

Hydrological landscape classification: investigating the performance of HAND based landscape classifications in a central European meso-scale catchment

S. Gharari¹, M. Hrachowitz¹, F. Fenicia², and H. H. G. Savenije¹

¹Delft University of Technology, Faculty of Civil Engineering and Geosciences, Water Resources Section, Delft, The Netherlands

²Public Research Center – Gabriel Lippmann, Belvaux, Luxembourg

Received: 26 April 2011 – Published in Hydrol. Earth Syst. Sci. Discuss.: 2 May 2011

Revised: 16 September 2011 – Accepted: 14 October 2011 – Published: 3 November 2011

Abstract. This paper presents a detailed performance and sensitivity analysis of a recently developed hydrological landscape classification method based on dominant runoff mechanisms. Three landscape classes are distinguished: wetland, hillslope and plateau, corresponding to three dominant hydrological regimes: saturation excess overland flow, storage excess sub-surface flow, and deep percolation. Topography, geology and land use hold the key to identifying these landscapes. The height above the nearest drainage (HAND) and the surface slope, which can be easily obtained from a digital elevation model, appear to be the dominant topographical controls for hydrological classification. In this paper several indicators for classification are tested as well as their sensitivity to scale and resolution of observed points (sample size). The best results are obtained by the simple use of HAND and slope. The results obtained compared well with the topographical wetness index. The HAND based landscape classification appears to be an efficient method to “read the landscape” on the basis of which conceptual models can be developed.

1 Introduction

Large scale hydrological behavior is still poorly understood, mainly as a result of the lack of realistically observable variables on the one hand and the complexity of catchment processes on the other. Catchment topography, readily available as digital elevation models (DEM), has the potential to provide important additional information on catchment processes, particularly due to its inherent co-evolution

and diverse feedback processes with hydrology and ecology (Savenije, 2010). A number of previous studies investigated the relationships between topography and hydrological behavior in the attempt to identify hydrologically different functional landscape units and to better characterize model structure, parameter sets as well as metrics of catchment similarity. For example, Winter (2001) classified the catchment into hydrological landscape units (upland, valley side and lowland) exploiting the combination of topographic, geological and climatic conditions. Based on this concept Wolock et al. (2004) classified hydrological units for the entire United States of America using GIS data. Topography, land use and geology have also been used to directly infer dominant runoff processes within a catchment (Flügel, 1995; Naef et al., 2002; Schmockler-Fackel et al., 2007; Hellebrand and van den Bos, 2008; Müller et al., 2009).

Another widely used indicator is the topographical wetness index (Beven and Kirkby, 1979) which is the basis of TOPMODEL and characterizes hydrological behavior based on upslope contributing area and local slope. The topographical wetness index was further modified by Hjerdt et al. (2004), who took into account downstream conditions considering how far a water particle needs to move to lose a certain amount of potential energy. Topography was also used to investigate the relationship of catchment transit times with numerous catchment characteristics such as flow path length, gradient and connectivity (McGuire et al., 2005; Jencso et al., 2009, 2010) or drainage density (Hrachowitz et al., 2009, 2010) using tracer techniques. Other tracer studies directly linked topography and hydrological behavior (Uhlenbrook et al., 2004; Tetzlaff et al., 2007). A wide range of additional topographical indices have been suggested, describing, among other aspects, the shape, age and stability of a catchment, such as the hypsometric integral (Ritter et al., 2002)



Correspondence to: S. Gharari
(s.gharari@tudelft.nl)

and its correlation with catchment processes (Singh et al., 2008). Other studies correlated topographical indices with soil type and hydrological behavior (Park and van de Giesen, 2004; Lin and Zhou, 2008; Pelletier and Rasmussen, 2009; Behrens et al., 2010; Detty and McGuire, 2010).

In spite of the rich information content of topography, its general usefulness for hydrology is controversial. It has been argued that climate and geology exert stronger influence on the rainfall runoff behavior of a catchment than topography (Devito et al., 2005). Furthermore, it was shown that flow patterns may be dominated by bedrock- rather than surface topography (McDonnell et al., 1996; Tromp-van Meerveld and McDonnell, 2006). According to McDonnell (2003) the “catchment hydrologist will need to develop hypotheses from non-linear theory that are testable on the basis of observations in nature. This will not come about via model intercomparison studies or DEM analysis”. These comments highlight the perception that DEM analysis alone may be of limited value for gaining deeper understanding of catchment processes and that this needs to be brought into a wider context, accounting for the subtle interplay of topography, geology, climate, ecology and hydrology.

In spite of the complexity of catchment processes and due to the frequent lack of data for bottom-up modeling approaches, relatively simple, lumped conceptual models can, due to the self-organizing nature of catchments, be efficient in identifying dominant flow generation processes and modeling stream flow (cf. Sivapalan et al., 2003; Savenije, 2010). However, even for these top-down models additional data, other than precipitation and stream flow, are desirable for enhancing physical significance of model parameters and evaluation (Nalbantis et al., 2011).

Recently, Rennó et al. (2008) formalized the Height Above the Nearest Drainage (HAND) metric and employed it for landscape classification. This metric may be more adequate to identify hydrologically different landscape units than the traditionally used elevation above mean sea level. HAND calculates the elevation of each point in the catchment above the nearest stream it drains to, following the flow direction. It thus extracts from the relatively uninformative topographic elevation the far more informative “hydrologic” elevation, thereby increasing the hydrologic information content of elevation data (Nobre et al., 2011). Nobre et al. (2011) showed that HAND is a stronger topographical descriptor than height above sea level by analyzing long term piezometer data (groundwater behavior). Based on hydrologically meaningful landscape analysis (Nobre et al., 2011), Savenije (2010) suggested that as topographical features are frequently linked to distinct hydrological functioning, they can be used to construct a conceptual catchment model perceived of hydrological units within a catchment.

Landscape classification based on HAND is potentially sensitive to different aspects, such as the definition of the threshold for channel initiation when deriving streams from a DEM, the seasonal fluctuations of the channel initiation,

and the resolution of the DEM. Furthermore, it is unknown to what extent local landscape features can introduce a bias and how robust HAND is to the resolution of observed points (sample size) and the locations of the observed calibration points. Hence, the application of HAND is still subject to considerable uncertainties. In addition, it is not well-understood how HAND relates to other landscape descriptors, such as the topographical wetness index.

The objectives of this paper are thus to (1) assess different hydrologically meaningful landscape classification tools based on the HAND metric and further parameters such as slope and the distance to the nearest drainage, (2) test the sensitivity of HAND-based landscape classification to small scale landscape features and resolutions of the DEM, (3) evaluate the effect of the sample size of the calibration data set on the robustness of HAND-based landscape classification and to (4) analyze the relation of HAND to the topographical wetness index in a mesoscale catchment in a temperate climate.

2 Study catchment

The study catchment is the Wark Catchment in the Grand Duchy of Luxembourg (Fig. 1); the catchment has an area of 82 km² with the catchment outlet (49.85° N, 6.10° E) located downstream of the town of Ettelbrück at the confluence with the Alzette River. With average annual precipitation of 850 mm yr⁻¹ and average annual potential evaporation of 650 mm yr⁻¹ the annual runoff is approximately 250 mm yr⁻¹. The geology in the northern part is dominated by schist while the southern part of the catchment is mostly underlain by sandstone and conglomerate. The dominant land uses are forest on hillslopes, agricultural land on plateaus and pastures in the valley bottoms. The elevation varies between 195 to 532 m with an average of 380 m a.s.l. (above sea level). The slope of the catchment varies between 0–200 %, with an average value of 17 %.

3 Methods

3.1 Terms

The HAND-based hydrologic landscape classification in this paper distinguishes three hydrologically, ecologically and morphologically different landscape units, which, in the following, will be referred to as wetland, hillslope and plateau (cf. Rennó et al., 2008; Nobre et al., 2011). The use of these terms might seem inconsistent as they originate from different disciplines – ecology (wetland), hydrology (hillslope) and morphology (plateau) – where they do have clear definitions. These terms were nevertheless deliberately chosen as they highlight distinct hydrological landscapes with different rainfall-runoff behavior (cf. Savenije, 2010). Note that in other physio-climatic regions more or different landscape

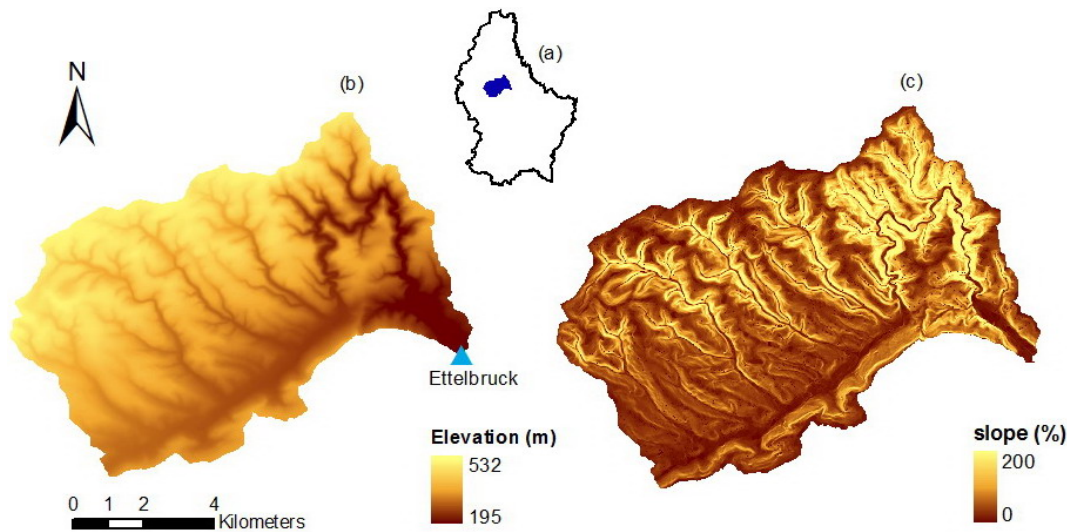


Fig. 1. (a) Location of the Wark Catchment in the Grand Duchy of Luxembourg, (b) digital elevation model (DEM) of the Wark Catchment with cell size of $5\text{ m} \times 5\text{ m}$ (m), (c) slope of the Wark Catchment with DEM resolution of $5\text{ m} \times 5\text{ m}$ (%).

categories may be necessary to adequately describe the landscape. The classification with proposed dominant runoff process in this paper is limited to the Wark catchment and may not be valid for other catchments. The terminology used in this paper is defined as follows:

Wetlands (W), are areas in which the groundwater level is expected to be high relative to the other two landscapes entities. In classical ecological terms they refer to the land where saturation with water is a dominant factor influencing the animal and plant species of that area (Cowardin et al., 1979). From a hydrological point of view, wetlands comprise a broader type of landscape units than the commonly used terms: riparian zones or valley bottom areas. They can be seen as areas which, due to the shallow depth of the water table, have limited residual storage capacity and therefore demonstrate a fast response to precipitation, independent from their location in the catchment. The term shallow in this regard means that in a normal wet season the groundwater table reaches the surface during heavy rainfall events. The predominant locations of wetlands, however, require a subdivision of this class into (a) flat wetlands (W_f), which are characterized by modest slopes, such as stream source areas and valley bottoms (b) sloped wetlands (W_s) in hollows close to streams where hillslopes end in valley bottoms or steep headwater regions, but which can nevertheless be characterized by considerably sloped terrain along the flow direction of the stream. Thus, while both wetland types exhibit relatively low HAND, they are distinguished by different slope angles. The dominant flow generation process for wetlands is saturation overland flow.

