EARTH SURFACE PROCESSES AND LANDFORMS
Earth Surf. Process. Landforms (2013)

Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/esp.3506

A comparison of wetness indices for the prediction of observed connected saturated areas under contrasting conditions

Geneviève Ali,[1]* Christian Birkel,[2] Doerthe Tetzlaff,[2] Chris Soulsby,[2] Jeffrey J. McDonnell[2,3] and Paolo Tarolli[4]
[1] Department of Geological Sciences, University of Manitoba, Manitoba, Canada
[2] Northern Rivers Institute, School of Geosciences, University of Aberdeen, Aberdeen, UK
[3] Global Institute for Water Security, University of Saskatchewan, Saskatchewan, Canada
[4] Department of Land, Environment, Agriculture and Forestry, University of Padova, Padova, Italy

Received 4 December 2012; Revised 11 November 2013; Accepted 12 November 2013

*Correspondence to: Geneviève Ali, Department of Geological Sciences, University of Manitoba, Manitoba, Canada. E-mail: Genevieve.Ali@umanitoba.ca.

ESPL
Earth Surface Processes and Landforms

ABSTRACT: For lack of other widely available spatial information, topography is often used to predict water fluxes and water quality in mesoscale watersheds. Such data have however proven to be misleading in many environments where large and flat valley bottoms and/or highly conducive soil covers determine water storage and water transport mechanisms. Also, the focus is generally on the prediction of saturation areas regardless of whether they are connected to the catchment hydrographic network or rather present in isolated topographic depressions. Here soil information was coupled with terrain data towards the targeted prediction of connected saturated areas. The focus was on the 30 km^2 Girnock catchment (Cairngorm Mountains, northeast Scotland) and its 3 km^2 sub-catchment, Bruntland Burn in which seven field surveys were done to capture actual maps of connected saturated areas in both dry and humid conditions. The 1 km^2 resolution UK Hydrology of Soil Types (HOST) classification was used to extract relevant, spatially variable, soil parameters. Results show that connected saturated areas were fairly well predicted by wetness indices but only in wet conditions when they covered more than 30% of the whole catchment area. Geomorphic indices including information on terrain shape, steepness, aspect, soil texture and soil depth showed potential but generally performed poorly. Indices based on soil and topographic data did not have more predictive power than those based on topographic information only: this was attributed to the coarse resolution of the HOST classification. Nevertheless, analyses provided interesting insights into the scale-dependent water storage and transport mechanisms in both study catchments. Copyright © 2013 John Wiley & Sons, Ltd.

KEYWORDS: wetness indices; observed saturation areas; connectivity; stream network; data resolution

Introduction

Spatial patterns of soil water content are key elements towards the understanding of the geographic sources of runoff, nutrients and sediments found in streams and rivers. Recent hydrological studies (Grayson *et al*., 1997; Western *et al*., 2001; McNamara *et al*., 2005; Tromp-Van Meerveld and McDonnell, 2006; James and Roulet, 2007) have therefore been focusing on 'lots of points' sampling strategies in order to collect information about surface saturation areas, shallow soil moisture and water table elevation above soil–bedrock interfaces. The drawbacks associated with such field mappings of soil water content are that they are not only labour-intensive but also costly and uneasily applicable over large study areas (Grabs *et al*., 2009). As a consequence, a wide range of topography-based wetness indices have been developed as plausible indicators of the location of saturation areas. Early wetness indices ensued from the variable source area concept (Cappus, 1960; Dunne and Black, 1970a, 1970b) and rely on the assumption that topography and soil properties are first- and second-order controls on shallow soil water content. Originally, riparian zones at valley bottoms and topographic depressions were thought to be the ones to first reach saturation and act as major runoff contributors (Kirkby, 1975; Beven and Kirkby, 1979; O'Loughlin, 1986). Later studies then showed that locations exhibiting long low-angled hillslopes and low saturated hydraulic conductivity values were the most prone to the formation of variable saturation areas (Ogden and Watts, 2000; Aryal *et al*., 2003; Yair and Raz-Yassif, 2004). Primary topographic attributes such as surface slope, upslope area or curvature have therefore been used individually or in combination towards the prediction of soil water content (Güntner *et al*., 2004). Other factors known to influence soil water content such as soil characteristics or available energy from solar radiation have also been integrated in the formula of certain topographic indices (Beven, 1986; Moore *et al*., 1991; Gómez-Plaza *et al*., 2001; Güntner *et al*., 2004).

