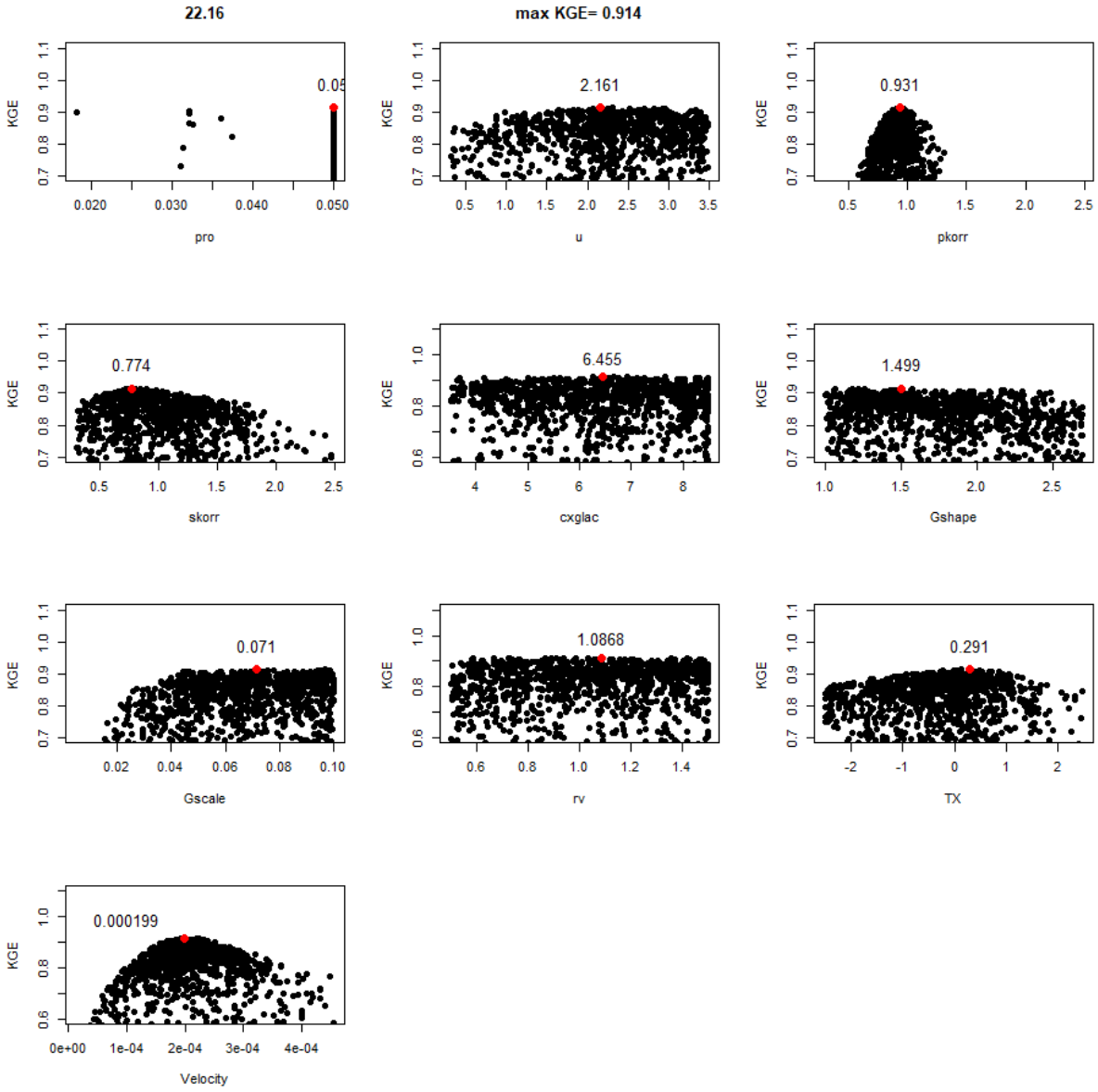


Figure 112: Dottyplot for station 22.16 (Myglevatn).

Figure 12 show the dottyplot for a coastal catchment. Here the *skorr* parameter shows a less defined curve than in the other dottyplots. This is likely due to the fact that snowfall is a less hydrologically relevant event in the coastal regions of Norway. Most of the precipitation falls as rain, and thus the *pkorr* curve is more well-defined. The other plots look similar to the preceding dottyplots.

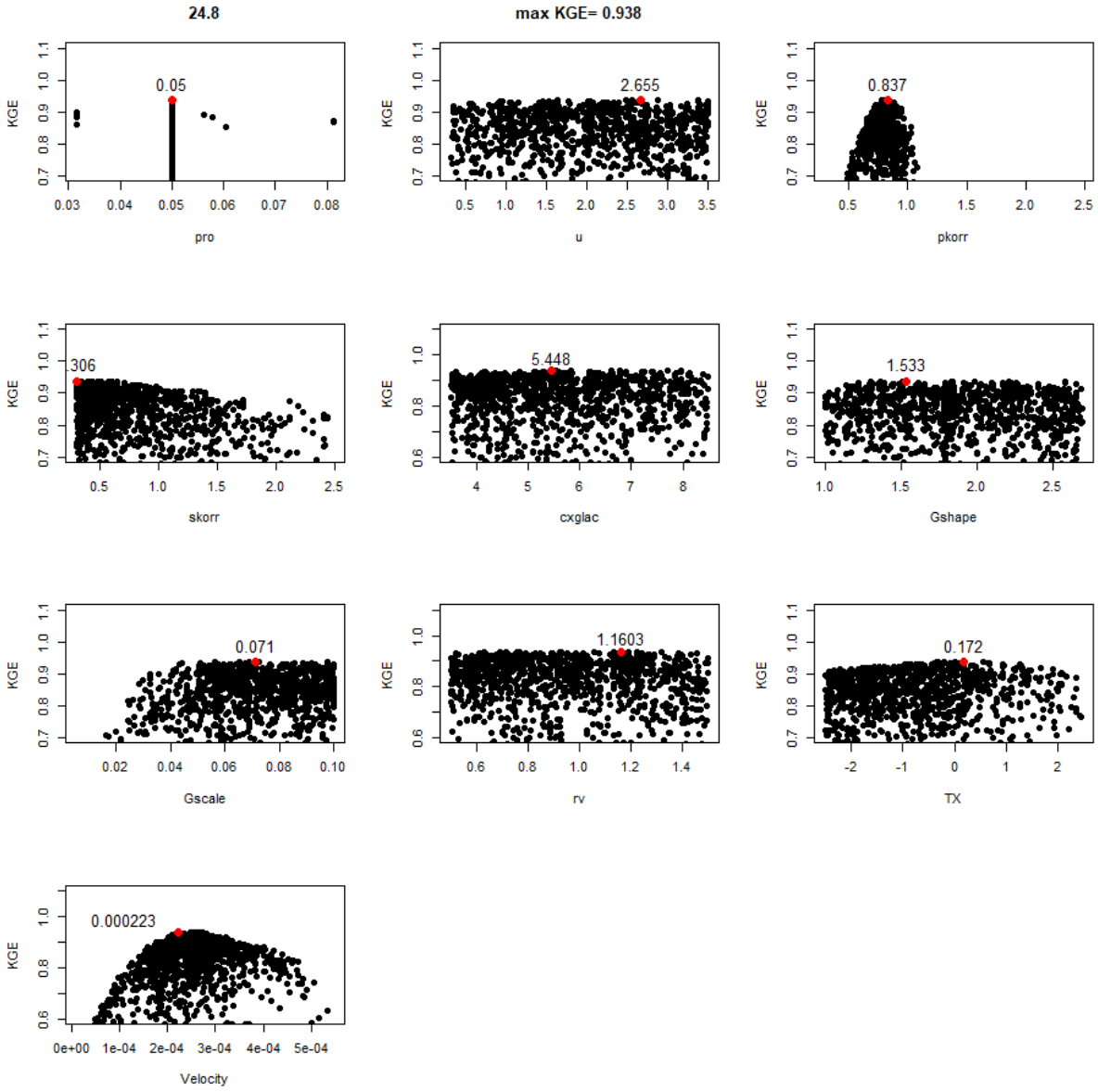


Figure 12: Dottyplot for station 24.8 (Moska).

Figure 13 shows the dottyplot for a catchment in the transition region. Again, the *skorr* parameter shows a less defined curve, this is most likely due to the precipitation falling as rain in this region. *Pkorr* is again the most defined curve, with all other plots looking similar to the preceding ones.

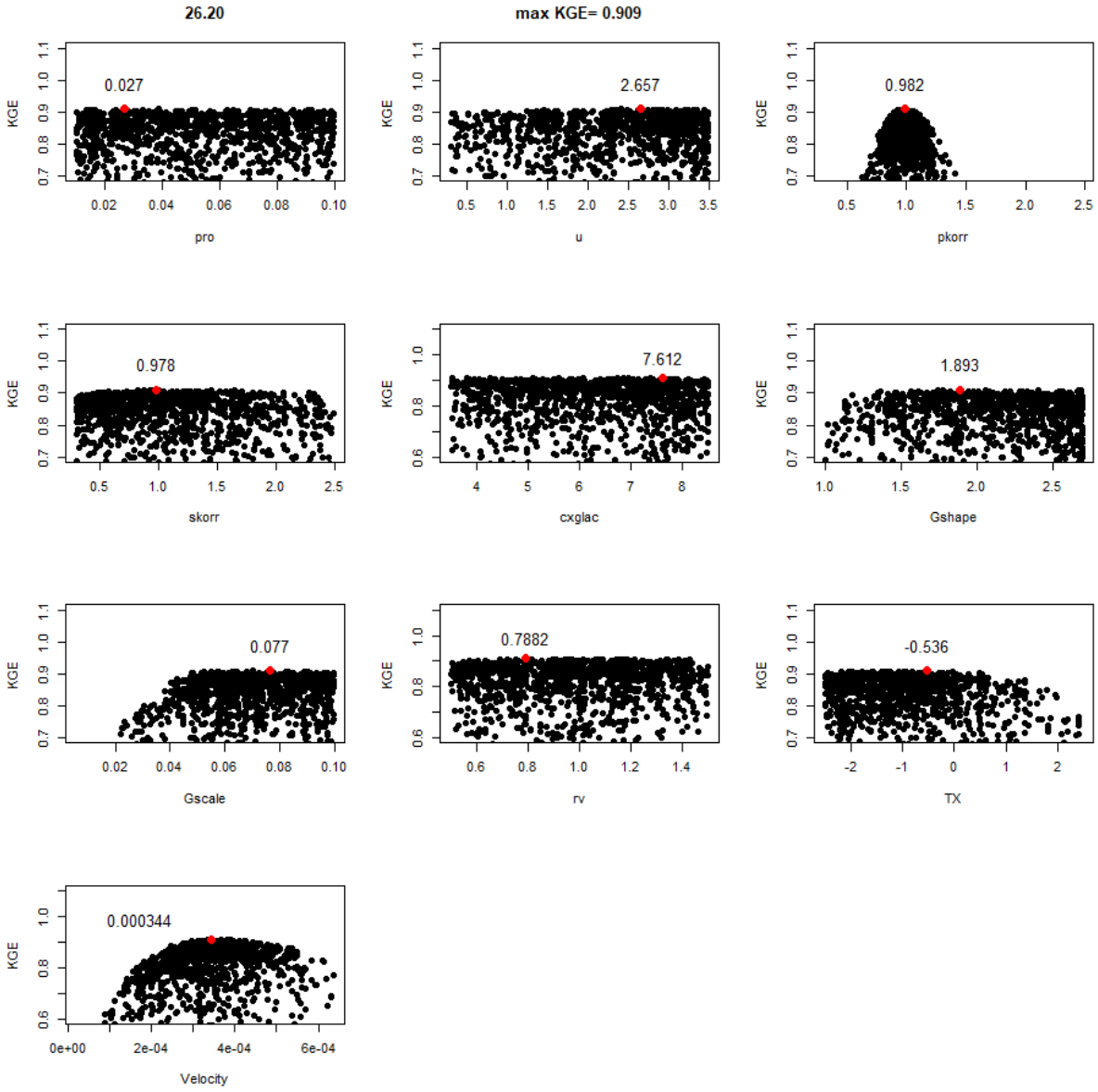


Figure 13: Dottyplot for station 26.20 (Aardal).

