Hydrol. Earth Syst. Sci. Discuss., 12, 11129–11171, 2015
www.hydrol-earth-syst-sci-discuss.net/12/11129/2015/
doi:10.5194/hessd-12-11129-2015

Hydrology and Earth System Sciences
Open Access
Discussions

Estimating catchment scale groundwater dynamics from recession analysis – enhanced constraining of hydrological models

T. Skaugen and Z. Mengistu

Department of Hydrology, Norwegian Water Resources and Energy Directorate, Oslo, Norway

Received: 11 September 2015 – Accepted: 14 October 2015 – Published: 30 October 2015

Correspondence to: T. Skaugen (ths@nve.no)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

In this study we propose a new formulation of subsurface water storage dynamics for use in rainfall–runoff models. Under the assumption of a strong relationship between storage and runoff, the temporal distribution of storage is considered to have the same shape as the distribution of observed recessions (measured as the difference between the log of runoff values). The mean subsurface storage is estimated as the storage at steady-state, where moisture input equals the mean annual runoff. An important contribution of the new formulation is that its parameters are derived directly from observed recession data and the mean annual runoff and hence estimated prior to calibration. Key principles guiding the evaluation of the new subsurface storage routine have been (a) to minimize the number of parameters to be estimated through the, often arbitrary fitting to optimize runoff predictions (calibration) and (b) maximize the range of testing conditions (i.e. large-sample hydrology). The new storage routine has been implemented in the already parameter parsimonious Distance Distribution Dynamics (DDD) model and tested for 73 catchments in Norway of varying size, mean elevations and landscape types. Runoff simulations for the 73 catchments from two model structures; DDD with calibrated subsurface storage and DDD with the new estimated subsurface storage were compared. No loss in precision of runoff simulations was found using the new estimated storage routine. For the 73 catchments, an average of the Nash–Sutcliffe Efficiency criterion of 0.68 was found using the new estimated storage routine compared with 0.66 using calibrated storage routine. The average Kling–Gupta Efficiency criterion was 0.69 and 0.70 for the new and old storage routine, respectively. Runoff recessions are more realistically modelled using the new approach since the root mean square error between the mean of observed and simulated recessions was reduced by almost 50 % using the new storage routine.

HESSD
12, 11129–11171, 2015

1 Introduction

The movement of groundwater to streams is an important component of catchment hydrology and simulating its movement is key to accurately reproducing the hydrograph. Unfortunately, at the spatial scale of interest for studying the dynamics of hydrological systems, the catchment scale, we are not able to actually see and learn how water is transported in the subsurface. Hence, for many decades the (subsurface) storage–runoff relationship has been the basis for countless hydrological model concepts. The subsurface water storage, hereafter denoted subsurface storage or storage, is to be understood as the dynamic storage, i.e. the variation in storage between wet and dry periods (Kirchner, 2009)

The linear reservoir, often visualised as a straight-sided bucket with a hole in the bottom (Dingman, 2002; Beven, 2001), has served as the most commonly used basic storage–runoff relationship. Such a reservoir has an exponentially declining outflow, and is the basis for the exponential unit hydrograph (UH). A single linear reservoir is too simple, however, for describing the variability and non-linearity of hydrological response. Therefore, most conceptual models use a system of several, possibly modified, linear reservoirs to describe the soil moisture accounting and runoff dynamics. The system may vary in complexity (and hence in the inclusion of calibration parameters), but the linear reservoir remains the basic building block of conceptual models. Examples of such models are the UH models of Nash (1957) and Dooge (1959) and the explicit soil-moisture accounting (ESMA) models, of which the work-horse of operational Nordic hydrology, the HBV model (Bergström, 1992) serves as an example (see Beven (2001) for a discussion on the evolution of rainfall–runoff models). In addition to the fundamental role the linear reservoir has played in simple conceptual rainfall–runoff models, some groundwater models conceptualise the stream- aquifer interactions as the drainage of an infinite number of independent linear reservoirs (Rupp et al., 2009; Bidwell et al., 2008; Pulido-Velasquez et al., 2005; Sloan, 2000). This comes as a result of solving the linearized Dupuit–Boussinesq equation for

HESSD
12, 11129–11171, 2015

saturated flow as an eigenvalue and eigenfunction problem. Despite the popularity of linear reservoirs, the non-linear relationship between storage, S and runoff Q has long been recognised and simple solutions for manipulating a single reservoir for taking into account non-linearity have been put forward. In Lindström et al. (1997) the upper zone (the reservoir responsible for quick response) of the HBV model was formulated as a non-linear reservoir, $Q = \vartheta S^{1+\delta}$ where ϑ and δ are calibrated constants. For $\delta = 0$, this is, of course an ordinary linear reservoir. The recently published rainfall–runoff model DDD (Distance Distribution Dynamics, Skaugen and Onof, 2014; Skaugen et al., 2015) is also based upon a high dependence between runoff and storage and uses linear reservoirs as its primary building block. In this model, the dynamics of runoff are modelled using linear storages arranged in parallel, a principle which resembles the stream- aquifer interaction model described by for example Bidwell et al. (2008). The non-linearity of the response in DDD comes from exponential UHs of different temporal scale.

Brutsaert and Nieber (1977) discuss several theoretical models from the soil sciences as a basis for describing the non-linearity of storage–runoff relationships and investigate this relationship using data from runoff recession events. Recession data have often been used to derive the storage–runoff relationship and Lamb and Beven (1997) developed a tool that used recession data to parameterize non-linear storage–runoff relationships but were not always able to fit single analytical functions. In Kirchner (2009), runoff is assumed to depend solely on the amount of water stored in the catchment and very carefully selected recession events are used to parameterize the storage–runoff relationship. The recession events were selected such that the possible contaminating effect of precipitation and evapotranspiration on the recession data was minimized. For two rivers in the UK, highly non-linear relationships between storage and runoff were found using this approach. As in Brutsaert and Nieber (1977) and Kirchner (2009), recession data are fundamental in DDD for describing the runoff dynamics. The temporal scales of the UHs are estimated assuming that the recessions provide the parameters of exponential UHs, which, together with the distribution

HESSD
12, 11129–11171, 2015

describing the distances from points in the catchment to their nearest river reach, are used to derive celerities and, hence, the temporal scale of the UHs. The linearity of the parallel reservoirs is not assumed but is dictated by the empirical distance distributions, which for Norwegian hillslopes can usually be modelled as exponential (Skaugen and Onof, 2014; Skaugen et al., 2015). The UHs are turned on and off according to the level of saturation (or storage) in the catchment.

The DDD model was developed with the aim of investigating how far one could parameterize a rainfall–runoff model using information obtained from maps and runoff records prior to model calibration. The result was a model that had no loss in precision or detail when compared with the HBV model, although the number of calibration parameters in the subsurface and dynamic modules was reduced from 7 (HBV) to 1 (DDD). This study is a continuation of that approach, and the aim is to investigate how storage dynamics can be related to runoff dynamics within the DDD framework. The recession data continue to play a crucial role in the model formulation and using these data together with the distance distributions and the mean annual runoff (MAR), we attempt to estimate all parameters of the subsurface and runoff dynamics prior to model calibration.

2 Methods

2.1 Hydrological model

The DDD model (Skaugen and Onof, 2014; Skaugen et al.2015) is a rainfall–runoff model written in the programming language R (www.r-project.org) and currently runs operationally at daily and 3 hourly time steps at the Norwegian flood forecasting service. The DDD model introduces new concepts in its description of the subsurface and of runoff dynamics. Input to the model is precipitation and temperature. In the subsurface module (see Fig. 1), the volume capacity of the subsurface water reservoir M (mm) is shared between a saturated zone with volume S (mm), called the

HESSD
12, 11129–11171, 2015

groundwater zone and an unsaturated zone with volume D (mm), called the soil water zone. The actual water volume 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 and snowmelt, $G(t)$,represent an excess of water over the field capacity, R, which is fixed at 30 % ($R = 0.3$) of $D(t)$ (Grip and Rohde, 1985, p. 26; Colleuille et al., 2007). If so, excess water $X(t)$ is added to $S(t)$. To summarize:

Excess water:
$$X(t) = \mathrm{Max}\left\{\frac{G(t)+Z(t)}{D(t)} - R, 0\right\} D(t). \quad (1a)$$

Groundwater:
$$\frac{\mathrm{d}S}{\mathrm{d}t} = X(t) - Q(t). \quad (1b)$$

Soil water content:
$$\frac{\mathrm{d}Z}{\mathrm{d}t} = G(t) - X(t) - \mathrm{Ea}(t). \quad (1c)$$

Soil water zone:
$$\frac{\mathrm{d}D}{\mathrm{d}t} = -\frac{\mathrm{d}S}{\mathrm{d}t}, \quad (1d)$$

where $Q(t)$ is runoff. Actual evapotranspiration, $\mathrm{Ea}(t)$, is estimated as a function of potential evapotranspiration and the level of storage. Potential evapotranspiration is estimated as $\mathrm{Ep} = \theta_{\mathrm{cea}} \times T$ ($\mathrm{mm\,day^{-1}}$), where θ_{cea} ($\mathrm{mm\,^{\circ}C^{-1}\,day^{-1}}$) is the degree-day factor which is positive for positive temperatures (T) and zero for negative temperatures. Actual evapotranspiration thus becomes $\mathrm{Ea} = \mathrm{Ep} \times (S+Z)/M$, and is drawn from Z.

In the current version of DDD, M is a calibrated parameter and is divided into storage levels i, which are all assigned different wave velocities, or celerities, v_i ($\mathrm{m\,s^{-1}}$) (see next section). Experience using the DDD model shows that the subsurface water reservoir M largely controls the variability of the hydrograph. Low values of M increase the amplitude of the hydrograph, since the entire range of celerities is engaged, and vice versa.

