Unraveling abiotic and biotic controls on the seasonal water balance using data-driven dimensionless diagnostics

Simon Paul Seibert[1], Conrad Jackisch[1], Uwe Ehret[1], Laurent Pfister[2], and Erwin Zehe[1]

[1]Karlsruhe Institute of Technology (KIT), Institute for Water and River Basin Management, Chair of Hydrology, Kaiserstrasse 12, 76131 Karlsruhe, Germany

[2]Luxembourg Institute of Science and Technology, Department Environmental Research and Innovation, Catchment and Eco-hydrology research group, 5 avenue des Hauts-Fourneaux, L-4362 Esch/Alzette, Luxembourg

Correspondence to: Simon P. Seibert (simon.seibert@kit.edu)

Abstract. Understanding runoff production and the underlying physiographic controls is important – but rarely fully understood in catchment hydrology. In order to advance on this matter we propose to use data-driven diagnostic signatures. Specifically, we present dimensionless double mass curves (dDMCs) which allow studying runoff formation and the water balance at the seasonal and annual time scale. By separating the vegetation and winter period dDMCs furthermore provide information on the role of biotic and abiotic controls on seasonal runoff formation.

A key aspect we address in this paper is the derivation of dimensionless fluxes which ensure the comparability of the signatures in space and time. We achieve this by using the limiting factors of a hydrological process as scaling reference. We show that different references result in different diagnostics. As such we define two kinds of dDMCs which allow to derive seasonal runoff coefficients and to characterise dimensionless streamflow release as a function of the potential renewal rate of the soil store. We expect these signatures for storage controlled seasonal runoff formation to remain invariant, as long as the ratios of release over supply and supply over storage capacity develop similarly in different catchments.

We test the proposed methods by applying them to an operational data set comprising 22 catchments (12…166 km^2) from different environments in southern Germany and hydrometeorological data from 4 hydrological years. The diagnostics are used to compare the sites and to reveal the dominant controls on runoff formation.

The key finding are that dDMCs are meaningful signatures for catchment runoff formation at the seasonal to annual scale and that the type of scaling strongly influences the diagnostic potential of the dDMC. Adding discrimination between vegetation and winter period was of fundamental importance and easy to implement by means of a simple temperature-index model. More specifically, temperature aggregates explain over 70 % of the variability of the seasonal summer runoff coefficients. The results also show that the product of topographic gradient and saturated hydraulic conductivity is significantly correlated to winter runoff coefficients whereas the two variables alone are not. We conclude, that proxies for gradients and resistances should be interpreted as a pair. Last, the dDMC concept reveals memory effects between summer and winter runoff regimes that are not relevant in spring between the transition from winter to summer.

1 Introduction

Understanding catchment scale runoff formation and the underlying controls is the key for building hydrological models, which work for the right reasons Kirchner (2006), as well as for hydrological similarity assessment (Larsen et al., 1994; Barthold and Woods, 2015) in general.

The latter includes "forward" and "backward" classification approaches: Forward in a sense to postulate similarity of runoff generation based on similarity of catchment structural attributes. "Backward" as search for data-driven measures to detect similarity in hydrological response variables either based on suitable diagnostic fingerprints or dimensionless similarity indicators (Sivapalan et al., 1987; Wood et al., 1990). Prominent forward approaches include for instance hydrological response units (HRU) (Leavesley, 1973), the concept of hydrology of soil types (HOST) (Boorman et al., 1995), or the topographic index (Kirby, 1975) which describes similarity of points within a catchment with respect to event scale runoff formation (Beven and Kirkby, 1979). The common ground of forward approaches is that they are often used in connection with hydrological models to subdivide the catchment or hillslope into control volumes of similar runoff generation, which is represented by similar parameter sets or closure relations. Up to now, a large set of HRU-types and separation methods has been suggested (Wood et al., 1990; Peschke et al., 1999; Scherrer and Naef, 2003; Schmocker-Fackel et al., 2007; Pelletier and Rasmussen, 2009; Santhi et al., 2008). Also, the representative elementary watersheds (REW) concept (Reggiani et al., 1998, 2000; Varado et al., 2006; Zhang et al., 2006) can be seen as a mathematically rigorous and thermodynamically consistent interpretation of the HRU idea. More recently a hierarchy of more specific functional units, defined on the basis of similarity of terrestrial and atmospheric controls on driving gradients and resistance terms were proposed by Zehe et al. (2014) as a refinement of the HRU idea. The forward classification schemes are strongly model-dependent which leaves us with questions about conclusiveness and transferability to other landscapes and how to define similarity in catchment response behaviour independently from models.

These questions are scope of the backward approaches, which aim at detecting similarity in hydrological response variables based on suitable diagnostic fingerprints in a data-driven way. The use of backward approaches has also been promoted by the Prediction in Ungauged Basins (PUB) initiative (Sivapalan, 2003; Hrachowitz et al., 2013) which suggested diagnosing catchment functioning through runoff "signatures". The underlying assumption is that "...runoff variability can be broken up into several components, each of them a manifestation of catchment functioning, albeit at different time scales,..." (Blöschl et al., 2013, pp. 7). Today, runoff signatures are commonly accepted and usually defined as specific characteristics of the hydrograph such as autocorrelation, slope of or bias in the flow duration curve – or different segments thereof, rising limb density, peak distribution, and/or as flow statistics such as mean, variance, skewness or the coefficient of variation (Pokhrel and Yilmaz, 2012; Casper et al., 2012; Pfannerstill et al., 2014; Euser et al., 2013, 2015). Runoff signatures are widely used for similarity assessment but also for model evaluation. In the former, similarity in the signature values is interpreted as hydrological similarity among catchments which is also the foundation of many catchment intercomparison studies (Merz et al., 2006; Oudin et al., 2008; Sawicz et al., 2011; Wang and Wu, 2013; Viglione et al., 2013). In model evaluation, the signature values obtained from the observation are compared to the signature values obtained from the simulation (Vrugt and Sadegh, 2013; Höllering et al., 2016).

Ideally, we expect that similarity of data-driven diagnostics can be explained by similarity in the architecture of the catchment and hence, that forward and backward approaches yield consistent results. However, such comparisons often fail to be conclusive. One possible explanation is that similarity in hydrograph-based signatures may indeed be caused by several reasons. For instance characteristics of the flow duration curve such as low slope values of the 33rd to 66th streamflow percentiles, implying damped runoff response, can for instance arise from persistent year-round rainfall regimes or from the dominance of groundwater contribution to streamflow. In consequence groupings of catchments based on either signatures or physiographic properties can be rather inconsistent as highlighted in the catchment intercomparison study of Ali et al. (2012) where "...catchment groupings obtained using physical properties only did not match those obtained using flow indices, mean transit times or storage estimates". Similar findings are reported by Oudin et al. (2010) and Ley et al. (2011). Albeit the potential of signatures for similarity assessment and diagnosing functional similarity is beyond question, the existing approaches to define and to detect similarity differ considerably with respect to the underlying assumptions, methods and proposed measures. As commonly agreed upon definitions for "signatures" and "hydrological, functional or behavioural similarity" are still missing we can not yet distil a convergence of approaches.

In the present study we add bits and pieces to this puzzle by proposing and testing data-driven dimensionless backward signatures to diagnose and characterise seasonal runoff formation. As depicted in Fig. 1 these signatures shall unravel the influences i) of the hydrometeorological forcing, i.e. radiation and precipitation, ii) of key catchment structural attributes, specifically of biotic (functional vegetation) and abiotic controls (e.g. topographic gradient, subsurface hydraulic conductivity) and iii) of the state of the catchment. The relative importance of these concurring influences on runoff generation depends on the time scale. Intuitively one might expect the wetness state of the catchment to be of the highest importance for runoff generation during events, while vegetation has surely the strongest impact on seasonal runoff production.

– Figure 1: SKETCH RUNOFF FORMATION –

So what could suitable data-driven diagnostic signatures look like? Inspired by the diagnostic approach to model evaluation (Gupta et al., 2008), McMillan et al. (2011) proposed to combine recession analysis, soil moisture dynamic analyses, and event-scale water balance investigations to obtain insights into different runoff response timescales and to assess the impact of pre-storm wetness conditions on catchment runoff dynamics. Hrachowitz et al. (2011) suggested combining event runoff coefficients and different tracers to infer knowledge of hydrological processes in small catchments. Tracer data are also widely used to estimate transit time characteristics which are considered as useful similarity index for process-based catchment classification (Soulsby et al., 2010; Capell et al., 2012). Therefore, breakthrough or flushing of either contaminants (Gassmann et al., 2013), artificial tracers (Wienhöfer et al., 2009), sediments (Martínez-Carreras et al., 2010) or even diatoms as smart tracers (Martínez-Carreras et al., 2015; Klaus et al., 2015) are used. More recently, McMillan et al. (2014) proposed a "targeted analyses of catchment response data" and combined hydrograph-based signatures with signatures that evaluate characteristics of the water balance, recession analysis and hydrological thresholds to examine the extent to which hydrological behaviour varies within a 50 km^2 catchment. The common ground of the mentioned approaches is that they combine multiple sources of information from rather different observations.

This is however only possible in well instrumented research catchments. As our focus is on providing diagnostic fingerprints for comparative hydrology (Sivapalan et al., 2003) we constrain our work to operational hydrometeorological data sets. These commonly consist of meteorological variables like precipitation, air temperature, humidity, etc. and discharge data. Information on the catchment "state" in terms of ground water level or soil moisture data is partly monitored too, but usually only in coarse spatial resolution. The advantage of operational data sets is that they allow us to include sufficient catchments with rather different physiographic and climate characteristics (end-members) for our analysis. A drawback of such data sets is that they mostly exist only for scales of 50 km^2 and larger and that specific data such as tracers, piezometric heads or soil hydraulic properties are usually not available.

In this study we propose and test dimensionless diagnostic signatures to characterise seasonal runoff formation based on rainfall runoff-data, which shall ultimately separate the terrestrial controls on runoff formation from the meteorological forcing. The dimensionless formulation of these diagnostics is of key importance to ensure comparability of the signatures in space and time. The proposed concept hooks back on the idea of Wagener et al. (2007) that catchments have the three main functions, i.e. to partition, store and release water. Essentially we suggest that the limiting factors of these functions provide the key for a proper dimensionless formulation of the desired signature. In line with Black (1997) we treat catchments as lumped terrestrial filters and following Kirchner (2009) we assume that annual runoff formation is essentially a monotonously increasing function of water storage in the catchment. Accordingly, accumulated runoff is expected to be limited by accumulated water supply (rainfall minus evaporation). Water storage is controlled and limited by water supply due to rainfall, the available subsurface storage volume as well as by its recharge and drainage properties. One might wonder whether a diagram that plots accumulated water release scaled with total annual supply against accumulated water supply scaled by the available storage capacity would be a suitable means to detect similarity in catchment seasonal runoff formation. Such a signature for storage controlled seasonal runoff formation is deemed to remain invariant during simple scaling operations as further explained below.

In line with these thoughts we hypothesize that dimensionless and season-specific double mass curves (dDMCs), which relate dimensionless accumulated runoff to dimensionless accumulated precipitation are suitable to discriminate differences in how a catchment releases water as function of the accumulated water supply. We test this hypothesis by addressing the following research questions within a catchment intercomparison:

- **Q1**: How can we obtain proper dimensionless double mass curves using site specific characteristics to ensure comparability among catchments based on operational data sets?
- **Q2**: Are temperature indices from vegetation ecology superior compared to calendrical definitions for separating summer and winter regimes?
- **Q3**: Can we identify physiographic and ecological properties which explain differences in seasonal runoff behaviour revealed by differences in the dDMCs?

The paper is structured as follows: In the method section in chapter 2 we introduce the data set, the concept of dDMCs, the separation of seasons, and the underlying statistical methods we apply to address Q3. After presenting the results in section 3 we close with discussion and conclusions in sections 4.

2 Concept, methods and study area

2.1 Study area

We propose and explore the potential of signatures to characterise seasonal runoff generation using operational data of 130 catchments located in the Bavarian part of the Danube basin. In this chapter we introduce the data set and detail the differences in the climate and physiographic setting of our test catchments. Before that, we will briefly discuss the quality of the database, which in fact was in most catchments so poor that the majority of the sites had to be excluded from the analysis.

2.1.1 Data quality and selection of headwater catchments

In our analysis we focus on lower mesoscale headwater catchments (< 200 km2) as routing effects are still small (Robinson et al., 1995) and because the geological and pedological setup is often still fairly homogeneous at this scale. In addition, non-convective storms often cover the majority of the catchment area and we may hence assume that the hydrometeorological forcing is well-observed by the given station and that it is fairly uniform within the entire headwater. We hence select all gauged headwaters of this size within the Bavarian part of the Danube basin. For this, hourly hydrometeorological time series from the period 01.11.1999 until 31.10.2004 are available. The data base in the resulting 130 catchments is analysed according to a set of quality criteria: We only include catchments where i) at least one meteorological station was closer than 20 km (which is less than in the Mopex data set (Schaake et al., 2000; Duan et al., 2006)), ii) the total absolute water balance error was smaller than 5 %, iii) the amount of missing and/or implausible meteorological data was < 5 % and iv) where the streams are not subject to any severe regulation. This screening resulted in 22 catchments being classified as suitable for the analysis. The sites are spread across the Bavarian part of the Danube basin (Fig. 2) and belong to different hydrological "regions" which we roughly classified based upon geology, climate and elevation. Moving from the North-West to the South we differentiate TRI (Triassic), JUR (Jurasic), BFO (Bavarian Forest), MOL (Faulted Molasse), AFO (Alpine Foreland) and ALP (Alpine) landscapes. The catchment identifiers reflect these units and include an index of the sample size e.g MOL1,..., MOL7. Catchment locations, real gauge and corresponding stream names are provided in Table 4 in the appendix.

– Figure 2: MAP STUDY SITE –

2.1.2 Data and physiographic site properties

The Bavarian part of the Danube basin exhibits considerable differences in the hydrometeorological regimes. To illustrate them we plot regime curves for four different catchments in different environmental settings Fig. 3. The catchment TRI1 (Fig. 3, top left) receives a fairly constant input in precipitation (P) throughout the year, but releases discharge (Q) with a strong seasonality and pronounced minimum during summer. Compared to the other sites the interannual variation in evapotranspiration (E) is rather large. The catchment AFO4 (Fig. 3, top right) in contrast shows seasonality in P but a fairly constant output in Q. ALP4 (Fig. 3, bottom left) and ALP2 (Fig. 3, bottom right), which are both Alpine sites, show a pronounced minimum in Q during February due to snow storage. ALP2 however shows a very large range in both P and Q during summer, which suggests

little buffering and a high reactivity. In contrast, ALP4 has a much more damped response to P during summer and a more pronounced seasonality in E.