Figure 14 shows a dottyplot for a catchment in the Trøndelag region. The *pkorr* parameter is most defined, with a less defined curve for the *skorr* parameter. All other parameters look similarly flat, with lots of different values reaching a high KGE value.

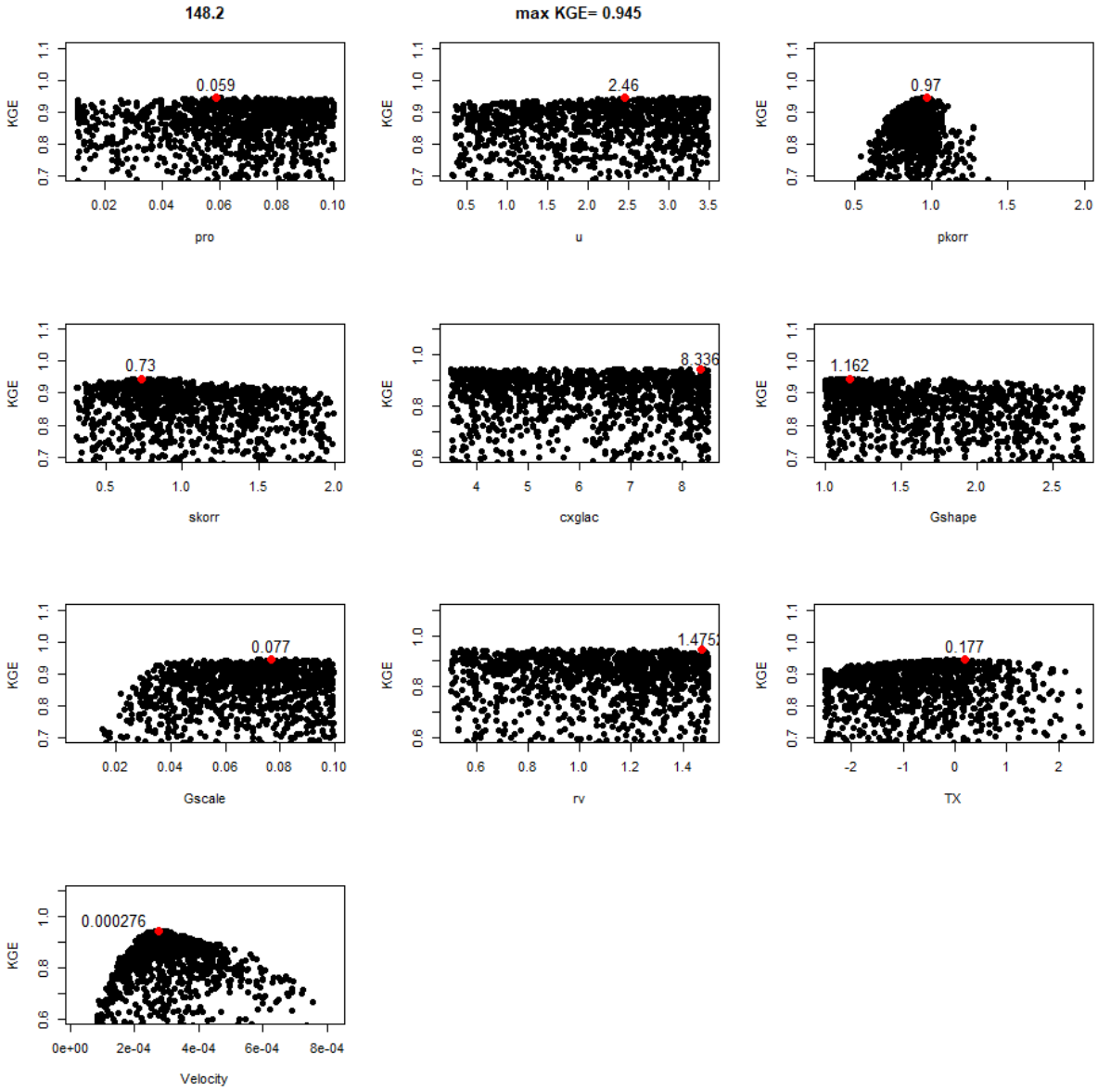


Figure 14: Dottyplot for station 148.2 (Mevatnet).

The dottyplots were quite similar in a lot of ways. The biggest difference is the *skorr* parameter, which had a noticeable curve for the regions with the biggest temperature difference, the inland and mountain regions, and a less definable curve for the warmer regions. This is due to the fact that less of the precipitation falls as snow in the warmer regions, and thus the pkorr parameter is more important for model performance.

4.3 Regionalization

4.3.1 Multiple Regression

The multiple regression equations were regressed from the model parameters and catchment characteristics of the 107 donor catchments. The multiple regression equations are shown in equations 41-48, with the subsequent coefficient of determination (R2) values in Table 3. The R2 score is defined as the proportion of the variance in the dependent variable that is predictable from independent variables (Hogg and Ledolter, 1992). The multiple regression equations contain two, three, or in one case four catchment characteristics.

$$GscInt = \exp\left(-3.13273 - 0.29724\log(Feltl) - 0.106\log(EL.) + 0.83762\log(P)\right) \quad (41)$$

$$GshInt = 0.142 + 0.0103\log(SpQ) - 0.03311\log(EL.) - 0.0003912\log\left(hfelt\right) \quad (42)$$

$$Pkorr = \exp(-0.98214 + 0.44176\log(SpQ) - 0.51571\log(P)) \quad (43)$$

$$Skorr = \exp\begin{matrix}(0.0838134 - 0.0003421(hfelt) + 0.449802(EL.) + 0.0040812\log\left(Sn_{Fj}\right) \\ -0.0717631\log\left(P\right))\end{matrix} \quad (44)$$

$$U = -0.6437 + 0.83661\log(SpQ) + 0.13944\log(Myr.) - 0.12529\log\left(Skog\right) \quad (45)$$

$$Pro = \exp(-4.02088 + 0.29397\log(Feltl) - 0.14403\log(Sn\_fj)) \quad (46)$$

$$TX = -1.941724 + 0.014716(SpQ) + 0.011695\left(Skog\right) \quad (47)$$

$$CGlac = 8.8296 + 0.1951log(EL.) - 0.6755log(Feltl.) \quad (48)$$

Table 4 shows the R2 score and the interval for the regressed equations. The R2 score is generally quite low, but comparable to Skaugen et al. (2015). Since the values of the catchment characteristics have distinct values for every single catchment, it is a good measure of the robustness of the equations to check the intervals they put out. If the multiple regression equation produces parameters that are behavioral, i.e., no illegal values for the parameters such as, for example, negative values for the gamma distribution, the multiple regression equations are more robust. The interval for the regressed parameters looks good for every parameter, with sensible values for the highest and lowest value for each parameter.

Table 4: Model parameters and their R2 score for the multiple regression equations.

Model Parameter	R2-score	Interval for the regressed parameters
GscInt	0.47	0.017 – 0.155
GshInt	0.43	1.444 – 1.467
Pkorr	0.51	0.832 – 1.254

Skorr	0.41	0.681 – 1.777
U	0.40	1.422 – 3.871
Pro	0.26	0.012 – 0.052
TX	0.34	-1.540 – 0.583
CGLAC	0.27	5.170 – 7.863

Table 5 shows the calibrated values and the values predicted with the multiple regression method. The multiple regression method yields good KGE values, even though the R2 score of the equations themselves are not the greatest.

As stated in Thiemig et al. (2013), 0.75 ≤ KGE < 0.9 is good, 0.5 ≤ KGE < 0.75 is intermediate and 0.0 ≤ KGE < 0.5 is poor. 15 of the catchments are in the good category, 10 are in the intermediate category and no catchments have KGE values under 0.5. The difference between the calibrated KGE and the KGE from the regression method is less than 0.1 for 15 catchments.

The bias is also generally good, with most catchments being within the 10 % interval. However, catchment 15.79 has a bias of 1.29 and 75.23 has a bias of 0.82. These two stations also have the lowest KGE values.

The catchments in bold are the catchments selected for the regional study. As one can see all the representatives from each region have KGE scores over 0.75 and thus all the catchments scores good on the regression method. This means that the multiple regression method is not spatially varied. The inland region (22.16) scores the worst with a KGE score of 0.83, while the Trøndelag region (148.2) and the transition region (26.20) scores the best, both with a KGE score of 0.89.

Table 5: Model results with both calibration and multiple regression, values in KGE, with Bias in parentheses.

Station ID	Calibration	Regression
2.32	0.87 (1.00)	0.61 (0.85)
2.463	0.90 (1.00)	0.87 (1.08)
12.70	**0.92 (0.97)**	**0.86 (1.12)**
12.215	0.89 (0.99)	0.85 (0.89)
15.79	0.90 (1.01)	0.59 (1.29)
16.75	0.91 (1.03)	0.83 (1.13)
16.193	0.82 (0.98)	0.73 (1.03)
20.2	0.90 (1.01)	0.84 (0.97)
22.16	**0.91 (0.99)**	**0.83 (0.91)**
24.8	**0.94 (0.99)**	**0.86 (1.10)**
24.9	0.91 (1.02)	0.89 (1.04)

25.24	0.87 (0.99)	0.60 (1.08)
25.32	0.78 (0.98)	0.72 (1.03)
26.20	**0.91 (1.00)**	**0.89 (0.95)**
27.16	0.84 (0.99)	0.69 (1.11)
35.16	0.89 (1.00)	0.81 (1.12)
42.2	0.79 (1.02)	0.62 (0.86)
48.5	0.89 (0.99)	0.88 (1.05)
55.5	0.79 (0.99)	0.67 (0.97)
62.5	0.92 (1.02)	0.80 (0.91)
75.23	0.83 (1.01)	0.66 (0.82)
122.11	0.90 (1.00)	0.81 (0.96)
148.2	**0.95 (1.01)**	**0.89 (1.08)**
165.6	0.77 (1.00)	0.66 (1.00)
212.27	0.92 (0.99)	0.77 (1.04)
Average	***0.88***	***0.77***

4.3.2 Physical Similarity

Physical similarity lists were created for each target catchment. For each target catchment the list consists of all the donor catchments and the physical similarity index value. The five catchments that are most similar (lowest similarity index) are chosen to be used in the predictions. The plots created were quite similar for all catchments, a few very similar catchments and quite a few that have a high similarity index.