HESSD
12, 11129–11171, 2015

2.2 Runoff dynamics

The runoff dynamics are completely parameterized from observed catchment features derived using a Geographical Information System (GIS) and runoff recession analysis. Central for the formulation of runoff dynamics for a catchment is the distance distribution derived using GIS. The distances, d, from points in the catchments to the nearest river reach are calculated for each catchment and for more than 120 studied catchments in Norway the exponential distribution describe the distribution of distances well. Figure 2 shows the empirical and exponential distributions for two Norwegian catchments and although the mean distance $\overline{d}$ is different, the exponential distribution is a good fit for both catchments. The parameter γ of the exponential distribution

$$f(d) = \gamma e^{-\gamma d}, \tag{2}$$

equals $\gamma = 1/\overline{d}$. The distance distributions (Fig. 2) express the areal fraction of the catchment as a function of distance from the river network.

In Fig. 3 the information of the distance distribution is visualised differently. Here, for the same two catchments as in Fig. 2, the consecutive fractional areas for each distance interval Δd are plotted against the distance to the river network, and the ratio, κ between consecutive fractional areas is a constant and it has been showed (Skaugen, 2002) that the parameter γ of the exponential distribution relates to κ as

$$\gamma = -\log(\kappa)/\Delta d. \tag{3}$$

If we assume that a uniform moisture input (i.e. excess rainfall or snowmelt) is transported through the hillslope to the river network with a constant velocity v, (or celerity, see Skaugen and Onof, 2014; Beven, 2006), then Δd is the distance travelled by water during a suitable time step, Δt, i.e., $\Delta d = v\Delta t$. When d in Eq. (2) is replaced with d/v, the distance distribution hence becomes a travel-time distribution with mean equal to $\frac{\overline{d}}{v}$ and parameter

$$\xi = -\log(\kappa)/\Delta t, \tag{4}$$

HESSD
12, 11129–11171, 2015

which constitutes a unit hydrograph (Maidment, 1993; Bras, 1990, p. 448). The variable κ, is now the ratio between volumes of water drained pr. time step, i.e. the volume of water drained into the river network is reduced by κ for each time step.

A linear reservoir has this same property of consecutive runoff values having a constant ratio. This can be seen if we compute successive volumes and runoff values according to a linear reservoir in recession with runoff coefficient ϑ, i.e. $Q(t) = \vartheta S(t)$ The ratio between consecutive values of runoff, $Q(t+1)/Q(t)$ remains constant and equal to $1-\vartheta$. Hence, a catchment with an exponential distance distribution and a constant celerity is equivalent to a linear reservoir with a runoff coefficient equal to $(1-\kappa)$, i.e.

$$Q(t) = (1-\kappa)S(t). \tag{5}$$

Furthermore, from Eqs. (4) and (5) we see that the runoff coefficient of a linear reservoir relates to the parameter of the travel time distribution as:

$$\vartheta = 1 - e^{-\xi \Delta t} \tag{6}$$

and since the mean of the travel-time distribution is $\frac{1}{\xi} = \frac{\overline{d}}{v}$, the runoff coefficient relates to the mean of the distance distribution as:

$$\vartheta = 1 - e^{-(v/\overline{d})\Delta t} \tag{7}$$

and the celerity can hence be formulated as:

$$v = \frac{-\log(1-\vartheta)\overline{d}}{\Delta t} = \frac{-\log(\kappa)\overline{d}}{\Delta t}. \tag{8}$$

This brief discussion on the distance distribution and linear reservoirs is relevant because it suggest that if a catchment exhibits an exponential distance distribution the linear reservoir comes as a natural choice for modelling the interaction between

HESSD
12, 11129–11171, 2015

hillslopes and the river network. Furthermore, the distance distribution suggest a geometrical configuration of the hillslope (or aquifer) (Fig. 3) and the linear reservoir model is partly parameterised from the parameter of the distance distribution Eq. (7). These latter statements assumes, of course, that the topographical catchment area and that of the aquifer are equal, an assumption that does not always hold (Bidwell et al., 2008).

In the DDD model, water is conveyed through the soils to the river network by waves with celerities determined by the actual storage, $S(t)$ in the catchment. The celerities associated with the different storages are estimated by assuming exponential recessions with parameter Λ, in $Q(t) = Q_0 \Lambda e^{-\Lambda(t-t_0)}$, where Q_0 is the peak discharge immediately before the recession starts (Nash, 1957). We can determine the parameter $\Lambda(t)$ from the difference, $\log(Q(t)) - \log(Q(t+\Delta t))$, at any time t, during the recession due to the lack of-memory property of the exponential distribution (Feller, 1971, p. 8),

$$\Lambda(t) = \log(Q(t)) - \log(Q(t+\Delta t)). \tag{9}$$

The parameter Λ is thus the slope per Δt of the recession in the log-log space and we see the relation between the variable $\kappa = Q(t+\Delta t)/Q(t)$ and Λ as:

$$\Lambda = -\log(\kappa). \tag{10}$$

From Eq. (8) we have that the celerity v as a function of Λ is:

$$v = \frac{\Lambda \overline{d}}{\Delta t}. \tag{11}$$

If we sample Λs from recession events according to Eq. (9), we find that they have a distribution which can be fitted to a gamma distribution. This is a development from the exponential model used in Skaugen and Onof (2014) and is based on more detailed analysis of a much larger number of runoff records. For the 73 catchments used in this study, the gamma distribution was a good fit for all catchments. In Fig. 4 we have plotted

HESSD
12, 11129–11171, 2015

the empirical and the gamma distribution of Λ for 6 catchments, and it is clearly seen that the flexibility of the gamma distribution is needed in order to model the observed quantiles (see for example the middle and bottom panels on the right side).

The capacity of the subsurface reservoir M, is divided into storage levels i corresponding to the quantiles of the distribution of Λ under the assumption that the higher the storage, the higher the values of Λ. Each level is further assigned a celerity $v_i = \frac{\lambda_i \overline{d}}{\Delta t}$ (see Eq. 11), where λ_i is the parameter of the unit hydrograph for the individual storage level i, and estimated such that the runoff from several storage levels will give a UH equal to the exponential UH with parameter Λ_i, i.e.:

$$\Lambda_i e^{-\Lambda_i (t-t_0)} = \varpi_1 \lambda_1 e^{-\lambda_1 (t-t_0)} + \varpi_2 \lambda_2 e^{-\lambda_2 (t-t_0)} + \ldots + \varpi_i \lambda_i e^{-\lambda_i (t-t_0)}, \quad (12)$$

where ϖ are the weights associated with the discharge from each level estimated by $\varpi_i = \Lambda_i / \sum_{k=1}^{i} \Lambda_k$. From Eq. (12) the λ_i are solved successively for increasing i under the assumption that $\lambda_1 = \Lambda_1$ (see Skaugen and Onof, 2014).

The quantiles of Λ are mapped to a uniform distribution of S, $F(\Lambda) = \frac{S}{M}$, which implies that all storage levels are equally probable and that the equally-spaced storage levels have equal capacity of water, i.e. if $M = 100$ mm and we use 10 storage levels, $i = 1 \ldots 10$, each level has a capacity of 10 mm. In Skaugen and Onof (2014) no increase in precision in daily runoff simulations using more than 5 storage levels was found.

Calibrated model parameters are hereafter denoted by θ with subscripts (e.g. θ_{M}), in order to clearly distinguish between estimated and calibrated parameters.

2.3 Reformulation of the subsurface of DDD

An obvious problem of the approach described above is that we attempt to estimate an extreme value, the maximum catchment scale storage θ_{M}, a task which is obviously associated with more uncertainty than estimating the mean catchment scale storage, m_{s}. Another problem is the assumption of a uniform distribution of storage levels.

HESSD
12, 11129–11171, 2015

A quick investigation of observed groundwater level fluctuations suggests that this is not the case. Figure 5 shows histograms of observed groundwater levels from three observation boreholes located in a small catchment (the Groset catchment, 6.33 km^2) in southern Norway. The figure clearly illustrates that fluctuations in storage and groundwater levels are spatially variable and should ideally be treated as such in rainfall–runoff models (Rupp et al., 2009; Sloan, 2000). This is a consequence of the differences in water level fluctuations depending on the location of the borehole relative to the river, i.e. top of a hillslope vs. adjacent to a river, and also of catchment topographic variability. It is therefore very difficult to parameterize the distribution of the catchment-scale groundwater fluctuations from such single observation points (Kirchner, 2009). In addition, the distribution is unlikely to be uniform as none of the individual histograms exhibits such a behaviour.

To overcome the problems identified above, we attempt to develop a subsurface model that differs from the previous model in that the groundwater reservoir is parameterised by its mean storage, m_s, as opposed to the maximum storage, θ_M. In addition, regarding the practical problems associated with the non-observability of catchment scale fluctuations of storage, we make the assumption that recession and its distribution carries information on the distribution of catchment-scale storage. More precisely, we assume that the temporal distribution of catchment- scale storage can be considered as a scaled version to that of Λ. Consequently, the subsurface reservoir no longer increases linearly with the quantiles (which is the case with storage levels of equal capacity), but rather, increases non-linearly according to the shape of the distribution of Λ. The assumption of equal shape for the distributions of Λ and S is, of course, difficult to verify as no direct observations of S are at hand. However, if we use the equation for the linear reservoir Eq. (5) and express the runoff coefficient as a function of $\Lambda(t)$ Eq. (10), we can, for observed values of $Q(t)$ and $\Lambda(t)$, calculate the corresponding values of $S(t)$ and compare the distributions of $\Lambda(t)$ and (the scaled) $S(t)$.

HESSD
12, 11129–11171, 2015

$$S(t) = \frac{Q(t)}{1 - e^{-\Lambda(t)}} \tag{13}$$

Figure 6 shows such a comparison for two catchments, and, except for the highest quantiles, the distributions of $\Lambda(t)$ and (scaled) $S(t)$ are almost identical and hence supporting our assumption. The high frequency of high $S(t)$ values present in Fig. 6, also seen for several other catchments (not shown), is the result of the combination of high $Q(t)$ values and low values of $\Lambda(t)$, i.e. very modest recession for situations with high runoff values. Such events are probably not representative for describing recessions, and by sampling $\Lambda(t)$ and estimating $S(t)$ under the condition that precipitation at the day of $(t + \Delta t)$ could not exceed a low threshold (for example 0, 2 and 5 mm) we found that the frequency of very high values of $S(t)$ were reduced. Hence, the very high values of $S(t)$ did not represent recession events. Moreover, the distribution of Λ was insensitive to such conditioning, implying that Eq. (9) is a robust estimate of recession characteristics, whereas the distribution of $S(t)$ was highly sensitive.