The increasing availability of digital elevation models (DEMs) has led to the development of more sophisticated topographic indices, the most common being the topographic wetness index (TWI) developed by Beven and Kirkby (1979) within the rainfall–runoff model TOPMODEL. The TWI is defined as $\ln(a/\tan\beta)$ where a is the specific upslope area (i.e. the upslope area per unit contour length) and tan β is the local surface slope; a indicates the amount of water flowing towards a certain location while the local slope angle β is assumed to reflect subsurface lateral transmissivity. Given its computation from easily accessible data, the TWI has been used for various purposes, for instance to identify sources of subsurface flow (Robson *et al.*, 1992), estimate the hydrological, physical and chemical properties of soils (Western *et al.*, 1999; Seibert *et al.*, 2007), characterize vegetation patterns (Moore *et al.*, 1993) or investigate scaling effects (Sivapalan *et al.*, 1990). Many variants of the TWI exist, and the main differences between them are the flow direction algorithm chosen to route the accumulated upslope area downwards (algorithm variants), or the appreciation of terrain gradient, the representation of stream channels and the inclusion of soil information into the index formulae (conceptual variants). Comprehensive reviews of such variants can be found in Barling *et al.* (1994); Borga *et al.* (2002); Güntner *et al.* (2004); Sorensen *et al.* (2006), and Grabs *et al.* (2009) among others. More complex variants of the TWI include additional information such as soil depth D and saturated hydraulic conductivity K_s so that the soil-topographic wetness index (STWI) can be expressed as $\ln(a/(D \cdot K_s \cdot \tan\beta))$. The logarithm function in the TWI formula that illustrates an exponential decline of the soil transmissivity with depth can also be changed so as to portray linear or parabolic declines (Ambroise *et al.*, 1996a; Duan and Miller, 1997). However, it was previously argued that higher complexity wetness indices do not necessarily lead to better soil water content predictions (Güntner *et al.*, 2004); this calls for a systematic and exhaustive evaluation of wetness indices in a wide variety of physical and hydroclimatic conditions in order to adequately consider the spatiotemporal variability of processes that control soil moisture.

The overall objective of this paper is to add to the body of work previously described by testing the ability of a range of wetness indices to predict the dynamics of connected saturation areas in two nested Scottish catchments. The special contribution of this paper is multifold. Firstly, in contrast to previous studies, we compare a range of wetness indices not against a single, but against seven maps of actual saturation areas observed in the field following various hydroclimatic conditions. The availability of such an information-rich dataset is indeed critical towards assessing when and where different wetness indices can serve as plausible indicators of the location of saturation areas. Secondly, it is worth noting that in earlier work on wetness indices, the focus was on the prediction of saturation areas regardless of whether they were connected to the catchment hydrographic network or rather present in isolated topographic depressions (Franks *et al.*, 1998; Güntner *et al.*, 2004). Here the rationale is that only saturation areas physically connected to the stream and the catchment outlet can be assumed to significantly contribute to streamflow in a landscape where surface flow processes are dominant (Ambroise, 2004). We hypothesize that catchment areas with a high potential to be connected to the catchment outlet should exhibit topographic characteristics which are significantly different from those of areas which hardly ever transmit water to the stream network. Sensitive wetness indices should therefore be able to capture those topographic differences. Thirdly, this study couples terrain data with soil characteristics extracted from the Hydrology of Soil Types (HOST) database (Boorman *et al.*, 1995) in an attempt to improve the prediction of connected saturation areas. Given that the HOST classification was developed to help differentiate between fast and slow water flow paths and understand subsequent implications for flood generation or baseflow maintenance, its potential for hydrological modelling has always been implicit but rarely tested in a spatially distributed framework (Dunn and Lilly, 2001). Building upon previous work reported in the literature, simple topographic attributes and various combinations of flow routing algorithms, slope definitions and soil characteristics are evaluated as wetness indices in this paper. In addition, indices inspired from geomorphologic landscape classifications [the topographic position index (TPI, Jenness, 2006) or the topographic relative moisture index (TRMI, Parker, 1982)] but which have not been tested against maps of observed saturation areas are also evaluated here. Three specific questions guide this comparative work, namely:

1. Does the ability of wetness indices to predict connected saturation areas remain the same through wet and dry periods?
2. Does the inclusion of soil information improve the predictive power of wetness indices?
3. What influence does topographic and soil data (spatial) resolution bear on the prediction of connected saturation areas?