Table 6 shows the results from the output average method and Table 7 shows the results from the parameter average method. The two methods are very similar, with the output average method averaging a slightly higher KGE value for increasing donor catchments. Nevertheless, the results are not significantly better than the multiple regression method for either of the physical similarity methods.
It is also worth noting that catchments with higher KGE values for the output average method, generally have quite high KGE for the parameter average method too.
The model performance, KGE, will on average increase with increasing number of donor catchments.
For the catchments chosen in the regional study, there are quite large differences in the KGE score. For both the output average method and the parameter average method, the mountain region (12.70) starts out with a high KGE value of about 0.8 for the most physically similar station, before the KGE starts decreasing to only an intermediate KGE score of about 0.58 when all the five physical similarity stations are included. For the other regions it is usually the opposite, starting with a decent value, before increasing to a good KGE value for

the five most physically similar stations. Especially the Trøndelag region (148.2) has a very high KGE score, with all but one of the donor catchments' physical similarity being over 0.9 KGE score.

Table 6: Model results from the physical similarity method, output average with increasing station numbers.

Station ID	1	2	3	4	5
2.32	0.20	0.24	0.18	0.33	0.39
2.463	0.12	0.55	0.62	0.71	0.78
12.70	**0.80**	**0.71**	**0.61**	**0.53**	**0.58**
12.215	0.80	0.78	0.80	0.79	0.84
15.79	0.71	0.75	0.75	0.81	0.80
16.75	0.55	0.64	0.49	0.62	0.71
16.193	0.79	0.80	0.79	0.77	0.77
20.2	0.84	0.84	0.87	0.86	0.84
22.16	**0.79**	**0.72**	**0.80**	**0.83**	**0.85**
24.8	**0.73**	**0.84**	**0.86**	**0.89**	**0.89**
24.9	0.67	0.79	0.86	0.88	0.84
25.24	0.77	0.80	0.73	0.75	0.70
25.32	0.72	0.74	0.73	0.65	0.69
26.20	**0.69**	**0.77**	**0.90**	**0.91**	**0.89**
27.16	0.63	0.76	0.69	0.61	0.64
35.16	0.80	0.78	0.72	0.75	0.77
42.2	0.57	0.61	0.64	0.64	0.67
48.5	0.73	0.76	0.81	0.89	0.91
55.5	0.37	0.43	0.51	0.48	0.48
62.5	0.72	0.88	0.80	0.69	0.73
75.23	0.55	0.55	0.65	0.65	0.62
122.11	0.55	0.62	0.63	0.76	0.79
148.2	**0.89**	**0.87**	**0.82**	**0.87**	**0.91**
165.6	0.32	0.43	0.52	0.54	0.52
212.27	0.56	0.36	0.30	0.27	0.43
Average	***0.64***	***0.68***	***0.68***	***0.70***	***0.72***

Table 7: Model results from the physical similarity method, parameter average with increasing station numbers.

Station ID	1	2	3	4	5
2.32	0.20	0.19	0.09	0.14	0.20
2.463	0.12	0.83	0.77	0.83	0.81
12.70	**0.80**	**0.71**	**0.62**	**0.52**	**0.57**
12.215	0.80	0.81	0.82	0.82	0.88
15.79	0.71	0.75	0.75	0.81	0.78
16.75	0.55	0.61	0.46	0.55	0.65
16.193	0.79	0.79	0.79	0.77	0.76
20.2	0.84	0.81	0.83	0.83	0.81
22.16	**0.79**	**0.73**	**0.83**	**0.86**	**0.88**
24.8	**0.73**	**0.84**	**0.85**	**0.88**	**0.89**
24.9	0.67	0.77	0.84	0.86	0.82

25.24	0.77	0.78	0.66	0.69	0.63
25.32	0.72	0.73	0.72	0.66	0.70
26.20	**0.69**	**0.85**	**0.81**	**0.82**	**0.87**
27.16	0.63	0.75	0.67	0.61	0.65
35.16	0.80	0.81	0.72	0.72	0.74
42.2	0.57	0.62	0.64	0.64	0.67
48.5	0.73	0.76	0.82	0.88	0.89
55.5	0.37	0.43	0.51	0.48	0.49
62.5	0.72	0.91	0.87	0.81	0.83
75.23	0.55	0.60	0.70	0.69	0.67
122.11	0.55	0.62	0.59	0.66	0.69
148.2	**0.89**	**0.87**	**0.78**	**0.83**	**0.91**
165.6	0.32	0.46	0.53	0.56	0.54
212.27	0.56	0.12	0.13	0.12	0.17
Average	***0.64***	***0.68***	***0.67***	***0.68***	***0.70***

Figure 15 shows a comparison between the output average method and the parameter average method. The two methods are very comparable and show little to no difference for the five donor catchments used for each target catchment. The output average method is slightly better for almost every donor station added to the physical similarity method, except the second one added, but the difference between the results is small.

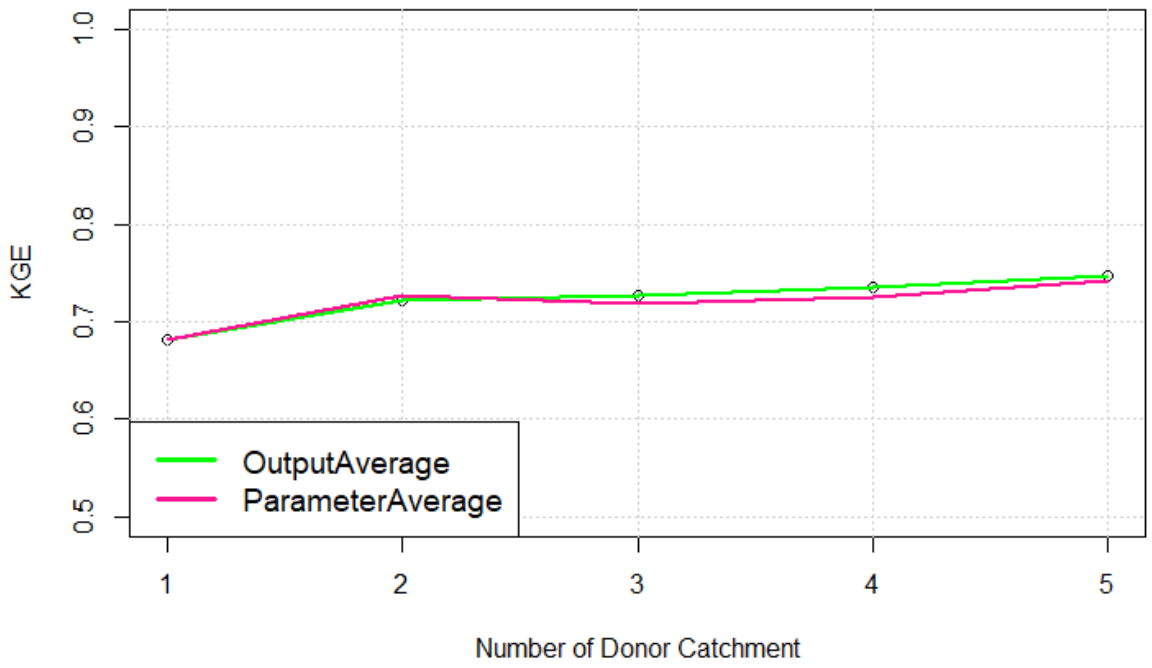


Figure 15: Comparison of the output average method and the parameter average method.

4.4 Comparison of Regionalization Methods

Table 8 shows a full comparison of all regionalization methods. The calibration has the highest KGE, followed by regression and then the two physical similarity methods, output average and parameter average. Out of the 25 target catchments in this study, 12 performed best when using the multiple regression method, five performed best using the output average method, four performed best using the parameter average method and four catchments had a tie.

All the catchments that represent their region has very high KGE values, except for the physical similarity for the mountain region (12.70). The best results come for the Trøndelag region, with all KGE values around or over 0.9. For the mountain region, the highest performance came with the multiple regression method, and for the transition region the best performance came with both the multiple regression and output average. For the Trøndelag, inland and coastal regions the best performance came with the physical similarity method. For the inland region, the output average performed the best, but for the Trøndelag and coastal it was a tie as both physical similarity methods performed equal.

Table 8: Comparison of all methods, output average and parameter average for the 5th station.

Station ID	Calibration	Regression	Output Average	Parameter Average
2.32	0.87	0.61	0.39	0.2
2.463	0.9	0.87	0.78	0.81
12.70	**0.92**	**0.86**	**0.58**	**0.57**
12.215	0.89	0.85	0.84	0.88
15.79	0.9	0.59	0.8	0.78
16.75	0.91	0.83	0.71	0.65
16.193	0.82	0.73	0.77	0.76
20.2	0.9	0.84	0.84	0.81
22.16	**0.91**	**0.83**	**0.85**	**0.88**
24.8	**0.94**	**0.86**	**0.89**	**0.89**
24.9	0.91	0.89	0.84	0.82
25.24	0.87	0.6	0.7	0.63
25.32	0.78	0.72	0.69	0.7
26.20	**0.91**	**0.89**	**0.89**	**0.87**
27.16	0.84	0.69	0.64	0.65
35.16	0.89	0.81	0.77	0.74
42.2	0.79	0.62	0.67	0.67
48.5	0.89	0.88	0.91	0.89
55.5	0.79	0.67	0.48	0.49
62.5	0.92	0.8	0.73	0.83

75.23	0.83	0.66	0.62	0.67
122.11	0.9	0.81	0.79	0.69
148.2	**0.95**	**0.89**	**0.91**	**0.91**
165.6	0.77	0.66	0.52	0.54
212.27	0.92	0.77	0.43	0.17
Average	***0.88***	***0.77***	***0.72***	***0.70***

Table 9 shows the spatial distribution of the regionalization methods. For every region, except the Trøndelag region, the multiple regression method is the best performing method. Behind the multiple regression method, the output average method and the parameter average method follows, respectively. The one exception to this, the Trøndelag region, only includes one catchment and will therefore be a result with a lot more uncertainty.

Table 9: Spatial distribution of regionalization methods.

Regions	Calibration	Regression	Output Average	Parameter Average
Mountain	0.89	0.75	0.69	0.64
Inland	0.87	0.78	0.73	0.72
Coastal	0.86	0.76	0.73	0.73
Transition	0.86	0.78	0.72	0.71
Trøndelag	0.95	0.89	0.91	0.91

Figure 16a shows the observed discharge for station 12.70, with the simulated discharge from the three regionalization methods from the start of 2000 to the end of 2002. As one can see the discharge is very similar for all the methods, however especially in the 2001 spring floods one can see a difference in the methods. The parameter average method peaks at about 120 m^3/s, while the multiple regression method does not even break 100 m^3/s. The peak in discharge is almost always overshot by the parameter average method. On the other hand, for the recession after a peak, very clear at the end of 2000, all methods will slightly undershoot until base flow is reached.