Since the distribution of Λ is modelled as a two parameter gamma distribution, we can write

$$f(\Lambda) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)}\Lambda^{\alpha-1}\exp(-\Lambda/\beta), \quad \alpha > 0, \beta > 0 \tag{14}$$

where α and β are the shape and scale parameters respectively and estimated from observed Λs (using Eq. 9)

The distribution of S is hence also modelled as a two-parameter gamma distribution:

$$f(S) = \frac{1}{\eta^{\alpha}\Gamma(\alpha)}S^{\alpha-1}\exp(-S/\eta), \quad \alpha > 0, \eta > 0 \tag{15}$$

where the scale parameter, η, is

$$\eta = \beta/c \tag{16}$$

HESSD
12, 11129–11171, 2015

and c is a constant and equal to

$$c = \overline{\Lambda}/m_s, \tag{17}$$

where $\overline{\Lambda}$ is the mean value of Λ, estimated from the parameters of the fitted gamma distribution and representing the mean recession characteristic. Note that since the distribution of S is a scaled version of Λ, the shape parameter α is equal for the two distributions.

In order to model the storage as a two-parameter gamma distribution we need to estimate the mean storage, m_S. We can then determine the constant c from Eq. (17), and finally, the scale parameter η using Eq. (16).

If we assume that the mean value of the sampled Λs, $\overline{\Lambda}$, represents the slope of recession in a state of mean storage in the catchment, then the associated unit hydrograph (UH) is,

$$u_{\overline{\Lambda}}(t) = \overline{\Lambda} e^{-\overline{\Lambda}(t-t_0)}. \tag{18}$$

The temporal scale of the UH in Eq. (18) is $t_{\mathrm{h,\,max}} = d_{\mathrm{max}}/\overline{v}_{\mathrm{h}}$, where d_{max} is the observed maximum distance of the hillslope distance distribution and $\overline{v}_{\mathrm{h}}$ is the celerity associated with $\overline{\Lambda}$ through $\overline{v}_{\mathrm{h}} = \frac{\overline{\Lambda}\overline{d}}{\Delta t}$ (see Eq. 11). Let $t_{\mathrm{h,\,max}}$ be divided into suitable time intervals, Δt, then the number of time intervals it takes to drain the hillslope is $J = t_{\mathrm{h,\,max}}/\Delta t$. When Eq. (18) is integrated over successive time intervals we obtain weights, μ_j, which if multiplied by the excess moisture input, $X(\Delta t)$, give the response (the water entering the river network) for the different time intervals. The weights are calculated as:

$$\mu(\overline{\Lambda})_j = \int_{(j-1)\Delta t}^{(j)\Delta t} u_{\overline{\Lambda}}(t)\,\mathrm{d}t, \quad j = 1 \ldots J, \sum \mu(\overline{\Lambda})_j = 1, \tag{19}$$

HESSD
12, 11129–11171, 2015

and scaled so that the sum of weights equals 1. The runoff at time interval j is calculated as

$$Q(j\Delta t) = X(\Delta t)\mu(\overline{\Lambda})_j. \quad (20)$$

2.3.1 Estimating the mean storage m_S

From runoff observations, we can calculate the mean annual runoff, MAR, which corresponds to a daily excess moisture input X of

$$X\left[\mathrm{mm\,day^{-1}}\right] = (1000 \times \mathrm{MAR}\left[\mathrm{m^3\,s^{-1}}\right] \times 86\,400\,[\mathrm{s}])/A\left[\mathrm{m^2}\right], \quad (21)$$

where A is the catchment area.

After J successive days of input X, routed with the UH of Eq. (18), we reach a steady state where the volume of the input equals the output (MAR). The total sum of moisture input after J days is

$$J \times X = S_{SS} + Q_{SS} \quad (22)$$

where total runoff, Q_{SS}, after J days is

$$Q_{SS} = \sum_{k=1}^{J}\sum_{j=1}^{k} X \times \mu(\overline{\Lambda})_j, \quad (23)$$

and k is the number of days and the subscript denotes "steady state". The water left in the soils, S_{SS}, at steady state (after J time intervals) and hence which is assumed to represent the mean storage m_S, is $S_{SS} = J \times X - Q_{SS}$, which can also be calculated as:

$$S_{SS} = \sum_{k=1}^{J-1}\sum_{j=k+1}^{J} X \times \mu(\overline{\Lambda})_j = m_S. \quad (24)$$

HESSD
12, 11129–11171, 2015

With an estimate of the mean storage, m_S, we can use Eq.(17) to estimate the scale parameter, η, of the distribution of S. The shape parameter, α, is already determined and equal to that of the distribution of Λ. The gamma distributed storage levels $S_i i$ are calculated as quantiles of the gamma distributed storage:

$$\frac{S_i}{\theta_M} = \int_0^{S_i} \frac{1}{\eta^\alpha \Gamma(\alpha)} S^{\alpha-1} \exp(-S/\eta) \mathrm{d}S \tag{25}$$

where θ_M is now estimated as the 99 % quantile of the distribution of S.

2.4 Test of new subsurface routine

We will test the performance of the new calibration-free formulation for the subsurface. This will be carried out by replacing the formulation of the subsurface where θ_M is a calibrated parameter and storage is uniformly distributed with a formulation where storage is gamma distributed with parameters, η and α, derived from recession data and MAR.

It is seldom straightforward to evaluate new algorithms for, at times, heavily parameterized rainfall–runoff models. Due to the tendency for calibration parameters to compensate for the models' structural errors (Kirchner, 2006; Beven and Binley, 1992), it may be difficult to identify the effect of changing an algorithm or parameterization from the calibrated results of the model (Clarke, 2011). Kirchner (2006) points out the need to develop models that are minimally parameterized, and which therefore stand a chance of failing the tests they are subjected to, which is exactly the problem faced when assessing the introduction of new algorithms with fewer parameters. Gupta et al. (2014) propose the use of large sample hydrology as a means for the testing of hypothesis and model structures, in order to (a) arrive at conclusions more general than can be achieved using a single catchment, (b) establish a range of applicability, and (c) ensure sufficient information to enable the identification of statistically significant relationships. These are certainly valid points, but, in addition, a minimal use of

HESSD
12, 11129–11171, 2015

calibration parameters should increase the efficiency in isolating and demonstrating the effect of new algorithms and parameterizations (Kirchner, 2006). To summarize, large sample hydrology *and* parameter parsimonious models are necessary tools to investigate the suitability of new model algorithms and structures.

The new parametrization of the subsurface is tested for 73 catchments distributed across Norway (see Fig. 7). The catchments vary in latitude, size, elevation and surface cover (see histograms of selected catchment characteristics in Fig. 8) and constitute thus a varied, representative sample of Norwegian catchments.

The following procedure was used for testing: the models were initially calibrated using long time series of precipitation and temperature to simulate runoff using an R-based Monte Carlo Marko Chain method (Soetart and Petzhold, 2010). The time series for precipitation and temperature are mean areal catchment values extracted from the current, operational meteorological grid (1 km × 1 km) which provides daily values of precipitation and temperature for Norway from 1957 to the present day (see www.senorge.no). This meteorological grid is denoted V1.

Recently, a new improved meteorological grid was developed, denoted V2, (Lussana and Tveito, 2014a, b) which reduced much of the positive bias in precipitation characteristic of V1 (see Saloranta, 2012). The new meteorological grid (V2) allows us to obtain reasonable simulated values of runoff without the need for a calibrated correction of the amount of precipitation (θ_{Pc}, see Table 1 for parameters of the DDD model). Areal averages of precipitation and temperature values are extracted for ten elevation zones (which constitutes the semi-distributed nature of the surface part of DDD) which makes it possible to eliminate calibrated precipitation and temperature gradients (θ_{Plr} and θ_{Tlr}). Three parameters associated with snow accumulation and melt (the correction factor for solid precipitation (θ_{Sc} = 1.0), the threshold temperature for snowmelt (θ_{Ts} = 0°C) and the threshold temperature for solid and liquid precipitation (θ_{TX} = 0.5°C) were fixed, thereby reducing the number of calibration parameters from 11 to 5. For the remaining 5 parameters, the calibrated values (from using V1 as input) are retained for 4 parameters ($\theta_{Ws}, \theta_{CX}, \theta_{cea}$ and θ_{CV}), whereas the parameter

HESSD
12, 11129–11171, 2015

of interest for this study θ_M, is recalibrated using V2 as input data. In using such a procedure, we assume that the 4 parameters which are calibrated using the V1 data (and, most likely, not optimal for the V2 data as input) will not favor either of the two compared model structures (calibrated- and estimated subsurface reservoir). When recalibrating the θ_M with V2 data, we attempt to make it as difficult as possible to accept the new subsurface routine. If we calibrated all 5 parameters using V2, we could risk that errors associated with the two subsurface routines were compensated by the other 4 parameters, such that we would not isolate and evaluate the effect of implementing the alternative subsurface routine.

With the procedure described above, we can compare the performances of the DDD model with calibrated storage (DDD_θ_M) and the DDD model with estimated storage (DDD_m_S) in a straightforward manner.

3 Results

Figure 9a–e show different skill scores obtained for the simulations for the 73 catchments with DDD_θ_M (calibrated storage with parameter θ_M, skill is shown with red crosses) and for DDD_m_S (estimated storage with parameter m_S, skill is shown with blue circles). Figure 9a shows the Nash–Sutcliffe Efficiency criterion (NSE, Nash and Sutcliffe, 1970), Fig. 9b the Kling–Gupta Efficiency criterion (KGE, Gupta et al., 2009; Kling et al., 2012) and Fig. 9c–e the three components of the KGE, correlation, bias and variability error, respectively. The variability error is given by the ratio of the coefficients of variation of simulated and observed runoff as suggested in Kling et al. (2012). The mean values of the skill scores for DDD_θ_M and DDD_m_S are shown as straight lines in the plots. We see from the Fig. 9 that no precision is lost in the results for DDD_m_S. The mean value of NSE is slightly better for DDD_m_S, whereas the mean value of KGE for DDD_m_S is slightly worse due to a lower correlation between simulated and observed (Fig. 9c). The results for the bias and variability errors favor DDD_m_S. Overall, the

HESSD
12, 11129–11171, 2015

differences in skill between DDD_m_S and DDD_θ_M are very small. Mean values of the skill scores for DDD_m_S and DDD_θ_M are shown in Table 2.