Hillslopes (H) are areas which connect concave and convex landscapes (Chorley et al., 1984). The widespread perception that floods are mainly generated on hillslopes (cf.

Beven, 2010) makes them a crucial element in landscape analysis. The co-evolution of ecology and hydrology, and thus the presence of preferential flow paths (Weiler and McDonnell, 2004), such as root canals, animal burrows, fissures and cracks, makes rapid subsurface flow the most effective and dominant runoff process of hillslopes as it fulfills the two functions essential for developing and maintaining their topographical appearance, i.e. drainage and moisture retention (Savenije, 2010).

Plateaus (P) are flat or undulating landscape units relatively high above streams. Due to the low gradients and comparably deep groundwater levels, plateaus mainly fulfill storage (both soil and surface) and evaporation functions, with mainly vertical flow processes, in particular deep percolation (Savenije, 2010).

3.2 Data

Landscape classification in the Wark catchment is based on a $5\text{ m} \times 5\text{ m}$ DEM with a vertical resolution of 0.01 m (Fig. 1). The flow direction network has been derived from the DEM using a D8 algorithm (O'Callaghan and Mark, 1984; Jenson and Domingue, 1988). Although HAND is critically sensitive to the stream initiation threshold, the threshold upslope contributing area has been fixed at a value of 10 ha . This value has been selected to maintain a close correspondence between the derived stream network and the mapped stream network. The value is also in the range of stream initiation thresholds reported by others (e.g. Montgomery and Dietrich, 1988). The relative height, i.e. HAND, was then calculated from the elevation of each raster cell above nearest grid cell flagged as stream cell following the flow direction. Similarly, distance to the nearest drainage was also computed

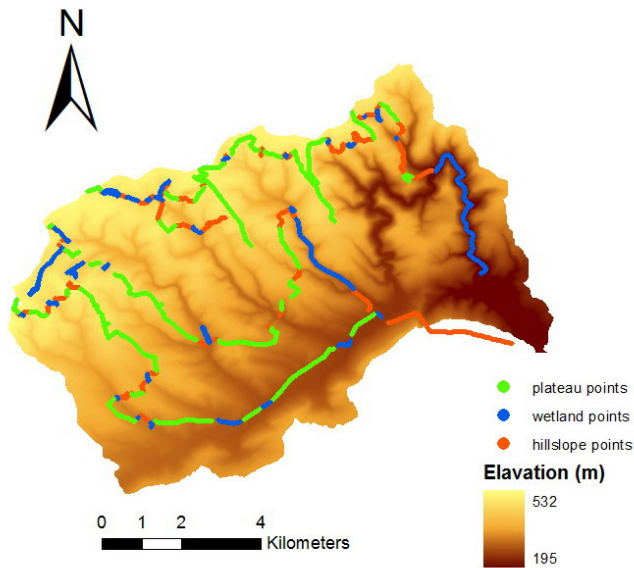


Fig. 2. Observation points in the Wark Catchment and their corresponding landscape classes.

along the flow path to the nearest stream cell. The slope of each grid cell was calculated using the average maximum technique (Burrough and McDonnell, 1998).

During a field campaign (16–20 November 2010), 5611 points in the catchment, hereafter referred to as sampling points, were mapped using GPS waypoints along various paths throughout the catchment and in-situ visually classified into the three landscape units – wetland, hillslope and plateau – in order to establish a “ground truth” according to expert knowledge of hydrological dominant behavior (Fig. 2). The resolution of the observed points is 5 m along the walking paths. The points were collected walking along downhill to uphill transects. The reason for this strategy was to have continuous HAND values from the lowest, near river, to the highest on the plateau. Transects were selected in different parts of the catchment: in the headwaters, in steep valley bottoms and subtle sloped areas in the southern part. Vegetation was also helpful to indicate each class, for example in the Wark catchment the valley bottoms are covered with grass, hillslopes are covered with forest and plateaus are mostly used for agriculture.

3.3 HAND-based landscape classification

The landscape units have been classified according to HAND (H), slope (S) and distance to the nearest drainage (D). A cell with a steep slope was classified as hillslope or sloped wetland and a cell with a low slope was classified either as flat wetland or plateau, depending on HAND or distance to the nearest drainage (Table 1). To separate HAND, slope and distance to the nearest drainage into high or low categories, thresholds had to be introduced. The

Table 1. Criteria for land classification using HAND and slope.

	Low HAND	High HAND
Low Slope	Wetland (flat)	Plateau
High Slope	Wetland (sloped)	hillslope

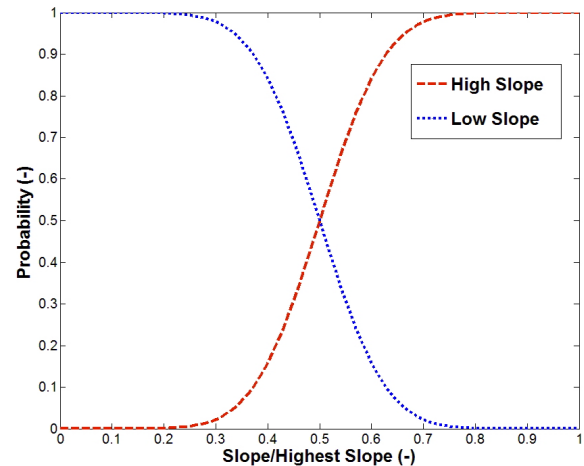


Fig. 3. An example of fuzzy classification for high and low slope. In the central part of the graph the classification is uncertain while at the extremes the uncertainty is low.

thresholds were adjusted in a way that the modeled landscape classes corresponded sufficiently well with the landscape classes of the observed sampling points.

In reality, the boundary between the different landscape units may not be sharply defined. The transition from one landscape category to another may have to be determined by fuzzy thresholds. The fact that transition is not sudden, reflects, similar to fuzzy set theory, the modeler’s and observer’s “degree of belief” (cf. Bárdossy et al., 1990) that a point belongs to a certain landscape unit, as shown in the illustrative example in Fig. 3. In the present case this results in three different percentages for one cell indicating to what extent a cell belongs to a landscape unit. Here, the fuzzy nature of the parameters is considered by using a two parameter cumulative Gaussian distribution function (F):

$$F(x|\mu, \sigma) = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{x - \mu}{\sqrt{2} \sigma} \right) \right] \quad (1)$$

where μ is the mean (i.e. HAND, slope, distance to the nearest drainage) and σ is standard deviation. Both parameters μ and σ are introduced as free calibration parameters in the landscape classification model. Note, that for very small standard deviations the model can be considered as “crisp” or deterministic.

Three landscape classification models have been tested using HAND, slope, distance to the nearest drainage and a combination of them:

1. The first classification model, based on HAND and slope (Model ID: M_{SH} (Rennó et al., 2008)) uses the four fuzzy threshold parameters μ_H , σ_H , μ_S and σ_S . The classification rules for the models are as below:

The probability (Q) of having high values of slope:

$$Q = F(S|\mu_S, \sigma_S). \quad (2)$$

The probability (R) of having high values of HAND:

$$R = F(H|\mu_H, \sigma_H). \quad (3)$$

Thus, the probability (P_H) of being hillslope is the same as the probability of high slope values and high HAND values:

$$P_H = QR. \quad (4)$$

Likewise, the probability (P_P) of being a plateau is defined as the probability of high HAND values and low values for slope:

$$P_P = (1 - Q)R. \quad (5)$$

Similarly the probability (P_W) of being wetland is defined as the probability of low HAND values and high (sloped wetland) or low (flat wetland) slope values.

$$P_W = Q(1 - R) + (1 - Q)(1 - R) = (1 - R) \quad (6)$$

where the first term reflects the probability of being sloped wetland P_{W_s} and the second the probability of flat wetland P_{W_f} .

2. The second model defines Q and R based on slope (μ_S , σ_S) and distance (μ_D , σ_D) to the nearest drainage (Model ID: M_{SD}) to classify the landscape:

The probability (Q) of having high values of slope:

$$Q = F(S|\mu_S, \sigma_S). \quad (7)$$

The probability (R) of having high distance to the nearest drainage:

$$R = F(D|\mu_D, \sigma_D). \quad (8)$$

3. For the third landscape classification model a combination of HAND and distance to the nearest drainage is used. HAND is normalized (H^*) to range from 0 to 1 by dividing the HAND value of each grid cell by the maximum HAND value. The same is done for distance to the nearest drainage (D^*). The multiplication of these two

matrices results in a third matrix, the I_{HD} index (Model ID: M_{SHD}). The values for this new matrix are low and the distribution is highly skewed with more than 86 % of the raster cell showing a value below 0.1. In order to normalize this raster it has been power transformed with an exponent between 0 and 1. The procedure is as below briefly:

$$H^* = \frac{H}{H_{max}} \quad (9)$$

$$D^* = \frac{D}{D_{max}} \quad (10)$$

$$(H^* D^*)^{0 < t < 1} = I_{HD}. \quad (11)$$

The power of the matrix has been chosen by manual calibration and kept constant at $t=0.1$. I_{HD} is a new matrix and its value is multiplication of normalized HAND values (H^*) and distance values (D^*) which was power transformed.