Figure 16b shows the mean monthly hydrograph for all methods and the observed runoff. The first few winter months are very similar for all methods, as for a catchment in a mountainous region this is the time for snow accumulation, and little to no runoff is seen. Coming into April one can see that all methods increase in runoff, with the observed runoff having the lowest increase. This continuous to the peak, in May, when all methods reach their peak. The observed runoff peaks at about 45 m^3/s, while the multiple regression method reaches about 50 m^3/s. The two physical similarity methods both peak at over 60 m^3/s, and thus shows much the same as in Figure 16a. For the spring floods, all methods therefore

overshoots. When the spring flood is over and the runoff decreases, the runoff decreases for all regionalization methods as well as the observed runoff. All methods drop down to an acceptable level, before slightly increasing again in the month of August. This slight increase in runoff in August is not as strong in the observed runoff. Towards the end of the year all methods follow the observed runoff, before ending the year with similar runoff to the observed runoff.

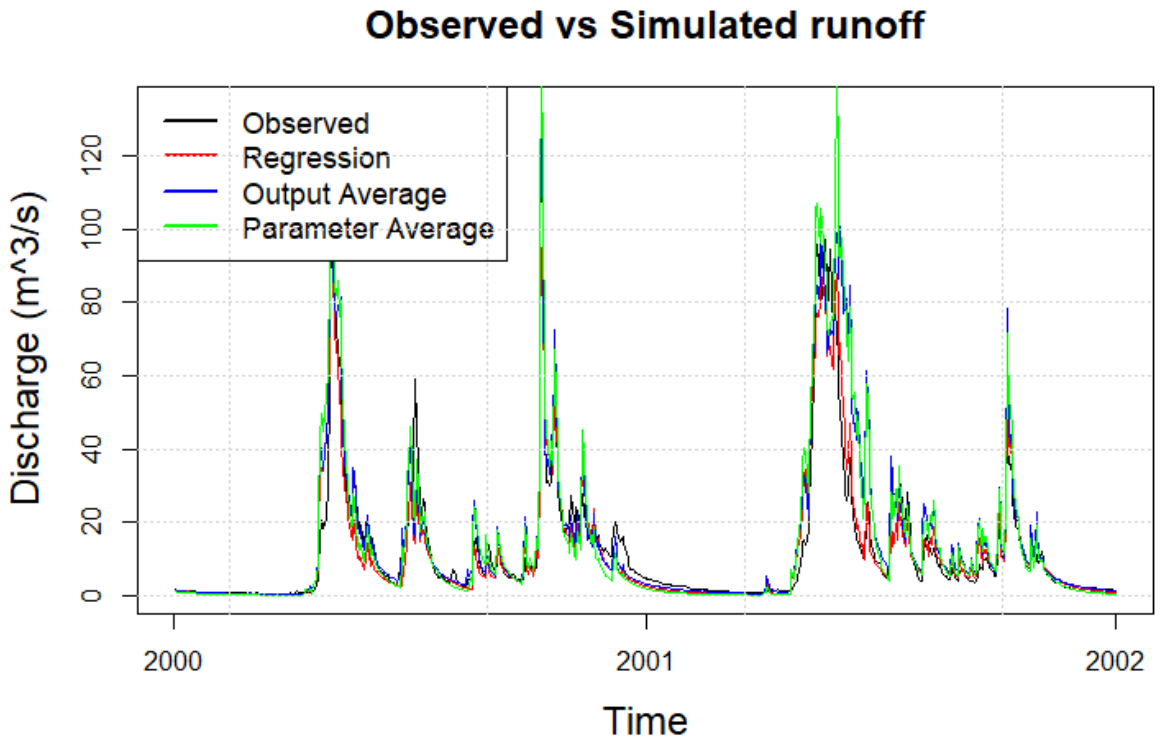


Figure 16a: Comparison of the 3 regionalization methods and the observed discharge for station 12.70 , mountain region.

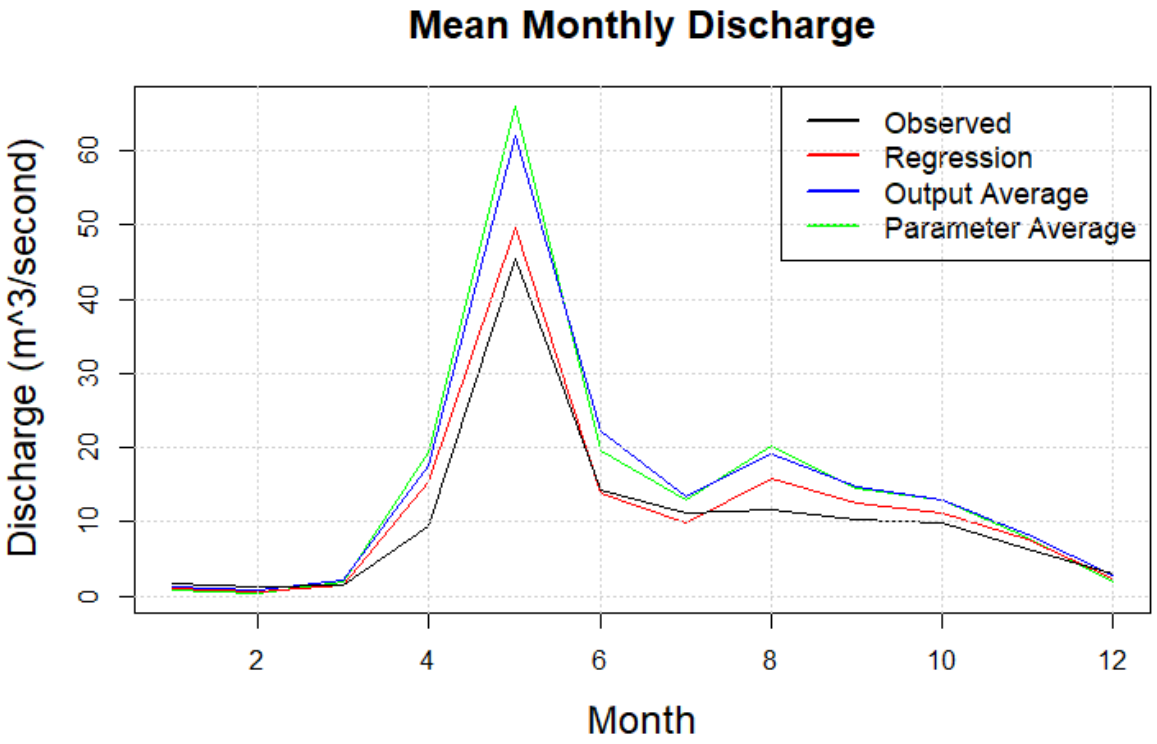


Figure 16b: Mean monthly hydrograph for station 12.70, for all regionalization methods and observed runoff.

Figure 17a shows the observed runoff for station 22.16, with the simulated discharge from the three regionalization methods from the start of 2000 to the end of 2002. The runoff is very similar for all the methods, with only one clear error. During the spring of 2001, the observed runoff is quite a lot higher than for any of the regionalization methods. On the contrary, all other peaks are similar for the methods and the observed runoff. The parameter average method will also for this catchment slightly overshoot some peaks. This is especially clear in the middle parts of 2001, with several smaller peaks in runoff being noticeably too big for the parameter average method.

Figure 17b shows the mean monthly hydrograph for all methods and the observed runoff. Generally, the multiple regression method slightly undershoots the observed runoff. For the two physical similarity methods, they both perform better. For the first three months of the year, all methods perform similarly, missing the observed runoff slightly when the runoff increases in February and March. For the first peak in April, the parameter average method is nearly spot on the observed runoff, while the output average method is just slightly smaller. The multiple regression method misses this peak by quite a lot. When the runoff decreases nearing the summer, all methods decrease by the right amount. For the trough in June, both of the physical similarity methods matches this well, while the multiple regression method is too low again. The increase in runoff towards the rain flood in early winter is well matched by all methods. For the second, and larger, peak in November, the two physical similarity

methods are very similar and just slightly more runoff than the observed runoff. The multiple regression method again undershoots.

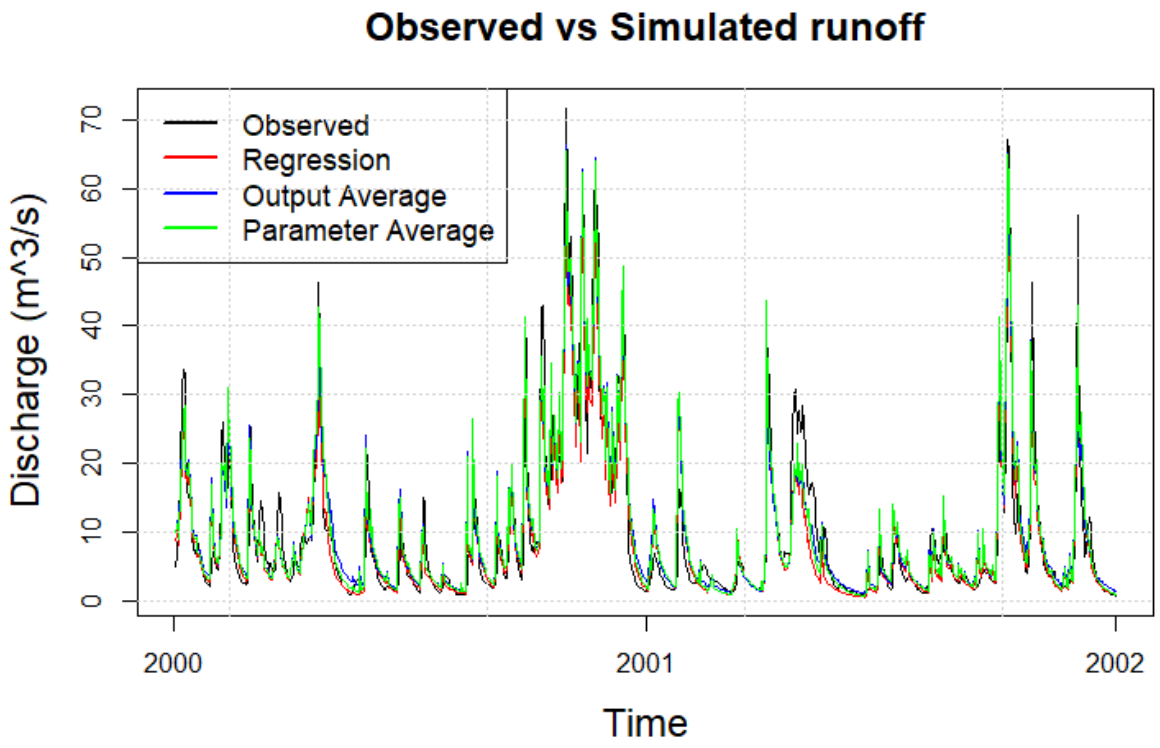


Figure 17a: Comparison of the 3 regionalization methods and the observed discharge for station 22.16, inland region.