The observed distribution of Λ has been identified, in this and in previous studies, as being crucial for both the estimation of the subsurface celerities and the estimation of m_S. If the distribution of simulated Λ, denoted $\dot{\Lambda}$, is similar to that of the observed, this suggests that recessions are well simulated and hence, that the dynamics of the model are realistic. Figure 10 shows scatter plots of the mean and standard deviation of observed Λ and simulated $\dot{\Lambda}$ for DDD_m_S (blue circles) and DDD_θ_M (red crosses). The root mean square error (RMSE) of the mean $\dot{\Lambda}$ is clearly less for DDD_m_S whereas the RMSEs of standard deviation of $\dot{\Lambda}$ for DDD_m_S and DDD_θ_M are similar (see Table 3).

Figure 11 shows histograms of simulated storage from DDD_θ_M (a) and DDD_m_S (b) with empirical CDFs (c) of the observed Λ (black line) and simulated $\dot{\Lambda}$ (DDD_θ_M, red line and DDD_m_S, blue line) for a specific catchment. The CDF of $\dot{\Lambda}$ simulated with DDD_m_S is clearly in better agreement with that of the observed. The shape of the histograms of storage fluctuations are very different, and as we have no way of actually knowing the true empirical distribution of storage at the catchment scale we cannot claim that the fluctuations simulated with DDD_m_S are closer to the truth than those simulated by DDD_θ_M. However, since the parameters of the subsurface- and dynamic module of DDD_m_S are estimated prior to calibration and that the recessions are demonstrably better simulated, it is not unreasonable to suggest that the storage fluctuations simulated with DDD_m_S are closer to the truth.

4 Discussion

The reduction of calibration parameters from the subsurface and dynamic module of the DDD model has attractive implications for the problem of predictions in ungauged basins (PUB) (see e.g. Sivapalan, 2003; Parajka et al., 2013; Hrachowitz, 2013; Blöschl et al., 2013; Skaugen et al., 2015). In Skaugen et al. (2015), 7 model parameters of the

HESSD
12, 11129–11171, 2015

DDD model (including θ_M and the parameters for the distribution of λ) were estimated from catchment characteristics (CCs) using multiple regression analysis. All model parameters were found to correlate significantly with the CCs and the median NSE for 17 catchments was found to be 0.66 and 0.72 for two timeseries when DDD was run with model parameters estimated from CCs. The change in the model structure of DDD presented in this paper with respect to predictions in ungauged basins implies that we need an estimate of MAR and the parameters for the distribution of Λ. The estimation of θ_M is, however, no longer needed. In Skaugen et al. (2015) and in Skaugen and Væringstad (2005), MAR, was found to be a parameter varying smoothly in space for southern Norway, depending largely on regional scale climatic conditions and hence quite straightforward to estimate for an ungauged catchment. The parameters of the distribution of λ were estimated in Skaugen et al. (2015) as functions of the mean of the distance distribution, $\overline{d}$, percentage of lake, percentage of bare rock and catchment length. Since λ is a function of Λ (see Eq. 12), a similar dependency between Λ and catchment characteristics is expected, but is to be investigated. Provided that MAR and the parameters of the distribution of Λ are as easy to estimate for an ungauged catchment as expected we may hope for improved predictions for ungauged basins. In a future paper we will redo the analysis of Skaugen et al. (2015), using the new model structure presented in this paper and compare results.

In Kirchner (2009) the storage–runoff relationship is assumed to be a single-valued function, i.e. S is a single valued function of Q. This leads to a very simple model with regards to the number of states in the subsurface, namely one. The number of states in DDD can, however, be very high. If we consider Eq. (24), the number of summations (time-steps) constituting S_{SS} can be viewed as a number of subsurface states since each summation represents a volume water that will sooner or later propagate into the river network. Equation (24) describes the subsurface using only one (mean) UH. When running the DDD model, the number of storage levels is fixed to 5, and the UHs constituting the storage levels all have the same shape (exponential) but have different temporal scales. The temporal scale (number of discretizations) of the UHs

HESSD
12, 11129–11171, 2015

vary according to their associated celerity, and the slowest (lowest) storage level may be discretisized such that hundreds of time steps are necessary for the complete attenuation of the UH. Such a system actually provides a 2-D representation of the subsurface (Rupp et al., 2009; Sloan, 2000) and gives numerous subsurface states. It is hence entirely possible to have different configurations of states associated with the same runoff. Figure 12 shows a snapshot of how DDD models the storage S. The catchment is represented as one hillslope where the x axis shows the distance (in metres) from the river network (at the right hand-side) to the top of the hillslope (at the left hand side). The y axis shows the different storage levels. We see the outline of boxes (especially for the higher storage levels) which represents the temporal discretisation of the UHs. Each box represents an area according to the distance distribution and the associated celerity that will drain pr time interval. The higher the celerity, the more of the catchment area is represented by each box. The darker the blue colour, the more water is present in the box. Figure 12 can be seen together with Fig. 3, which illustrates how the distance distribution (and the travel time distribution) determines the fractional areas that drain pr time interval for a given celerity. It can also be noted in Fig. 12 that it is more or less dry at the top of the hillslope and saturated near the river. This is consistent with the wetting up of a catchment from the riparian zone outwards and up the hillslope (Myrabø, 1997; Kirkby, 1978, p. 275; Dunne and Black, 1970).

Figure 13 shows simulated storage S, plottet against observed runoff Q, for two catchments of different size (50 and 1833 km^2). It is quite clear that the relationship between Q and S is not single valued. The variability of Q for the same S (and vice versa) is to be expected given the multitude of possible configurations of the subsurface states (i.e. the discretisations of the UHs). The shape of the clouds of points resemble those found for observations of groundwater vs. runoff (Rupp et al., 2009; Laudon et al., 2004; Myrabø, 1997). The points in Fig. 13, however, do not level off to the same degree as does for groundwater observations. This can probably be explained by the fact that storage in DDD is simulated for an entire catchment, and it is more

HESSD
12, 11129–11171, 2015

unlikely that an entire catchment will reach full saturation than individual groundwater boreholes, located relatively close to the river (Myrabø, 1997; Laudon et al., 2004).

The new formulation for the subsurface gives good results, and it is promising that the replacement of a calibrated routine with an algorithm with parameters estimated prior to calibration produces runoff simulations which are as precise and robust as those produced by the calibrated routines. In addition, the simulated recessions $\dot{\Lambda}$, are much closer to those observed, suggesting a more realistically modelled storage–runoff relationship (i.e. the non-linearly increasing storage capacity). The parameter of the subsurface and the dynamical modules of the DDD model are now all estimated prior to any calibration and we see this as a necessary development if we are to effectively test new algorithms for snow distribution, snowmelt, evapotranspiration etc. at the scale that matters for practical applications, i.e. the catchment scale (Clarke, 2011). Multi-variable parameter estimation (Bergström et al., 2002) has been put forward as a means to increase confidence in hydrological modelling and models. Although we agree that such procedures indeed narrows the parameter-space (although not its number of dimensions), the interaction and compensating nature of the calibration parameters makes it almost impossible to reject flawed model structures so that we can concentrate on building models that work well for the right reasons. In this paper, and in previous ones (Skaugen and Onof, 2014; Skaugen et al., 2015), easily derived information, such as GIS-derived distance distributions functions and runoff records have proved sufficient for parameterising algorithms describing basic hydrological processes. This approach will be continued, and algorithms with a minimum of calibration parameters for describing the spatial distribution of snow and for snowmelt will be implemented in the DDD model structure and tested with the approach of large-sample hydrology as proposed by Gupta et al. (2014).

HESSD
12, 11129–11171, 2015

5 Conclusions

In this paper a new formulation of the subsurface in the DDD model is presented. In the new formulation, the subsurface storage capacity increases non-linearly with saturation, following a two-parameter gamma distribution for which parameters are estimated from observed runoff recession data and the mean annual runoff. The new subsurface formulation has been tested for 73 catchments in Norway of varying size, mean elevations and landscape types, with no loss in precision. In addition, more realistic runoff recessions are found using the new subsurface routine. An important contribution of the new formulation is that its parameters are estimated solely from observed recession data and the mean annual runoff (i.e. not through calibration).

With increased parameter parsimony in the already parameter parsimonious DDD model, important tasks such as predictions in ungauged basins (PUB) and testing new algorithms for different hydrological processes are more feasible. Large-sample hydrology has proven useful for revealing and quantifying the effects of the introduction of a new formulation of the hydrological subsurface.

Data availability

The precipitation, temperature and runoff data used in this study are available by contacting the corresponding author.

Acknowledgements. This work was conducted in the projects VANN- Evolutionary ecology and hydrology, FloodQ and ExPrecFlood, all funded by the Norwegian Research Council. We also wish to thank colleagues at NVE for their valuable comments.

HESSD
12, 11129–11171, 2015

References

Bergström, S.: The HBV model – its structure and applications. SMHI Reports Hydrology No. 4, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden, 1992.

Bergström, S., Lindström, G., and Petterson, A.: Multi-variable parameter estimation to increase confidence in hydrological modelling, Hydrol. Process., 16, 413–421, doi:10.1002/hyp.332, 2002.

Beven, K. J.: On the generalized kinematic routing method, Water Resour. Res., 15, 1238–1242, 1979.

Beven, K. J.: Rainfall–runoff modelling: the primer, Wiley, Chichester, UK, 2001.

Beven, K. J.: Benchmark papers in Streamflow Generation Processes, IAHS Press, Wallingford, UK, 2006.

Beven, K. J. and Binley, A.: The future of distributed models: model calibration and uncertainty prediction, Hydrol. Process., 6, 279–298, 1992.

Bidwell, V. J., Stenger, R., and Barkle, G. F.: Dynamic analysis of groundwater discharge and partial-area contribution to Pukemanga Stream, New Zealand, Hydrol. Earth Syst. Sci., 12, 975–987, doi:10.5194/hess-12-975-2008, 2008.