Study Sites and Connected Saturation Area Surveys

Two nested study sites were chosen in the Cairngorms National Park, Scotland: the 30.4 km^2 Girnock Burn and the 3.4 km^2 Bruntland Burn catchments (Figure 1A and 1B). The Bruntland Burn is a tributary of the Girnock Burn which drains into the River Dee. Detailed descriptions of both sites are given elsewhere (Tetzlaff *et al.*, 2007; Birkel *et al.*, 2010). The Bruntland and Girnock catchments have mean altitudes of about 360 m above sea level (a.s.l.) and 407 m a.s.l., respectively (Table I). Annual precipitation is mainly generated by westerly frontal systems and is 1059 mm, with the summer months (May–August) generally being the driest. Snow makes up less than 10% of annual precipitation and melts rapidly below 700 m. As for specific mean daily discharge, it is higher at the outlet of the Bruntland catchment in comparison to the Girnock catchment (Table I). Both catchments showcase typical features of the Scottish Highlands (Birkel *et al.*, 2010), namely a combination of steep and rolling hillslopes and over-widened valley bottoms (Figure 1C). Land-use is dominated by heather moorland (*Calluna vulgaris*), with smaller areas of rough grazing and forestry on the lower hillslopes. Higher areas of the landscape are underlain by granite while lower elevation areas are underlain by schists and other metamorphic rocks. Glacial drift deposits of various thickness and permeability are superimposed on the solid geology.

Previous studies have shown that in both catchments, fast-responding near-surface processes dominate the storm hydrograph (Tetzlaff *et al.*, 2007; Birkel *et al.*, 2010) and that these processes are directly related to the expansion and contraction of soil saturation areas in the vicinity of the stream (Birkel *et al.*, 2010). Those findings suggest that a field conceptualization of saturation area dynamics is crucial to assess how the topology of soils, and not topography alone, determine dominant stormflow processes. For that purpose, the UK HOST system is heavily relied on as it classifies soils according to dominant hydrological processes (Boorman *et al.*, 1995). The most extensive soils at both study sites have large peat content (peaty gleys, deep peats, Figure 1D). Gleys (HOST 14 and 24) and peaty gleys (HOST 15) are saturated for much of the year due to low permeability drift deposits in valley bottoms; they generate substantial amounts of saturation excess overland flow and shallow lateral flow in organic surface horizons (Tetzlaff *et al.*, 2007), especially in the Girnock catchment (Figure 1D). Deeper peats (HOST 27) and rankers (HOST 22) are rather present in the Bruntland catchment (Figure 1D), well connected to the stream channels and are responsible for the