– FIGURE 3 REGIME CURVES –

To characterise the structural setup of the different catchments we employ various landscape analyses techniques. Using these methods we derive nearly 30 different physiographic characteristics for each of the catchments. They are summarized in Table 2 and 3 in appendix A. Topographic information was extracted from a digital elevation model (DEM) with a resolution of 25 m. We use it to calculate the median topographic gradient (ϕ) (-) of each catchment according to McGuire et al. (2005) as the flow path length from each pixel to the stream divided by the corresponding difference in height using the Whitebox geographic analysis toolbox (Lindsay, 2014). Areal shares of various surface cover and soil properties like average sand, silt or skeleton contents, rootzone effective field (eFC) (mm) and air capacities (AC) (mm) among others were derived from CORINE land use data as of 2006 and the digital soil map of Germany (scale: 1:1,000,000) (Hartwich et al., 1995). A rough approximation of the average saturated hydraulic conductivity (Ks) (m s^{-1}) is estimated for each catchment based on available grain size distributions using *Rosetta's* pedo-transfer functions (Schaap et al., 2001). Since we assume that the topographic gradient of the catchment and the conductivity of the control volume act in group we also include the product of them (τ) (m s^{-1}):

$$\tau = \phi \cdot Ks \tag{1}$$

Last but not least, we employed the conceptual hydrological "Large Area Runoff Simulation Model" (LARSIM) (Ludwig and Bremicker, 2006) to obtain time series of areal estimates of E, P and snow water equivalent (W) for each catchment. For this purpose a gridded version of the LARSIM model (spatial resolution: 1 km^2, time step: 1 h), which is calibrated and in operational use at the Bavarian flood forecasting agency, was made available by the LfU. We forced the model using observed standard hydrometeorological time series like temperature (T), wind speed, humidity, radiation (R) and P, among others. LARSIM simulates E using the Penman-Monteith equation. Model input is based on interpolated station data (grid point method, NOAA, 1972). It is important to note that we use LARSIM exclusively as a uniform interpolator for P and E and that we do not use any other model output. Note that Table 2 also provides hydrometeorological catchment characteristics like four year mean annual precipitation ($\overline{P}$), discharge ($\overline{Q}$), runoff coefficient ($\overline{CR}$) and streamflow coefficient of variation ($\nu_{\overline{Q}}$). These quantities are all calculated based upon the available observables, except of E. Additional information on the Bavarian part of the Danube basin, the hydrometeorological data and the LARSIM model can be obtained from Seibert et al. (2014).

2.2 Dimensionless double mass curves to characterise runoff formation and the water balance

Essentially, we suggest that dimensionless double mass curves (dDMC), describing dimensionless accumulated release as function of dimensionless accumulated rainfall supply, are feasible to characterise storage controlled runoff formation in a scale invariant way. In this chapter we provide the corresponding principles (section 2.2.1) and introduce two different approaches to derive dDMCs (section 2.2.2) which allow us to address Q1. To address research questions Q2 and Q3 we employ

a temperature-index method to obtain season-specific dDMCs and relate their properties to the available catchment descriptors. The corresponding methods are described in chapter 2.3 and 2.4, respectively.

2.2.1 Theoretical principles

Pfister et al. (2002) introduced double mass curves as an easy to calculate measure for comparing seasonal runoff formation and runoff coefficients among catchments situated in different geologies in the Alzette basin. The double mass curve quantifies cumulated runoff as function of cumulated precipitation within the hydrological year. These diagrams are also helpful to learn about the interannual variability of the interplay of stream flow release and rainfall supply (Jackisch, 2015) and as a diagnostic signature to improve the development of perceptual models (Wrede et al., 2015). A straightforward means to assure comparability of double mass curves within the same clean geological and climate setting is to relate fluxes instead of the flows. This implies independence of a simple scaling of the respective flows with catchment area. Jackisch (2015) and Pfister et al. (2002) showed that the interannual variation of double mass curves and their shape is rather invariant within a given geological setting but reveals distinct differences among different geologies. The key question is, however, whether these differences reflect indeed non-trivial differences in terrestrial controls of runoff formation or whether these are (partly) caused by volumetric "scale effects", i.e. differences in storage volume, which arise from different vertical extents and not from differences in specific soil and bedrock porosity.

A nice example that underpins that a dimensionless formulation removes such simple scale effects is the dimensionless breakthrough curve (BTC). BTCs are used in soil physics to reveal (dis)similarities in transport and adsorption properties of a soil "filter" independently from the spatial extent of the probe (Hillel, 2004). Conceptually, BTCs are plots of scaled cumulated solute outflow against cumulated infiltration scaled by the pore volume of the sample. The BTC is hence invariant for a given solute as long as the ratio of accumulated water supply and storage capacity remains the same. The key to scale independence is to use the ratio of cumulated irrigation and storage volumes instead of using the cumulated irrigation alone.

The essential difference between a breakthrough experiment and the catchment water balance is, however, that the catchment releases water vapour to the atmosphere and liquid water to the stream at two different interfaces and that the subsurface extent of the storage volume and the flow path length to the riparian zone are largely unknown. Within our analogy, we regard liquid water release as the analogue to solute breakthrough, because both processes are controlled by gravity driven water fluxes through the subsurface. In contrast, E feeds largely from soil water which is stored against gravity. This part of soil water is largely immobile similar to adsorbed solutes (Zehe and Jackisch, 2016). The second difference is that the underlying controls on Q and particular on E do not only depend on properties of the soil but are also a function of time-dependent system states and boundary conditions like the meteorological forcing or the pressure difference between the groundwater system and the stream. Whereas the driving force on Q, i.e. gravity, is constant, the radiative forcing and the impact of vegetation on E follow a pronounced seasonal cycle – at least in temperate environments.

2.2.2 Derivation of dimensionless double mass curves

The key to obtain scale invariant dimensionless quantities is to divide a state variable of interest – for instance a force, velocity or length by a characteristic quantity of the system (Blöschl and Sivapalan, 1995). A popular example is the Reynolds number in fluid mechanics which relates inertial forces to viscous forces. It is defined as flow velocity times a characteristic length divided by the dynamic viscosity of the fluid, which is suited to compare turbulence in open channel flow independent of the channel width. The best known example of a dimensionless response measure in hydrology is the runoff coefficient (CR) (-) defined as specific discharge $(L\,T^{-1})$ over specific rainfall $(L\,T^{-1})$. The latter is often used as "diagnostic" variable to detect scale invariant differences in generation of runoff volume (Merz et al., 2006; Graeff et al., 2012).

As the double mass curve shall relate the accumulated dimensionless water release to dimensionless accumulated water supply of the catchment, the dependent and the independent variable need to be scaled properly. We define dimensionless accumulated water release (Q^*) (-) of the catchment based on the ratio of accumulated specific discharge $(cum.\,Q)$ (mm) and annual totals of P ($\sum P_{yr}$) (mm) of the corresponding hydrological year:

$$Q^* = \frac{cum.\,Q}{\sum P_{yr}} \tag{2}$$

We are aware that this definition is, while being close to the definition of a runoff coefficient, not exactly the desired ratio of accumulated release to total annual supply. The latter, i.e. infiltration needs to be determined based on the difference of accumulated P and accumulated E. As evaporation is however not easy to observe Eq. (2) might still be a suitable data-driven estimate.

Removing the dimension of the abscissa of the double mass curve, i.e. $cum.\,P$ (mm) is less trivial. An easy way is to use $\sum P_{yr}$ (mm) again to scale accumulated rainfall supply (Eq. 3). This nicely scales both axes to the range of $[0 \ldots 1]$ and allows to compare multiple sites and years.

$$P^* = \frac{cum.\,P}{\sum P_{yr}} \tag{3}$$

An advantage of plotting Q^* against P^* is that it yields curves that are fairly intuitive to interpret as this scaling preserves the shape of the original double mass curve:

- In snow-free periods with negligible E catchment "filter" properties naturally are expected to cause dDMCs with overall slopes smaller than unity. Step changes and deviations from the overall pattern towards smaller slopes suggest either activation of additional storage or that water leaves the catchment in form of evaporation, or both. Differentiation among these options is unfortunately not straightforward. A possible remedy is to focus on summer and winter periods individually, because snow is not relevant during summer and because E is deemed to be small during winter (see also chapter 2.3). To indicate the activation of storage in the snow pack we highlight dDMC periods with temperatures < 0° Celsius.

- Segments with slopes of 1 indicate that release equals supply and that water is neither stored in the system nor that significant amounts of water leave the catchment in the vapour phase. Segments with a slope of one or close to it are expected to occur during intensity controlled runoff formation, i.e. Hortonian overland flow (Horton, 1933) or preferential

flow (Beven and Germann, 2013). These mechanisms are not controlled by storage capacity but by infiltrability and conductivity of the subsurface (Struthers and Sivapalan, 2007). dDMC segments with slopes close to one may however also indicate fast responding catchments e.g. due to dominance of poorly developed shallow soils, impermeable substratum or strong topographic gradients.

– Segments of the dDMCs that expose slopes higher than one indicate that release is larger than supply. Provided that routing effect and runoff concentration times are negligible, these phenomena should be restricted to periods of significant snow melt. They should hence only occasionally occur in the winter and coincide with periods of snow accumulation that are followed by warming periods.

A considerable disadvantage of removing dimensions of the abscissa using $\sum P_{yr}$ (mm), named as type I dDMC here after, is however that same ordinate values do not reflect the same mass input and output. A comparison of the double mass curves defined like this among different sites and years might thus suggest greater similarity than there actually is.

Our alternative approach to scale the accumulated water supply is motivated by the dimensionless breakthrough curve and may be derived from the widely used HBV beta store (Bergstroem, 1976). The latter conceptualizes storage controlled formation of direct runoff (Q_d) as function of the relative saturation of the soil store, actual precipitation P(t) and a form parameter (β) (Eq. 4).

$$Q_d(t) = P\left(\frac{S_m(t)}{S_{max}}\right)^{\beta} \tag{4}$$

Considering the mass balance and Q_d as a surrogate for rainfall driven stream flow generation it can be shown from Eq. 4 that accumulated Q_d scales with the ratio of annual precipitation amount (cum. P) and S_{max} as one key factor, Eq. 5. The second factor that impacts the integral of Q_d is sum of the scaled storage difference minus the integral of infiltration minus evapotranspiration. Whereas both factors are strongly dependent on the unknown form parameter β, the second factor also depends on E. In appendix B we derive Eq. 5 from the integral of Eq. 4 and illustrate the role of the different terms using the case $\beta = 1$ (compare Eq. B7).

$$\int_{t_0}^{t^*} Q_d\, dt = \frac{cum.P}{S_{max}}\left(\frac{(S_m(t^*))^{\beta}}{(S_{max})^{\beta-1}} - \frac{(S_m(t_0))^{\beta}}{(S_{max})^{\beta-1}} - \int_{t_0}^{t^*}\left(\beta\left(\frac{S_m(t)}{S_{max}}\right)^{\beta-1}\left[P(t)\left(1-\left(\frac{S_m(t)}{S_{max}}\right)^{\beta}\right) - E(t)\right]\right)dt\right) \tag{5}$$

In accordance with the analytical proof we define the type II dDMC as accumulated stream flow release, scaled by the annual precipitation amount plotted against cum. P scaled by Smax. We expect this curve to be well suited to separate differences which arise from a different scaling of cum. P/S_{max} from differences arising due to differences in the second factor. Those might for instance arise by differences in how contribution areas in a catchment grow with relative saturation of the store. We expect those to be identifiable during the wet period. When increasing β to values larger than 1 the second factor gets non-linear which yields an increasingly convex function, which implies increasingly smaller winter slopes in the type II dDMC. When decreasing β to values smaller than unity the second factor gets non-linear and increasingly concave, which implies

increasingly larger type II dDMCs winter slopes. Differences which stem from differences in E, which might either arise from inter-annual climate variability ("warm or cold" summers) or in the climate setting are expected to be identifiable during the vegetation dominated summer period.

Deriving type II dDMCs requires however an estimate of S_{max}, i.e. information about the characteristic vertical extend of the storage volume, which is largely not observable. Standard data for soil storage that are available in Germany (Hartwich et al., 1995) include estimates of specific effective field capacity (eFC) and air capacity (AC) (both in mm/dm) which we multiplied by root zone depth to obtain (mm). Though these data are uncertain, particularly at the catchment scale we use them to scale cum P.

$$P^* = \frac{cum.\ P}{eFC + AC} \tag{6}$$

2.3 Season-specific evaluation of the double mass curve

During summer E often makes up the largest portion of the water balance. A possible way to isolate abiotic controls on runoff formation from the impact of the vegetation is to focus on data from frost-free periods of the "dormant" fall and winter season. Accordingly we derive season-specific double mass curves to unravel the impact of biotic and abiotic controls on the water balance. A season-specific evaluation however requires a meaningful separation of the dormant, i.e. the "winter period" and the vegetation dominated "summer period". Such a separation of seasons is however not at all trivial or not straight forward – particular for intercomparison studies. Conventional definitions like hydrological years or the astronomical equinoxes (varying between 19 and 21 March and 22 and 24 September in central Europe) are easy to implement and thus frequently used. However, they often introduce uncertainty as they fail in separating the dormant period from the active summer period in a meaningful way because the period of active vegetation does not only depend on the location of a catchment with respect to latitude but also with respect to altitude.

As an alternative we test temperature aggregates that are considered meaningful predictors of bulk ecological activity in plant physiology (Mäkelä et al., 2004). Specifically we use a temperature index model (Menzel et al., 2003) to define the onset and the duration of the period of vegetation in an ecologically meaningful, and data-driven way. In this approach the growing season is defined as the period between the last spring frost and the first fall frost in which the mean temperature constantly equals or exceeds 5 degree Celsius. The proposed method solely requires daily data of mean and minimum temperatures which are widely accessible and usually available within operational data sets. We evaluate the potential of this method visually by plotting the results into the double mass curves and by comparing them to the Gregorian definition of the beginning of spring which is March 20th in our climate. This way we address the second research question (Q2).

The separation of seasons essentially allows to derive average summer and winter runoff coefficients, CR_S (-) and CR_W (-) respectively, for the type I dDMCs, by calculating the slope of linear regression lines which we fit to the summer and winter segments of the type I dDMCs. For the type II dDMCs we evaluate only the winter regimes as we expect them to be primarily sensitive to the shape parameter β, and not to E, as suggested by the proof in appendix B. Analogously to the type I dDMCs we fit linear regression to the winter segments of the type II dDMCs. Here, the regression slope (m_W) (-) characterises how fast relative release increases with the potential renewal rate of the soil stock.

2.4 Catchment intercomparison and statistical evaluation of the double mass curves

To explain difference in seasonal runoff formation and hence to address Q3 we analyse properties of the dDMCs of the 22 different catchments. To this end we compare the resulting average m_W (-), CR_S (-), CR_W (-) and CR_{YR} (-) values and corresponding interannual variations, i.e. σ_{m_W}, σ_{CR_S}, σ_{CR_W} and $\sigma_{CR_{yr}}$ respectively. Specifically we compare the spatial patterns of the different quantities and test if they coincide with the groupings of our catchments which is based on hydrological regions as outlined in chapter 2.1.1. In a statistical evaluation we attribute differences in m_W (-), CR_S (-) and CR_W (-) to the structural and climatic properties of the catchments listed in Table 2 and Table 3 in the appendix. The key intention of this evaluation is to identify the underlying controls but also to evaluate the different scaling schemes of the dependent variable against the dimensionless formulation of accumulated supply. For this we use simple and multiple linear regression analyses and corresponding tests of significance which are implemented in the statistical computing platform R (R Core Team, 2015).