The probability (Q) of having high values of slope is thus:

$$Q = F(S|\mu_S, \sigma_S). \quad (12)$$

The probability (R) of having high I_{HD} can be written as:

$$R = F(I_{HD}|\mu_{I_{HD}}, \sigma_{I_{HD}}). \quad (13)$$

3.4 Model calibration

The model calibration procedure has been designed to minimize a set of objective functions, set-up as an objective matrix. The objective matrix is divided into two parts. The first part consists of the coordinates and landscape classes of the observed sample points and the second part of the modeled landscape classes for the respective points. The objective function allows to evaluate the goodness of fit based on the probability that a modeled point belongs to the same class as the respective observed point and is defined as:

$$O = \left[1 - \frac{\sum_{i=1}^{N_H} P_{H,i}}{N_H} \right] + \left[1 - \frac{\sum_{i=1}^{N_P} P_{P,i}}{N_P} \right] + \left[1 - \frac{\sum_{i=1}^{N_W} P_{W,i}}{N_W} \right] \quad (14)$$

where P_{Hi} , P_{Pi} and P_{Wi} are the probabilities of observed hillslope, plateau and wetland grid cells i , to be classified by the model as hillslope, plateau or wetland respectively. N_H , N_P and N_W are the numbers of observed grid cells for the hillslope, plateau and wetland classes.

For crisp models the probability of a certain point is 1 for one class and 0 for the two other classes. For fuzzy models the probability of a modeled point is divided into three classes summing up to unity, while leaving the observed sample points crisp, i.e. the observed points are clearly defined as wetland, hillslope or plateau. The idea behind this approach is to let the model decide about the hydrological behavior of a cell which may not be unique given the large number of crisp observed points. The fact that the objective function is made up of three parts helps to calibrate the model based on a normalized value where all the classes participate equally; even when the proportion of one landscape class is large compared to others, and the in-situ observed points are not the same number for all classes. Note, that although the maximum value of the objective function can be 3, in practice it will not exceed 2 because with an extremely unrealistic set of parameters the entire basin will be classified as one unit. As a result, the objective function for that unit will be zero and for the remaining classes will each sum up to unity.

Calibration of the models in this paper has been done using Monte-Carlo sampling, i.e. the parameters were sampled, in absence of further prior information, from uniform distributions, within predefined threshold ranges ($S(-) \in [0, 0.2]$, $H(m) \in [0, 20]$, $D(m) \in [0, 100]$, $I_{HD}(-) \in [0, 1]$) in 20 000 Monte-Carlo realizations. Similar to the idea behind the Generalized Likelihood Uncertainty Estimation (GLUE; Beven and Binley, 1992), it has been assumed that there is, due to equifinality, no single best model parameter set. A range of acceptable (i.e. behavioral) sample rates (ASR; Li et al., 2010) is tested in sensitivity analysis. The parameters are reported based on the best performance and their likelihood weighted 95 % (2.5–97.5 %) uncertainty interval (UI) for an ASR of 5 %, whereby the value of the objective function is used as likelihood measure. For sensitivity analysis ASR between 1–10 % was used.

3.5 Optimal effective DEM window size and resolution

The 5 m × 5 m resolution of the DEM allowed a relatively accurate representation of the catchment topography in detail. However, high resolution DEMs can introduce a bias in the results as hydrologically negligible local landscape features, such as steep, small scale rock outcrops, can cause certain grid cells to be inappropriately classified. To reduce this problem the DEM has been smoothed using a Gaussian filter with mean μ_{SM} and standard deviation σ_{SM} (hereafter referred to as characteristic smoothing scale). Applied as a moving window with $\mu_{SM}=0$ and different values of $\sigma_{SM}=0.5, 1, 1.5, 2, 5, 10$ grid cells (equivalent to 2.5, 5, 7.5, 10, 25, 50 m) this allowed the removal of “noise” in the landscape while keeping the high DEM resolution by assigning each grid cell the Gaussian weighed elevation of the neighboring cells (truncated at a radius of $3\sigma_{SM}$) (cf. Hrachowitz and Weiler, 2011). The optimal effective window size, which

is two times the radius of $3\sigma_{SM}$, is the one minimizing the objective function.

Furthermore the effect of lower DEM resolutions (10, 20, 50 or 100 m) on the model parameters and performances has been investigated, to test which DEM resolution is necessary to provide acceptable model results. For model runs with lower resolution no filter was used as it was assumed that local landscape features would automatically average out in the process of resampling the DEM at lower resolutions.

3.6 Sensitivity to calibration point sample size

The effect of different calibration point sample sizes on the robustness and predictive power of the models was assessed by cross-validation. More specifically, repeated random sub-sampling validation (Vapnik, 1998) was used to investigate how best fit parameter sets change for calibration point sample sizes of 2806 (s_1), 1122 (s_2), 561 (s_3), 281 (s_4), 112 (s_5), 56 (s_6) and 28 (s_7) points (i.e. 50, 20, 10, 5, 2, 1 and 0.5 % of the complete of 5611 calibration points which consisted of 1501 (26.8 %), 1385 (24.6 %), 2725 (48.6 %) points for wetland, hillslopes and plateau respectively). 100 random sub-samples for each of the sample sizes s_1 – s_7 were drawn from the complete set of 5611 calibration points. The best parameter set for each of the 100 sub-samples was then estimated by 500 Monte-Carlo realizations. Thus, a central parameter estimate together with a spread around that central value was obtained from the 100 sub-samples for each of the samples sizes s_1 – s_7 . The objective function for the remaining 5611– s_1 – s_7 points not used for calibration (validation points) was then predicted using the 100 individual parameter sets. The mean and spread of the deviation of the validation point objective function from the calibration point objective function was used as an indicator for the predictive power of models with different calibration point sample sizes, i.e. the closer the validation objective function is to the calibration objective function the higher is the predictive power of the models at a given calibration sample size. Likewise, the robustness of the models was further assessed by relating the 100 central parameter estimates and their spreads to the respective sample sizes, i.e. the higher the spread in the parameter estimates, the less robust or the more sensitive the model is to the chosen calibration points, indicating a too small calibration point sample size.

3.7 Sensitivity to the location of calibration points

As the topography of the Wark catchment sharply changes from undulating hills in the western part to plateaus above steep, incised valleys in the eastern part (Fig. 1), this allowed assessing the robustness of the landscape classification models to changing landscape structures. That is, the ability of the model to correctly predict landscape classes when it was calibrated in a structurally different landscape. Here this was done by splitting the Wark Catchment into four zones; North,

East, West and South, by using mean latitude and longitude (the mean of maximum and minimum of latitude within the catchment and the same procedure for longitude). While the eastern part of the catchment has very pronounced landscape features with sharp hillslopes and narrow valleys, the western part is characterized by a comparably subdued profile with wider valleys. The models were subsequently calibrated using observed points from one zone, while the observed points in the remaining zones were predicted. The changes in objective functions and parameter sets were then used as indicators of the model sensitivity to changing landscapes.

3.8 Comparison between the topographical wetness index and different landscape classes

As mentioned above the land classification aims at categorizing the catchment into hydrologically similar zones. For this study the land classification was based on visual observation. In reality it is expected that the position of the groundwater table can provide a more objective selection criterion as the groundwater for wetlands can be assumed to be shallower than the groundwater for plateaus and hillslopes. To see how well the model predicts the likely position of the groundwater table, hereafter referred to as indicator of “wetness” of each landscape, the models and their result were compared to the Topographical Wetness Index (I_{TW}), which is the base for TOPMODEL (Beven and Kirkby, 1979). The I_{TW} is defined as follows:

$$I_{TW} = \ln(A/\tan\beta) \quad (15)$$

where A is the upstream contributing area and β is local slope. The principle behind TOPMODEL is that locations with similar wetness indices are considered to have similar hydrological behavior.

4 Results and discussion

4.1 Comparing the performance of different models for the original DEM

In order to identify the most adequate landscape classification model, the three models (M_{SH} , M_{SD} , M_{SHD}) were run with the original $5\text{ m} \times 5\text{ m}$ DEM. Model M_{SH} , which is equivalent to the original HAND-based model (Rennó et al., 2008), is found to be the most adequate model with an objective function (Eq. 14) value of $O = 0.527$, while the objective function values for the M_{SD} and M_{SHD} models are moderately higher with values of 0.702 and 0.584, respectively. For the model M_{SH} the best fit threshold values for slope (S) and HAND (H) are found to be $\mu_S = 0.129$ (95 % UI: 0.096–0.166) with $\sigma_S = 0.002$ (95 % UI: 0.001–0.039) and $\mu_H = 5.9\text{ m}$ (95 % UI: 3.2–8.9 m) with $\sigma_H = 0.23\text{ m}$ (95 % UI: 0.05–2.9 m). Correspondingly, for model M_{SD} the threshold values for slope (S) and distance to the nearest drainage (D)

are $\mu_S = 0.127$ (95 % UI: 0.102–0.150) with $\sigma_S = 0.001$ (95 % UI: 0–0.026) and $\mu_D = 62.6\text{ m}$ (95 % UI: 42.6–84.5 m) with $\sigma_D = 2.80$ (95 % UI: 0.3–22.5 m). While for M_{SHD} , the slope and the normalized metric of combined H and D (I_{HD}) are, $\mu_S = 0.135$ (95 % UI: 0.092–0.183) with $\sigma_S = 0.004$ (95 % UI: 0–0.044) and $\mu_{I_{HD}} = 0.512$ (95 % UI: 0.454–0.585) with $\sigma_{I_{HD}} = 0$ (95 % UI: 0.001–0.075). Since in M_{SH} the σ values for the Gaussian distribution are very low, these results suggest that all grid cells with $S < 0.129$ and $H < 5.9\text{ m}$ are to be classified as flat wetlands, while grid cells with $S > 0.129$ and $H < 5.9\text{ m}$ are classified as sloped wetlands. Grid cells with $S > 0.129$ and $H > 5.9\text{ m}$ are defined as hillslopes while those with $S < 0.129$ and $H > 5.9\text{ m}$ represent plateaus.