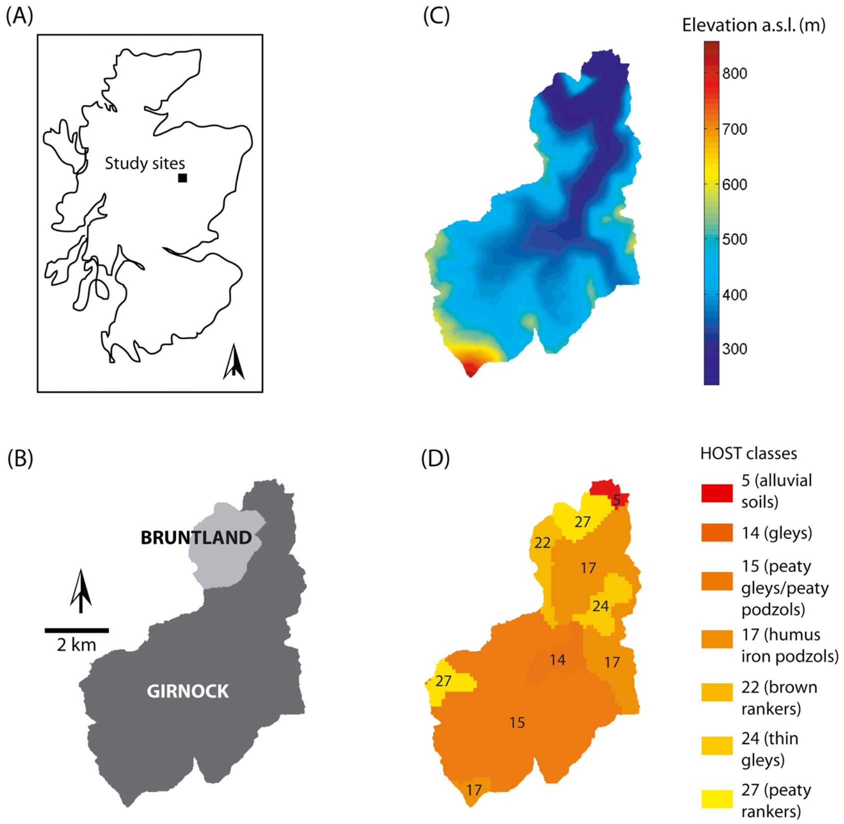

Figure 1. (A) Study sites location within Scotland; (B) Girnock and Bruntland catchments; (C) digital elevation model and (D) Hydrology of Soil Types (HOST) classification maps for both study sites. The spatial resolution of the maps is 10 m in panels C and D. This figure is available in colour online at wileyonlinelibrary.com/journal/espl

flashy hydrological regime driven by saturation excess overland flow. The second most common soil units are peaty podzols (HOST 15), humus iron podzols (HOST 17), and alluvial soils (HOST 5) which do not favour near-surface saturation but rather facilitate groundwater recharge through vertical water movement. These generally unsaturated soils are especially present in higher altitude areas, on steeper slopes and near fractured bedrock outcrops. It is believed that a portion of groundwater recharge moves quickly through shallow fracture systems or freely draining drift deposits to discharge in valley bottom areas (Soulsby *et al.*, 2005). This groundwater emerges either as return flow back to the surface of gleyed and peat soils (Shand *et al.*, 2006) or through the bed and banks of the streams (Malcolm *et al.*, 2006).

To further characterize near-surface stormflow dynamics in both catchments, global positioning system (GPS) mapping was used to delineate saturation areas connected to the stream under different antecedent conditions (Birkel *et al.*, 2011). The 'squishy-boot' method was used to target superficial water saturation, which corresponds to areas where: (i) a squelchy noise can be heard when stepping on the ground, even in the absence of ponding water; (ii) water squeezes out of the topsoil when stepping on it with a boot; or (iii) water is present on the soil surface. The three qualitative criteria that we used for the definition of 'superficial water saturation' are aligned with the recent definition of 'qualitative wetness classes' suggested by Rinderer *et al.* (2012). One single operator was responsible for all the mapping and used the same GPS unit from one survey to the other to ensure that personal and instrument-driven biases would be minimized. Saturation areas that were not spatially connected to the stream were not mapped as they could not be assumed to be surface contributing areas. Seven field surveys of connected saturation areas were achieved in the Bruntland catchment and one field survey was done in the whole Girnock catchment (Table II). In the Bruntland, the expansion and contraction dynamics of the connected saturation areas were important as superficial water was found to cover between 3 and 35% of the catchment area depending on antecedent moisture conditions (Table II). The goal of the current study was to test the ability of different wetness indices to simulate these expansion and contraction dynamics.

Wetness Indices

Theory and computation

Three types of data were used in this study: (i) digital elevation models (DEMs) at a 10 m resolution for both catchments; (ii) 1 km^2 HOST maps downsampled at a 10 m resolution for consistency purposes; (iii) generic values of the soil porosity, soil depth and saturated hydraulic conductivity associated with each HOST class; these generic values were originally published for rainfall–runoff modelling (Moore *et al.*, 2007). A number of wetness indices were derived and can be classified in several sub-categories: steady-state versus quasi-dynamic indices, DEM derivatives versus compound indices, and variants of the TWI versus geomorphic indices (Figure 2, Appendix A). Simple DEM derivatives (e.g. slope, curvature) were obtained using traditional algorithms implemented in the ArcGIS software, version 9.3 (ESRI, 2008) and the TauDEM plug-in, version 5.0 (Tarboton, 2010).

For DEM derivatives but also for variants of the TWI, several options were available for handling flow directions, terrain slope and stream channels. With regards to flow routing, traditionally a single direction algorithm was used to restrict the flow from a given cell to be transferred only to one of its eight neighbours along the steepest downslope direction. Tarboton (1997) slightly modified this algorithm using triangular facets

Table I. Summary of catchment properties for the Bruntland subcatchment and the whole Girnock catchment.