Last but not least we compare what we learn from the dDMCs concept against what to be learned from a benchmark diagnostic signature. The latter is chosen as the well known and widely used flow duration curve (FDC) (Ye et al., 2012; Euser et al., 2013; Viglione et al., 2013). FDCs describe the distribution of streamflow magnitudes with respect to exceedence probability. According to Sawicz et al. (2011) the slope of the flow duration curve (sFDC) can be considered as a proxy for average flow variability. Specifically we compare CR_S (-) and CR_W (-) against sFDC between the 33rd and 66th percentile within a linear regression analysis. Since exceedence probabilities are not sensitive to seasonality, we also derive season-specific FDCs based on the temperature-indexed derived seasons. Using the seasonal FDCs we calculate inter-percentile slopes of the dormant period ($sFDC_W$) and that of the period of vegetation ($sFDC_S$) and include them in the analysis.

3 Results

3.1 Dimensionless double mass curves

3.1.1 Type I dDMCs and temperature-index based season separation

The season-specific evaluation of the type I dDMCs revealed a surprisingly consistent pattern with similar mass curves in all catchments in terms of a fairly steep increase during winter and a clear regime shift towards much flatter, partly zero regression slopes in the vegetation period. In fact the slopes of the dDMCs for different years of a catchment are almost constant, just shifted in parallel, during the period of vegetation at many sites and for many years as depicted in Fig. 4. As indicated by Table 1 the seasonal winter runoff coefficients ($CR_W = 0.67$) exceed the average summer runoff coefficients ($CR_S = 0.32$) by a factor of 2 on average which is in line with the findings of (Hellebrand et al., 2008) – with exception of the two Alpine sites ALP2 and ALP3. The interannual variation of the seasonal runoff coefficients are twice as large during winter ($\sigma_{CR_W} = 0.1$) as during summer ($\sigma_{CR_S} = 0.06$) indicating a rather strong homogeneity during summer.

As depicted in Fig. 4 the temperature-based estimates for the beginning and end of winter and summer period (Menzel et al., 2003) nicely coincide with the points where the slope of the dDMCs changes (in terms of $cum.P/\sum P$). This applies

for different years and the entire range of physiographic conditions of our catchments, except for the Alpine region which is shown in the bottom left panel of Fig. 4. In the Alpine settings the accumulation and melting of snow has a stronger impact on the seasonal water balance than the period of vegetation. The spring and fall equinoxes (drawn as hatched areas in Fig. 4) and the temperature-based estimates of the tipping points deviate by -38 ± 17 and -26 ± 15 days, respectively. This corroborates that the use of conventional definitions for the onset and end of the period of vegetation may yield significantly different results compared to temperature-based estimates. We conclude that temperature aggregates can be included easily in hydrological analyses and suggest that they may have a higher explanatory power than static conventional definitions for the beginning and end of the period of vegetation such as the equinoxes.

– FIGURE 4: DIMENSIONLESS DOUBLE MASS CURVES TYPE I –

3.1.2 Season-specific evaluation of type II double mass curves

The type II dDMCs, which scale accumulated supply by the specific storage volume of the soil, in Fig. 5 reveal a similar shape as the type I dDMCs in Fig. 4. They exhibit an obvious regime shift between steep winter and flat summer regression slopes but expose a larger spread and a less linear progression. More importantly these curves reveal differences among the different hydrological years as dry and wet years are nicely separated as indicated by the end points of the individual curves. The type II dDMC is also much more suited to reveal differences among catchments, in terms of how "fast" they release a fraction of rainfall input in terms of how often the annual rainfall supply could potentially renew the soil water stock. This is particularly visible when comparing the ALP3 catchment in the lower left panel of Fig. 5, with a potential renewal rate of 30 to 40 soil pore volumes, against the catchments from the other settings with potential renewal rates almost one order of magnitude lower. Another advantage is the visibility of differences in winter slopes, which are more uniform in the non-Alpine catchments compared to type I dDMCs. Average m_W values and corresponding interannual variations (σ_{m_W}) are provided in Table 1.

With respect to the shape of the winter segment of the type II dDMCs we find that it is fairly linear for catchments located in the Faulted Molasse, Triassic and pre-Alpine regions (MOL2, TRI3 and AFO2) (Fig. 5, red lines). Step changes towards dDMC segments with smaller slopes indicate activation of additional storage or a larger beta parameter. This is particular visible in the winter periods of the Alpine catchments (ALP3) but also in the Bavarian Forest (BFO1) due to storage in a snow pack. This is also suggested by temperature data as these segments coincide with frost periods and are followed by segments of warming periods with much steeper slope. Clear step changes towards steeper slopes in the frost-free period are visible in the Alpine settings (ALP3) and to a smaller degree in AFO2 and BFO1 suggesting activation of rapid flow paths connected to the gauge. In the Jurassic site (JUR1), where the subsurface is karstified, the likely rapid runoff formation can neither distilled by the type II nor the type I dDMC. A possible explanation is that dDMCs are tailored for the seasonal to annual scale and that the detection of intensity controlled processes such as the formation of Hortonian overland flow and rapid subsurface flow require a signature that is tailored to the event scale.

– FIGURE 5: TYPE II DDMCS –

3.1.3 Spatial patterns and intercomparison of seasonal runoff formation

With respect to the different physiographic settings the type I dDMCs reveal distinct seasonal and spatial patterns. As indicated by the values in Table 1 the highest average winter runoff coefficients (CR_W=0.8-0.9) occur in the north-eastern catchments (BFO1, BFO2, BFO3) which are rather densely forested, but also in the Alpine ALP1 catchment. ALP4, ALP3 and ALP2, which are also Alpine catchments and located on similar altitudes, show much smaller CR_W values of 0.64 to 0.71 on average, most likely due to storage in the snow pack. The smallest winter runoff coefficients (0.35 and 0.55) occur in MOL7 and TRI3, respectively. With respect to the interannual winter variance we encounter small mean absolute deviations ≤ 0.05 in low-land sites of the Molasse and glacial drift areas e.g. MOL5, AFO4, AFO2, AFO1. High mean absolute deviations $\geq$0.15 occur in different geologies, including the sites TRI2, BFO1, JUR1 and BFO3. Please also note that $CR_{YR} > CR_W$ in a few cases where snow exhibits a strong control on winter runoff regimes (e.g. ALP3 or ALP2, see Table 1). Here, fitting linear regressions to the double mass curves is not suitable for estimating seasonal winter runoff coefficients.

The summer season is characterised by an opposite spatial pattern. $CR_S \geq$0.8 occur in the snow-dominated Alpine catchments of ALP3 and ALP2. The smallest CR_S with values between 0.07 and 0.12 are encountered at the Triassic sites (TRI3, TRI2, TRI1). In contrast to these end-members the variations within and among the Bavarian Forest, Molasse and Alpine Foreland regions differ between 0.20 and 0.35 indicating similar seasonal regimes at many sites. This implies that the large scale geological setting which we used for the naming of the sites (chapter 2.1) does not coincide with the spatial patterns in CR_S and CR_W which we obtain from the evaluation of the type I dDMCs. It is also important to note that for several low-land sites the CR_S shows very little interannual variance as indicated by mean absolute deviations $\leq$ 0.03 (e.g. MOL5, TRI3, TRI2, MOL6, MOL2, MOL4, JUR1, MOL3 and others) (Table 1). In these catchments the slopes of the dDMCs are fairly constant throughout different hydrological years indicating a very strong control of evapotranspiration on the water balance during summer. At these sites the curves of the dDMCs in summer have nearly identical slopes which is due to their state at the onset of vegetation activity in spring.

– TABLE 1: STATISTICS AND SLOPES OF THE DDMCS –

3.1.4 Regression analysis of dDMC characteristics with catchment characteristics

To better understand the physiographic controls on the seasonal runoff regimes we regress summer and winter runoff coefficients against 24 different site-specific variables. A key finding in this evaluation is that temperature sums, which can be considered as bulk parameter for ecological activity, explain 71 % of the variance of the CR_S with respect to the entire range of physiographic setting in our data set. During winter, temperature aggregates are not significant and without predictive power (Fig. 6). We thus state that the period of the vegetation exerts first-order controls on runoff formation as it dominates the seasonal interplay in storage and release in all of our physiographic and climatic settings except for the snow dominated Alpine region.

– FIGURE 6: SCATTERPLOT TEMPERATURE SUMS VS. SEASONAL RUNOFF COEFFICIENTS –

With respect to CR_W, which correspond to the winter slopes of the type I dDMCs, we did not find any single variable or combination of variables which explained more than 30 % of the variance in the CR_W. The best predictors for the latter are sand content ($r^2 = 0.29$), the product of topographic gradient and Ks τ ($r^2 = 0.22$) and silt content ($r^2 = 0.22$) which are followed by forest coverage ($r^2 = 0.16$), skeleton content ($r^2 = 0.15$), number of frost days ($r^2 = 0.14$), effective field capacity ($r^2 = 0.13$) and absolute sum of negative temperatures ($r^2 = 0.12$). All other variables have coefficients of determination $r^2 \leq 0.10$. In several multiple linear regressions based on the above mentioned variables the best result is obtained for a combination of τ, forest cover and absolute sum of negative temperatures (multiple $r^2 = 0.30$, p-value<0.001). Summer temperature sums, length and/or end of the period of vegetation from the previous hydrological year do not improve the prediction of the actual CR_W. The key finding of this analysis is that τ yields a $r^2 = 0.22$, whereas ϕ and Ks alone do only explain 0.02 and 0.08 % of the variance in the CR_W, respectively. This corroborates that surrogates for gradients and resistances act as team, as their product jointly controls the water flux in Darcy's Law. With respect to CR_W their joint impact is, despite of the coarse spatial resolution of the underlying soil and topographic map, even detectable at the lower mesoscale. We hence suggest that physically meaningful combinations of catchment descriptors should gain more attention in catchment intercomparison studies.

To further explore runoff formation during winter, we evaluate average m_W values of the type II double mass curves within a similar regression exercise. The latter characterises how fast accumulated dimensionless water release grows with the potential renewal rate of the soil store. The most explanatory variables for m_W are ϕ ($r^2 = 0.44$), relative fraction of rock-outcrops ($r^2 = 0.34$), length of the previous period of vegetation and corresponding temperature sums ($r^2 = 0.34$, $r^2 = 0.33$), total modelled snow water equivalent ($r^2 = 0.30$), skeleton content ($r^2 = 0.22$), and number of frost days ($r^2 = 0.19$). All correlations are negatively except for the two variables that relate to the previous period of vegetation. Average root depth, top soil air capacity (AC) and effective field capacity (eFC) also showed high coefficients of determination ($r^2 = 0.39$, 0.35 and 0.26, respectively) but non-trivial spurious fractions of correlation (Pearson, 1987; Kenney, 1982) make these variables difficult to interpret and hence, of little use. With respect to all sites and the four hydrological years winter potential renewal rates of the soil store differ by more than one order of magnitude (minimum=1.1, maximum=19.2). Average and standard deviation of all sites are 3.8 and 4.0 respectively.

There are two main findings in this evaluation. The first is that properties of the preceding summer period control the speed of relative water release in winter. This implies a long term catchment memory, which is most likely caused by the baseflow regimes, at some sites. In contrast, this is not the case for the period of vegetation where antecedent conditions from the winter period seem to have little or no impact on the growth of dimensionless water release to the stream. The second important finding is that ϕ is the most explanatory variable for m_W while τ is of little use here. This might suggests that capacity controlled runoff formation in the riparian zones might be dominant in most of the areas and that relative water release to the stream grows with accumulated water supply into these wetlands. However, the correlation of m_W and ϕ, fraction of rock-outcrops, modelled snow water equivalent, skeleton content, and number of frost days may also be a signal for the impact of physiographic and hydrometeorlogical conditions. In our data set all of these variables increase with altitude and have their maximum in the Alpine region.

In summary, we state that dimensionless and season-specific double mass curves provide manifold insights into the interplay of accumulated rainfall forcing and release among mesoscale catchments. As expected the type I dDMC, which is based on a trivial scaling of the abscissa, revealed a much more homogeneous picture of seasonal runoff formation than the type II dDMC, which uses proper scaling by considering the capacity of the soil store as a characteristic quantity for storage driven runoff formation. We hence state that the type II dDMC is the signature of choice but require more information.

3.2 Seasonal flow duration curves and comparison with dDMC properties

Season-specific evaluations reveal that seasonality in the rainfall-runoff formation is only weakly reflected by sFDC. As shown in Table 1 the slopes of the flow duration curve between the 33^{rd} and 66^{st}-percentiles are smaller during the period of vegetation ($sFDC_S$) than during the dormant period ($sFDC_W$). Absolute differences between $sFDC_W$ and $sFDC_S$ amounted 0.24, averaged over all sites and the four different years. In the most extreme case at site AFO4, which is depicted in Fig. 7, $sFDC_W$ and $sFDC_S$ differ by a factor of 2.

– FIGURE 7: SEASONAL FLOW DURATION CURVES –

Correlation analysis between $sFDC_W$ and $sFDC_S$ yields a coefficient of correlation of 0.64 ($r^2 = 0.40, p = 0.0014$). The slopes of the "all-data" flow duration curve sFDC and $sFDC_W$ are more narrowly correlated ($r^2 = 0.82, p < 0.001$) than sFDC and $sFDC_S$ ($r^2 = 0.52, p < 0.001$) implying that sFDC does rather reflect winter conditions than summer conditions. Comparing sFDC to seasonal coefficients of summer and winter flow variation ($\nu_{Q,S}$ and $\nu_{Q,W}$ respectively) yields coefficients of determination of $r^2 = 0.58$ and $r^2 = 0.47$ respectively. In contrast, we find only weak and insignificant correlations (all $r < 0.38$) among properties of the dDMC, i.e. CR_W, CR_S and CR_{YR}, and the slopes of the (seasonal) flow duration curves, $sFDC_W$, $sFDC_S$ and sFDC. This highlights that seasonality in the rainfall-runoff regimes is poorly reflected by FDCs regardless whether they are derived season-specific or for longer periods of time. We hence state that dDMCs provide insights into seasonal runoff dynamics that cannot be extracted from FDCs.

The spatial evaluations suggest that the geological groupings that have been used for the naming of the sites (chapter 2.1) do (again) not coincide with spatial patterns of the FDC properties. For instance TRI2, AFO1 and ALP1, which do belong to different geological units, share similar and fairly high $sFDC_W$ values of -1.24, -1.23 and -1.21, respectively, implying responsive regimes. The disagreement in the spatial patterns is also reflected by the $sFDC_S$ values. At the MOL-sites the $sFDC_S$ values differ between -0.22 and -0.99. This almost corresponds to the range of the $sFDC_S$ values observed at all other sites.

4 Discussion and conclusions

The rationale of this study was to derive data-driven dimensionless diagnostics for seasonal runoff formation and the water balance. Motivated by the notion of the three main hydrological functions of catchments proposed by Wagener et al. (2007), we tested the feasibility of different types of dimensionless double mass curves for this purpose. Conceptually we designed these

curves to characterise annual patterns of dimensionless water release to the stream as function of dimensionless water supply to the catchment store. Their feasibility was tested using operational data from 22 lower mesoscale catchments. Within this test we compared two different approaches to obtain dDMCs and put special emphasis on detecting and explaining a) seasonal differences in annual runoff formation and b) differences among the different catchments, which we a priori classified into five major geological and climatological settings. Our results provide evidence that dDMCs are straightforward to implement and well suited for characterizing the seasonal pattern of dimensionless water supply and release to stream flow as discussed in the following.