The classified landscapes are illustrated in Fig. 4. The worst performance was obtained with model M_{SD} . This model cannot mimic flat wetland and especially headwater, narrow valley bottoms and wide valleys simultaneously. For headwaters and wide valleys the model needs to use a high distance from the stream to correctly model the observed point, however for narrow valley bottoms the distance should be as little as possible not to overlap with neighboring hillslopes. This causes a poor performance of M_{SD} . The model which used HAND performs the best; it can predict the headwater as well as wide and narrow valley bottoms better than M_{SD} .

One problem which is obvious in Fig. 4 is the noise within a specific landscape. Some raster cells with very high resolution have completely different characteristic from their neighboring cells. For example a cell (which may be a road or other human interference) may have a zero slope and be classified as plateau while its neighboring cells having steep slopes are classified as hillslope.

The relatively low spread for both parameters, HAND and slope, in the M_{SH} highlights that the landscape units can be classified with a surprisingly low fuzziness, i.e. there is only limited uncertainty if a landscape element belongs to one class or to another and it shows that a crisp model with $\sigma_{\text{slope}} = \sigma_{\text{HAND}} = 0$ (Model ID: $M_{SH\text{crisp}}$) would produce results very close, in terms of model performance and parameter estimates, to those from the fuzzy approach.

The results of M_{SH} furthermore suggest that HAND is a better indicator for landscape classification than distance to the nearest drainage or than a combination of distance and HAND, as used in M_{SD} and M_{SHD} models. It shows that additional or similar parameters do not necessarily lead to equally good representations of landscape units as shown in Fig. 4, where several areas of obvious landscape misclassification can be seen, especially for M_{SD} . This underlines the potential of HAND to meaningfully characterize landscapes as it originates, other than elevation, directly from the feedback processes between water and topography and as it is, other than distance to the nearest drainage, directly linked to the dominant driver of storage-discharge relationships, which has co-evolved with the landscape: the hydraulic head.

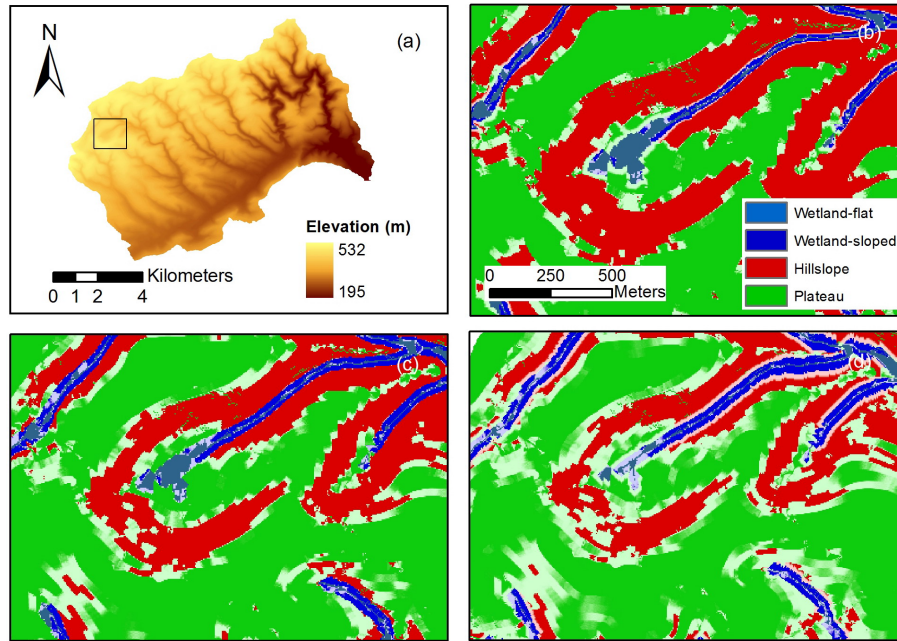


Fig. 4. Comparison between different models for land classification in a headwater of one of the tributaries of the Wark. **(a)** The location of the headwater in the Wark Catchment; **(b)** model using HAND and slope (M_{SH}); **(c)** model using distance to the nearest drainage and slope (M_{SD}); **(d)** model using HAND, distance to the nearest drainage and slope (M_{SHD}). White areas represent areas which are classified in more than one class with high uncertainty.

4.2 Effect of smoothing on model performance and parameters

Relatively prominent, though small scale, landscape features, such as rock outcrops or hollows, can be present in landscapes of any type. However, up to a certain size they do not significantly change the appearance of the overall landscape or the associated dominant runoff process. Thus they should be smoothed out in order to reduce noise in the resulting landscape classification. Here it is found that, with increasing characteristic smoothing scale from $\mu_{SM} = 0$ (original $5\text{ m} \times 5\text{ m}$ DEM) to $\sigma_{SM} = 10\text{--}25\text{ m}$, equivalent to an actual window size of $60\text{--}150\text{ m}$ (truncating the normal distribution at cumulative probabilities of 0.005 and 0.995 or 3σ) the model performance of all three models M_{SH} , M_{SD} and M_{SHD} significantly increased and sharply declined thereafter (Fig. 5). With an objective function value of $O = 0.491$, the crisp (i.e. $\sigma_{\text{slope}} = \sigma_{\text{HAND}} = 0$) $M_{SH\text{crisp}}$ with $\sigma_{SM} = 10\text{ m}$, $H = 4.7\text{ m}$ (95 % UI: 3.5–7.1 m) and $S = 0.113$ (95 % UI: 0.103–0.140) clearly outperformed all other models. The parameters H and S developed contrarily with increasing size of the smoothing window. While H did not show any consistent relationship and a rapid increase for window size of 300 m , S decreased generally with increasing characteristic smoothing scale (Fig. 6a and b). The improved model performance with smoothed landscapes, however, comes at the price of a considerable trade-off with parameter identifiability. As the smoothing implies an assimilation of landscape

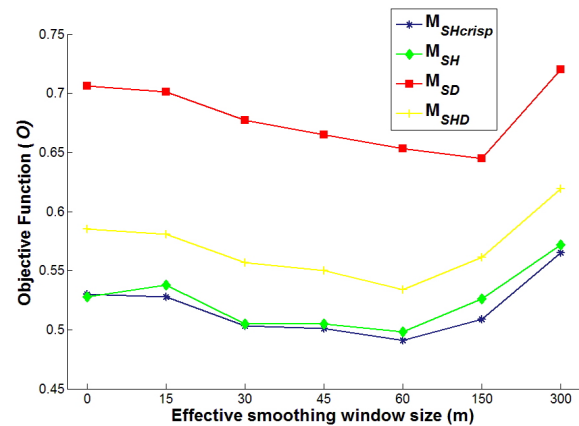


Fig. 5. Performance of different classification models for different effective smoothing window sizes.

features, clear distinctions between them are lost and a wider range of parameter combinations can lead to the same model results. This is shown using the parameter range for different ASR. As the smoothing window size increases the parameter identifiability for H decreases and for the largest window size of 300 m as the ranges decrease for lower value of ASR (Fig. 6c). As the smoothing window size increases, S shows a general decrease up to window size of 150 m before it increases for the largest window size of 300 m (Fig. 6d).

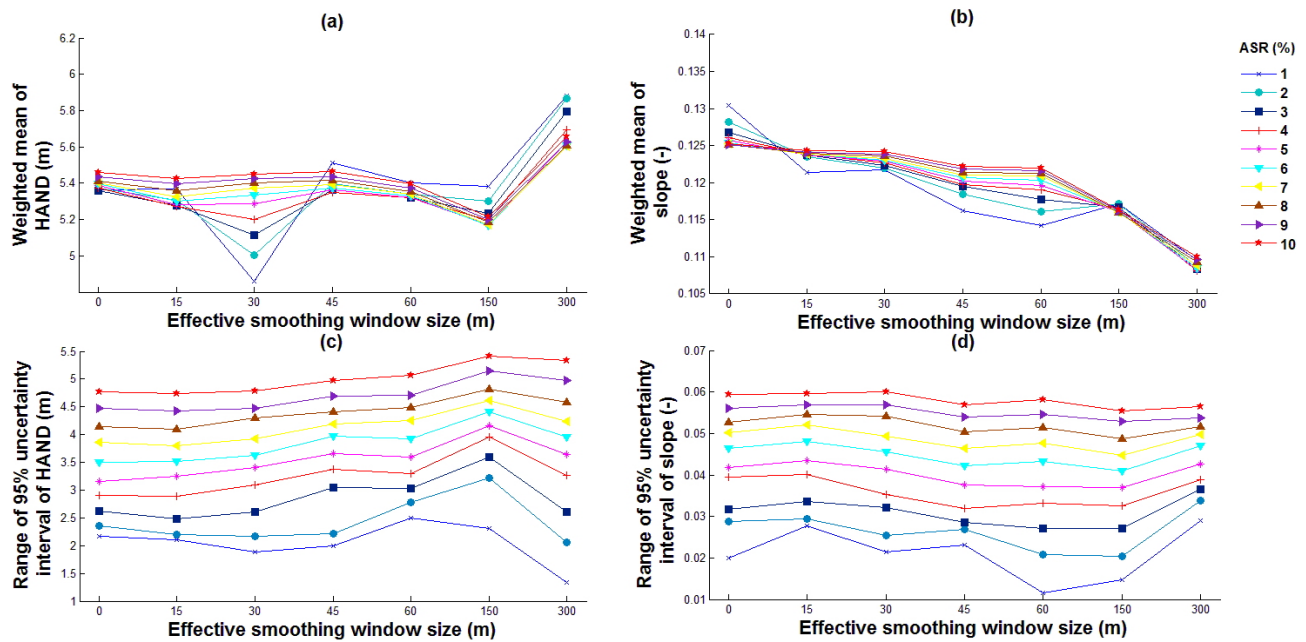


Fig. 6. (a) Likelihood weighted mean values of HAND for the modeling results retained as behavioral vs. different effective smoothing window sizes for acceptable sample rate (ASR) of 1–10%; (b) likelihood weighted mean values of slope for the modeling results retained as behavioral vs. different effective smoothing window sizes for acceptable sample rate (ASR) of 1–10%; (c) parameter range for 95% (2.5–97.5%) uncertainty interval of HAND (d) parameter range for 95% (2.5–97.5%) uncertainty interval for slope. Note that for comparative reasons and due to the inherent subjectivity in the choice of a threshold for defining behavioral parameter sets, the sensitivity of the parameter uncertainty ranges to varying thresholds here is illustrated by showing the parameter ranges for best 1–10% of the acceptable sample rate (ASR).