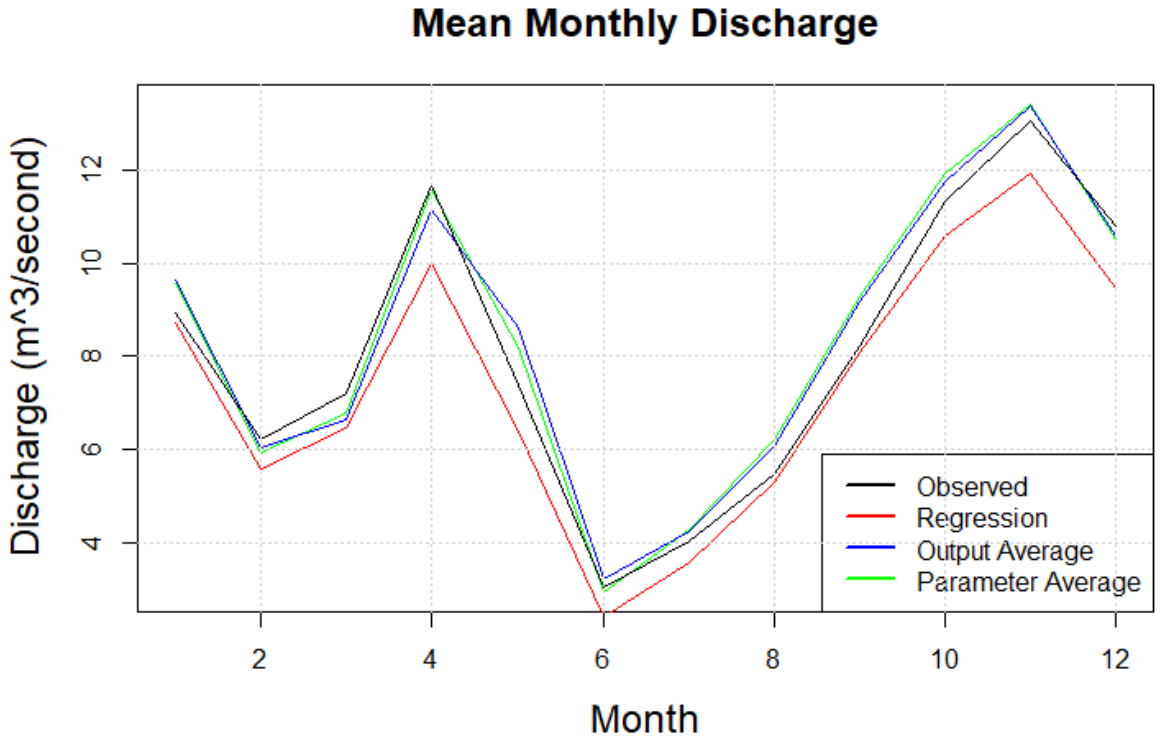


Figure 17b: Mean monthly hydrograph for station 22.16, for all regionalization methods and observed runoff.

Figure 18a shows the observed runoff for station 24.8, with the simulated runoff from the three regionalization methods from the start of 2000 to the end of 2002. Again, all the methods look pretty good, with only slight misses of the peaks in early 2000 and in late 2001. It is also worth noting that for a runoff event in early 2001, all methods miss both the increase in runoff as well as the recession. The next runoff event is then very similar again.
Figure 18b shows the mean monthly hydrograph for all methods and the observed runoff. For this monthly hydrograph, the methods seem to slightly overshoot the observed runoff. For the first two months, all the regionalization methods are very similar, while the observed runoff is higher than the methods. For the next months, the multiple regression method increases more than the other methods, and noticeably more than the observed runoff. In April, the methods show a peak, while the observed runoff continuous the drop off in runoff. The peak is bigger in the multiple regression method than the two physical similarity methods. The observed runoff shows a trough in the summer, when the evapotranspiration is so large that the runoff is almost non-existent. The observed runoff does however drop down to a lower level than for any of the regionalization methods. The methods are very similar for the trough, as well as for the increase in runoff come autumn. Nevertheless, the observed runoff is noticeably lower than the regionalization methods. When a new peak is reached in November all methods are similar to the observed runoff, before decreasing more than the observed runoff in December.

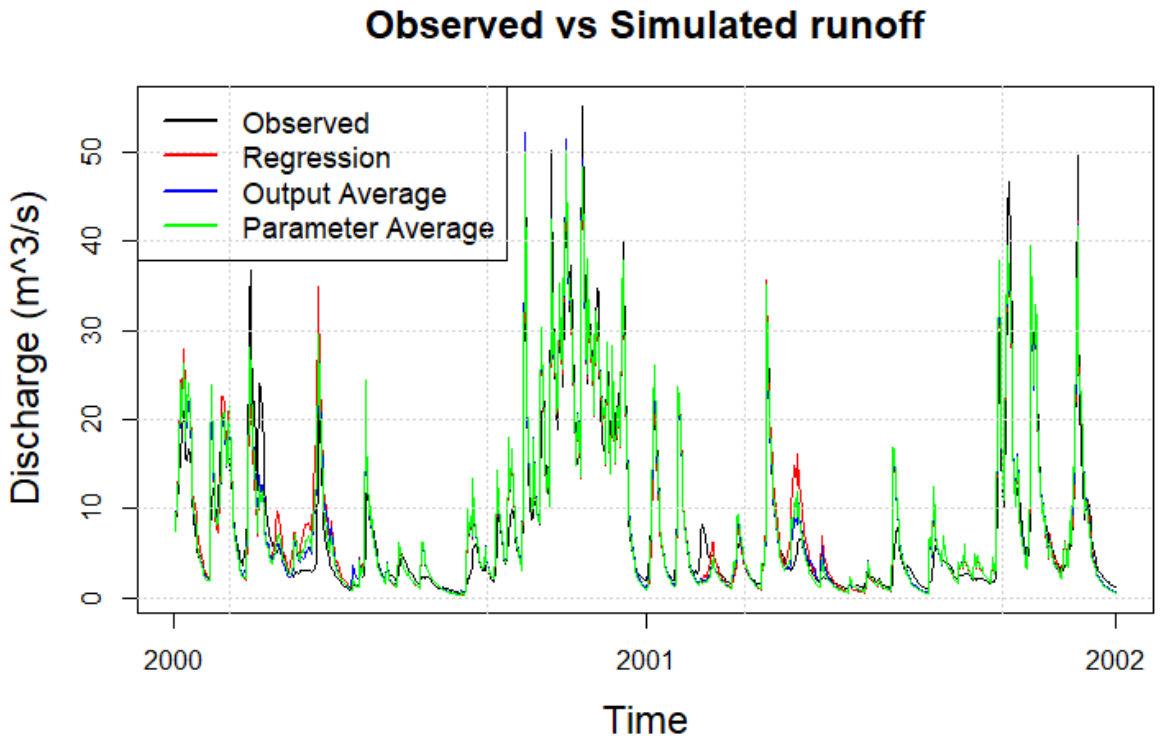


Figure 18a: Comparison of the 3 regionalization methods and the observed discharge for station 24.8, coastal region.

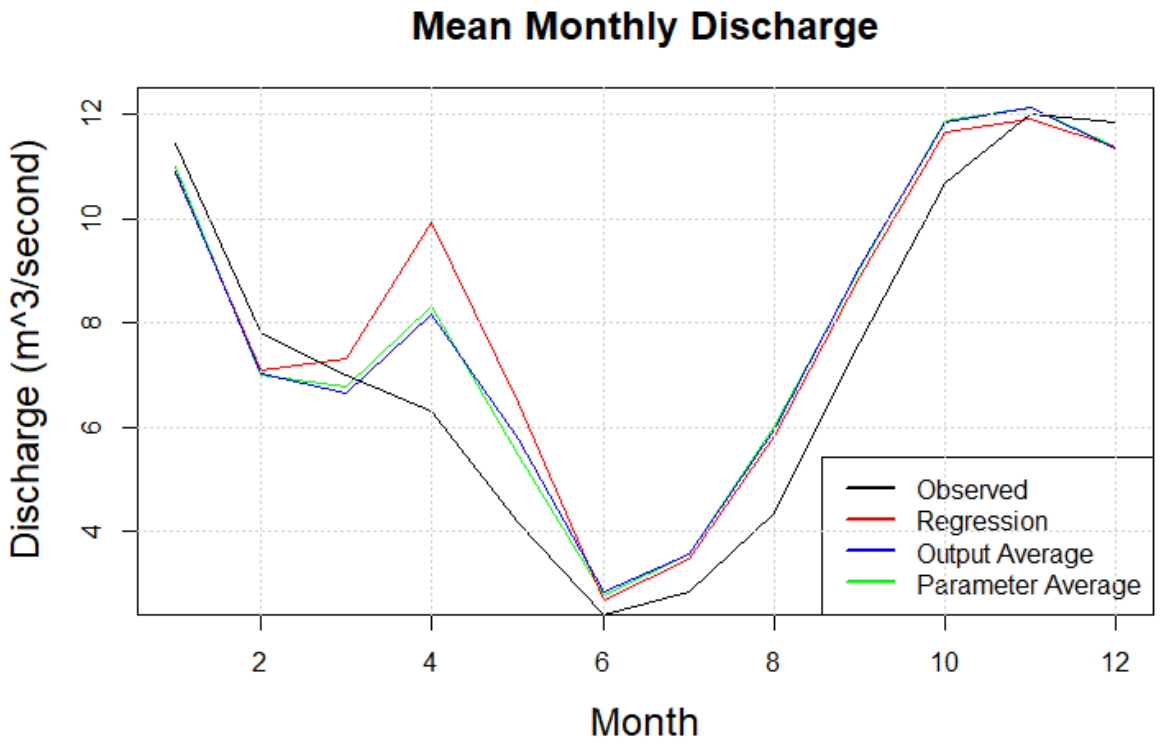


Figure 18b: Mean monthly hydrograph for station 24.8, for all regionalization methods and observed runoff.

Figure 19a shows the observed runoff for station 26.20, with the simulated runoff from the three regionalization methods from the start of 2000 to the end of 2002. All methods look fairly good, again the parameter average method has problem with overshooting the biggest peak discharge. A noticeable event happens in the middle part of 2000, with the observed runoff dropping down to almost 0 m^3/s, while the regionalization methods keep a steady runoff going.

Figure 19b shows the mean monthly hydrograph for all methods and the observed runoff. For the first three months, the regionalization methods follow the observed runoff very closely, with a slight undershooting for the February and March months. For the peak in April, all methods overshoot the observed runoff, with the parameter average method being the method with the biggest runoff peak. For the trough in the summer months, all methods are similar to the observed runoff, with the multiple regression method having a trough lower than the other methods and the observed runoff. For the increase in runoff after the summer, the multiple regression method has the best simulated runoff, being very similar to the observed runoff. The two physical similarity methods are both showing a slightly too high simulated runoff. For the peak in November, the multiple regression method shows a very similar runoff to the observed, while the two physical similarity methods have a slightly too large peak.