4.1 Potential and limitations of dimensionless double mass curves

We showed that a dimensionless formulation and season-specific evaluation of the traditional double mass curves is helpful to separate those influences which relate to simple scaling of the total supply from those which relate to the catchment specific way contribution areas grow with relative saturation of the catchment. This is a step ahead compared to the findings of Pfister et al. (2002) and Jackisch (2015). In line with their findings, we also found characteristic regime shifts in dDMCs between "steep" winter and "flat" summer conditions, however our test catchments are one to two orders of magnitude larger than those used in their studies. Our results revealed also that neither seasonal FDCs nor usual FDCs, which are frequently used to characterise runoff regimes in catchment intercomparison studies (Oudin et al., 2010; Sawicz et al., 2011; Casper et al., 2012; Viglione et al., 2013), yielded similar information. We hence conclude that season-specific dDMCs are a well suited fingerprint for characterizing seasonal runoff formation in meso-scale catchments of temperate environments and that dDCMs are suited for intercomparison studies.

A drawback of dDMCs is however, that accumulated dimensionless water supply is at best an estimator for total relative storage in the catchment. We have neither information on where the water is stored nor on whether it is subject to strong, weak or no capillary or osmotic forces. Even more important, dDMCs do not characterise the entire water balance. We cannot infer whether the mass balance residual of a hydrological year is partly stored within the catchment or whether it leaves the system as E. The real object of desire is, hence, a dimensionless triple mass curve (dTMC) which adds dimensionless E as third dimension to the double mass curve. As suggested by Seibert et al. (2016b) we expect that the resulting 3-D curves could also serve as an exhaustive diagnostic fingerprint for similarity in the annual and seasonal water balance. In combination with an analysis of the autocorrelation of the mass balance residuals, dTMCs could serve as indicator for long term memory effects. The drawback is here, however, that dTMCs are not entirely based on observables as measurements of E are commonly not available. Data-driven estimates on E can at best be provided in terms of maximum potential evaporation e.g. by dividing net radiation (R_{net}) with the latent heat of vaporization (λ).

Last not least we suggest that dDMCs can be used for model evaluation. Specifically, we propose that the model based reproduction of the tipping point between winter and summer (regime shift) can serve as a powerful benchmark for the accuracy of a model with respect to timing which is important but rarely done in hydrology (Seibert et al., 2016a). The reproduction of this signature is particularly relevant for models which are used for climate change studies where coping with the most likely non-stationary role of biotic controls (Milly et al., 2008) is of utmost importance.

4.2 Appropriate scaling of catchment input and output

Proper scaling means to divide a state variable of a system or a flux leaving the system by a time-invariant characteristic quantity (often a length or time) which limits the processes, i.e. the system dynamic of interest. The Reynolds number scales driving inertial forces, which scale with velocity and a characteristic length of the system, with counteracting viscous forces to characterise turbulence in a scale independent manner (Reynolds, 1883). The Damköhler number relates residence time in a reactor to the first-order reaction constant, to characterise its efficiency in a scale independent manner (Fogler, 1999; Oldham et al., 2013). We may also formulate cumulated input-output relations in a scale independent form, e.g. the dimensionless breakthrough curve characterise transport and adsorption properties of a soil probe in a scale invariant manner (Hillel, 2004). In analogy to this curve, we designed the double mass curve to characterise accumulated dimensionless water release as function of dimensionless water supply. Our choice of this input–output relation, and particular our approaches to scale input and output fluxes, are based on implicit assumptions about the dominant runoff generation processes.

Our scaling of the accumulated output by the annual rainfall depth essentially assumes that primarily the amount of rainfall limits the annual amount of stream flow generation and not so much the temporal pattern within the year or rainfall totals at the event scale (among other assumptions like a well defined catchment area and negligible exchange across subsurface boundaries). Scaled cum. P can however also be interpreted as characteristic time for accumulated direct runoff formation, and time stands still in case of no precipitation. Here we essentially assume that it is only the precipitation amount that counts, but not its intensity. This assumption might become inappropriate in areas were Hortonian overland flow dominates runoff generation (Mualem et al., 1990; Cerdan et al., 2002) because the intensity of the precipitation input crucially determines whether runoff leaves the system or not (Reaney et al., 2014). In such areas the same annual rainfall total can hence cause different amounts of annual release, depending on the subscale rainfall intensity pattern.

In line with this, we assume water release to the stream to increase monotonically with relative water storage. If this was true, accumulated release should scale with the ratio of cum. P over S_{max} as shown in Eq. B6. By estimating the potential renewal rate of the soil store based on the ratio of cum. P over S_{max}, we furthermore assume that transport through the subsurface is not limiting. Due to the absence of detailed information about active subsurface storage volumes in the target area, we scale cum. P with the specific storage volume of the soil (AC+eFC) in the root zone to obtain what we call the type II dDMC. We are aware that this is not the best choice as estimates of AC and eFC are uncertain (particular at the catchment scale) and refer to the root zone instead of to active storage – but this is what is available from operational data sets in Southern Germany. In any case we suggest testing additional scaling approaches by replacing the denominator in Eq. 6 by a more meaningful surrogate that relates to active storage. Originally, we scaled the annual accumulated rainfall with the annual rainfall depth of the respective hydrological year (Seibert et al., 2016b). This so called type I dDMC, preserves the shape of the non-scaled double mass curve and surrogates similarity among the different hydrological years in the same catchment and among different catchment as well. This is because both axes of the type I dDMC are scaled to unity. However, as the same step on the abscissa does not imply the same amount of mass input into the system, this similarity is in fact pseudo similarity, dry and wet years appear as the same.

In line with Blöschl and Sivapalan (1995) we hence conclude that proper scaling requires considering a suitable characteristic length and that the annual rainfall amount cannot be regarded as characteristic length scale for supply in this respect.

In contrast, we found that the type II dDMC are well suited to distinguish interannual differences between dry and wet years within the same catchment. They furthermore nicely visualize that the growth of relative water release with the potential renewal rate of the store was rather different between the different geological settings, while being typical within those settings. We hence conclude that the type II dDMC is well suited to discriminate seasonal patterns of runoff formation among different landscape settings and that its shape allows evaluating whether the above mentioned assumptions are met or not. A linear growth of dimensionless water release with the ratio of cum. P over S_{max} indicates that runoff generation grows rather linearly with relative saturation, i.e. if the form parameter in the beta store is close to unity. This can be for instance due to the dominance of saturation excess runoff formation in the riperian zone and is consistent with our finding that the topographic gradient was the best predictor for the winter slopes of the type II dDCMs within the catchment intercomparison. Deviations towards steeper slopes indicate that runoff generation is to a lesser extend growing with potential renewal of the store. This might for instance reflect importance of intensity controlled runoff formation such a Hortonian overland flow or rapid preferential flow. In summary we state the type I dDMCs focus on the "plain water balance". Their strength is in characterizing seasonal regimes, whereas type II dDMCs relate to storage referenced states.

It is also worth to shortly reflect on meaningful separations and definition of seasons and years. We argue that concepts like the hydrological year or conventional definitions like the equinoxes which are often used as a proxy for the beginning and end of the period of vegetation or the filling of the catchment store are fairly limited. This is in line with Hellebrand et al. (2008) who state that a meaningful separation of seasons is not straightforward to determine as hydrologically meaningful definitions of these seasons beyond a calendaric ones are not available. The same applies for instance for dynamic storage (dS) (Sayama et al., 2011). A comparison of dS time series would require that integration starts at the same relative storage amount in all catchments – and not at the same date. Significant dry or wet periods when subsurface wetness can be deemed as being either near saturation or near the minimum could for instance help to identify these points in time. These considerations differ from the dimensionless time used for instance by Woods (2003) who defined dimensionless time relative to the length of a seasonal cycle which was one year in his study for sites outside the tropics.

4.3 Physiographic and ecological controls on seasonal runoff formation

The analysis of physiographic and ecological controls revealed several important findings. The first is that temperature data prove to be good predictors for CR_S, independent from the physiographic and climatological conditions represented by our data set. This is in line with approaches used in plant and tree physiology where air temperature aggregates are widely used, for instance to estimate photosynthetic capacity (Mäkelä et al., 2004) or as harvest and hence biomass predictors (Perry et al., 1993; Rawson and Gomez-Macpherson, 2000). The applied temperature sum definitions however differ with respect to the process of interest. In Finland effective temperature sums are for instance defined as the sum of positive differences between the diurnal mean temperature and 5 °C (Solantie, 2004). Holdridge (1967) includes accumulated mean daily temperature (in ° C) above zero, starting on January 1^{st} among other variables to measure vegetative development in the tropics. In our analysis

we employ hourly temperature sums for the dormant period and the period of vegetation. It remains open for future research to find out if different definitions of temperature aggregates can further increase the prediction of CR_S. In this context we also raise the question whether characteristic ecological time scales are adequately represented in hydrological models. The parametrization of variables such as albedo, leaf-area-index, stomata and cuticula resistance which govern plant and canopy roughness still prevails in the form of time-invariant annual look-up tables which have typically been derived within a specific hydro-climatic setting – albeit they exhibit pronounced seasonal and spatial dynamics (Reichenau et al., 2016). This even applies for so called "physical" models such as Catflow (Maurer, 1997; Zehe et al., 2001) or WaSiM-ETH (Schulla, 1997) and it is in particular relevant for climate change studies (Milly et al., 2008). For further reading please refer to Loritz et al. (2016) who conduct additional studies on this topic and among others confront modelled E estimates with sap-flow data and both Gregorian and temperature-based estimates for the beginning of the vegetation period.

The second important finding is that τ, the product of the topographic gradient and the saturated hydraulic conductivity is an important predictor for the average winter runoff coefficients while the two variables alone are not. Because gradients and resistances exert first-order controls it is not surprising that both variables are often interpreted as independent predictors (McGuire et al., 2005; Santhi et al., 2008; Sayama et al., 2011; Li et al., 2012), while they in fact play as a team. The latter is essentially also supported by the gradient–flux relationship: Picturing a synthetic hillslope we intuitively expect that a doubled ϕ and a half Ks would yield the same Darcy flux. We hence suggest that physically meaningful combinations of catchment descriptors should gain more attention in catchment intercomparison studies. In this context it is also important to note that seasonal runoff coefficients are independent of the dDMCs. They only depend on the separation of seasons. Please note also that the areal share of impermeable substratum within the catchment is considered a good predictor for CR_W (Hellebrand et al., 2008). We did not test this variable as quantitative estimates were operationally not available.

The last important aspect is that type II dDMCs provide means to analyse how fast relative release increases with the potential renewal rate of the soil stock. dDMCs with a linear "first-order type" shape suggest that streamflow release from a catchment works similar to how saturation excess systems operate and that the form parameter in the beta store is close to 1. This is in line with the concept proposed by Kirchner (2009). Step changes and deviations of the dDMCs towards "flatter" and "steeper" segments may indicate the activation of additional storage, e.g. in the snow pack in winter respective snow melt, and relevance of intensity controlled processes in summer. Care must be taken if snow dominates parts of the dDMCs as the interpretation of these periods is ambiguous. We also want to highlight, that type II dDMCs are able to reveal memory effects between winter runoff conditions and the previous period of vegetation that can not be extracted from type I dDMCs. This corroborates that scaling the abscissa using a storage proxy is more meaningful than the use of total annual rainfall. To increase confidence in the proposed signatures we suggest applying them i) to a larger number of catchments and ii) to a (nested) set of small and densely instrumented catchments with more homogeneous physiognomic predictors.

4.4 Conclusions

In summary, we conclude that data-driven dDMCs turn out to be an easy-to-compute, yet very powerful diagnostic signature to characterise seasonal runoff formation and the water balance – also with respect to comparative analyses. They allow separating

terrestrial controls on runoff formation from that of the meteorological forcing, proper scaling and a season-specific evaluation provided. This holds also true as comparison of dDMCs and (seasonal) flow duration curves show that the former provide information that can not be extracted from the latter.

The use of cumulated precipitation to estimate supply can also be interpreted as characteristic time for accumulated direct runoff formation, and time stands still in case of no precipitation. Here we essentially assume that it is only the precipitation amount what counts, but not its intensity. We overall conclude that dimensionless double mass curves in combination and ecological temperature index is well suited to unravel biotic and abiotic controls on seasonal runoff formation as long as runoff formation does monotonously increase with storage. They are particularly suited to depict shifts between the winter and summer regimes, and where either runoff generation our evapotranspiration dominates water release.

Nevertheless, it is clear that seasonal scale is only one perspective on runoff formation and that understanding functional similarity, i.e. the behaviour of catchments (Schaefli et al., 2011) requires considering diagnostics for different temporal scales of runoff formation but also diagnostics for threshold processes.

5 Appendix

Appendix A: Soil and physiographic catchment properties

– TABLE 2: physiographic properties –

– TABLE 3: Soil properties –

Appendix B: Derivation of the type II dDMCs

In this proof we show that accumulated direct runoff (Q_d) scales with the ratio of cumulate precipitation (cum. P) over storage capacity of the soil store (S_{max}). It is based on the integral of Eq. (4) from t_0 to t^* which equals accumulated Q_d in a hydrological year Eq. B1:

$$\int_{t_0}^{t^*} Q_d \, dt = \int_{t_0}^{t^*} \left[P(t) \left(\frac{S_m(t)}{S_{max}} \right)^{\beta} \right] dt \tag{B1}$$

As product Eq. B1 needs to be evaluated using the formula for partial integration (which can be found in any text book).

$$\int_{t_0}^{t^*} u'(t) \, v \, dt = [u(t)v(t)]_{t_0}^{t^+} - \int_{t_0}^{t^*} u(t)v'(t)dt \tag{B2}$$

While u'(t) denotes the temporal derivative of u(t). Here we set P(t)=u'(t), implying that u(t) corresponds to the integral of P(t) over time + a constant $u(t) = \int P(t)dt + c$. Hence, v(t) is equal to $v(t) = \left(\frac{S_m(t)}{S_{max}}\right)^{\beta}$ and v'(t) is $v'(t) = \beta \left(\frac{S_m(t)}{S_{max}}\right)^{\beta-1} \frac{1}{S_{\max}} \left[\frac{dS}{dt}\right]$ according to the chain rule. Inserting these terms yields:

$$\int_{t_0}^{t^*} Q_d \, dt = \int_{t_0}^{t^*} P(t) \, dt \left[\left(\frac{S_m(t^*)}{S_{max}}\right)^{\beta} - \left(\frac{S_m(t_0)}{S_{max}}\right)^{\beta} \right] - \int_{t_0}^{t^*} P(t) \, dt \int_{t_0}^{t^*} \left[\beta \left(\frac{S_m(t)}{S_{max}}\right)^{\beta-1} \frac{1}{S_{max}} \frac{dS}{dt} \right] dt \tag{B3}$$

Note that the integral of P(t) from from t_0 to t^* is equal to the cumulated precipitation amount in this period (cum. P). Using this definition of cum. P we obtain Eq. B4.

$$\int_{t_0}^{t^*} Q_d \, dt = cum.P \left[\left(\frac{S_m(t^*)}{S_{max}}\right)^{\beta} - \left(\frac{S_m(t_0)}{S_{max}}\right)^{\beta} \right] - cum.P \int_{t_0}^{t^*} \left[\beta \left(\frac{S_m(t)}{S_{max}}\right)^{\beta-1} \frac{1}{S_{max}} \frac{dS}{dt} \right] dt \tag{B4}$$