Blöschl, G., Sivaplan, M., Wagener, T., Viglione, A., and Savenije, H. (Eds.): Runoff Prediction in Ungauged Basins: Synthesis across Processes, Places and Scales, Cambridge University Press, NY, USA, 2013.

Bras, R. L.: Hydrology – An introduction to Hydrological Science, Addison-Wesley, Reading, Mass., USA, 1990.

Brutsaert, W. and Nieber, J. L.: Regionalized drought flow hydrographs from a mature glaciated plateau, Water Resour. Res., 13, 637–643, 1977.

Clarke, M. P., Kavetski, D., and Fenicia, F.: Pursuing the method of multiple working hypotheses for hydrological modelling, Water Resour. Res., 47, W09301, doi:10.1029/2010WR009827, 2011.

Colleuille, H., Haugen, L. E., and Øverlie, T.: Vann i jord – Simulering av vann og energibalansen på Groset markvannsstasjon, NVE report No. 19, The Norwegian Water Resources and Energy Directorate (NVE), Telemark, Norway, 2007 (in Norwegian).

Dingman, S. L.: Physical hydrology, Prentice Hall, New Jersey, USA, 2002.

Dooge, J. C. I.: A general theory of the unit hydrograph, J. Geophys. Res., 64, 241–256, doi:10.1029/JZ064i002p00241, 1959.

HESSD
12, 11129–11171, 2015

Dunne, T. and Black, R. D.: Partial area contributions to storm runoff in a small New England watershed, Water Resour. Res., 6, 1296–1311, 1970.
Feller, W.: An Introduction to Probability Theory and its Applications, Wiley, New York, USA, 1971.
Grip, H. and Rohde, A.: Vattnets väg från regn til bäck, Forskningsrådens Förlagstjänst, Karlshamn, Sweden, 156 pp., 1985 (in Swedish).
Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling, J. Hydrol., 377, 80–91, doi:10.1016/j.jhydrol.2009.08.003, 2009.
Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M., and Andréassian, V.: Large-sample hydrology: a need to balance depth with breadth, Hydrol. Earth Syst. Sci., 18, 463–477, doi:10.5194/hess-18-463-2014, 2014.
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., and Cudennec, C.: A decade of Predictions in Ungauged Basins (PUB) – a review, Hydrolog. Sci. J., 58, 1–58, doi:10.1080/02626667.2013.803183, 2013.
Kirchner, J. W.: Getting the right answer for the right reasons: linking measurements, analysis, and models to advance the science of hydrology, Water Resour. Res., 42, W03S04, doi:10.1029/2005WR004362, 2006.
Kirchner, J. W.: Catchments as simple dynamical systems: Catchment characterization, rainfall–runoff modelling, and doing hydrology backwards, Water Resour. Res., 45, W02429, doi:10.1029/2008WR006912, 2009.
Kirkby, M. J.: Hillslope hydrology, Wiley, Chichester, UK, 1978.
Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper danube basin under an ensemble of climate change scenarios, J. Hydrol., 424, 264–277, doi:10.1016/j.jhydrol.2012.01.011, 2012.
Lamb, R. and Beven, K.: Using interactive recession curve analysis to specify a general catchment storage model, Hydrol. Earth Syst. Sci., 1, 101–113, doi:10.5194/hess-1-101-1997, 1997.

HESSD
12, 11129–11171, 2015

Laudon, H., Seibert, J., Köhler, S., and K. Bishop: Hydrological flow paths during snowmelt: Congruence between hydrometric measurements and oxygen 18 in meltwater, soil water, and runoff, Water Resour. Res., 40, W03102, doi:10.1029/2003WR002455, 2004.
Lindström, G., Johansson, B., Persson, M., Gardelin, M., and Bergström, S.: Development and test of the distributed HBV-96 hydrological model, J. Hydrol., 201, 271–288, 1997.
Lussana, C. and Tveito, O.-E.: Spatial Interpolation of precipitation using Bayesian methods, Unpublished research note, The Norwegian Meteorological Institute, Oslo, Norway, 2014a.
Lussana, C. and Tveito, O.-E.: Spatial Interpolation of temperature using Bayesian methods, Unpublished research note, The Norwegian Meteorological Institute, Oslo, Norway, 2014b.
Nash, J. E.: The form of the instantaneous unit hydrograph. C. R. et Rapports, Assn. Internat. Hydrol. IUGG, Toronto, 1957.
Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models, Part 1 – a discussion of principles, J. Hydrol., 10, 282–290, 1970.
Maidment, D.: Developing a spatially distributed unit hydrograph by using GIS, in: Proc. Int. Conf. HydroGIS 93, Application of Geographic Information Systems in Hydrology and Water resources, Vienna, Austria, April 1993, IAHS Publ. No. 211, 181–192, 1993.
Myrabø, S.: Temporal and spatial scale of response area and groundwater variation in till, Hydrol. Process., 11, 1861–1880, 1997.
Parajka, J., Viglione, A., Rogger, M., Salinas, J. L., Sivapalan, M., and Blöschl, G.: Comparative assessment of predictions in ungauged basins – Part 1: Runoff-hydrograph studies, Hydrol. Earth Syst. Sci., 17, 1783–1795, doi:10.5194/hess-17-1783-2013, 2013.
Pulido-Velazquez, M. A., Sahuquillo-Herraiz, A., Camilo Ochoa-Rivera, J. and Pulido-Velazquez, D.: Modeling of stream-aquifer interaction: the embedded multireservoir model, J. Hydrol., 313, 166–181, 2005.
Rupp, D. E., Schmidt, J., Woods, R. A. and Bidwell, V. J.: Analytical assessment and parameter estimation of a low-dimensional groundwater model, J. Hydrol., 377, 143–154, doi:10.1016/j.jhydrol.2009.08.018, 2009.
Sælthun, N. R.: The "Nordic" HBV model, Description and documentation of the model version developed for the project Climate Change and Energy Production, NVE, Oslo, Publication No. 7–1996, 26 pp., 1996.
Saloranta, T. M.: Simulating snow maps for Norway: description and statistical evaluation of the seNorge snow model, The Cryosphere, 6, 1323–1337, doi:10.5194/tc-6-1323-2012, 2012.

HESSD
12, 11129–11171, 2015

Sivapalan, M.: Prediction in ungauged basins: a grand challenge for theoretical hydrology, Hydrol. Process., 17, 3163–3170, 2003.
Skaugen, T.: A spatial disaggregating procedure for precipitation, Hydrolog. Sci. J., 47, 943–956, 2002.
Skaugen, T. and Onof, C.: A rainfall–runoff model parameterized form GIS and runoff data, Hydrol. Process., 28, 4529–4542, doi:10.1002/hyp.9968, 2014.
Skaugen, T. and Væringstad, T.: A methodology for regional flood frequency estimation based on scaling properties, Hydrol. Process., 19, 1481–1495, 2005.
Skaugen, T., Peerebom, I. O., and Nilsson, A.: 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.
Sloan, W. T.: A physics-based function for modeling transient groundwater discharge at the watershed scale, Water Resour. Res., 36, 225–242, 2000.
Soetart, K. and Petzholdt, T.: Inverse modelling, sensitivity and Monte Carlo analysis in R using package FME, Journal of Statistical Software, 33, 1–28, www.jstatsoft.org/article/view/v033i03/v33i03.pdf (last access: 29 October 2015), 2010.

HESSD
12, 11129–11171, 2015

Table 1. Parameters of the DDD model with comments and method of estimation. Some parameters (denoted with a $*$) have fixed values obtained through experience in calibrating DDD for gauged catchments in Norway. These values are within the recommended range for the HBV model (Sælthun,1996). Other parameter values are assigned standard values as suggested in the literature. The GIS analyses are carried out using the national 25 m × 25 m DEM (www.statkart.no).

Parameter	Comment	Method of est.	Value	Ref
Hypsograpic curve	11 values describing the quantiles 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100	GIS		
θ_{Ws} [%]	Max liquid water content in snow	Calibrated (V1)	5	
Hfelt	Mean elevation of cathment	GIS		
θ_{Tlr} [°C 100 m^{-1}]	Temperature lapse rate for (pr 100 m)	Standard value	0.0	
θ_{Plr} [mm 100 m^{-1}]	Precipitation gradient (mm per 100 m)	Standard value	0.0	
θ_{Pc}	Correction factor for precipitation	Standard value	1.0	
θ_{Sc}	Correction factor for precipitation as snow	Standard value	1.0	
θ_{TX} [°C]	Threshold temperature rain/snow	Standard value	0.5	
θ_{TS} [°C]	Threshold temperature melting/freezing	Standard value	0.0	
θ_{CX} [mm °C^{-1} day^{-1}]	Degree-day factor for melting snow	Calibrated (V1)		
C_{Glac} [mm °C^{-1} day^{-1}]	Degree-day factor for melting glacier Ice	*	1.5 × θ_{CX}	Sælthun (1996)
CFR [mm °C^{-1} day^{-1}]	Degree-day factor for freezing	*	0.02	Sælthun (1996)
Area [m^2]	Catchment area	GIS		
maxLbog [m]	Max of distance distribution for bogs	GIS		
midLbog [m]	Mean of distance distribution for bogs	GIS		
Bogfrac	Fraction of bogs in catchment	GIS		
Zsoil	Areal fraction of zero distance to the river network for soils	GIS		
Zbog	Areal fraction of zero distance to the river network for bogs	GIS		

HESSD
12, 11129–11171, 2015

Table 1. Continued.