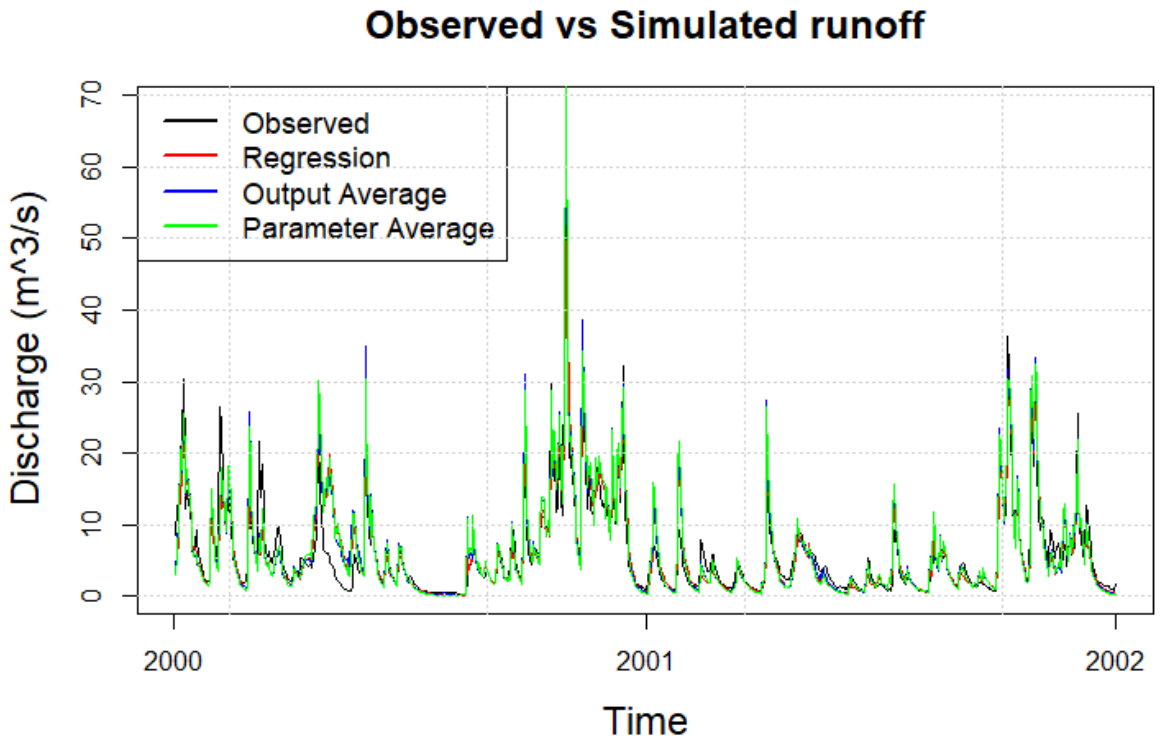


Figure 19a: Comparison of the 3 regionalization methods and the observed discharge for station 26.20, transition region.

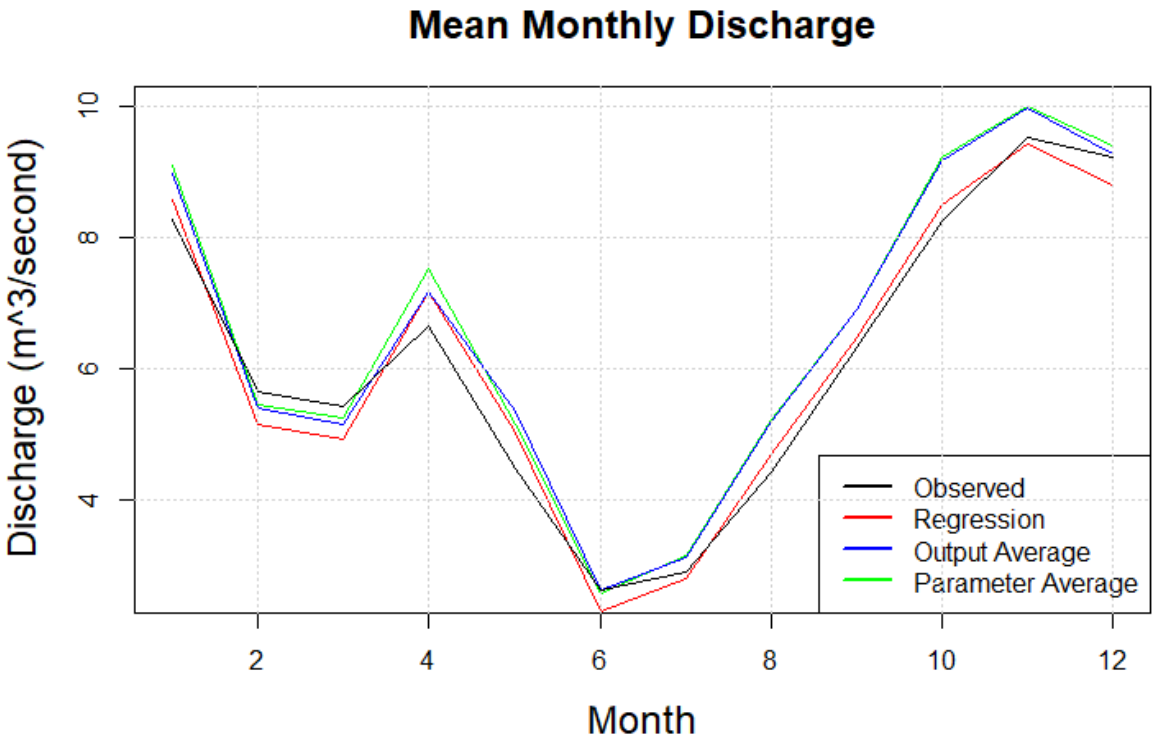


Figure 19b: Mean monthly hydrograph for station 26.20, for all regionalization methods and observed runoff.

Figure 20a shows the observed runoff for station 148.2, with the simulated runoff from the three regionalization methods from the start of 2000 to the end of 2002. Due to the fact that this catchment has the best KGE values of all the catchments chosen for their specific region, the plot also looks the best here. All methods catch all peaks and troughs and generally looks very good. Though the output average method shows a larger peak in the early 2000, and a general recession which is too slow.

Figure 20b shows the mean monthly hydrograph for all methods and the observed runoff. Starting off in January, the two physical similarity methods show a slightly smaller simulated runoff, than the observed runoff, while the multiple regression method is very similar to the observed runoff. For the peak in April however, the two physical similarity performs slightly better than the multiple regression method, with the output average method showing a very similar simulated runoff to that of the observed runoff. All methods look good on the recession toward the trough in July, with both the simulated runoff from the multiple regression method and the output average method hitting the trough from the observed runoff. The simulated runoff from the parameter average method slightly undershoots the observed runoff. In August, all methods show a sharp increase in simulated runoff, while the observed runoff is still quite low. The regionalization methods all follow the observed runoff up until the peak in September, with all methods slight overshooting the peak. Here the output average methods are closest, followed by the parameter average method and the

multiple regression method, respectively. Towards the end of the year all regionalization methods show a decrease in simulated runoff, while the observed runoff keeps on slightly increasing. The two physical similarity methods show the biggest drop in simulated runoff, while the multiple regression methods' simulated runoff is the closest to the observed runoff.

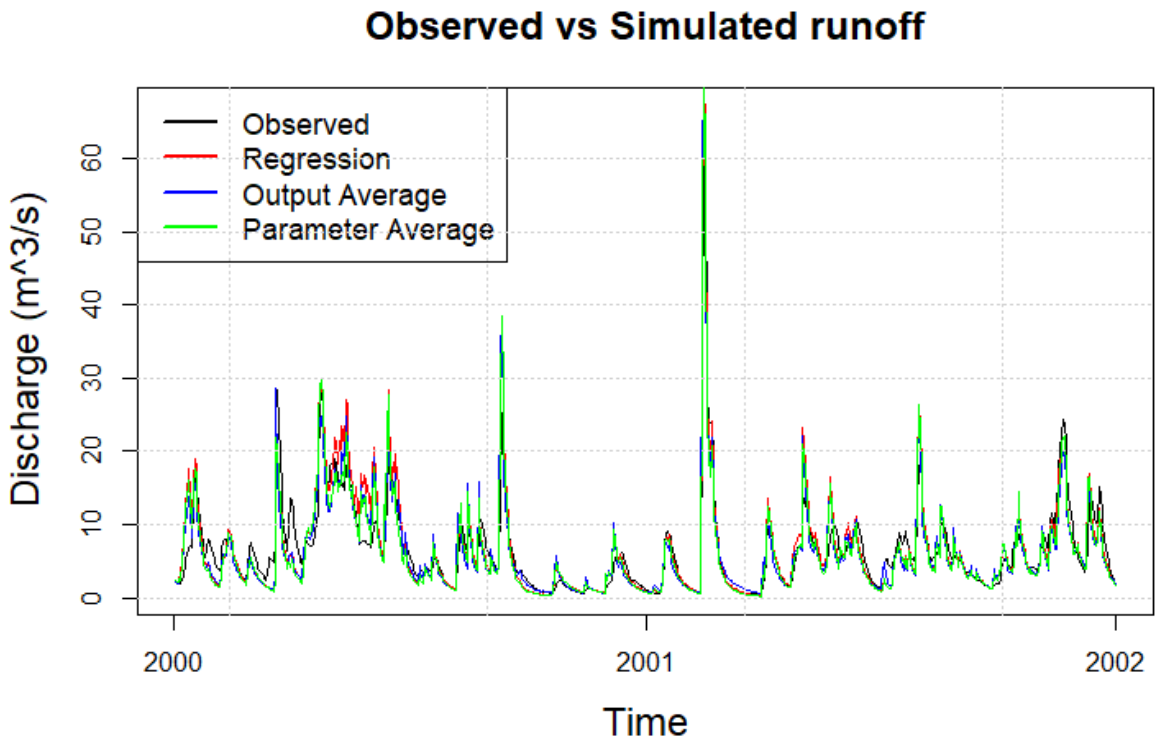


Figure 20a: Comparison of the 3 regionalization methods and the observed discharge for station 148.2, Trøndelag region.

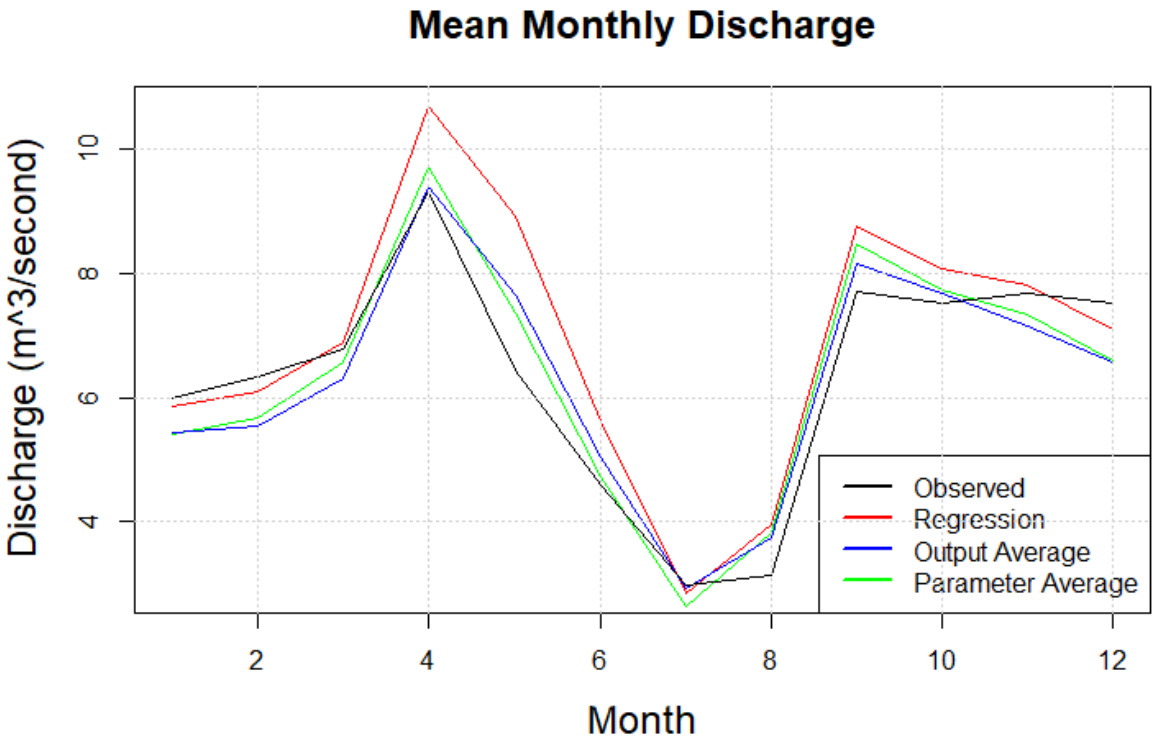


Figure 20b: Mean monthly hydrograph for station 148.2, for all regionalization methods and observed runoff.