Accumulated rainfall driven water release in a hydrological year is equal to the product of cum. P with a) the difference in relative saturation to the power of β at t_0 and t^* and b) the integral of the relative storage change scaled with the relative saturation of the store to the power of $\beta - 1$. The second summand is, depending on the form parameter β a strongly non-linear function and depends implicitly also on E(t) as the latter affects dS/dt. This is obvious if we insert the HBV equation, i.e. Eq. 4, into the water balance equation $\frac{dS}{dt} = P(t) - Q_d(t) - E(t)$:

$$\frac{dS}{dt} = \left[P(t) \left(1 - \left(\frac{S_m(t)}{S_{max}}\right)^{\beta} \right) - E(t) \right] \tag{B5}$$

In the last step we insert Eq. B5 into the second summand of Eq. B4 and factoring out cum. P/S_{max} we finally obtain:

$$\int_{t_0}^{t^*} Q_d \, dt = \frac{cum.P}{S_{max}} \left(\frac{(S_m(t^*))^{\beta}}{(S_{max})^{\beta-1}} - \frac{(S_m(t_0))^{\beta}}{(S_{max})^{\beta-1}} - \int_{t_0}^{t^*} \left(\beta \left(\frac{S_m(t)}{S_{max}}\right)^{\beta-1} \left[P(t) \left(1 - \left(\frac{S_m(t)}{S_{max}}\right)^{\beta} \right) - E(t) \right] \right) dt \right) \tag{B6}$$

Equation B6 shows that accumulated rainfall driven water release in a hydrological year scales with cum. P/S_{max} as one key factor. The second factor is sum of the scaled storage difference minus the integral of infiltration minus evapotranspiration; both are strongly dependent on the unknown form parameter β. The second summand also crucially depends on E. In case of a linear store, i.e. $\beta = 1$, the second factor is linear and equal to the storage difference ($\Delta S_m = S_m(t^*) - S_m(t_0)$) minus the integral of infiltration minus evapotranspiration.

$$\int_{t_0}^{t^*} Q_d \, dt = \frac{cum.P}{S_{max}} \left(\Delta S_m - \int_{t_0}^{t^*} \left(\left[P(t) \left(1 - \frac{S_m(t)}{S_{max}} \right) - E(t) \right] \right) dt \right) \tag{B7}$$

Appendix C: Link table

– TABLE 4: Link table –

Acknowledgements. We gratefully acknowledge the three anonymous reviewers whose comments helped us to improve the quality of the manuscript considerably. Furthermore, we acknowledge data provision by the Bavarian Environmental Agency (LfU) and the German (DWD) and Austrian weather services (ZAMG). Funding was partly provided by the LfU which is gratefully acknowledged as well. We also thank Hoshin Gupta for the inspiring discussion on the triple mass curve concept several years ago. We furthermore acknowledge support by Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of Karlsruhe Institute of Technology.

References

Ali, G., Tetzlaff, D., Soulsby, C., McDonnell, J. J., and Capell, R.: A comparison of similarity indices for catchment classification using a cross-regional dataset, Advances in Water Resources, 40, 11–22, doi:10.1016/j.advwatres.2012.01.008, 2012.

Barthold, F. K. and Woods, R. A.: Stormflow generation: A meta-analysis of field evidence from small, forested catchments, Water Resources Research, 51, 3730–3753, doi:10.1002/2014WR016221, 2015.

Bergstroem, S.: Development and application of a conceptual runoff model for Scandinavian catchments, Tech. rep., Swedish Meteorological and Hydrological Institute (SMHI), Report RHO 7, Norrkoping, 1976.

Beven, K. and Germann, P.: Macropores and water flow in soils revisited, Water Resources Research, 49, 3071–3092, doi:10.1002/wrcr.20156, 2013.

Beven, K. and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology, Hydrological Sciences Bulletin, 24, 43–69, 1979.

BGR and SGD: Hydrogeological spatial structure of Germany (HYRAUM). Digital map data v3.2. German Federal States Geological Surveys (SGD) and Federal Institute for Geosciences and Natural Resources (BGR)., 2015.

Black, P. E.: Watershed Functions, Journal of the American Water Resources Association, 33, 1–11, doi:10.1111/j.1752-1688.1997.tb04077.x, 1997.

Blöschl, G. and Sivapalan, M.: Scale issues in hydrological modelling: A review, Hydrological Processes, 9, 251–290, doi:10.1002/hyp.3360090305, 1995.

Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., and Savenije, H. H. G.: Runoff prediction in ungauged basins: synthesis across processes, places and scales, Cambridge University Press, Cambridge, 2013.

Boorman, D. B., Hollis, J. M., and Lilly, A.: Hydrology of soil types: a hydrologically-based classification of the soils of United Kingdom., Tech. Rep. 126, Institut of Hydrology, Wallingford, 1995.

Capell, R., Tetzlaff, D., Hartley, A. J., and Soulsby, C.: Linking metrics of hydrological function and transit times to landscape controls in a heterogeneous mesoscale catchment, Hydrological Processes, 26, 405–420, doi:10.1002/hyp.8139, 2012.

Casper, M. C., Grigoryan, G., Gronz, O., Gutjahr, O., Heinemann, G., Ley, R., and Rock, A.: Analysis of projected hydrological behavior of catchments based on signature indices, Hydrology and Earth System Sciences, 16, 409–421, doi:10.5194/hess-16-409-2012, 2012.

Cerdan, O., Souchère, V., Lecomte, V., Couturier, A., and Le Bissonnais, Y.: Incorporating soil surface crusting processes in an expert-based runoff model: Sealing and Transfer by Runoff and Erosion related to Agricultural Management, Catena, 46, 189–205, doi:10.1016/S0341-8162(01)00166-7, 2002.

Clark, M. P., Slater, A. G., Rupp, D. E., Woods, R. a., Vrugt, J. a., Gupta, H. V., Wagener, T., and Hay, L. E.: Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, Water Resources Research, 44, 1–14, doi:10.1029/2007WR006735, 2008.

Duan, Q., Schaake, J., Andr??assian, V., Franks, S., Goteti, G., Gupta, H. V., Gusev, Y. M., Habets, F., Hall, A., Hay, L., Hogue, T., Huang, M., Leavesley, G., Liang, X., Nasonova, O. N., Noilhan, J., Oudin, L., Sorooshian, S., Wagener, T., and Wood, E. F.: Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops, Journal of Hydrology, 320, 3–17, doi:10.1016/j.jhydrol.2005.07.031, 2006.

Euser, T., Winsemius, H. C., Hrachowitz, M., Fenicia, F., Uhlenbrook, S., and Savenije, H. H. G.: A framework to assess the realism of model structures using hydrological signatures, Hydrology and Earth System Sciences, 17, 1893–1912, doi:10.5194/hess-17-1893-2013, 2013.

Euser, T., Hrachowitz, M., Winsemius, H. C., and Savenije, H. H.: The effect of forcing and landscape distribution on performance and consistency of model structures, Hydrological Processes, 29, 3727–3743, doi:10.1002/hyp.10445, 2015.

Farmer, D., Sivapalan, M., and Jothityangkoon, C.: Climate, soil and vegetation controls upon the variability of water balance in temperate and semi-arid landscapes: Downward approach to hydrological prediction, Water Resources Research, 39, 1035, 2003.

Fogler, S. H.: Elements of Chemical Reaction Engineering, Prentice Hall, New Jersey, 1999.

Gassmann, M., Stamm, C., Olsson, O., Lange, J., Kümmerer, K., and Weiler, M.: Model-based estimation of pesticides and transformation products and their export pathways in a headwater catchment, Hydrology and Earth System Sciences, 17, 5213–5228, doi:10.5194/hess-17-5213-2013, 2013.

Graeff, T., Zehe, E., Blume, T., Francke, T., and Schröder, B.: Predicting event response in a nested catchment with generalized linear models and a distributed watershed model, Hydrological Processes, 26, 3749–3769, doi:10.1002/hyp.8463, 2012.

Gupta, H. V., Wagener, T., and Liu, Y.: Reconciling theory with observations: Elements of a diagnostic approach to model evaluation, Hydrological Processes, 22, 3802–3813, doi:10.1002/hyp.6989, 2008.

Hartwich, R., Behrens, J., Eckelmann, W., Haase, G., Richter, A., Roeschmann, G., and Schmidt, R.: Bodenübersichtskarte der Bundesrepublik Deutschland 1 : 1 000 000. Karte mit Erläuterungen, Textlegende und Leitprofilen, 1995.

Hellebrand, H., Van Den Bos, R., Hoffmann, L., Juilleret, J., and Pfister, L.: The potential of winter stormflow coefficients for hydrological regionalization purposes in poorly gauged basins of the middle Rhine region, Hydrological Sciences Journal, 53, 773–788, doi:10.1623/hysj.53.4.773, 2008.

Hillel, D.: Introduction to Environmental Soil Physics, Academic Press, Elsevier, San Diego, California, 2004.

Holdridge, L. R.: Life zone ecology, Tropical Science Center, San Jose, Costa Rica, 1967.

Höllering, S., Ihringer, J., Samaniego, L., and Zehe, E.: An integrated multi-fingerprint sensitivity-nested approach for regional model parameter estimation and catchment similarity assessment, Hydrology and Earth System Sciences Discussions, pp. 1–41, doi:10.5194/hess-2016-249, 2016.

Horton, R. E.: The role of infiltration in the hydrologic cycle, Transactions of the American Geophysical Union, 14, 446–460, 1933.

Hrachowitz, M., Bohte, R., Mul, M. L., Bogaard, T. A., Savenije, H. H. G., and Uhlenbrook, S.: On the value of combined event runoff and tracer analysis to improve understanding of catchment functioning in a data-scarce semi-arid area, Hydrology and Earth System Sciences, 15, 2007–2024, doi:10.5194/hess-15-2007-2011, 2011.

Hrachowitz, M., Savenije, H. H. G., Bloschl, 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, Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 58, 1198–1255, doi:10.1080/02626667.2013.803183, 2013.

Jackisch, C.: Linking structure and functioning of hydrological systems. How to achieve necessary experimental and model complexity with adequate effort., Ph.D. thesis, Institut für Wasser und Gewässerentwicklung, Bereich Hydrologie, Karlsruher Institut für Technologie (KIT), doi:10.5445/IR/1000051494, 2015.

Kenney, B. C.: Beware of spurious self-correlations!, Water Resources Research, 18, 1041–1048, doi:10.1029/WR018i004p01041, 1982.

Kirby, M.: Hydrograph modelling strategies, in: Progress in Physical and Human Geograph, edited by Peel, R., pp. 69–90, Heinemann, London, 1975.

Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, Water Resources Research, 42, 5 pp, doi:10.1029/2005WR004362, 2006.

Kirchner, J. W.: Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward, Water Resources Research, 45, W02 429, doi:10.1029/2008WR006912, 2009.

Klaus, J., Wetzel, C. E., Martinez-Carreras, N., Ector, L., and Pfister, L.: A tracer to bridge the scales: On the value of diatoms for tracing fast flow path connectivity from headwaters to meso-scale catchments, Hydrological Processes, 29, 5275–5289, doi:10.1002/hyp.10628, 2015.

Larsen, J. E., Sivapalan, M., Coles, N. A., and Linnet, P. E.: Similarity analysis of runoff generation processes in real-world catchments, Water Resources Research, 30, 1641–1652, doi:10.1029/94WR00555, 1994.

Leavesley, G. H.: A mountain watershed simulation model, Ph.D. thesis, Colorado State University, Fort Collins, 1973.

Ley, R., Casper, M. C., Hellebrand, H., and Merz, R.: Catchment classification by runoff behaviour with self-organizing maps (SOM), Hydrology and Earth System Sciences, 15, 2947–2962, doi:10.5194/hess-15-2947-2011, 2011.

Li, H., Sivapalan, M., and Tian, F.: Comparative diagnostic analysis of runoff generation processes in Oklahoma DMIP2 basins: The Blue River and the Illinois River, Journal of Hydrology, 418-419, 90–109, doi:10.1016/j.jhydrol.2010.08.005, 2012.

Li, H.-Y., Sivapalan, M., Tian, F., and Harman, C.: Functional approach to exploring climatic and landscape controls of runoff generation: 1. Behavioral constraints on runoff volume, Water Resources Research, 50, 9300–9322, doi:10.1002/2014WR016307, 2014.

Lindsay, J. B.: The Whitebox Geospatial Analysis Tools project and open-access GIS, in: Proceedings of the GIS Research UK 22nd Annual Conference, p. 8, doi:10.13140/RG.2.1.1010.8962, 2014.

Loritz, R., Hassler, S. K., Jackisch, C., Allroggen, N., van Schaik, L., Wienhöfer, J., and Zehe, E.: Picturing and modelling catchments by representative hillslopes, Hydrology and Earth System Sciences Discussions, pp. 1–56, doi:10.5194/hess-2016-307, 2016.

Ludwig, K. and Bremicker, M.: The Water Balance Model LARSIM, Tech. Rep. 22, Institut für Hydrologie der Universität Freiburg i.Br., 2006.

Mäkelä, A., Hari, P., Berninger, F., Hänninen, H., and Nikinmaa, E.: Acclimation of photosynthetic capacity in Scots pine to the annual cycle of temperature., Tree physiology, 24, 369–76, doi:10.1093/TREEPHYS/24.4.369, 2004.

Martínez-Carreras, N., Krein, A., Gallart, F., Iffly, J. F., Pfister, L., Hoffmann, L., and Owens, P. N.: Assessment of different colour parameters for discriminating potential suspended sediment sources and provenance: A multi-scale study in Luxembourg, Geomorphology, 118, 118–129, doi:10.1016/j.geomorph.2009.12.013, 2010.

Martínez-Carreras, N., Wetzel, C. E., Frentress, J., Ector, L., McDonnell, J. J., Hoffmann, L., and Pfister, L.: Hydrological connectivity inferred from diatom transport through the riparian-stream system, Hydrology and Earth System Sciences, 19, 3133–3151, doi:10.5194/hess-19-3133-2015, 2015.

Maurer, T.: Physikalisch begründete zeitkontinuierliche Modellierung des Wassertransports in kleinen ländlichen Einzugsgebieten, Phd thesis, Mitteilungen des Instituts für Hydrologie und Wasserwirtschaft, Bd. 61, Universitat Fridericiana zu Karlsruhe (TH), 1997.

McGuire, K. J., McDonnell, J. J., Weiler, M., Kendall, C., McGlynn, B. L., Welker, J. M., and Seibert, J.: The role of topography on catchment-scale water residence time, Water Resources Research, 41, W05 002, doi:10.1029/2004WR003657, 2005.

McMillan, H., Gueguen, M., Grimon, E., Woods, R., Clark, M., and Rupp, D. E.: Spatial variability of hydrological processes and model structure diagnostics in a 50 km 2 catchment, Hydrological Processes, 28, 4896–4913, doi:10.1002/hyp.9988, 2014.

McMillan, H. K., Clark, M. P., Bowden, W. B., Duncan, M., and Woods, R. A.: Hydrological field data from a modeller's perspective: Part 1. Diagnostic tests for model structure, Hydrological Processes, 25, 511–522, doi:10.1002/hyp.7841, 2011.

Menzel, A., Jakobi, G., Ahas, R., Scheifinger, H., and Estrella, N.: Variations of the climatological growing season (1951-2000) in Germany compared with other countries, International Journal of Climatology, 23, 793–812, doi:10.1002/joc.915, 2003.