Parameter	Comment	Method of est.	Value	Ref
NOL	Number of storage levels	Standard value	5	Skaugen and Onof (2014)
θ_{cea} [mm °C^{-1} day^{-1}]	Degree day factor for evapotranspiration	Calibrated (V1)		
R	Ratio defining field capacity	Standard value	0.3	Skaugen and Onof (2014)
α	Shape parameter of gamma distributed recession characteristic λ	Estimated from recession		
β	Scale parameter of gamma distributed recession characteristic λ	Estimated from recession		
θ_{CV}	Coeff. of variation for spatial distribution of snow	Calibrated (V1)		
v_r [m s^{-1}]	Mean celerity in river.	Standard value	1.0	Beven (1979)
m_{Rd} [m]	Mean of distance distribution of the river network	GIS		
s_{Rd} [m]	Standard deviation of distance distribution of the river network	GIS		
Rd_{max} [m]	Max of distance distribution in river network	GIS		
θ_M/m_S (mm)	Max subsurface water reservoir/Mean of subsurface water reservoir	Calibrated (V2)/Estimated from recession		
$\overline{d}$ [m]	Mean of distance distribution for hillslope	GIS		
d_{max} [m]	Max of distance distribution for hillslope	GIS		
Glacfrac	Fraction of bogs in catchment	GIS		
m_{Gl} [m]	Mean of distance distribution for glaciers	GIS		
s_{Gl} [m]	Standard deviation of distance distribution for glaciers	GIS		
Areal fraction of glaciers in elevation zones	10 values	GIS		

HESSD
12, 11129–11171, 2015

Table 2. Mean values of skill scores obtaind with simulating with DDD_m_S and DDD_θ_M for 73 catchments. KGE_r measures correlation, KGE_b, the bias error and KGE_g the variability error. All skill scores have an ideal value of 1.

	NSE	KGE	KGE_r	KGE_b	KGE_g
DDD_m_S	0.68	0.69	0.83	0.85	1.0
DDD_θ_M	0.66	0.70	0.85	0.83	1.04

HESSD
12, 11129–11171, 2015

Table 3. Root mean square error (RMSE) values for the mean and standard deviation of simulated $\dot{\Lambda}$ for the 73 catchments

	RMSE mean Λ	RMSE SD Λ
DDD_m_S	0.275	0.389
DDD_θ_M	0.511	0.392

HESSD
12, 11129–11171, 2015

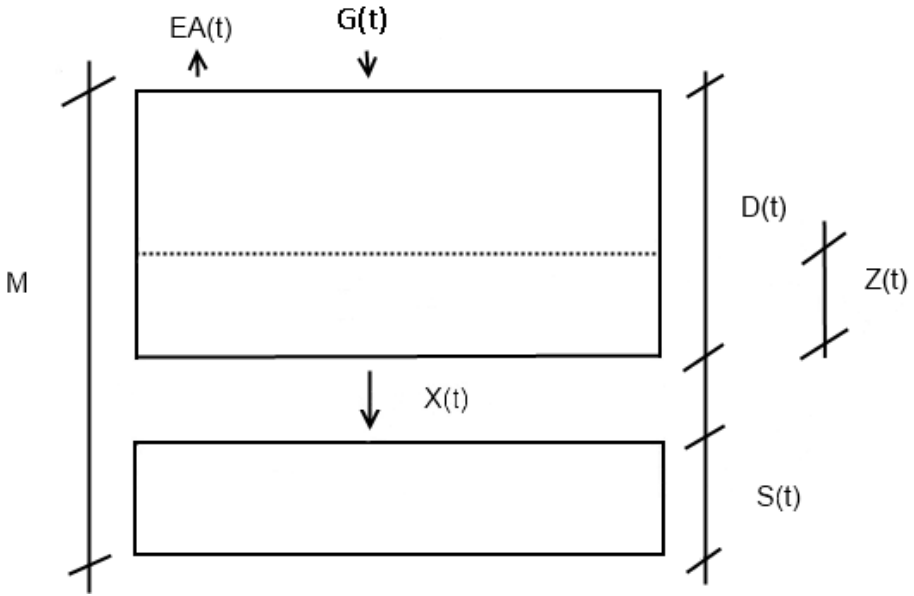


Figure 1. Schematic of the subsurface water reservoir M of DDD. $G(t)$ represents moisture input, rain and snowmelt. The dotted horizontal line is the actual level Z, of soil moisture in D. The ratio $(G(t)+Z(t))/D(t)$ controls the release of excess water to S and hence to runoff. Note that D, S and Z are functions of time, whereas M is fixed.

HESSD
12, 11129–11171, 2015

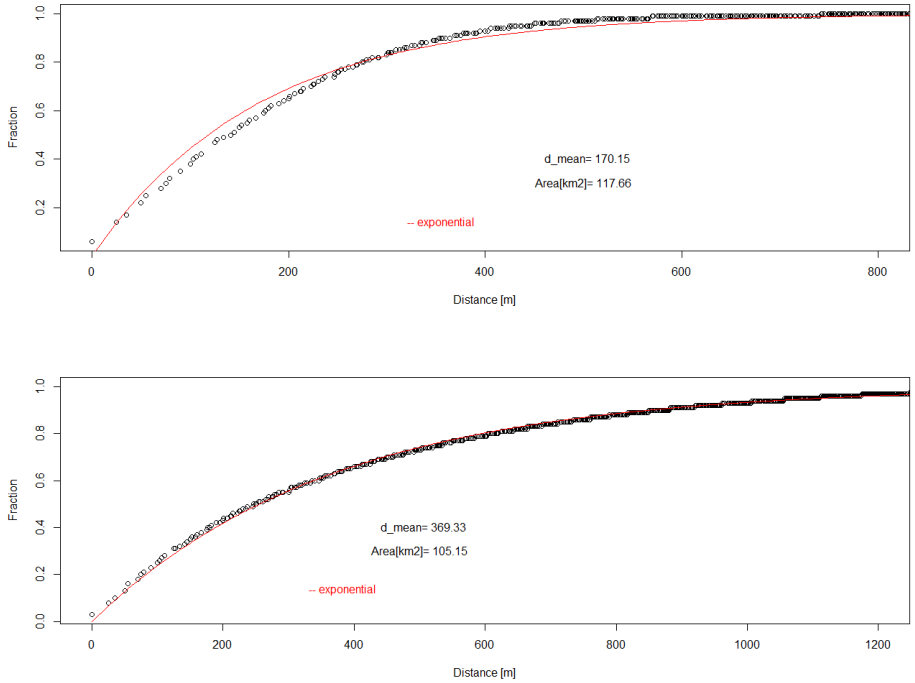
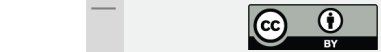


Figure 2. Empirical and fitted (exponential, red line) CDFs of distances from a point in the catchment to the nearest river reach for two Norwegian catchments. The mean distance (denoted d_mean) and catchment size differ, but the shape of the distribution is similar.

HESSD
12, 11129–11171, 2015



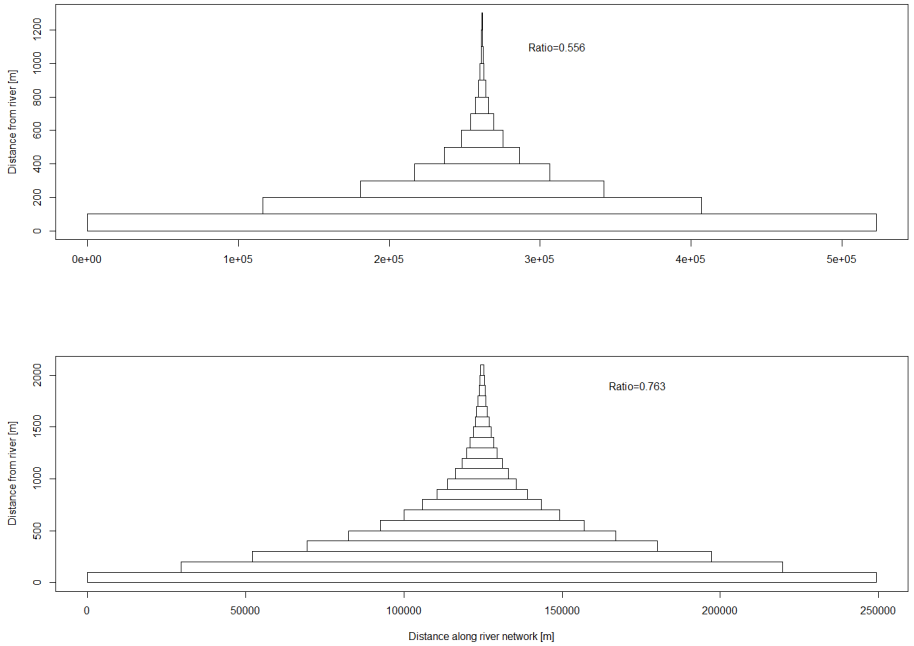


Figure 3. Fractional catchment area as a function of distance from the river network for the same two catchments as in Fig. 2. The ratio κ, between consecutive areas is shown as "Ratio".

HESSD
12, 11129–11171, 2015

Estimating catchment scale groundwater dynamics

T. Skaugen and Z. Mengistu

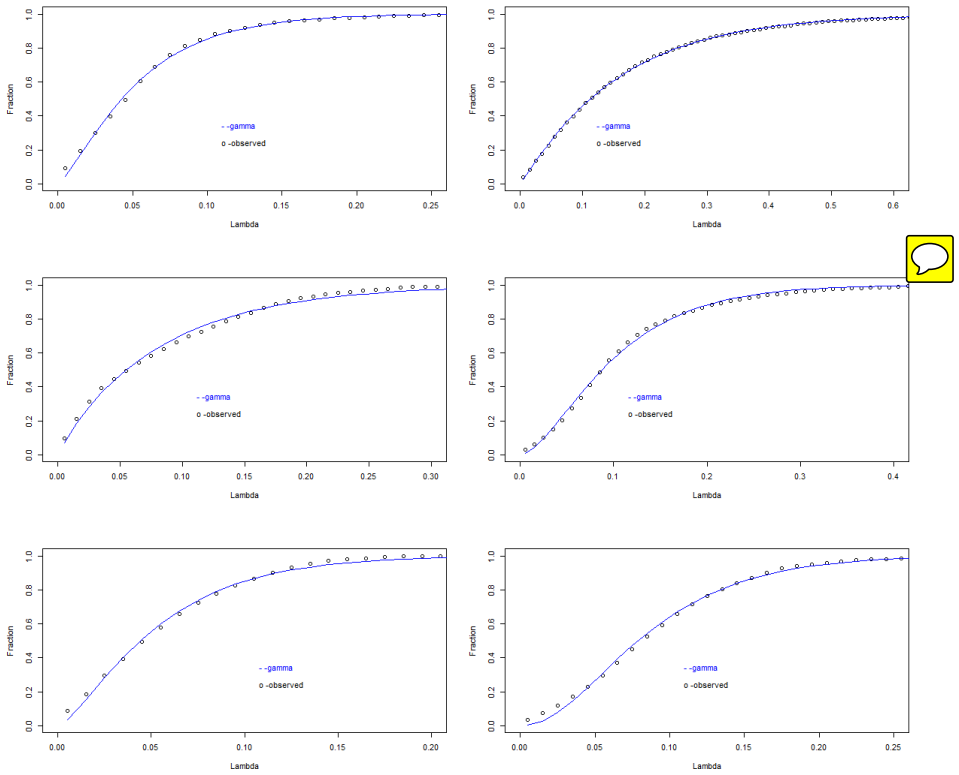


Figure 4. Empirical and fitted (gamma, blue line) CDFs of Λ for 6 Norwegian catchments. The Λ are sampled using Eq. (9) for all observed recession events.