Using the best model, M_{SH} , and comparing the resulting landscape classes derived from a smoothed DEM ($\sigma_{SM} = 10$ m, Fig. 7c), to the map obtained from the raw, high resolution DEM ($\sigma_{SM} = 0$ m, Fig. 7b) it can be seen that much of the scattered small scale noise and obvious misclassifications disappeared in favor of a more consistent and smooth representation of hydrologically dominant landscape classes. Hillslope and plateau are the two landscapes which are classified more uniformly with less scatter for the larger smoothing window. On the other hand by increasing the size of smoothing window ($\sigma_{SM} = 50$ m, Fig. 7d), the small valley bottoms are smoothed out and classified as plateau.

From these results it can be inferred that the characteristic scale of landscape features that determine landscape classes is in the order of approximately 50 m in this study area and landscape features larger than that do significantly change the appearance of the landscape and its associated dominant runoff processes based on visual observation. However, note that this characteristic landscape feature size should be treated as site specific as it can potentially vary in other regions, where different or additional landscape classes are present.

4.3 Effect of DEM resolution on model performance and parameters

Frequently, only DEMs with resolutions coarser than $5\text{ m} \times 5\text{ m}$ are available. Therefore the robustness and sensitivity of the landscape classification models M_{SH} , M_{SD} and M_{SHD} was assessed with several re-sampled, coarser DEMs, similar to what was done earlier by Zhang and Montgomery (1994), who tested the effect of DEM resolution on topographic wetness index and slope. Again, the crisp model $M_{SHcrisp}$ (i.e. $\sigma_{slope} = \sigma_{HAND} = 0$) is generally the best performing one (Fig. 8) with objective function values between $0.515 < O < 0.993$, depending on the DEM resolution. It is found that the performance first slightly improves up to a resolution of 10 m and deteriorating thereafter, implying that DEMs with resolutions higher than 20 m show sufficient detail to effectively produce results close to those obtained from a $5\text{ m} \times 5\text{ m}$ DEM. DEMs with resolutions lower than 20 m, on the other hand, gradually lose important detail, causing a relatively sharp increase in the objective function. Such low resolutions do not contain sufficient fine detail of the landscape and in particular fail to correctly represent narrow but incised, deep valleys or small head waters, thus introducing considerable error in the HAND as well as in the slope threshold values. For coarse resolution DEMs an additional

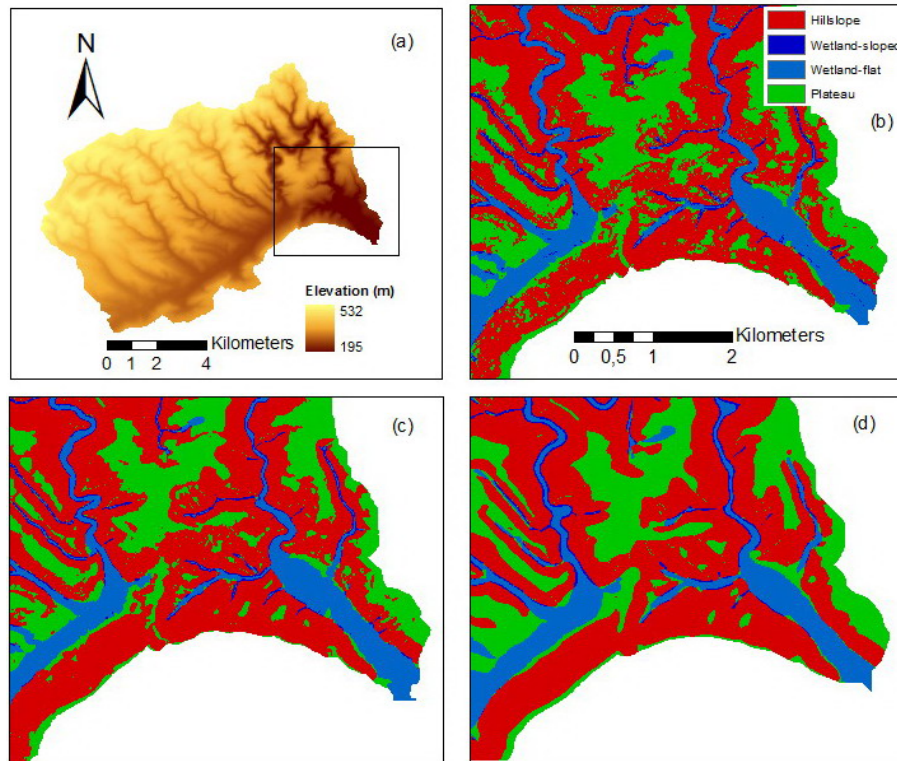


Fig. 7. Smoothing window effect on the land classification for the model M_{SHopt} . (a) Location of selected area in the Wark catchment; (b) classified landscapes for original DEM with resolution of $5\text{ m} \times 5\text{ m}$ without using a smoothing window; (c) classified landscapes using a smoothing window of 60 m ($\sigma = 10\text{ m}$); (d) classified landscapes using smoothing window of 300 m ($\sigma = 50\text{ m}$).

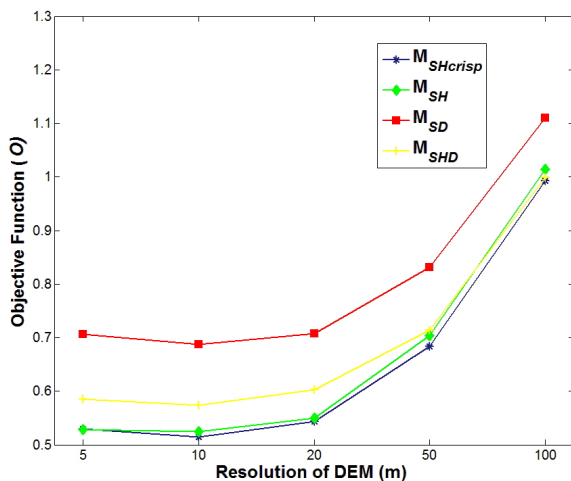


Fig. 8. Performance of different classification models for different DEM resolutions.

source of error was identified. The models were calibrated to observed, clustered sample points, which were, within the clusters, generally located at distances $< 10\text{ m}$ from each other. It was thus possible that observed points next to each other represent different landscape classes. For coarse DEM

resolutions, several classes could thus be contained within a single DEM cell (e.g. $100\text{ m} \times 100\text{ m}$). For this reason it was possible that one cell could represent all three observed classes at the same time which clearly increases the objective function. The parameters H and S developed differently with decreasing DEM resolution. While H shows a slight convex relation to resolution with minimum values for resolution of 10 m , slope shows a strong convex behavior, decreasing to resolution of 50 m and increasing for resolution of 100 m (Fig. 9a and b). The parameter identifiability, however, decreased with decreasing DEM resolution for H , which was not true for parameter S (Fig. 9c and d).

From the analysis in the previous three sections it was found that the most adequate landscape classification in this study could be obtained by the use of the highest resolution DEM ($5\text{ m} \times 5\text{ m}$), smoothed with a Gaussian filter with a characteristic smoothing scale $\sigma_{SM} = 10\text{ m}$, which is equivalent to an effective window size of 60 m , and a crisp (i.e. $\sigma_S = 0$, $\sigma_{HAND} = 0$) model set-up $M_{SHcrisp}$, with landscape classification threshold parameters $H = 4.7\text{ m}$ (95 % UI: $3.5\text{--}7.1\text{ m}$) and $S = 0.113$ (95 % UI: $0.102\text{--}0.140$), which resulted in an objective function value of $O = 0.491$. This model set-up has been used for comparative analysis in the remainder of this study and is hereafter referred to as M_{SHopt} . A summary of the different landscape classes and

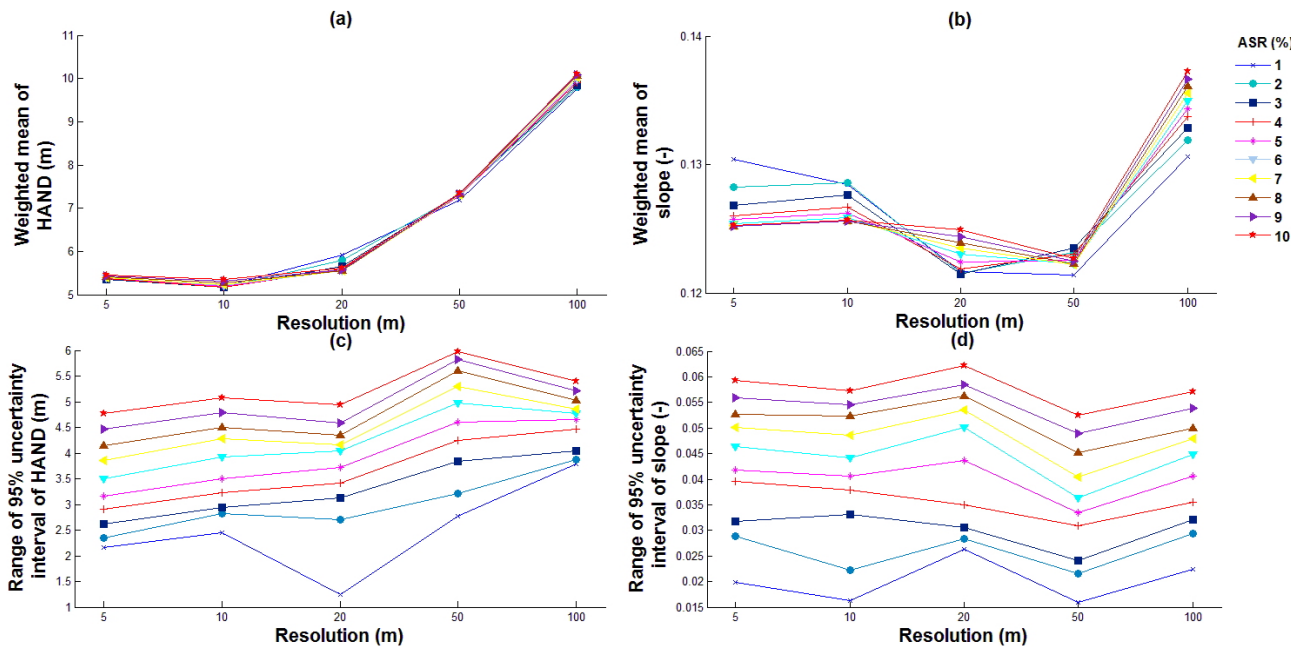


Fig. 9. (a) Likelihood weighted mean values of HAND for the modeling results retained as behavioral vs. different DEM resolutions for an acceptable sample rate (ASR) of 1–10 %; (b) likelihood weighted mean values of slope for the modeling results retained as behavioral vs. different DEM resolutions for an acceptable sample rate (ASR) of 1–10 %; (c) parameter range for 95 % (2.5–97.5 %) uncertainty interval of HAND for different resolutions (d) parameter range for 95 % (2.5–97.5 %) uncertainty interval for slope for different resolutions. Note that for comparative reasons and due to the inherent subjectivity in the choice of a threshold for defining behavioral parameter sets, the sensitivity of the parameter uncertainty ranges to varying thresholds is illustrated by showing the parameter ranges for best 1–10 % of the acceptable sample rate (ASR).