Merz, R., Blöschl, G., and Parajka, J.: Spatio-temporal variability of event runoff coefficients, Journal of Hydrology, 331, 591–604, doi:10.1016/j.jhydrol.2006.06.008, 2006.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., and Stouffer, R. J.: Stationarity Is Dead: Whither Water Management?, Science, 319, 573–574, doi:10.1126/science.1151915, 2008.

Mualem, Y., Assouline, S., and Rohdenburg, H.: Rainfall induced soil seal (A) - A critical review of observations and models, CATENA, 17, 185–203, doi:10.1016/0341-8162(90)90008-2, 1990.

NOAA: National Weather Service River Forcast System Forecast Procedures, Tech. rep., NWS-Hydro-14, NOAA Technical Memorandum. US Department of Commerce, Washington, D.C., 1972.

Oldham, C. E., Farrow, D. E., and Peiffer, S.: A generalized Damköhler number for classifying material processing in hydrological systems, Hydrology and Earth System Sciences, 17, 1133–1148, doi:10.5194/hess-17-1133-2013, 2013.

Oudin, L., Andreassian, V., Perrin, C., Michel, C., and Le Moine, N.: Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments, Water Resour. Res., 44, W03 413, doi:10.1029/2007WR006240, 2008.

Oudin, L., Kay, A., Andreassian, V., and Perrin, C.: Are seemingly physically similar catchments truly hydrologically similar?, Water Resources Research, 46, W11 558, doi:10.1029/2009wr008887, 2010.

Pearson, K.: Mathematical Contributions to the Theory of Evolution. - On a Form of Spurious Correlation Which May Arise When Indices Are Used in the Measurement of Organs, Proceedings of the Royal Society of London, 60, 489–498, doi:10.1098/rspl.1896.0076, 1987.

Pelletier, J. D. and Rasmussen, C.: Geomorphically based predictive mapping of soil thickness in upland watersheds, Water Resources Research, 45, doi:10.1029/2008WR007319, 2009.

Perry, K., Sanders, D., Granberry, D., Thomasgarrett, J., Decoteau, D., Nagata, R., Dufault, R., Deanbatal, K., and McLaurin, W.: Heat units, solar radiation and daylength as pepper harvest predictors, Agricultural and Forest Meteorology, 65, 197–205, doi:10.1016/0168-1923(93)90004-2, 1993.

Peschke, G., Etzenberg, C., Müller, G., Töpfer, J., and Zimmermann, S.: Das wissensbasierte System FLAB - ein Instrument zur rechnergestützten Bestimmung von Landschaftseinheiten mit gleicher Abflussbilung, Tech. rep., Internationales Hochschulinstitut Zittau, IHI-Schriften, BD. 10, Zittau, 1999.

Pfannerstill, M., Guse, B., and Fohrer, N.: Smart low flow signature metrics for an improved overall performance evaluation of hydrological models, Journal of Hydrology, 510, 447–458, doi:10.1016/j.jhydrol.2013.12.044, 2014.

Pfister, L., Iffly, J. F., and Hoffmann, L.: Use of regionalized stormflow coefficients with a view to hydroclimatological hazard mapping, Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 47, 479–491, doi:10.1080/02626660209492948, 2002.

Pokhrel, P. and Yilmaz, K. K.: Multiple-criteria calibration of a distributed watershed model using spatial regularization and response signatures, Journal of Hydrology, 418, 49–60, doi:10.1016/j.jhydrol.2008.12.004, 2012.

R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, 2015.

Rawson, H. M. and Gomez-Macpherson, H.: Irrigated wheat - managing your crop, Food and Agriculture Organization of the United Nations (FAO), Rome, 2000.

Reaney, S. M., Bracken, L. J., and Kirkby, M. J.: The importance of surface controls on overland flow connectivity in semi-arid environments: results from a numerical experimental approach, Hydrological Processes, 28, 2116–2128, doi:10.1002/hyp.9769, 2014.

Reggiani, P., Sivapalan, M., and Majid Hassanizadeh, S.: A unifying framework for watershed thermodynamics: balance equations for mass, momentum, energy and entropy, and the second law of thermodynamics, Advances in Water Resources, 22, 367–398, doi:10.1016/S0309-1708(98)00012-8, 1998.

Reggiani, P., Sivapalan, M., and Hassanizadeh, S. M.: Conservation equations governing hillslope responses: Exploring the physical basis of water balance, Water Resources Research, 36, 1845, doi:10.1029/2000WR900066, 2000.

Reichenau, T. G., Korres, W., Montzka, C., Fiener, P., Wilken, F., Stadler, A., Waldhoff, G., and Schneider, K.: Spatial heterogeneity of Leaf Area Index (LAI) and its temporal course on arable land: Combining field measurements, remote sensing and simulation in a Comprehensive Data Analysis Approach (CDAA)., PLoS ONE, 11, 1–24, doi:10.1371/journal.pone.0158451, 2016.

Reynolds, O.: An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels, Philosophical Transactions of the Royal Society of London, 174, 935–982, doi:10.1098/rstl.1883.0029, 1883.

Robinson, J. S., Sivapalan, M., and Snell, J. D.: On the relative roles of hillslope processes, channel routing, and network geomorphology in the hydrologic response of natural catchments, Water Resources Research, 31, 3089–3101, doi:10.1029/95WR01948, 1995.

Samuel, J. M., Sivapalan, M., and Struthers, I.: Diagnostic analysis of water balance variability: A comparative modeling study of catchments in Perth, Newcastle, and Darwin, Australia, Water Resources Research, 44, n/a–n/a, doi:10.1029/2007WR006694, 2008.

Santhi, C., Allen, P. M., Muttiah, R. S., Arnold, J. G., and Tuppad, P.: Regional estimation of base flow for the conterminous United States by hydrologic landscape regions, Journal of Hydrology, 351, 139–153, doi:10.1016/j.jhydrol.2007.12.018, 2008.

Sawicz, K., Wagener, T., Sivapalan, M., Troch, P. A., and Carrillo, G.: Catchment classification: Empirical analysis of hydrologic similarity based on catchment function in the eastern USA, Hydrology and Earth System Sciences, 15, 2895–2911, doi:10.5194/hess-15-2895-2011, 2011.

Sayama, T., McDonnell, J. J., Dhakal, A., and Sullivan, K.: How much water can a watershed store?, Hydrological Processes, 25, 3899–3908, doi:10.1002/hyp.8288, 2011.

Schaake, J., Duan, Q., Smith, M., and Koren, V.: Criteria to select basins for hydrologic model development and testing, 15th Conference on Hydrology (P1.8), AMS, January 9-14. Long Beach, CA, 2000.

Schaap, M. G., Leij, F. J., and van Genuchten, M. T.: Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions, Journal of Hydrology, 251, 163–176, doi:10.1016/S0022-1694(01)00466-8, 2001.

Schaefli, B., Harman, C. J., Sivapalan, M., and Schymanski, S. J.: HESS Opinions: Hydrologic predictions in a changing environment: Behavioral modeling, Hydrology and Earth System Sciences, 15, 635–646, doi:10.5194/hess-15-635-2011, 2011.

Scherrer, S. and Naef, F.: A decision scheme to indicate dominant hydrological flow processes on temperate grassland, Hydrological Processes, 17, 391–401, doi:10.1002/hyp.1131, 2003.

Schmocker-Fackel, P., Naef, F., and Scherrer, S.: Identifying runoff processes on the plot and catchment scale, Hydrology and Earth System Sciences, 11, 891–906, doi:10.5194/hess-11-891-2007, 2007.

Schulla, J.: Hydrologische Modellierung von Flussgebieten zur Abschätzung der Folgen von Klimaäderungen, Phd thesis, Zürcher Geographische Schriften, Bd. 69, ETH Zürich, doi:10.3929/ethz-a-001763261, 1997.

Seibert, S. P., Skublics, D., and Ehret, U.: The potential of coordinated reservoir operation for flood mitigation in large basins - A case study on the Bavarian Danube using coupled hydrological-hydrodynamic models, Journal of Hydrology, 517, 1128–1144, doi:10.1016/j.jhydrol.2014.06.048, 2014.

Seibert, S. P., Ehret, U., and Zehe, E.: Disentangling timing and amplitude errors in streamflow simulations, Hydrology and Earth System Sciences, 20, 3745–3763, doi:10.5194/hess-20-3745-2016, 2016a.

Seibert, S. P., Jackisch, C., Pfister, L., Ehret, U., and Zehe, E.: Exploring the interplay between state, structure and runoff behavior of lower mesoscale catchments, Hydrol. Earth Syst. Sci. Discuss., pp. 1–51, doi:10.5194/hess-2016-109, 2016b.

Sivapalan, M.: Prediction in ungauged basins: a grand challenge for theoretical hydrology, Hydrological Processes, 17, 3163–3170, doi:10.1002/hyp.5155, 2003.

Sivapalan, M., Beven, K., and Wood, E. F.: On hydrologic similarity: 2. A scaled model of storm runoff production, Water Resour. Res., 23, 2266–2278, 1987.

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V., Karambiri, H., Lakshmi, W., Liang, X., McDonnell, J., Mendionde, E. M., O'Connel, P., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S., and Zehe, E.: IAHS Decade on Predictions in Ungauged Basins (PUB), 2003-2012: shaping an exciting future for the hydrological sciences, Hydrol. Sci. J., 48, 857–880, 2003.

Solantie, R.: Daytime temperature sum - a new thermal variable describing growing season characteristics and explaining evapotranspiration, Boreal environment research, 9, 319–333, 2004.

Soulsby, C., Tetzlaff, D., and Hrachowitz, M.: Are transit times useful process-based tools for flow prediction and classification in ungauged basins in montane regions?, Hydrological Processes, 24, 1685–1696, doi:10.1002/hyp.7578, 2010.

Struthers, I. and Sivapalan, M.: A conceptual investigation of process controls upon flood frequency: role of thresholds, Hydrology and Earth System Sciences, 11, 1405–1416, doi:10.5194/hess-11-1405-2007, 2007.

Tian, F., Li, H., and Sivapalan, M.: Model diagnostic analysis of seasonal switching of runoff generation mechanisms in the Blue River basin, Oklahoma, Journal of Hydrology, 418-419, 136–149, doi:10.1016/j.jhydrol.2010.03.011, 2012.

Varado, N., Braud, I., Galle, S., Le Lay, M., Séguis, L., Kamagate, B., and Depraetere, C.: Multi-criteria assessment of the Representative Elementary Watershed approach on the Donga catchment (Benin) using a downward approach of model complexity, Hydrology and Earth System Sciences, 10, 427–442, doi:10.5194/hess-10-427-2006, 2006.

Viglione, A., Parajka, J., Rogger, M., Salinas, J. L., Laaha, G., Sivapalan, M., and Blöschl, G.: Comparative assessment of predictions in ungauged basins - Part 3: Runoff signatures in Austria, Hydrology and Earth System Sciences, 17, 2263–2279, doi:10.5194/hess-17-2263-2013, 2013.

Vrugt, J. A. and Sadegh, M.: Toward diagnostic model calibration and evaluation: Approximate Bayesian computation, Water Resour. Res., 49, 4335–4345, doi:10.1002/wrcr.20354, 2013.

Wagener, T., Sivapalan, M., Troch, P., and Woods, R.: Catchment Classification and Hydrologic Similarity, Geography Compass, 1, 901–931, doi:10.1111/j.1749-8198.2007.00039.x, 2007.

Wang, D. and Wu, L.: Similarity of climate control on base flow and perennial stream density in the Budyko framework, Hydrology and Earth System Sciences, 17, 315–324, doi:10.5194/hess-17-315-2013, 2013.

Wienhöfer, J., Germer, K., Lindenmaier, F., Färber, A., and Zehe, E.: Applied tracers for the observation of subsurface stormflow at the hillslope scale, Hydrology and Earth System Sciences, 13, 1145–1161, doi:10.5194/hess-13-1145-2009, 2009.

Wood, E. F., Sivapalan, M., and Beven, K.: Scale and similarity in catchment storm response, Rev. Geophys., 28, 1–18, 1990.

Woods, R.: The relative roles of climate, soil, vegetation and topography in determining seasonal and long-term catchment dynamics, Advances in Water Resources, 26, 295–309, doi:10.1016/S0309-1708(02)00164-1, 2003.

Wrede, S., Fenicia, F., Martinez-Carreras, N., Juilleret, J., Hissler, C., Krein, A., Savenije, H. H. G., Uhlenbrook, S., Kavetski, D., and Pfister, L.: Towards more systematic perceptual model development: A case study using 3 Luxembourgish catchments, Hydrological Processes, 29, 2731–2750, doi:10.1002/hyp.10393, 2015.

Ye, S., Yaeger, M., Coopersmith, E., Cheng, L., and Sivapalan, M.: Exploring the physical controls of regional patterns of flow duration curves - Part 2: Role of seasonality, the regime curve, and associated process controls, Hydrology and Earth System Sciences, 16, 4447–4465, doi:10.5194/hess-16-4447-2012, 2012.

Zehe, E. and Jackisch, C.: A Lagrangian model for soil water dynamics during rainfall-driven conditions, Hydrology and Earth System Sciences, 20, 3511–3526, doi:10.5194/hess-20-3511-2016, 2016.

Zehe, E., Maurer, T., Ihringer, J., and Plate, E.: Modeling water flow and mass transport in a loess catchment, Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere, 26, 487–507, doi:10.1016/S1464-1909(01)00041-7, 2001.

Zehe, E., Ehret, U., Pfister, L., Blume, T., Schröder, B., Westhoff, M., Jackisch, C., Schymanski, S. J., Weiler, M., Schulz, K., Allroggen, N., Tronicke, J., Van Schaik, L., Dietrich, P., Scherer, U., Eccard, J., Wulfmeyer, V., and Kleidon, A.: HESS Opinions: From response units to functional units: A thermodynamic reinterpretation of the HRU concept to link spatial organization and functioning of intermediate scale catchments, Hydrology and Earth System Sciences, 18, 4635–4655, doi:10.5194/hess-18-4635-2014, 2014.

Zhang, G. P., Savenije, H. H. G., Fenicia, F., and Pfister, L.: Modelling subsurface storm flow with the Representative Elementary Watershed (REW) approach: application to the Alzette River Basin, Hydrology and Earth System Sciences, 10, 937–955, doi:10.5194/hess-10-937-2006, 2006.

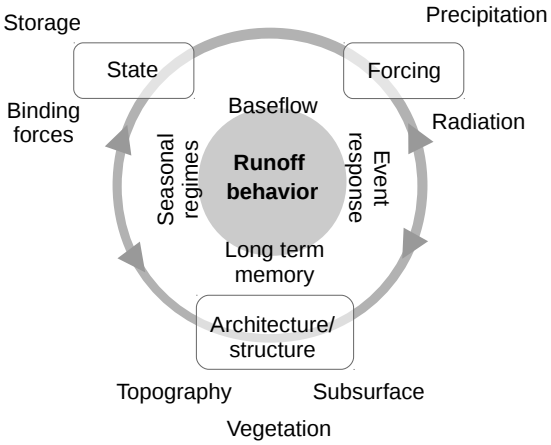

Figure 1. Conceptual sketch of the major temporal scales of runoff formation, i.e. long-term behaviour, seasonal regimes, generation of baseflow and event runoff production, and corresponding first-order controls.