HESSD
12, 11129–11171, 2015

Estimating catchment scale groundwater dynamics

T. Skaugen and Z. Mengistu

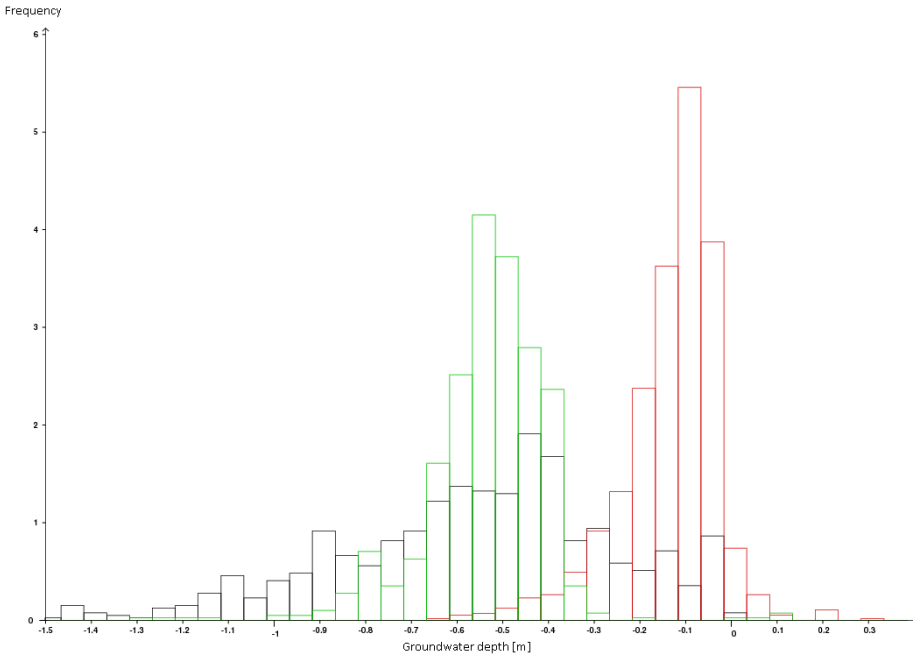


Figure 5. Histograms (in black, green, and red) of groundwater levels at three different locations in the Groset catchment (6.33 km^2) located in southern Norway.

HESSD
12, 11129–11171, 2015

Estimating catchment scale groundwater dynamics

T. Skaugen and Z. Mengistu

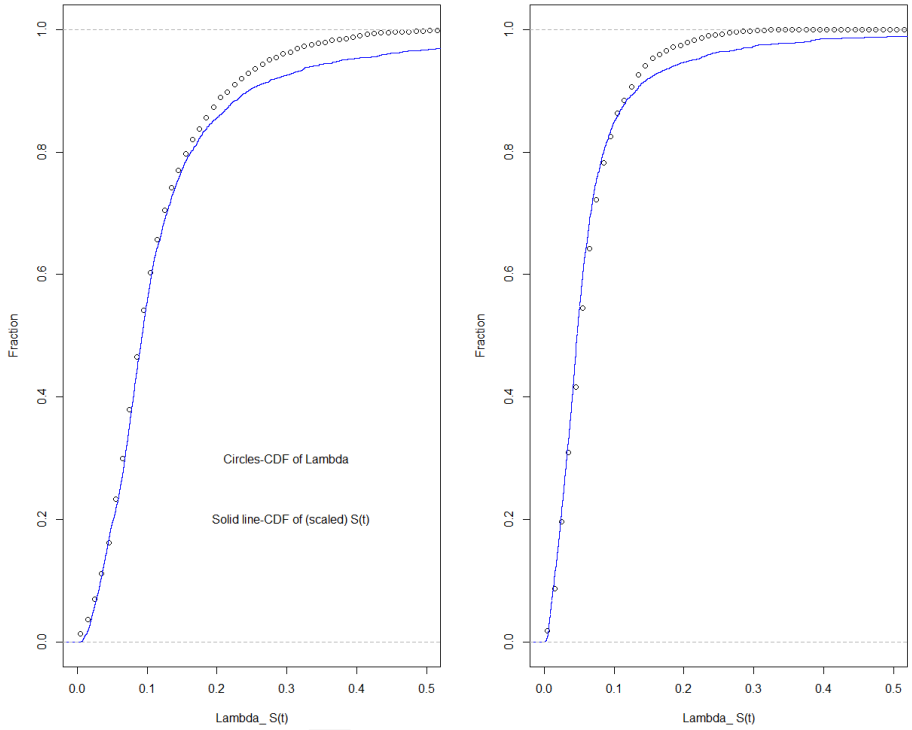


Figure 6. Empirical CDFs of Λ (circles) and scaled $S(t)$ (blue line) for two Norwegian catchments.

HESSD
12, 11129–11171, 2015

Estimating catchment scale groundwater dynamics

T. Skaugen and Z. Mengistu

Figure 7. Location of the 73 catchments used to evaluate the new subsurface routine

HESSD
12, 11129–11171, 2015

Estimating catchment scale groundwater dynamics

T. Skaugen and Z. Mengistu

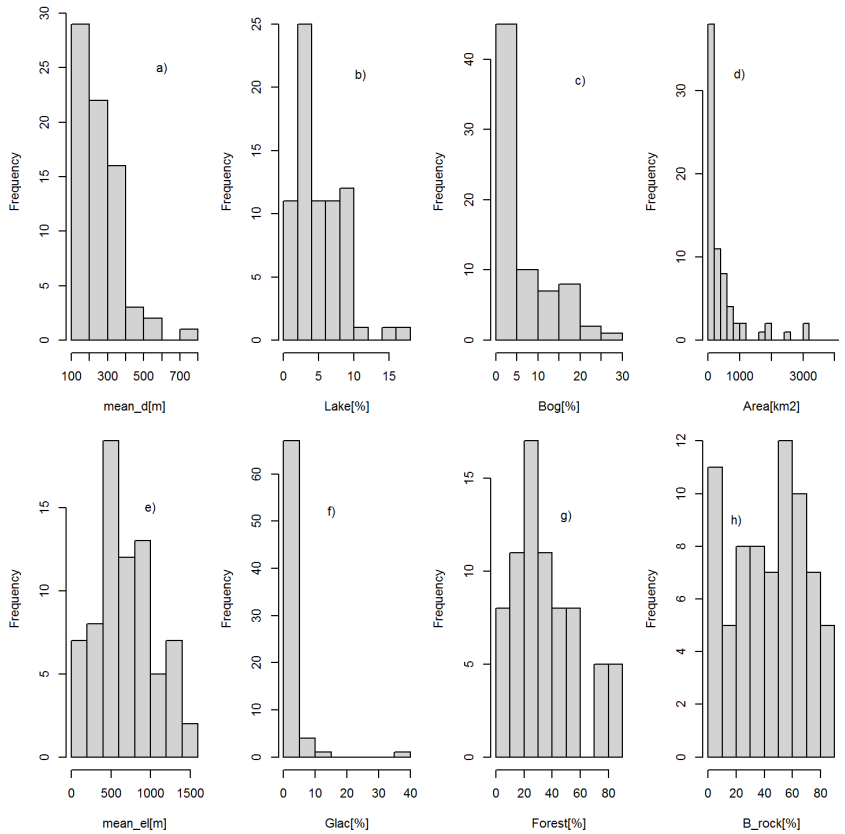


Figure 8. Histograms of catchment characteristics for the 73 catchments. **(a)** Mean of the hillslope distance distribution, $\overline{d}$, **(b)** areal percentage of lakes, **(c)** areal percentage of bogs, **(d)** catchment area, **(e)** mean elevation, **(f)** areal percentage of glaciers, **(g)** areal percentage of forests and **(h)** areal percentage of bare rock.

HESSD
12, 11129–11171, 2015

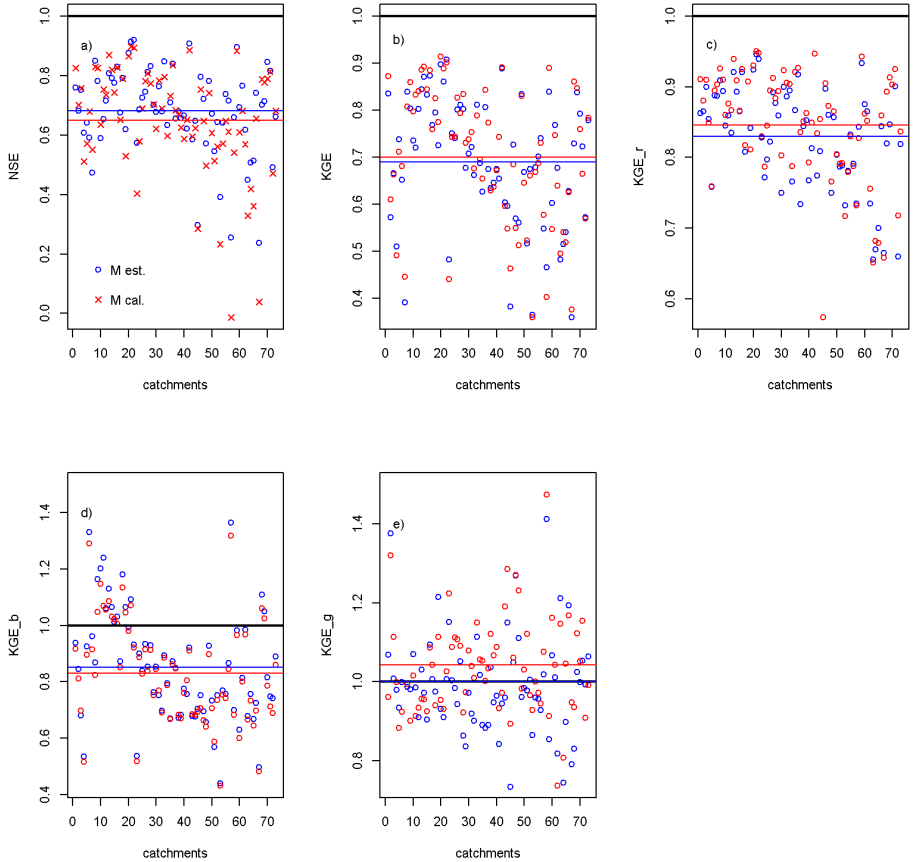


Figure 9. Skill scores for DDD_m_S (blue circles) and DDD_θ_M (red crosses) for 73 Norwegian catchments. Mean skill score values are shown in horizontal lines (same color code). **(a)** NSE, **(b)** KGE, **(c)** KGE_r (correlation), **(d)** KGE_b (bias) and **(e)** KGE_g (variability error).