5 Discussion

5.1 Correlation

As seen in Table 3, the correlation analysis of the relationship between the catchment characteristics and model parameters shows significant, but not very high correlations. The highest correlations are those between the *pkorr* model parameter and the geographic catchment characteristics, especially the positive bare rock percentage and negative forest percentage. This is expected as a dense forest cover will more than likely intercept some of the precipitation, and thus the negative correlation makes sense. All model parameters are significantly correlated to at least one catchment characteristics. This differs greatly from studies using parameter rich hydrological models such as the HBV-model, namely Seibert (1999) and Merz & Blöschl (2004) where a few to none of the model parameters was correlated with catchment characteristics. However, in Skaugen et al. (2009), which used the DDD model too, all model parameters were correlated to catchment characteristics.
As seen in Figure 7, the correlations between the catchment characteristics have quite a few significant correlations. All these correlations are reasonable, such as the strong positive correlation between the area of a catchment and the length of the catchment and the strong

negative correlation between percentage of forest and percentage of bare rock.
For the model parameters themselves, there is very few significant correlations. From Figure 8 one can see that the only two correlations that are significant are the positive correlation between the *skorr* and *pkorr* parameters and the negative correlation between *skorr* and *TX*. This is expected, if the *TX* value is high, less of the precipitation falls as snow and less precipitation in the form of snow needs to be corrected. If a catchment needs a lot of correction to its precipitation as rainfall, it is also quite logical that the catchment also needs that correction when the precipitation falls as snow.

5.2 Model Parameter Sensitivity

The model parameters which are most sensitive to changes in values are the *pkorr* and the *skorr* model parameters. As illustrated in Figures 9-14, one can see that these two parameters contribute the most to the changes in KGE, when the model parameters vary. However, one important thing to note is the fact that this is not the same for all regions in Norway. The *pkorr* parameter looked the same for every region, but the *skorr* parameter had a lot less impact on the model performance in the warmer regions near the coast. This is likely due to the fact that the western coast of Norway receives little to no snowfall during the winter, and thus changes in the correction factor for snow has little to say for the overall model performance. Additionally, it is important to remember that a lot of the mountainous catchments are dominated by glaciers, here the precipitation has less to say on the discharge and almost all of the discharge comes from meltwater. The meltwater will only flow out of the glacier when the energy input from the atmosphere is sufficient, as the glacier is always full of water.

5.3 Regionalization

5.3.1 Multiple Regression

The multiple regression method performs satisfactory in regionalizing the model parameters, even with quite low R2 score for the individual equations. The average KGE for the multiple regression method is 0.77 compared to the average calibrated KGE score of 0.88.
As seen in Table 4, the regressed model parameters perform well for every catchment. Therefore, it seems that the regression method does not work great at finding the individual parameters but works well on finding a set of parameters that gives a good result. From Figures 9-14, one can see that the pkorr and skorr parameters are the most sensitive, while the others have a range of parameters that can work well in the model. As seen in Table 3, the *pkorr* and *skorr* equations have higher R2 values than most other parameters, especially *pkorr* have a quite high R2. Thus, equations for *skorr* and *pkorr* being fairly robust helps the

multiple regression to good model results, even though some of the other equations are not as strong.

The five representative catchments, one for each region, that was selected perform quite well for the multiple regression method. With a range from 0.83 to 0.89 KGE, all regions performed well with the mountain region being the lowest score, and the Trøndelag and transition regions both obtain a KGE value of 0.89.

The performance of the multiple regression method in this study strengthens the statement that linking hydrological model parameters to catchment characteristics can be used to predict discharge at ungauged catchments(Magette et al., 1976). The good predictions by the multiple regression method also agrees with the recent results of Skaugen et al. (2015) and Tsegaw et al. (2019), both using the DDD model on Norwegian catchments.

5.3.2 Physical Similarity

The physical similarity method also performs relatively well. Tables 6 and 7 shows the KGE values of the model results from the two methods, while Figure 15 shows how the model results change when more donor catchments are added. The output average method performs slightly better for all numbers of donor catchments. This result contradicts many previous studies such as Parajka et al. (2005); Oudin et al. (2008); Zhang & Chiew (2009); Yang et al. (2018) and Yang et al. (2020), where the physical similarity method outperformed the multiple regression method, often by quite a lot. In this study the results are much closer, and the multiple regression method even exceeds the physical similarity method. The main difference between this study and the previous studies is the hydrological model used. The other studies all used models such as the HBV-model, while this study used a parsimonious model, the DDD model. The DDD model also has fewer model parameters to regionalize, with only 7 compared to other models which can have up to 17 parameters to regionalize (Yang et al., 2020).

This could be a good explanation as in Yang et al. (2020), different models with different number of parameters needing regionalization are compared. The difference in performance between the output average and parameter average method increased with increasing number of parameters. With the DDD model being a parsimonious model and has relatively few parameters to regionalize, the results do seem to make some sense.

There is no large, abrupt increase in KGE with increasing donor catchments, but relatively small increases with each donor catchment added. This supports the findings of Yang et al. (2020) in which small increases to the model performance was seen, but no significant increases to when exceeding 5 donor catchments was observed.

For the different regions, the physical similarity method showed divergent results. Catchment 12.70, the representative of the mountain region, showed a clear decrease in KGE score with increasing donor stations. On the other hand, all other regions showed an increasing KGE score when approaching five donor catchments, and always the best result with the full five donor catchments. The transition region, catchment 148.2, showed the best result, with a KGE score of 0.91 for both the output average method and the parameter average method.

5.4 Comparison of Regionalization Methods

Table 8 shows the model results from all the different regionalization methods and the calibration. The calibration unsurprisingly has the highest KGE value, followed by multiple regression and then the two physical similarity methods. The fact that the multiple regression method gives better model results than the physical similarity contrasts what is seen in Oudin et al. (2005) and Yang et al. (2018 & 2020), in which the multiple regression method performs the worst out of all the regionalization methods. This is most likely due to the choice of model structure and the number of and the physical realism of the model parameters, as the Yang study shows that the multiple regression method is closest to the physical similarity method in models with few and physically meaningful parameters.

However, in Tsegaw et al. (2019) which also uses the DDD model, the multiple regression method gives better model performance than the physical similarity method. This can be a further indication that for a parsimonious model, such as the DDD model, with fewer model parameters needing regionalization, the multiple regression method works the best.

This is further shown in Young (2006), which used a large dataset of 260 catchments in the United Kingdom to look at the regionalization in the Probability-Distributed Moisture (PDM) model. This model too had six model parameters to regionalize, and the study looked at two regionalization methods, relating model parameters to the catchment characteristics using regression and nearest neighbor method. The regression of the model parameters had the best results, and the fact that there were fewer model parameters that needed regionalization in that study too, could indicate that for models with fewer model parameters to regionalize, the regression method is the best.

Out of the 25 target catchments, 12 performed best using the multiple regression method, while only nine performed best for the physical similarity. The other four had a tied result. This once again indicates that the best performing method was the multiple regression method.

The five target catchment selected to represent their region all had very high KGE values for all methods, except for the mountain region, which performed very poorly using either of the physical similarity methods. One interesting thing to note is the fact that out of the remaining four target catchments, all had equal or higher KGE values for the physical similarity

methods.

The hydrographs and mean monthly hydrographs in Figures 16-20 shows that the regionalization methods perform quite well for catchments in all different regions. The biggest thing to notice is the physical similarity and their simulated runoff slightly missing some peak runoffs. This is especially clear in Figure 16b, for a mountainous region. Since the only real runoff comes from the melting of snow in the late spring, and the two physical similarity methods missing this peak by quite a lot, the KGE values for this catchment is relatively low. For the other catchments, with more peaks and troughs, this is not as noticeable. However, for both Figures 18 and 20, the multiple regression method is the method which overshoots the observed runoff.

For further studies into the predictions of ungauged basins in Norway, there are some parts that could be done better. One could use a machine learning algorithm to perfect the multiple regression equation. This would likely create even better multiple regression equations, such as seen in Wu et al. (2018). Even though that study is not related to hydrology, it shows that a machine learning algorithm can create very robust multiple regression equations.

Even though the module describing dynamics in the DDD model is primarily parametrized from observed features of the catchment, the DDD model is still, undoubtably, a simplification of reality. With abundant calibrated model parameters, that will to their best to compensate for errors, flawed structures are very difficult to identify (Kirchner, 2006). Reducing the number of model parameters to calibrate and will provide structural errors, that are easier to identify. Even though the number of model parameters that needs calibration in the DDD model is quite heavily reduced, there are still several model parameters which concerns the adjustment of input data, namely precipitation. That means that there is still room for a potentially better version of the DDD model.

6 Conclusion

This study aimed at evaluating the DDD model and its performance on predictions in ungauged basins, as well as comparing and evaluating the different regionalization methods on catchments in Norway. The comparison of methods was done at two levels, as a whole over Norway, as well as regionally over catchments in specific regions over Norway.

This study shows that the new and improved DDD model is very good at predicting hydrology in ungauged basins, with average KGE values ranging from 0.7 up to 0.77 for the different regionalization methods. As stated in Thiemig et al. (2013), $0.75 \leq KGE < 0.9$ is good, $0.5 \leq KGE < 0.75$ is intermediate and $0.0 \leq KGE < 0.5$ is poor, so the performance of the DDD model is in the high intermediate to the lower side of good performances.

The best regionalization method to use was the multiple regression method, in which the

average KGE value were 0.77, compared to 0.72 and 0.7 for the output average and parameter average, respectively.

7 References

Acreman, M. C., and C. D. Sinclair (1986), Classification of drainage basins according to their physical characteristics; an application for flood frequency analysis in Scotland, J. Hydrol., 84, 365–380.

Arnell, N., 2002. Hydrology- and Global Environmental change, Pearson Education Ltd. Prentice Hall, Harlow UK.

Arsenault, R., Poissant, D. & Brissette, F. (2015) Parameter dimensionality reduction of a conceptual model for streamflow prediction in Canadian, snowmelt dominated ungauged basins. Adv. Water Resour. 85, 27–44. doi:10.1016/ j.advwatres.2015.08.014.