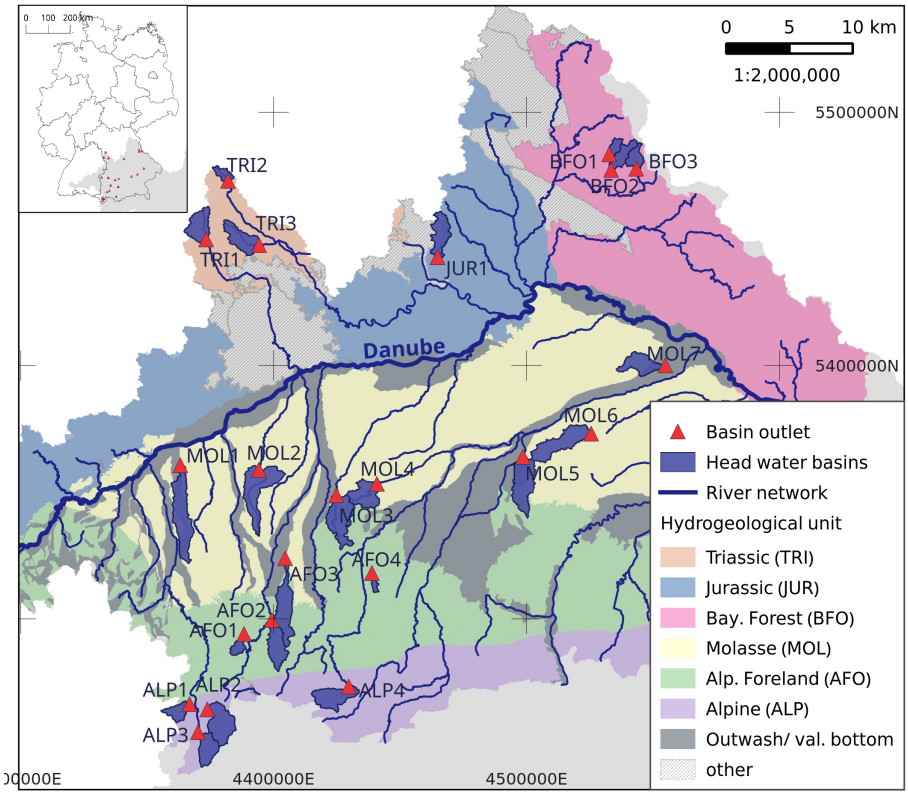

Figure 2. Upper Danube catchment in southern Germany with selected head water basins (blue polygons), corresponding gauges (red triangles) and major river network (blue lines). The site identifiers (IDs) refer to the corresponding (hydro)geological unit (color coded map in the background, adapted from BGR and SGD (2015)) and a single Arabic numeral. Moving from the North-West to the South we differentiate TRI (Triassic), JUR (Jurasic), BFO (Bavarian Forest), MOL (Faulted Molasse), AFO (Alpine Foreland) and ALP (Alpine). Table 4 provides links between the site IDs and the real gauge names. The inset in the upper left corner shows Germany's federal state boundaries, the individual head water outlets and the basin of the Danube (grey area). The grid coordinates refer to the Gauss-Kruger zone 4 projection (CRS identifier EPSG:31468).

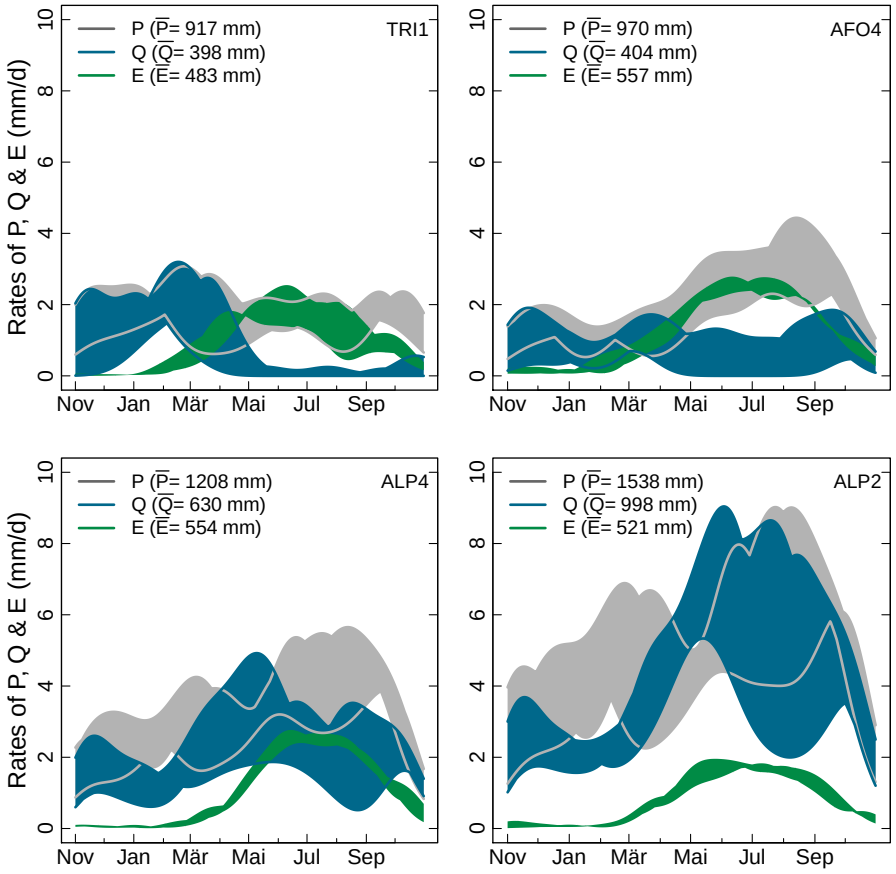

Figure 3. Regime curves of observed areal precipitation (P) (grey), discharge (Q) (blue) and calculated areal mean evapotranspiration (E, Penman-Monteith) (green) of four selected head water catchments. The width of the individual bands illustrates the interannual variation during the four year lasting period. The curves are based on kernel density estimates using identical kernels and bandwidths for variables of the same type. $\overline{P}$, $\overline{Q}$, and $\overline{E}$ provide the four year mean annual average of the different variables respectively.

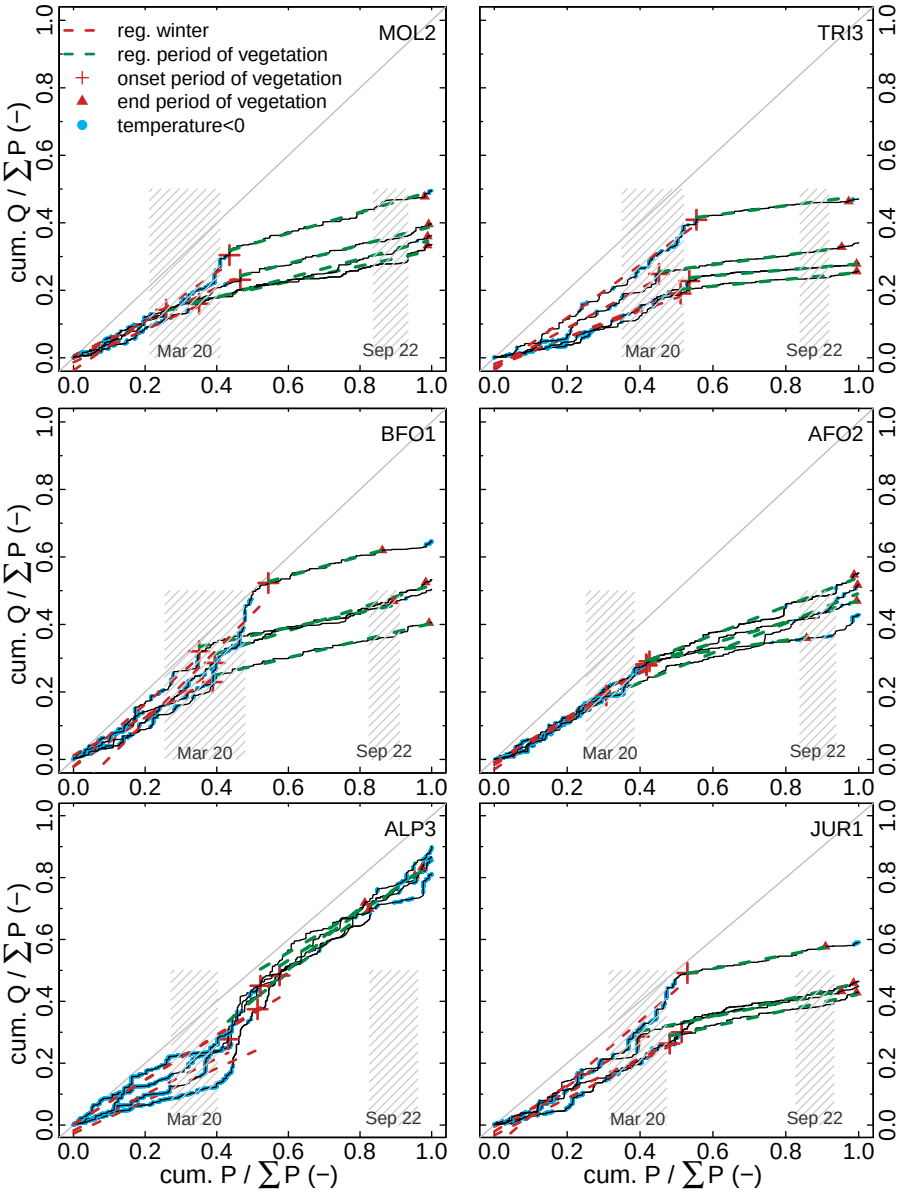

Figure 4. Type I dimensionless double mass curves for catchments of different geological units for the hydrological years 1999-2003. Onset and end of the period of vegetation are determined using a temperature index model. Regression lines are fitted to both periods (dotted lines in red/ green), their slopes are interpreted as seasonal runoff coefficients. Periods with temperatures < 0 °C are highlight in blue. Gregorian definitions for the first start of spring (Mar 20th) and last start of fall (Sep 22th) are added to the $cum.\,P/\sum P$ plane as hatched rectangles (to improve readability) to highlight their differences to temperature-based estimates on the onset and end of the period of vegetation. Statistical properties of the dDMCs are summarized in Table 1.

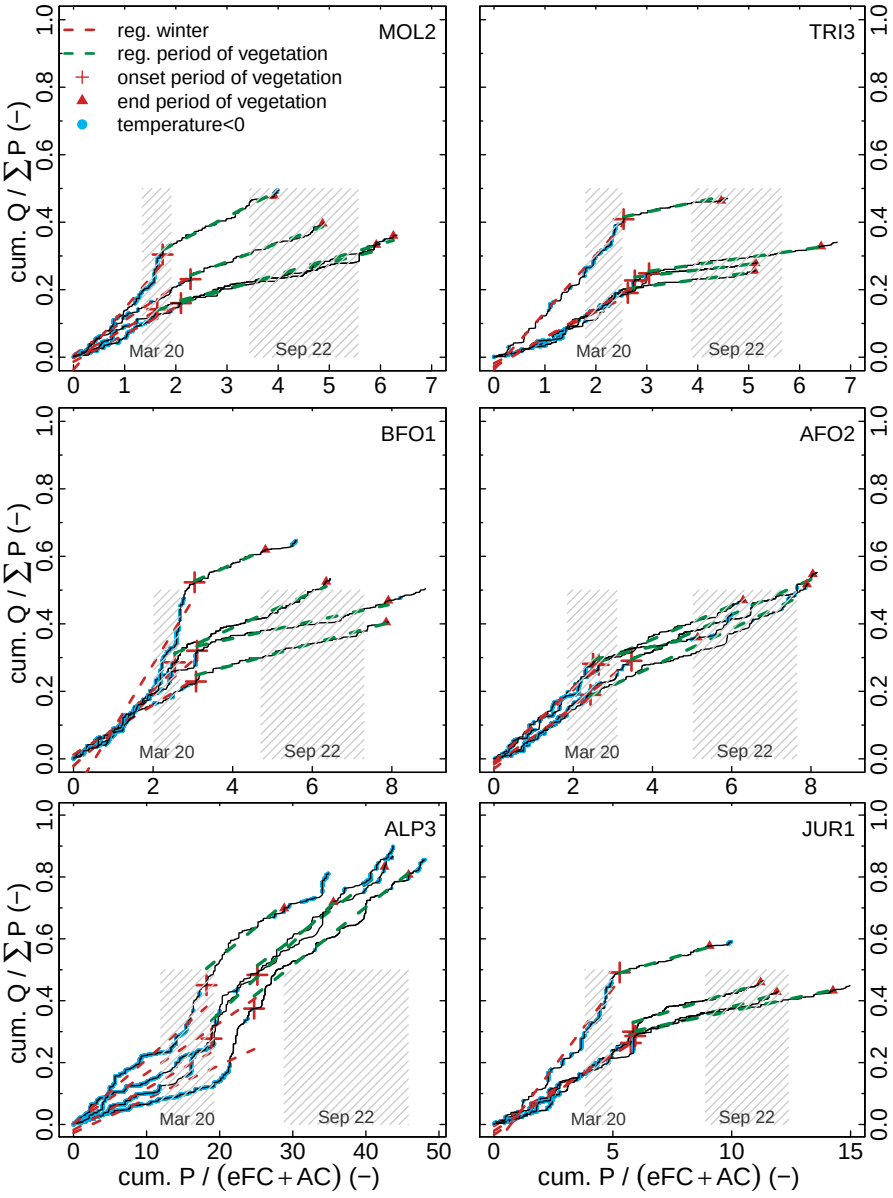

Figure 5. Type II dimensionless double mass curves for six catchments of different geological units. The dDMCs depicted here are the same as those depicted in Fig. 4 but with the difference that the abscissa is scaled using the sum of average effective field capacity (eFC) and air capacity (AC). All other properties are analogue to Fig. 4.

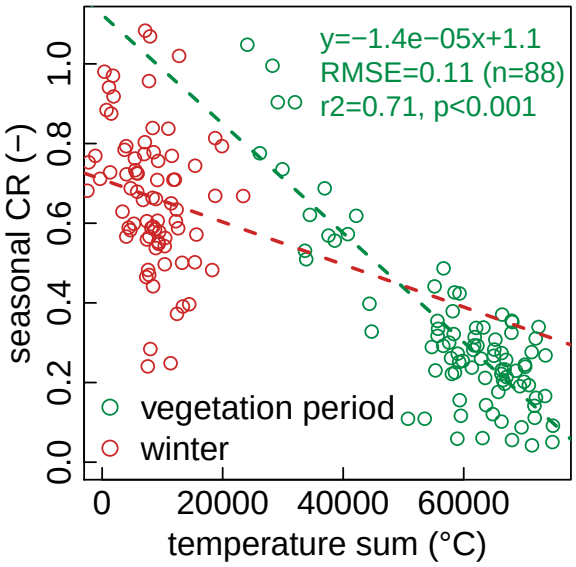

Figure 6. Hourly summer and winter temperature sums, calculated for the dormant period and the period of vegetation, plotted against seasonal summer (green) and winter (red) runoff coefficients for all sites (n=22) and years (n=4). The dotted lines are linear regressions. Statistical information on the summer model is plotted in green. During winter we did not find a significant statistical relation (r^2=0.04, p=0.062). Statistical properties of all dDMCs are summarized in Table 1.