their topographical indicators (slope, HAND and distance to the nearest drainage) for model M_{SHopt} is presented in Table 2. The classified map of the Wark Catchment resulting from M_{SHopt} is illustrated in Fig. 10.

4.4 Sensitivity to calibration point sample size

The landscape classification obtained from the above suggested that model results depend strongly on how the model was calibrated and on the robustness of the parameter estimates. Inadequate calibration strategies could thus cause considerable bias and inaccuracies in the results. The sensitivity of M_{SHopt} model results and threshold parameters to the sample size of calibration points was estimated by comparing the results obtained from different sub-sample sizes (s_1-s_7).

The results are summarized in Fig. 11. While clearly the objective function can occasionally take low values for small sets of calibration points, its mean value decreases with smaller sample size (Fig. 11c). This is, however, largely an effect of the reduced constraints to the model, as can be seen in the pattern of the objective function for the validation points (Fig. 11d). As the calibration sample set is reduced, the performance of the models in validation deteriorates. This is also illustrated in Fig. 11e, where the deviations from

Table 2. Mean and standard deviation values for HAND, slope and distance to the nearest drainage for each landscape class of the best model performance M_{SHopt} (threshold values $H = 4.7$ m, $S = 0.113$).

	HAND (m)		slope (–)		distance to the nearest drainage (m)	
	μ	σ	μ	σ	μ	σ
wetland	2.44	1.83	0.091	0.087	74.99	78.31
wetland (flat)	2.05	1.76	0.047	0.030	88.53	86.13
wetland (sloped)	3.45	1.60	0.206	0.083	39.35	31.54
hillslope	40.96	28.07	0.262	0.144	233.49	152.31
plateau	39.81	26.69	0.066	0.028	410.51	185.00

the line of perfect agreement of the relationship between calibration and validation objective functions are shown, i.e. the higher the deviation the more the objective functions in calibration and validation modes differ. For very robust models with high predictive power only small deviations in validation would be expected. It can also be seen that as the sample size of the calibration points increases, the range of the two

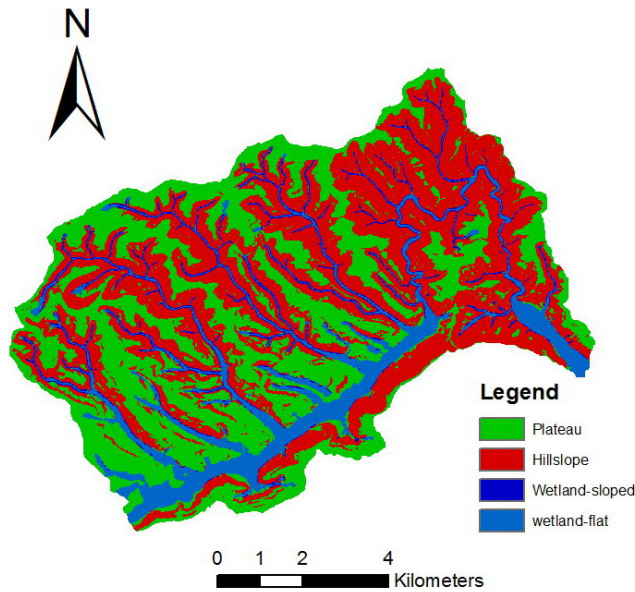


Fig. 10. Classified map of the Wark catchment resulting from the best model performance M_{SHopt} .

landscape classification threshold parameters H and S decreases exponentially (Fig. 11a and b). Hence, as additional calibration data are introduced, model parameters show less uncertainty. The smaller the calibration sample set the less representative it is for the landscape, resulting in considerably different parameter sets for each realization of each different sub-sample of size n (see wider range for parameters as well as objective functions of calibration and validation in Fig. 11a–d). This also implies better parameter identifiability for larger calibration sample sets. In general it can be said that the model stabilizes and remains relatively robust with calibration sample sets of at least $n = 560$ points (in this study approximately 10% of the available sample size), or more specifically with an average calibration sample density of 7 points per km^2 .

4.5 Sensitivity to the location of the calibration points

Subsequently we tested how well the calibration sample set represents the landscape features of the overall catchment in order to produce good and robust model performance. This is considered helpful, as it can potentially give modelers the possibility to a priori assess if there are landscape features with a higher, landscape classification relevant information content than others. This type of information can help to identify areas where it is most useful to collect calibration sample points.

The results of the four analyzed sub-sets of calibration points taken from four parts of the catchment have been compared for M_{SHopt} . These four parts of the catchment are summarized in Table 3 and it can be seen that using calibration points only from the northern and eastern parts results in the

Table 3. Objective function values (O) for calibration and validation in different parts of the Wark Catchment; N, S, W and E represents northern, southern, western and eastern part respectively.

Calibration	Validation
0.407 (N)	0.902 (S)
0.502 (S)	0.607 (N)
0.402 (E)	0.669 (W)
0.565 (W)	0.506 (E)

best model performances in the calibration mode ($O = 0.407$ and 0.402). These two parts are characterized by a very pronounced landscape profile, dominated by steep, incised valleys and narrow valley bottoms. In the remaining two parts, South and West, which are dominated by a more subdued landscape with undulating hills and wide valley bottoms the model performance in calibration mode is less good ($O = 0.520$ and 0.565). Clearly, the distinction between landscape units is more ambiguous in areas with subdued profile, as the transition between different landscape classes, such as wetland and hillslope, is much more subtle, which reduces a two-criterion classification to a single-criterion one. This consequently leads to uncertainties, misclassifications and thus a reduced model performance. However, the results are different in the validation mode. Calibration point subsets from the northern and eastern parts, characterized by a very pronounced profile, do not serve very well for predicting landscape classes in areas with gentle slopes and wide valley bottoms like southern and western parts respectively ($O = 0.902$ and 0.669). This is caused by the models inability to recognize subtle landscape transitions as these were not available for model conditioning. On the other hand, models calibrated to low profile landscapes like the southern and western parts, in spite of a less good calibration performance, show a significantly better performance in predicting landscape classes in areas with different topographical characteristics northern and eastern part respectively ($O = 0.607$ and 0.506). This is not entirely surprising as it may be assumed that a model conditioned to recognize subtle landscape differences will also recognize much clearer differences in the profile. Based on these findings an efficient strategy to choose calibration points would include a few points characterizing pronounced landscape features, such as incised valleys. The majority of sampling points, however, should cover parts of the catchment with subdued topography and rather subtle landscape features, where classification can be most ambiguous.

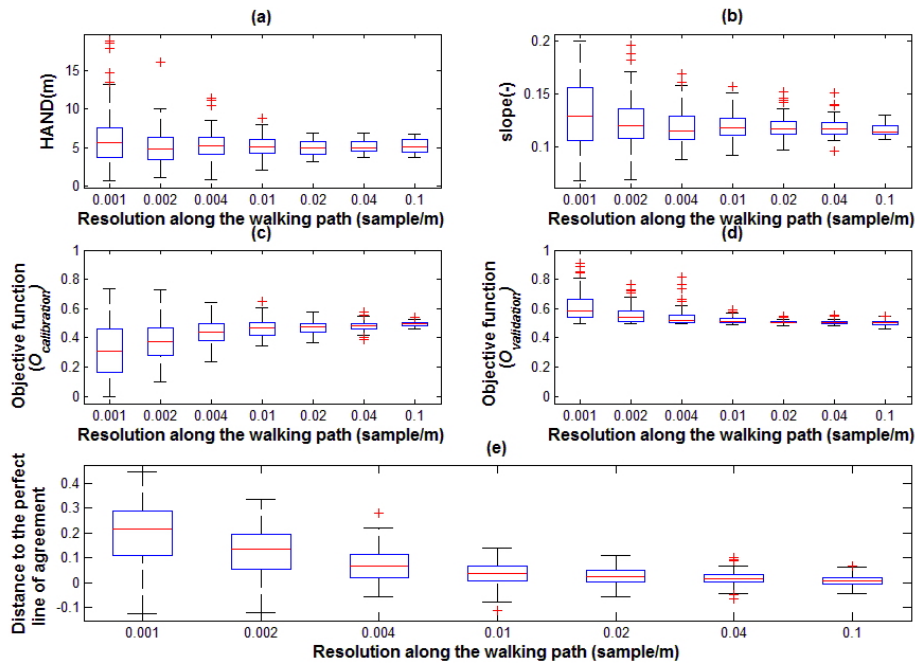


Fig. 11. Behaviour of parameters and objective functions depending on different calibration sample sizes; different sample sizes are shown based on resolution (sample/m) along the walking path (a) the behavior of HAND for different sample sizes (b) behavior of slope for different sample sizes (c) behavior of calibration objective function (O) and (d) behavior of validation objective function (O) (e) distance of validation points to the line of perfect agreement. The distance is positive for points above the line and negative for the points below the line. Whiskers represent the 1.5 times the interquartile range (IQR) and the red crosses show outliers.