	Bruntland	Girnock
Terrain features		
Area (km^2)	3.4	30.4
Perimeter (km)	7.8	30.2
Minimum elevation (m)	255.2	230.1
Maximum elevation (m)	542.0	861.0
Mean elevation (m)	358.7	406.8
Maximum slope (deg)	50.2	61.4
Mean slope (deg)	13.4	9.9
Drainage density (km/km^2)	0.6	0.8
Mean flow path length (km)	2.6	12.9
Coefficient β of Hack's law (–)	1.2	1.7
Longest stream length (km)	2.6	12.9
Soil properties		
Alluvial soils (%)	0.0	1.7
Humus–iron podzols, subalpine soils (%)	39.1	25.2
Brown forest soils (%)	0.0	0.0
Rankers (%)	60.8	12.3
Peaty podzols and peaty gleys (%)	0.1	52.0
Peat (%)	0.0	0.0
Eroded peat (%)	0.0	0.0
Gleysols (%)	0.0	8.8
Open water (%)	0.0	0.0
Freely draining soils (%)	39.0	43.0
Responsive soils (%)	61.0	57.0
Hydroclimatic characteristics		
Mean annual precipitation (mm)	1059	1059
Mean annual temperature (°C)	7.3	7.5
Mean annual wind speed ($m\,s^{-1}$)	5.9	6.2
Mean daily discharge ($l\,s^{-1}\,km^{-2}$)	29.4	19.7
Q95 (exceeded 95% of time) ($l\,s^{-1}\,km^{-2}$)	3.4	1.9
Q5 (exceeded 5% of time) ($l\,s^{-1}\,km^{-2}$)	57.2	71.3
Median MTT (days) (5–95%)	682 (345–1019)	582 (272–892)
Median storage (mm) (5–95%)	1308 (662–1954)	1051 (491–1610)

to account for the fact that the steepest slope might not follow one of the eight cardinal and diagonal directions. Other flow routing algorithms are said to be multi-directional since they allow the flow from a given cell to be distributed to all downslope neighbours proportionally to their respective slopes (Quinn *et al.*, 1991; Seibert and McGlynn, 2007), and these multi-directional algorithms tend to produce more realistic flow patterns than unidirectional ones. Another issue concerns the assumption of time-invariant upslope areas which is often challenged, especially in flat areas where poorly defined flow directions are likely to change with time (Grabs *et al.*, 2009). A few studies (Barling *et al.*, 1994; Borga *et al.*, 2002; Tarolli *et al.*, 2008, 2011) have therefore relaxed the time invariance assumption by defining quasi-dynamic upslope areas that can be significantly smaller than the steady-state area derived from DEM analysis depending on the drainage period considered. When it comes to the appreciation of terrain gradient, the question remains as to whether the tangent or the sine of the local slope angle should be used. Indeed, the sine of the ground-surface inclination is often said to be more physically correct to represent the total head gradient driving subsurface flow (O'Loughlin, 1986; Montgomery and Dietrich, 2002); while the difference between the sine and tangent functions is negligible for low angled areas, such is not the case for steeper hillslopes. The idea of using the local slope itself to approximate the downslope hydraulic gradient is criticized because the effects of downslope topography at a distance of more than one cell are not considered (Sorensen *et al.*, 2006). The problem is especially important in low relief areas where the downslope hydraulic gradient is overestimated by the local slope (Grabs *et al.*, 2009). In an attempt to solve this issue, Hjerdt *et al.* (2004) suggested the use of a downslope index $\tan_d$, which is the slope to the closest point that is *d* metres below the reference grid cell. The distance between the reference grid cell and this closest target point can be measured following the steepest direction either as a beeline or along theoretical flow paths, and the downslope index has proven to better estimate groundwater gradients. As for the stream channel initiation threshold area, its chosen value bears important implications for the TWI since no explicit routing of accumulated upslope area is needed within the hydrographic network (Sorensen *et al.*, 2006). All these options or algorithms for handling flow directions, terrain slope and stream channels, as well as different mathematical representations of the decline of transmissivity with depth, were used (Figure 2, Appendix A) so as to assess their influence on the performance of the wetness indices. The Generalized Quasi-dynamic Model (QDM) put forward by Tarolli *et al.* (2008) and generalized to describe surface and subsurface runoff propagation on various surface types and soil-mantled elements was notably used to derive time-variable upslope contributing areas for 25 different drainage times (Figure 2).