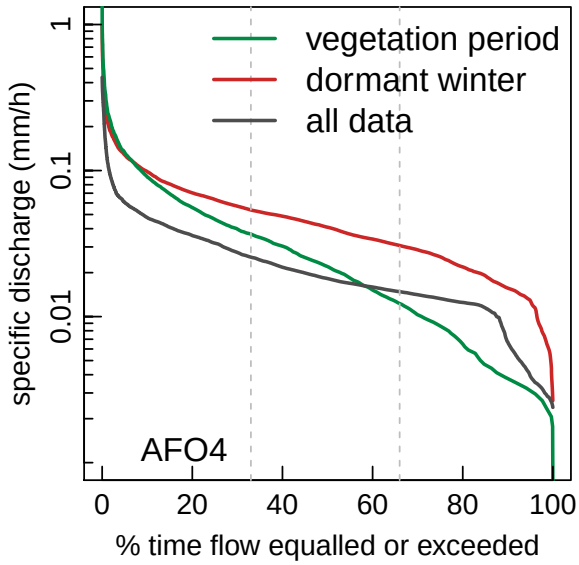

Figure 7. Seasonal flow duration curves for the site AFO4. They exhibit considerable differences in the slopes between the 33rd and 66th flow percentile (dashed vertical lines) which are widely used to characterise average flow conditions (Sawicz et al., 2011). The vegetation and dormant period are separated using the temperature index method proposed by Menzel et al. (2003).

Table 1. Mean seasonal winter (CR_W), summer (CR_S) and annual runoff coefficients (CR_{YR}) as indicated by the slope of linear regression lines fitted to the different segments of the type I dDMCs. The analogue for the type II dDMCs is m_W which describes how fast dimensionless relative release grows with the potential renewal rate of the soil stock during winter. The interannual variation of these quantities within the hydrological years ('99-'03) is quantified using the mean absolute deviation and given by σ_{CR_W}, σ_{CR_S}, $\sigma_{CR_{YR}}$, and σ_{m_W}, respectively. $sFDC_W$, $sFDC_S$ and $sFDC$ characterise the slope of seasonal flow duration curves between the 33rd and 66th percentile of the dormant period, the period of vegetation and all data, respectively. The separation of seasons was done using a temperature-index model (Menzel et al., 2003). All quantities are dimensionless.

Site	CR_W	CRS	CR_{YR}	m_W	σ_{CR_W}	σ_{CR_S}	$\sigma_{CR_{YR}}$	σ_{m_W}	$sFDC_W$	$sFDC_S$	sFDC
TRI1	0.72	0.12	0.43	0.098	0.058	0.030	0.031	0.026	-1.15	-0.88	-1.46
TRI2	0.70	0.07	0.37	0.111	0.198	0.027	0.105	0.033	-1.24	-1.04	-1.94
TRI3	0.55	0.12	0.34	0.105	0.133	0.024	0.069	0.040	-0.93	-0.65	-0.99
JUR1	0.73	0.25	0.48	0.063	0.150	0.017	0.054	0.039	-0.88	-0.49	-0.78
BFO1	0.82	0.28	0.52	0.120	0.169	0.037	0.068	0.035	-0.58	-0.48	-0.61
BFO2	0.85	0.30	0.54	0.114	0.143	0.020	0.067	0.034	-0.69	-0.39	-0.64
BFO3	0.93	0.29	0.57	0.126	0.173	0.034	0.089	0.035	-0.82	-0.46	-0.76
MOL1	0.60	0.24	0.38	0.121	0.103	0.040	0.021	0.009	-0.60	-0.54	-0.63
MOL2	0.56	0.27	0.40	0.113	0.080	0.022	0.049	0.028	-0.38	-0.22	-0.36
MOL3	0.62	0.34	0.46	0.121	0.069	0.019	0.045	0.010	-0.32	-0.24	-0.27
MOL4	0.56	0.23	0.37	0.101	0.084	0.021	0.051	0.020	-0.53	-0.31	-0.50
MOL5	0.69	0.22	0.41	0.134	0.055	0.028	0.026	0.027	-0.56	-0.36	-0.57
MOL6	0.60	0.20	0.36	0.152	0.066	0.018	0.025	0.005	-0.61	-0.42	-0.58
MOL7	0.35	0.27	0.31	0.110	0.139	0.089	0.031	0.003	-0.55	-0.99	-0.71
AFO1	0.72	0.35	0.51	0.135	0.031	0.120	0.042	0.027	-1.23	-1.25	-1.32
AFO2	0.68	0.34	0.49	0.098	0.036	0.099	0.043	0.003	-0.77	-0.58	-0.76
AFO3	0.56	0.24	0.38	0.115	0.130	0.068	0.031	0.016	-0.72	-0.84	-0.75
AFO4	0.66	0.22	0.39	0.136	0.050	0.090	0.030	0.035	-0.73	-1.42	-1.03
ALP1	0.89	0.50	0.75	0.045	0.098	0.088	0.031	0.043	-1.21	-0.93	-1.14
ALP2	0.71	0.82	0.83	0.017	0.067	0.161	0.047	0.037	-0.65	-0.76	-0.83
ALP3	0.64	0.84	0.86	0.015	0.082	0.109	0.025	0.013	-1.01	-0.66	-1.07
ALP4	0.66	0.53	0.64	0.024	0.064	0.066	0.032	0.059	-0.52	-0.30	-0.49
Mean	0.67	0.32	0.49	0.099	0.10	0.06	0.046	0.026	-0.76	-0.66	-0.83

Table 2. Physiographic catchment properties in terms of topography, landuse and hydro-meteorology. The columns contain site identifier (ID), catchment size (A), mean catchment elevation above sea level (elev), median topographic gradient (ϕ), median topographic gradient times average saturated hyd. conductivity (τ), relative land coverage ratios for infrastructure (infr), arable land (arab), pasture (past), forest (frst), wetlands (wet) and rock outcrops (rock), the 30 year mean annual precipitation (MAP), four year mean annual precipitation ($\overline{P}$), discharge ($\overline{Q}$), runoff coefficient ($\overline{CR}$) and streamflow coefficient of variation ($\nu_{\overline{Q}}$).

	topography				% land use coverage						hydrometeorological characteristics				
ID	A (km^2)	elev (m)	ϕ (-)	τ ($m\ s^{-1}$)	infr	arab	past	frst	wet	rock	MAP (mm)	$\overline{P}$ (mm)	$\overline{Q}$ (mm)	$\overline{CR}$ (-)	$\nu_{\overline{Q}}$ (-)
TRI1	88	481	2.8e-02	1.0e-06	0.04	0.54	0.20	0.21	0	0	802	919	0.045	0.43	2.00
TRI2	26	460	2.5e-02	9.8e-07	0	0.60	0.12	0.29	0	0	707	801	0.033	0.36	2.20
TRI3	93	468	3.8e-02	8.3e-07	0.02	0.62	0.07	0.30	0	0	738	829	0.032	0.33	1.80
JUR1	90	518	7.2e-02	2.0e-06	0.01	0.59	0.15	0.26	0	0	833	839	0.046	0.48	0.83
BFO1	25	620	6.9e-02	2.5e-06	0	0.35	0.06	0.60	0	0	889	933	0.054	0.51	0.89
BFO2	64	635	6.1e-02	1.3e-06	0.02	0.39	0.12	0.47	0.01	0	893	920	0.055	0.53	0.91
BFO3	58	624	7.9e-02	1.2e-06	0.01	0.24	0.21	0.55	0	0	908	825	0.052	0.56	1.10
MOL1	166	543	1.1e-02	1.6e-08	0.07	0.42	0.29	0.23	0	0	889	973	0.042	0.38	0.83
MOL2	163	515	4.0e-02	6.9e-08	0.03	0.28	0.37	0.32	0	0	901	1010	0.045	0.39	0.90
MOL3	163	558	1.4e-02	2.0e-08	0.05	0.69	0.10	0.15	0	0	933	1100	0.057	0.45	0.64
MOL4	97	517	2.5e-02	3.9e-08	0.04	0.77	0.03	0.15	0	0	888	1016	0.042	0.36	0.93
MOL5	133	473	2.0e-02	4.4e-08	0.05	0.81	0.05	0.09	0	0	883	1016	0.047	0.40	1.00
MOL6	146	484	4.2e-02	7.8e-08	0.02	0.79	0.04	0.15	0	0	856	721	0.029	0.35	1.60
MOL7	87	379	2.4e-02	3.4e-08	0.02	0.73	0.01	0.24	0	0	744	733	0.026	0.31	1.10
AFO1	45	840	3.4e-02	3.6e-08	0.01	0.11	0.27	0.62	0	0	1388	1243	0.073	0.51	1.90
AFO2	95	777	4.0e-02	9.1e-08	0.01	0.11	0.55	0.31	0.01	0	1292	1466	0.083	0.50	1.30
AFO3	136	751	2.2e-02	5.2e-08	0.05	0.12	0.62	0.22	0	0	1198	1015	0.045	0.38	0.72
AFO4	12	688	2.9e-02	3.2e-08	0.07	0.32	0.31	0.29	0	0	1114	1024	0.047	0.40	1.20
ALP1	47	1279	3.3e-01	1.8e-06	0.00	0.05	0.50	0.45	0	0	2212	2662	0.230	0.75	1.40
ALP2	127	1433	4.0e-01	7.7e-07	0.01	0.02	0.55	0.28	0	0.15	2315	2526	0.240	0.83	1.10
ALP3	76	1539	5.1e-01	7.6e-07	0.01	0.01	0.59	0.18	0	0.21	2438	2181	0.210	0.86	0.89
ALP4	114	1270	4.2e-01	5.1e-07	0.01	0.01	0.25	0.61	0.02	0.09	1826	1684	0.120	0.64	1.00

Table 3. Soil properties and climatological characteristics of the test catchments. Next to the site identifier (ID) the columns contain average clay, silt, sand and skeleton contents, root zone depth (h_{RZ}), effective field (eFC) and air capacity (AC) therein and corresponding average saturated hydraulic conductivity (Ks). All soil properties are provided by the national soil map of Germany (Hartwich et al., 1995) except of Ks which was estimated based on pedo-transfer functions (Schaap et al., 2001). The climatological variables include four year mean winter and summer temperature sums ($\sum \overline{T_W}$ and $\sum \overline{T_S}$), length of the period of vegetation ($\overline{t_{veg}}$) and number of frostdays ($\overline{n_{FD}}$).

	soil properties								climatological characteristics			
ID	clay (%)	silt (%)	sand (%)	skeleton (%)	h_{RZ} (m)	eFC (mm)	AC (mm)	Ks (m s^{-1})	$\sum \overline{T_W}$ (°C)	$\sum \overline{T_S}$ (°C)	$\overline{t_{veg}}$ (d)	$\overline{n_{FD}}$ (-)
TRI1	42.7	9.3	48.6	1.8	0.68	73.9	47.1	3.6e-05	15941	65040	174	45
TRI2	39.9	9.0	51.7	1.9	0.67	71.0	49.8	3.9e-05	10601	69717	191	45
TRI3	23.3	18.0	58.4	2.5	0.76	94.2	59.2	2.2e-05	9788	67822	188	49
JUR1	39.2	16.8	43.5	3.7	0.26	32.5	37.0	2.8e-05	8209	61808	178	61
BFO1	7.5	9.5	83.2	4.0	0.60	74.0	55.5	3.6e-05	3960	59262	176	79
BFO2	12.4	17.0	71.1	3.6	0.55	74.5	40.8	2.1e-05	3345	58475	176	82
BFO3	14.3	18.7	67.6	3.8	0.58	67.8	36.4	1.5e-05	3685	58760	176	81
MOL1	26.4	58.8	14.4	1.1	0.85	144.8	44.6	1.5e-06	10227	67692	192	48
MOL2	22.3	56.5	21.8	1.4	0.79	146.1	44.3	1.7e-06	9567	67994	192	50
MOL3	22.0	48.5	28.9	1.9	0.87	151.3	55.4	1.5e-06	8485	67990	190	56
MOL4	22.7	57.2	21.0	1.2	0.89	132.5	46.2	1.6e-06	8805	69246	192	53
MOL5	19.0	46.5	34.8	1.9	0.88	138.3	54.3	2.2e-06	8783	68689	190	53
MOL6	21.0	54.0	24.7	1.3	0.88	129.5	47.7	1.9e-06	7423	68895	190	58
MOL7	23.0	66.8	11.2	1.0	0.90	167.1	47.5	1.4e-06	7571	70519	192	55
AFO1	21.0	39.0	40.0	2.0	1.0	153.0	76.0	1.1e-06	8088	56910	176	65
AFO2	19.8	39.5	40.8	2.2	0.74	149.6	57.0	2.3e-06	11966	57158	187	50
AFO3	19.3	33.7	47.1	2.5	0.79	138.4	61.8	2.4e-06	14974	54775	180	50
AFO4	20.8	38.0	41.2	1.8	0.80	142.2	59.5	1.1e-06	10437	61874	182	58
ALP1	15.8	23.7	60.5	2.5	0.56	90.2	38.5	5.5e-06	243	39834	141	109
ALP2	28.2	28.8	44.5	4.2	0.28	42.2	18.8	1.9e-06	-4490	31694	123	129
ALP3	29.4	28.7	42.4	4.4	0.24	35.2	16.0	1.5e-06	-7934	30122	118	141
ALP4	37.5	24.4	39.3	4.3	0.29	41.1	18.7	1.2e-06	-220	37509	135	114

Table 4. Link table that relates the site identifiers (ID) introduced in section 2.1 to the corresponding gauge and stream names. Gauge locations are provided in Gauß-Krüger zone 4 coordinates (GKR and GKH, CRS identifier EPSG:31468).

ID	Gauge	Stream	GKR	GKH
TRI1	Reichenbach (REIB)	Wörnitz	4373327	5449863
TRI2	Binzwangen (BINZ)	Altmühl	4381996	5473002
TRI3	Bechhofen (BECH)	Wieseth	4394270	5447640
JUR1	Holnstein (HOLN)	Unterbürger Laber	4464800	5442860
BFO1	Gartenried (GART)	Murach	4532661	5483477
BFO2	Untereppenried (UEPR)	Ascha	4533425	5477338
BFO3	Tiefenbach (TIEF)	Bayerische Schwarzach	4543360	5477800
MOL1	Roth (ROTR)	Roth	4363140	5360723
MOL2	Fleinhausen (FLEI)	Zusam	4394141	5358887
MOL3	Mering (MERI)	Paar	4424840	5348870
MOL4	Odelzhausen (ODZH)	Glonn	4440860	5353360
MOL5	Appolding (APPO)	Strogen	4498575	5364071
MOL6	Dietelskirchen (DIKI)	Kleine Vils	4525540	5373175
MOL7	Wallersdorf (WALR)	Reißingerbach	4554850	5400160
AFO1	Unterthingau (alt) (UTHI)	Kirnach	4388313	5294058
AFO2	Hörmanshofen (HOER)	Geltnach	4399272	5299593
AFO3	Buchloe (BUCH)	Gennach	4404574	5323974
AFO4	Herrssigmang (HERR)	Kienbach	4438860	5318140
ALP1	Gunzesried (GZRI)	Gunzesrieder Ach	4366798	5266382
ALP2	Reckenberg (RECK)	Ostrach	4373822	5264305
ALP3	Oberstdorf (OBTR)	Trettach	4370128	5255320
ALP4	Oberammergau (OAMM)	Ammer	4429723	5273332