4.6 Comparison between the topographical wetness index and different landscape classes

As mentioned earlier the relation with different classified landscapes were compared with the topographical wetness index. To investigate how the wetness index differs for different landscapes, the I_{TW} was calculated for the entire catchment using the smoothed DEM with $\sigma_{SM} = 10$ m in order to allow comparison with the best performing model M_{SHopt} . Nobre et al. (2011) reported a relatively weak, inverse relation between HAND and I_{TW} in the Amazon region, showing that with increasing HAND the wetness index decreases. The scatter in the relationship (cf. Nobre et al., 2011) was reduced by classifying I_{TW} values of all cells in the study catchment into different bins, in order to facilitate clearer interpretation. Different bins or class sizes (5, 10 and 20 classes) show similar behavior regarding the proportions and changes of each landscape for each class.

As the wetness index for each class increases the proportion of plateau and wetland increases and the proportion of hillslope decreases; for the class with highest wetness index the proportion of plateau also decreases and proportion of wetland shows a rapid increase. From Fig. 12 (10 classes) it is clear that by increasing the wetness index the proportion of wetland in each class gradually increases and the proportion of hillslope is gradually less. It can be inferred from Fig. 12 that the locations which were defined as hillslope and

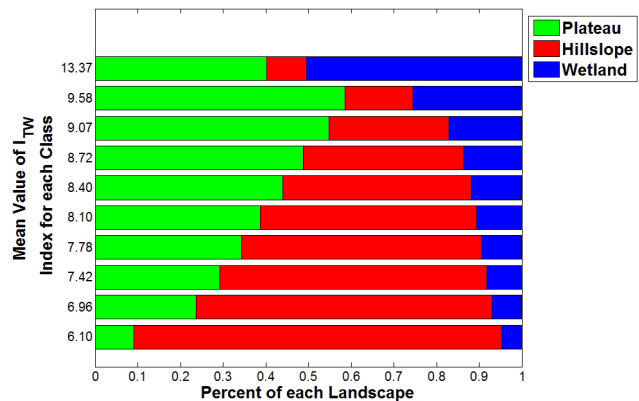


Fig. 12. The analysis for 10 classes of the topographical wetness index (I_{TW}) and landscape component of each class.

wetland are the driest and wettest areas respectively in the Wark catchment according to I_{TW} . The mean value for each class shows that wetland areas have the highest wetness index values (Table 4). Within the wetland class, flat wetlands have a higher wetness index than sloped wetlands.

Hillslopes are the driest classified landscape based on the I_{TW} . Although the aim of the land classification is not to predict the exact depth and behavior of the water table, since many factors play a role in the position of the groundwater

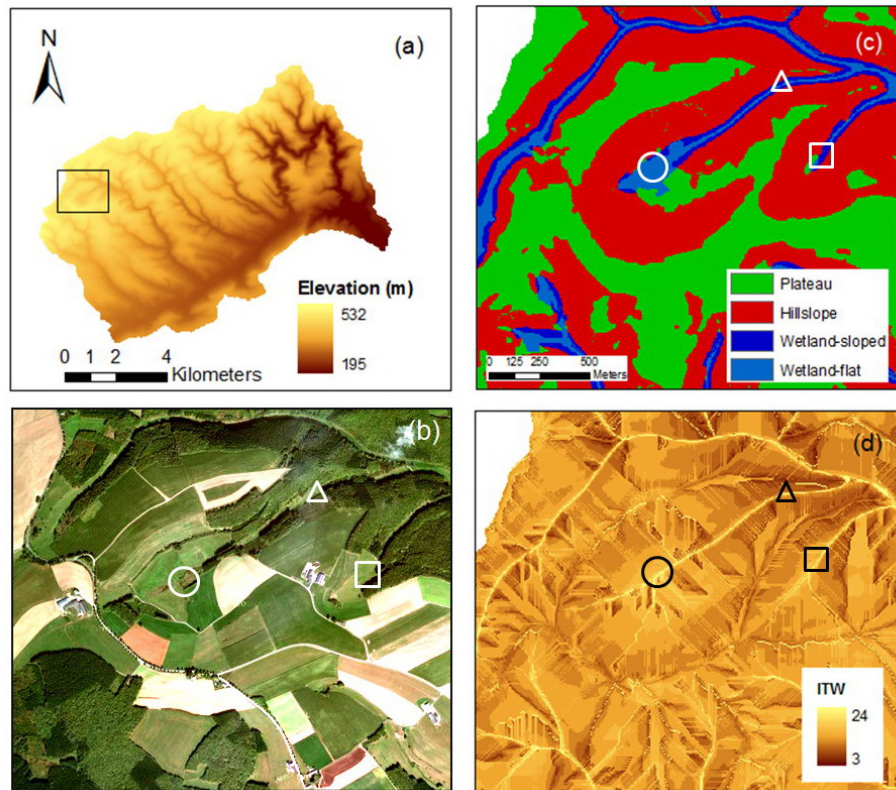


Fig. 13. (a) Location of a selected headwater of the Wark (b) aerial photo of the headwater (c) categorized landscapes using the best model, M_{SHopt} (d) topographical wetness index (I_{TW}). For better orientation, shapes (Δ , \circ , \square) indicate identical positions on the maps (b), (c) and (d).

Table 4. Mean and standard deviation of the topographical wetness index (I_{TW}) for each landscape class

	I_{TW}	
	μ	σ
Wetland	9.9	2.5
Wetland (flat)	10.5	3.6
Wetland (sloped)	8.5	1.6
Plateau	8.3	1.2
Hillslope	7.6	1.2

table such as the recharge, boundary conditions (e.g. Haitjema and Mitchell-Bruker, 2005) and bedrock topography, the classified landscape can potentially give a good estimate of groundwater depth.

One aspect regarding the I_{TW} is that in a GIS a stream in general has a width of one cell size. Adjacent cells with steep slopes and small contributing areas may be much drier than stream cells while in reality it is expected that the wetness of most of the cells close to a stream change gradually and

uniformly (Burrough and McDonnell, 1998). For example, a raster cell with the steepest slope in the catchment area, which is located near a cell flagged as stream and drains to it with a contributing area of one cell (the lowest contributing area possible) will exhibit the lowest I_{TW} compared to the rest of the cells. For the HAND based method this cell, however, will, arguably more realistically, be classified as sloped wetland because of low HAND index and steep slope. A visual comparison between the I_{TW} and classified map and aerial picture is presented in Fig. 13.

5 Conclusions

In this study we tested and assessed the applicability and sensitivity of a HAND based landscape classification framework in a meso-scale headwater catchment in Luxembourg, characterized by a temperate, humid climate. With this approach it was possible to classify landscape units into flat wetland, sloped wetland, hillslopes and plateaus, which are perceived to exhibit distinct dominant runoff generation processes. Three different model types, using different topographical variable combinations, such as HAND, slope, distance to the

nearest drainage and a combined HAND-distance parameter, were investigated. Best landscape classification results were obtained from the model based merely on HAND and slope. This implies that HAND is a stronger indicator for different dominant runoff processes than for example the distance to the nearest drainage or absolute elevation, as HAND links more directly to the hydraulic gradient, arguably the most dominant factor for any type of runoff generation. Based on experiments on sample size and observation density, it was furthermore shown that landscape classes and thus dominant runoff processes are determined by a characteristic landscape feature scale of approximately 50–100 m. Local landscape features smaller than that generally do not influence the overall landscape classification and thereby the dominant runoff processes. Landscape classification based on DEMs with resolutions of 20 m × 20 m and above can give sufficiently accurate results, whereas lower resolution DEMs lack the fine detail necessary to identify critical features, such as narrow valleys. As the landscape classification model needs to be calibrated to observed points, the sensitivity of the calibration point set was analyzed and it was found that sample density of 50 m along walking path can be assumed to be representative in this study, giving robust model results with high predictive power. It was also shown that calibration sample points from subdued landscapes, with subtle and frequently ambiguous transitions between landscape classes contain more information for model calibration than calibration points in clearly defined landscape classes. The classification model was compared with the topographical wetness index and a clear relation between classified landscape and groundwater table based on binned values of the topographical wetness index values was found.

The landscape classification results could in future work be refined by using additional information such as distributed soil moisture or groundwater data, for establishing a yet stronger link between landscape classes and runoff processes.

The resulting maps show a relatively realistic, high accuracy landscape classification presumably associated closely to the dominant runoff generation processes in the individual parts of the study catchment. Such results can in the future serve as basis for the development of conceptual hydrological models by assigning different model structures to the individual landscape classes, thereby potentially improving model realism without the need for further parameters.

Acknowledgements. The authors would like to thank the people who facilitated this research at Gabriel Lippmann Research Institute, especially Laurent Pfister and Lucien Hoffman. We would like to thank Dimitri Solomatine, Sebastian Wrede, Miriam Coenders and Jerome Juilleret for their constructive comments on this work. Special thanks to Bouke Kooreman who did a great job sampling points for the model calibration, walking and running the entire Wark Catchment in heavy rain and cold weather, and to all who created a friendly environment during the research at the Gabriel Lippmann Institute and Delft University of Technology.

The Authors would like to thank the referees, Ioannis Nalbantis, Antonio Donato Nobre and Erwin Zehe for their constructive comments on our work.