A few wetness indices derived from terrestrial geomorphology were also derived and evaluated in this study (Figure 2, Appendix A). For instance, the focal standard deviation of elevation over a circular window was computed. Circular windows of different radii were used, knowing that standard deviation provides a measure of local relief over small windows and landscape roughness over large windows. For each grid cell of the DEM, the TPI was calculated by comparing its elevation to the average elevation of its surrounding cells in a given radius (Weiss, 2001; Jenness, 2006). The TPI is therefore a measure of whether a point is on a hilltop, in a valley bottom, on an exposed ridge or on another kind of feature in the landscape. Locations that are higher than their surroundings have positive TPI values, while those which are lower have negative values. This assessment is however scale-specific as it depends on the radius of the neighbourhood used. Zero or near-zero TPI values can flag either flat areas or areas of constant surface gradient unless slope is explicitly taken into account. Thus, a slope position classification comprising six categories (valley, toe slope, flat, midslope, upper slope, and ridge areas) was also derived by discriminating locations not only based on how extreme their TPI values are but also based on their surface gradient (Weiss, 2001; Jenness, 2006). Furthermore, a landform classification was obtained by combining two TPI maps from different scales and a slope map to discriminate canyons, U-shaped valleys, plains and other landform types. An extension (Jenness, 2006) compatible with the ArcGIS software was used to compute TPI values and to allow the automatic classification of slope positions and the delineation of landforms for different neighbourhoods. Lastly, the TRMI (Parker, 1982) was computed to illustrate the potential of the soil at given location to be saturated with water based on four slope parameters: position (i.e. valley bottom, midslope), aspect (azimuth degrees), shape (i.e. concave, straight, convex) and steepness (in degrees). The first two parameters had classification scores ranging from 0 to 20 and the last two from 0 to 10, with the characteristics more favourable to soil saturation being associated with higher classification scores. In the end, for each location, the TRMI was obtained by summing up the scores assigned to each of the four parameters; xeric locations are characterized by TRMI

Table II. Characteristics of the connected saturation area surveys conducted in the Bruntland sub-catchment and the whole Girnock catchment.

	Survey date	02/05 2008	02/07 2008	04/08 2008	03/09 2008	01/10 2008	26/11 2008	21/01 2009	Spearman rank correlation coefficient r_s between the percentage of area saturated (Bruntland) and each AMC surrogate: r_s	p-Value
Percentage of catchment area saturated	Bruntland	34.7	3.3	9.6	8.6	6.2	13.6	32.1		
	Girnock	n/a	n/a	10.6	n/a	n/a	n/a	n/a		
AMC surrogates	AP1day (mm)	4.6	4.4	6.3	0.3	0.5	0.0	0.1	–0.11	0.84
	AP2day (mm)	13.7	4.4	7.3	0.5	1.1	3.1	0.9	0.25	0.59
	AP7day (mm)	30.3	16.4	26.2	2.4	2.3	26.2	37.3	0.82	0.03
	AP10day (mm)	30.9	19.7	26.7	3.7	2.4	27.6	42.5	0.86	0.02
	AP15day (mm)	30.9	25.7	26.9	28.3	17.3	34.2	47.1	0.82	0.03
	AP20day (mm)	41.5	32.0	39.4	48.5	40.8	51.8	52.1	0.64	0.14
	AP30day (mm)	101.4	35.5	66.5	85.2	69.4	75.8	54.7	0.46	0.30
	DSP_50th (d)	0	0	0	176	0	70	13	0.16	0.76
	DSP_75th (d)	0	0	0	187	1125	82	107	–0.22	0.64
	DSP_90th (d)	41	480	217	1152	1136	137	120	–0.86	0.02
	DSP_95th (d)	111	5967	250	1180	1389	150	126	–1.00	0.00
Bruntland catchment response	Daily discharge ($l\,s^{-1}\,km^{-2}$)	2.398	0.396	0.938	0.636	0.690	3.760	0.607	0.46	0.24

Note: 'n/a' indicates that no spatial survey was done in the whole Girnock catchment at a specific date. Antecedent moisture conditions (AMC) surrogates are computed using precipitation data only. 'AP*x*day' is the cumulated amount of precipitation from the *x* days preceding the survey. 'DSP_yth' is the number of days (elapsed) since precipitation intensity exceeded a certain value, in mm/d, corresponding to the yth percentile of historical precipitation in the catchment. From the 2000–2010 precipitation record, the 50th, 75th, 90th and 95th percentiles of daily rainfall intensities were computed as 0.75, 3.00, 7.25 and 11.25 mm/d, respectively. In the right portion of the table, the Spearman rank correlation coefficient and its associated *p*-value are reported to examine the links between the extent of connected saturation areas, AMCs and catchment discharge. Note that discharge values for the Bruntland catchment are in fact daily flows from the Girnock catchment prorated according to the difference in catchment drainage area.