HESSD
12, 11129–11171, 2015

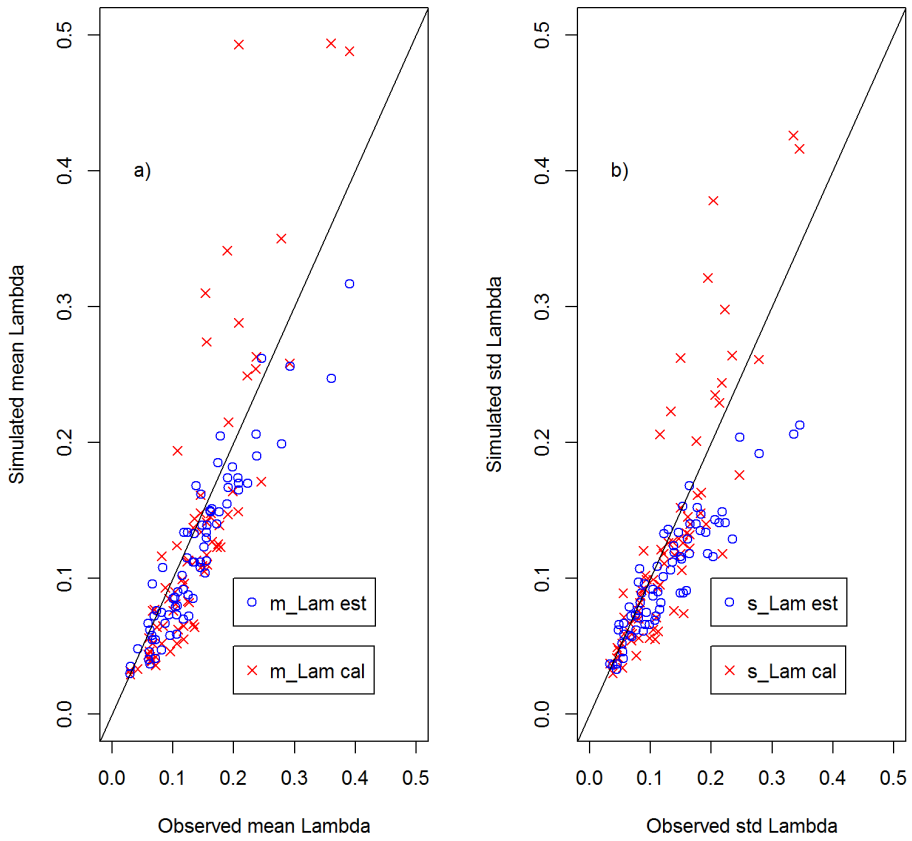


Figure 10. Scatterplot of mean **(a)** and standard deviation **(b)** of observed Λ and simulated with DDD_m_S (blue circles) and DDD_θ_M (red crosses) $\dot{\Lambda}$ for 73 catchments.

HESSD
12, 11129–11171, 2015

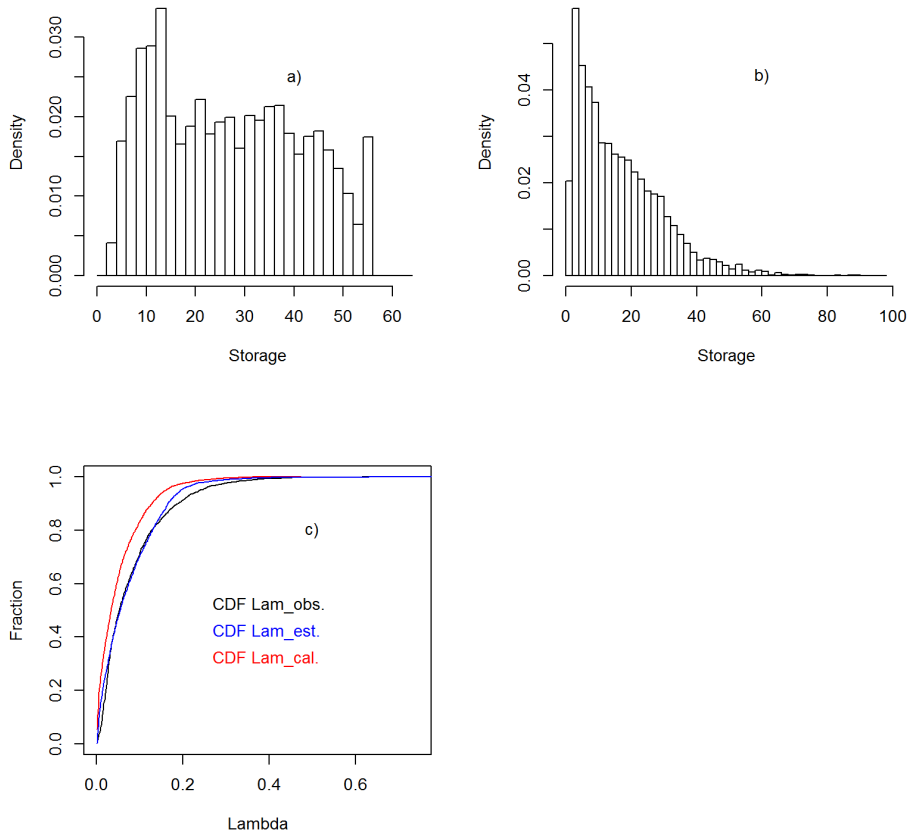


Figure 11. Histograms of storage simulations with DDD_θ_M **(a)** and DDD_m_S **(b)**. Empirical CDFs of observed Λ (black line) and simulated $\dot{\Lambda}$ with DDD_θ_M (red line) and DDD_m_S (blue line) are shown in **(c)**.

HESSD
12, 11129–11171, 2015

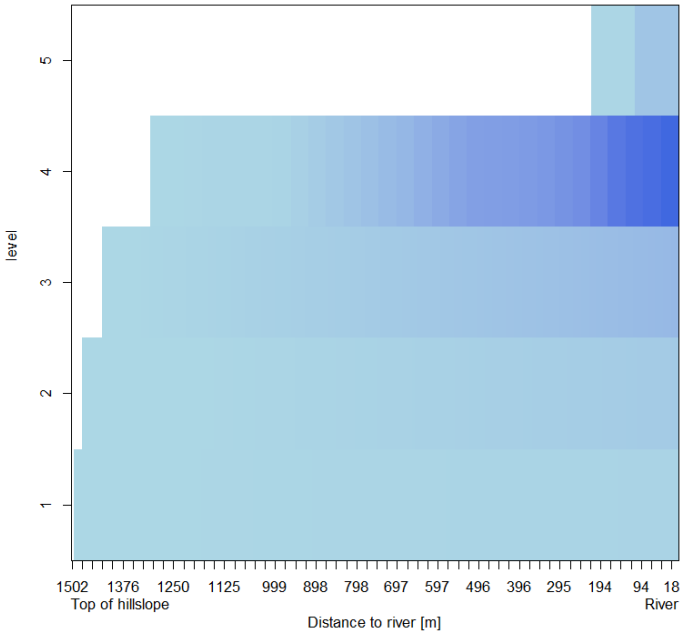


Figure 12. Snapshot of the saturated zone *S* of the DDD model. The catchment is represented as one hillslope. The *x* axis shows the distance from the river (right hand-side) to the top of the hillslope (left hand-side). The *y* axis show the storage levels. The darker the blue colour, the more water is present in the storage level.

HESSD
12, 11129–11171, 2015

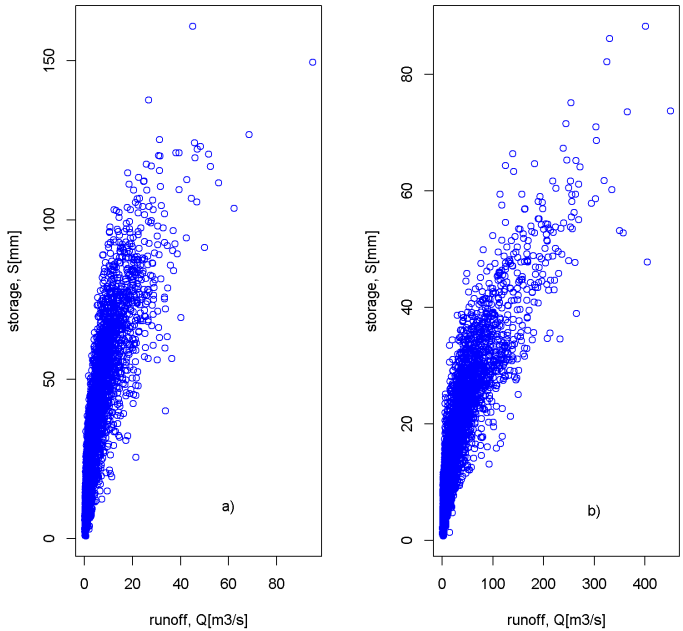


Figure 13. Simulated storage S plotted against observed runoff Q for a catchment of 50 km^2 **(a)** and a catchment of 1833 km^2 **(b)**.

HESSD
12, 11129–11171, 2015