Edited by: G. Blöschl

References

- Bárdossy, A., Bogardi, I., and Duckstein, L.: Fuzzy regression in hydrology, *Water Res. Res.*, 26, 1497–1508, doi:10.1029/WR026i007p01497, 1990.
- Behrens, T., Zhu, A.-X., Schmidt, K., and Scholten, T.: Multi-scale digital terrain analysis and feature selection for digital soil mapping, *Geoderma*, 155, 175–185, doi:10.1016/j.geoderma.2009.07.010, 2010.
- Beven, K. J.: Preferential flows and travel time distributions: defining adequate hypothesis tests for hydrological process models, *Hydrol. Process.*, 24, 1537–1547, doi:10.1002/hyp.7718, 2010.
- Beven, K. J. and Binley, A. M.: The future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Process.*, 6, 279–298, 1992.
- Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology, *Hydrolog. Sci. J.*, 24, 43–69, doi:10.1080/02626667909491834, 1979.
- Burrough, P. A. and McDonnell, R. A.: *Principles of Geographical Information Systems*, Oxford University Press, Oxford, 1998.
- Chorley, R. J., Schumm, S. A., and Sugden, D. E.: *Geomorphology*, Methuen & Co. Ltd, 1984.
- Cowardin, L. M., Carter, V., Golet, F., and LaRoe, E.: *Classification of wetlands and deepwater habitats of the United States*, U.S. Fish and Wildlife Service, Washington, DC, 1979.
- Detty, J. M. and McGuire, K. J.: Topographic controls on shallow groundwater dynamics: implications of hydrologic connectivity between hillslopes and riparian zones in a till mantled catchment, *Hydrol. Process.*, 24, 2222–2236, doi:10.1002/hyp.7656, 2010.
- Devito, K., Creed, I., Gan, T., Mendoza, C., Petrone, R., Silins, U., and Smerdon, B.: A framework for broad-scale classification of hydrologic response units on the Boreal Plain: is topography the last thing to consider?, *Hydrol. Process.*, 19, 1705–1714, doi:10.1002/hyp.5881, 2005.
- Flügel, W.-A.: Delineating hydrological response units by geographical information system analyses for regional hydrological modelling using PRMS/MMS in the drainage basin of the River Bröl, Germany, *Hydrol. Process.*, 9, 423–436, doi:10.1002/hyp.3360090313, 1995.
- Haitjema, H. M. and Mitchell-Bruker, S.: Are Water Tables a Subdued Replica of the Topography?, *Ground Water*, 43, 781–786, doi:10.1111/j.1745-6584.2005.00090.x, 2005.
- Hellebrand, H. and van den Bos, R.: Investigating the use of spatial discretization of hydrological processes in conceptual rainfall runoff modelling: a case study for the meso-scale, *Hydrol. Process.*, 22, 2943–2952, doi:10.1002/hyp.6909, 2008.
- Hjerdt, K. N., McDonnell, J. J., Seibert, J., and Rodhe, A.: A new topographic index to quantify downslope controls on local drainage, water resources research, *Water Resour. Res.*, 40, W05602, doi:10.1029/2004WR003130, 2004.

- Hrachowitz, M. and Weiler, M.: Uncertainty of Precipitation Estimates Caused by Sparse Gauging Networks in a Small, Mountainous Watershed, *J. Hydrol. Eng.-ASCE*, 16, 460–471, doi:10.1061/(ASCE)HE.1943-5584.0000331, 2011.
- Hrachowitz, M., Soulsby, C. D. T., Dawson, J. J. C., and Malcolm, I. A.: Regionalization of Transit Time Estimates in montane catchments by integrating landscape controls, *Water Resour. Res.*, 45, W05421, doi:10.1029/2008WR007496, 2009.
- Hrachowitz, M., Soulsby, C., Tetzlaff, D., Malcolm, I. A., and Schoups, G.: Gamma distribution models for transit time estimation in catchments: Physical interpretation of parameters and implications for time-variant transit time assessment, *Water Resour. Res.*, 46, W10536, doi:10.1029/2010WR009148, 2010.
- Jencso, K. G., McGlynn, B. L., Gooseff, M. N., Wondzell, S. M., Bencala, K. E., and Marshall, L. A.: Hydrologic connectivity between landscapes and streams: transferring reach- and plot-scale understanding to the catchment scale, *Water Resour. Res.*, 45, W04428, doi:10.1029/2008WR007225, 2009.
- Jencso, K. G., McGlynn, B. L., Gooseff, M. N., Bencala, K. E., and Wondzell, S. M.: Hillslope hydrologic connectivity controls riparian groundwater turnover: Implications of catchment structure for riparian buffering and stream water sources, *Water Resour. Res.*, 46, W10524, doi:10.1029/2009WR008818, 2010.
- Jenson, S. K. and Domingue, J. O.: Extracting topographic structure from digital elevation data for geographic information system analysis, *Photogramm. Eng. Remote S.* 54, 1593–1600, 1988.
- Li, L., Xia, J., Xu, C. Y., and Singh, V. P.: Evaluation of the subjective factors of the GLUE method and comparison with the formal Bayesian method in uncertainty assessment of hydrological models, *J. Hydrol.*, 390, 210–221, doi:10.1016/j.jhydrol.2010.06.044, 2010.
- Lin, H. and Zhou, X.: Evidence of subsurface preferential flow using soil hydrologic monitoring in the Shale Hills catchment, *Eur. J. Soil Sci.*, 59, 34–49, doi:10.1111/j.1365-2389.2007.00988.x, 2008.
- McDonnell, J. J.: Where does water go when it rains? Moving beyond the variable source area concept of rainfall-runoff response, *Hydrol. Process.*, 17, 1869–1875, doi:10.1002/hyp.5132, 2003.
- McDonnell, J. J., Freer, J., Hooper, R., Kendall, C., Burns, D., Beven, K. J., and Peters, J.: New method developed for studying flow on hillslopes, *EOS, Trans. AGU*, 77, 465–472, 1996.
- 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 Resour. Res.*, 41, W05002, doi:10.1029/2004WR003657, 2005.
- Montgomery, D. R. and Dietrich, W. E.: Where Do Channels Begin?, *Nature*, 336, 232–234, 1988.
- Müller, C., Hellebrand, H., Seeger, M., and Schobel, S.: Identification and regionalization of dominant runoff processes – a GIS-based and a statistical approach, *Hydrol. Earth Syst. Sci.*, 13, 779–792, doi:10.5194/hess-13-779-2009, 2009.
- Naef, F., Scherrer, S., and Weiler, M.: A process based assessment of the potential to reduce flood runoff by land use change, *J. Hydrol.*, 267, 74–79, doi:10.1016/S0022-1694(02)00141-5, 2002.
- Nalbantis, I., Efstratiadis, A., Rozos, E., Kopsiafti, M., and Koutsosoyannis, D.: Holistic versus monomeric strategies for hydrological modelling of human-modified hydrosystems, *Hydrol. Earth Syst. Sci.*, 15, 743–758, doi:10.5194/hess-15-743-2011, 2011.
- Nobre, A. D., Cuartas, L. A., Hodnett, M., Rennó, C. D., Rodrigues, G., Silveira, A., Waterloo, M., and Saleska, S.: Height above the Nearest Drainage, a hydrologically relevant new terrain model, *J. Hydrol.*, 404, 13–29, 2011.
- O’Callaghan, J. F. and Mark, D. M.: The extraction of drainage networks from digital elevation data: *Computer Vision, Graphics and Image Processing*, 28, 323–344, doi:10.1016/S0734-189X(84)80011-0, 1984.
- Park, S. and van de Giesen, N.: Soil-landscape delineation to define spatial sampling domains for hillslope hydrology, *J. Hydrol.*, 295, 28–46, doi:10.1016/j.jhydrol.2004.02.022, 2004.
- Pelletier, J. D. and Rasmussen, C.: Geomorphically based predictive mapping of soil thickness in upland watersheds, *Water Resour. Res.*, 45, W09417, 2009.
- Rennó, C. D., Nobre, A. D., Cuartas, L. A., Soares, J. V., Hodnett, M. G., Tomasella, J., and Waterloo, M. J.: HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rain-forest environments in Amazonia, *Remote Sens. Environ.*, 112, 3469–3481, doi:10.1016/j.rse.2008.03.018, 2008.
- Ritter, D. F., Kochel, R. C., and Miller, J. R.: *Process geomorphology*, McGraw Hill, Boston, 2002.
- Savenije, H. H. G.: *HESS Opinions* “Topography driven conceptual modelling (FLEX-Topo)”, *Hydrol. Earth Syst. Sci.*, 14, 2681–2692, doi:10.5194/hess-14-2681-2010, 2010.
- Schmocker-Fackel, P., Naef, F., and Scherrer, S.: Identifying runoff processes on the plot and catchment scale, *Hydrol. Earth Syst. Sci.*, 11, 891–906, doi:10.5194/hess-11-891-2007, 2007.
- Singh, O., Sarangi, A., and Sharma, C. M.: Hypsometric integral estimation methods and its relevance on erosion status of north-western lesser Himalayan watersheds, *Water Resour. Manage.*, 22, 1545–1560, doi:10.1007/s11269-008-9242-z, 2008.
- Sivapalan, M., Blöschl, G., Zhang, L., and Vertessy, R.: Downward approach to hydrologic prediction, *Hydrol. Process.*, 17, 2101–2111, doi:10.1002/hyp.1425, 2003.
- Tetzlaff, D., Soulsby, C., Waldron, S., Malcolm, I. A., Bacon, P. J., Dunn, S. M., Lilly, A., and Youngson, A. F.: Conceptualization of runoff processes using a geographical information system and tracers in a nested mesoscale catchment, *Hydrol. Process.*, 21, 1289–1307, doi:10.1002/hyp.6309, 2007.
- Tromp-van Meerveld, H. J. and McDonnell, J. J.: Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis, *Water Resour. Res.*, 42, W02411, doi:10.1029/2004WR003800, 2006.
- Uhlenbrook, S., Roser, S., and Tilch, N.: Hydrological process representation at the meso-scale: the potential of a distributed, conceptual catchment model, *J. Hydrol.*, 291, 278–296, doi:10.1016/j.jhydrol.2003.12.038, 2004.
- Vapnik, V.: *Statistical Learning Theory*, John Wiley & Sons, New York, 1998.
- Weiler, M. and McDonnell, J. J.: Virtual experiments: a new approach for improving process conceptualization in hillslope hydrology, *J. Hydrol.*, 285, 3–18, doi:10.1016/S0022-1694(03)00271-3, 2004.
- Winter, T. C.: The Concept OF Hydrologic Landscapes, *J. Am. Water Resour. Assoc.*, 37, 335–349, doi:10.1111/j.1752-1688.2001.tb00973.x, 2001.

Wolock, D. M., Winter, T. C., and McMahon, G.: Delineation and evaluation of hydrologic-landscape regions in the United States using geographic information system tools and multivariate statistical analyses, *Environ. Manage.*, 34, S71–S88, doi:10.1007/s00267-003-5077-9, 2004.

Zhang, W. and Montgomery, D. R.: Digital elevation model grid size, landscape representation and hydrologic simulation, *Water Resour. Res.*, 30, 1019–1028, doi:10.1029/93WR03553, 1994.