Artusi, R., Verderio, P & Marubini, E. (2002) Bravais-Pearson and Spearman correlation coefficients: meaning, test of hypothesis and confidence interval. The International Journal of Biological Markers, Vol. 17 no. 2, pp. 148-151.

Bao, Z., Zhang, J., Liu, J., Fu, G., Wang, G., He, R., Yan, X., Jin, J. & Liu, H. (2012) Comparison of regionalization approaches based on regression and similarity for predictions in ungauged catchments under multiple hydro-climatic conditions. J. Hydrol. 466–467, 37–46. doi:10.1016/j.jhydrol. 2012.07.048.

Bergström S, Lindström G, Petterson A. (2002). Multi-variable parameter estimation to increase confidence in hydrological modeling. Hydrological Processes 16: 413–421. DOI: 10.1002/hyp.332

Blöschl, G., Montanari, A., (2010). Climate change impacts-throwing the dice? Hydrol. Process 24, 374–381. https://doi.org/10.1002/hyp.7574

Burn, D. H. & Boorman, D. B. (1993) Estimation of hydrological parameters at ungauged catchments. J. Hydrol. 143, 429–454. doi:10.1016/0022-1694(93)90203-L.

Chanzy, A. and L. Bruckler, 1993. Significance of soil surface moisture with respect to daily bare soil 652evapotranspiration, Water Resources Research, 29, 4, 1113-1125 (3)

Dunne, T. (1976) Energy Balance Computations of Snowmelt in a Subarctic Area DOI: 10.1029/WR012i004p00686

Gottschalk L. (1985). Hydrological regionalization of Sweden. Hydrological Sciences Journal 30(1): 65–83, doi:10.1080/02626668509490972, URL https://doi.org/10.1080/02626668509490972.

Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., (2009) Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. J. Hydrol. 377 (1), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08. 003.

He, Y., Bárdossy, A., and Zehe, E (2011).: A review of regionalization for continuous streamflow simulation, Hydrol. Earth Syst. Sci., 15, 3539–3553, https://doi.org/10.5194/hess-15-3539-2011.

Hogg, R.V., Ledolter J. 1992. Applied Statistics for Engineers and Physical Scientists: Macmillan

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A.,

Gupta, H. V., Hughes, D. A., Hut, R.W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E. & Cudennec, C.(2013) A decade of predictions in ungauged basins (PUB)—a review. Hydrol. Sci. J. 58, 1198–1255. doi:10.1080/02626667.2013.803183

Kirchner JW. 2006. Getting the right answer for the right reasons: linking measurements, analysis, and models to advance the science of hydrology. Water Resources Research 42:W03S04. DOI: 10.1029/2005WR004362

Kjeldsen TR, Jones DA, Morris DG. (2014). Using multiple donor sites for enhanced flood estimation in ungauged catchments. Water Resources Research 50(8): 6646–6657, doi:10.1002/2013wr015203, URL https://doi.org/10.1002/2013wr015203.

Kokkonen TS, Jakeman AJ, Young PC, Koivusalo HJ. (2003). Predicting daily flows in ungauged catchments: model regionalization from catchment descriptors at the Coweeta Hydrologic Laboratory, North Carolina. Hydrological Processes 17(11): 2219– 2238, doi:10.1002/hyp.1329

Magette, W.L., Shanholtz, V.O., Carr, J.C., (1976). Estimating selected parameters for the Kentucky Watershed Model from watershed characteristics. Water Resour. Res. 12 (3), 472–476. https://doi.org/10.1029/WR012i003p00472.

McIntyre, N., Lee, H., Wheater, H., Young, A. & Wagener, T. (2005) Ensemble predictions of runoff in ungauged catchments. Water Resour. Res. 41, 1–14. doi:10.1029/2005WR004289

Merz, R. & Blöschl, G. (2004) Regionalization of catchment model parameters. J. Hydrol. 287, 95–123. doi:10.1016/j.jhydrol. 2003.09.028.

Merz, R., Blöschl, G. & Parajka, J. (2006) Regionalization methods in rainfall-runoff modelling using large samples. Lr. Model Param. Exp. IAHS Publ. 307, 117–125.

Moore RJ. (1985). The probability-distributed principle and runoff production at pint and basin scales. Hydrological Sciences Journal 30: 273–297

Nathan, R. J., and T. A. McMahon (1990), Identification of homogeneous regions for the purposes of regionalization, J. Hydrol., 121, 217–238.

Oudin, L., Andréassian, V., Perrin, C., Michel, C. & Le Moine, N. (2008) Spatial proximity, physical similarity, regression, and ungauged catchments: a comparison of regionalization approaches based on 913 French catchments. Water Resour. Res. 44, 1–15. doi:10.1029/2007WR006240

Parajka, J., Viglione, A., Rogger, M., Salinas, J. L., Sivapalan, M., and Blöschl, G. (2013): Comparative assessment of predictions in ungauged basins – Part 1: Runoff-hydrograph studies, Hydrol. Earth Syst. Sci., 17, 1783–1795, https://doi.org/10.5194/hess-17-1783-2013.

Parajka, J., Blöschl, G. & Merz, R. (2007) Regional calibration of catchment models: potential for ungauged catchments. Water Resour. Res. 43. doi:10.1029/2006WR005271.

Parajka J, Blöschl G. 2008. The value of MODIS snow cover data in validating and calibrating conceptual hydrological models. Journal of Hydrology 358: 240–258.

Parajka, J., Merz, R. & Blöschl, G. (2005) A comparison of regionalization methods for catchment model parameters. Hydrol. Earth Syst. Sci. Discuss. 2, 509–542. doi:10.5194/ hessd-2-509-2005.

Poissant D, Arsenault R, Brissette F. (2017) Impact of parameter set dimensionality and calibration procedures on streamflow prediction at ungauged catchments. Journal of 47 Hydrology: Regional Studies 12: 220–237, doi:10.1016/j.ejrh.2017.05.005

Priestley, C. H. B. & Taylor, R. J. (1972). On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters. Monthly Weather Review, 100 (2): 81-92. doi: 10.1175/1520-0493(1972)100<0081:Otaosh>2.3.Co;2.

Ramirez, J. A. (2000). Prediction and modeling of flood hydrology and hydraulics. In Wohl, E. (ed.) Inland Flood Hazards: Human, Riparian and Aquatic Communites Cambridge University Press.

Seibert J. (1999). Regionalization of parameters for a conceptual rainfall runoff model. Agricultural and Forest Meteorology 98-99: 279–293.

Shah, Viral B., Jeff Bezanson, Alan Edelman, Stefan Karpinski (2017) Julia: A Fresh Approach to Numerical Computing. SIAM Review, 59: 65–98. doi: 10.1137/141000671. pdf.

Skaugen, Thomas (2018) “The Distance Distribution Dynamics Model- Development for use in ungauged catchments” NCCS report no. 1/2018.

Skaugen, T. and Onof, C. (2014): A rainfall runoff model parameterized form GIS and runoff data, Hydrol. Process., 28, 4529–4542, doi:10.1002/hyp.9968, 2014

Skaugen, T., Peerebom, I. O., and Nilsson, A. (2015): Use of a parsimonious rainfall–runoff model for predicting hydrological response in ungauged basins, Hydrol. Process., 29, 1999–2013, doi:10.1002/hyp.10315, 2015.

Skaugen, T., Onof, C., (2014). A rainfall-runoff model parameterized from GIS and runoff data. Hydrol. Process. 28 (15), 4529–4542. https://doi.org/10.1002/hyp.9968.

Skaugen, Thomas & Lawrence, Deborah & Ortega, Rengifo. (2020). A parameter parsimonious approach for catchment scale urban hydrology – Which processes are important?. Journal of Hydrology X. 8. 100060. 10.1016/j.hydroa.2020.100060.

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J. J., Mendiondo, E. M., O’Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S. & Zehe, E. (2003) IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences. Hydrological Sciences Journal 48 (6), 857–880. https://doi.org/10.1623/hysj.48.6.857.51421

Stavang, A. E. (2019) Rainfall-runoff modelling with high temporal resolution on the arable catchment Skuterud - Assessing the performance of the Distance Distribution Dynamics model. http://hdl.handle.net/11250/2609963

Sælthun NR. (1996). The ``Nordic'' HBV model. Description and documentation of the model version developed for the project Climate Change and Energy Production, NVE Publication no. 7-1996, Oslo, 26 pp.

Tallaksen, L. M. (1995). A review of baseflow recession analysis. Journal of Hydrology, 165: 349-370.

Tegegne, G., Kim, Y., & Park, D. (2017) Comparison of hydrological models for the assessment of water resources in a data-scarce region, the Upper Blue Nile River Basin. https://doi.org/10.1016/j.ejrh.2017.10.002

Thiemig, V., Rojas, R., Zambrano-Bigiarini, M., De Roo, A., (2013). Hydrological evaluation of satellite-based rainfall estimates over the Volta and Baro-Akobo Basin. J. Hydrol. 499, 324–338. https://doi.org/10.1016/j.jhydrol.2013.07.012.

Thomas, B., Vogel, R., Famiglietti, J. (2015) Objective hydrograph baseflow recession analysis https://doi.org/10.1016/j.jhydrol.2015.03.028

Tsegaw, A., Alfredsen, K., Skaugen, T. & Muthanna, T. (2019). Predicting hourly flows at ungauged small rural catchments using a parsimonious hydrological modelhttps://doi.org/10.1016/j.jhydrol.2019.03.090

Vandekerckhove, J., Matzke, D., Wagenmakers, E. (2015) Model Comparison and the Principle of Parsimony DOI: 10.1093/oxfordhb/9780199957996.013.14

Wilkerson, J., Merwade (2010) Incorporating surface storage and slope to estimate Clark Unit Hydrographs for ungauged Indiana watersheds J. Hydrol. Eng., 15 (11) (2010), pp. 918-930

Wu, C., Patil, P., Gunaseelan, S. (2018) Comparison of Different Machine Learning Algorithms for Multiple Regression on Black Friday Sales Data DOI: 10.1109/ICSESS.2018.8663760

Yadav M, Wagener T, Gupta H. (2007). Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. Advances in Water Resources 30: 1756–1774.

Yang, X., Magnusson, J., Rizzi, J., & Chong-Yu, X. (2018) Runoff prediction in ungauged catchments in Norway: comparison of regionalization approaches Hydrology Research 49(2) DOI: 10.2166/nh.2017.071

Yang, X., Magnusson, J., Huang, S., Beldring, S., Xu, C. (2020) Dependence of regionalization methods on the complexity of hydrological models in multiple climatic regions. Journal of Hydrology Volume 582, March 2020, 124357.

Young A. (2006). Stream flow simulation within UK ungauged catchments using a daily rainfall- runoff model. Journal of Hydrology 320:

Zar, Jerrold H. (1972) Testing of the Spearman Rank Correlation Coefficient, Journal of American Statistical Association

Zhang, Y., Liu, Y., Yang, X. (2015) Parametric Sensitivity Analysis for Importance Measure on Failure Probability and Its Efficient Kriging Solution, https://doi.org/10.1155/2015/685826

Zhang, Y. & Chiew, F. (2009) Evaluation of regionalization methods for predicting runoff in ungauged catchments in Southeast Australia. In: 18th World IMACS/MODSIM Congress, Cairns, Australia, pp. 3442–3448