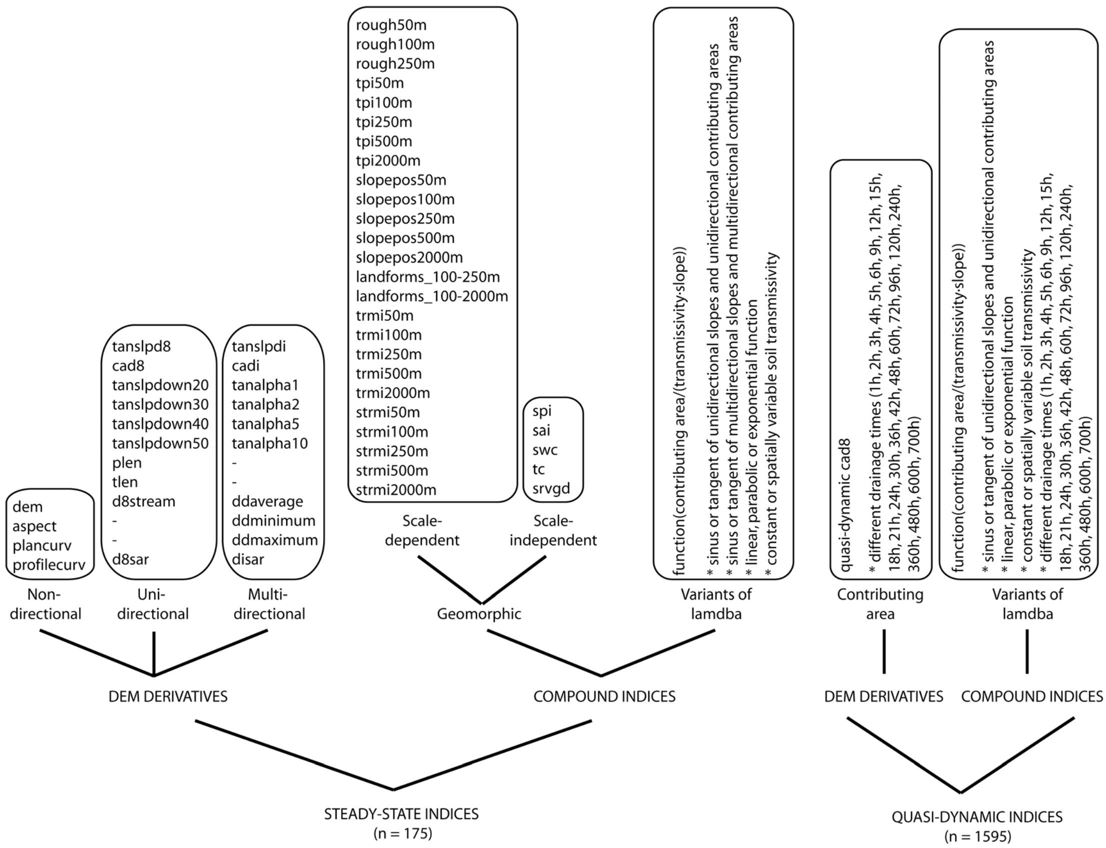

Figure 2. Summary of wetness indices evaluated in this study. Abbreviations are fully explained in Appendix A.

values near zero while mesic environments have TRMI values near 60. A modification of the TRMI, called the RSMI (relative site moisture index, Van de Grift, 2006), was also applied here by considering two additional parameters, soil depth (in centimetres) and soil texture (i.e. loam, silt, sandy loam) ranging from 0 to 10, also with higher classification scores illustrating saturation prone conditions.

Performance criteria

The combination of the different primary and secondary DEM attributes led to the evaluation of 1700 wetness indices in this study (Figure 2, Appendix A). Several post-processing steps were then carried out not only to compare the indices with one another but also to evaluate their relative ability to model the spatiotemporal variability of connected saturation areas in the Bruntland and Girnock catchments.

First, to assess the impact of spatial resolution on the results, for each of the 1700 wetness indices four different maps were compared: a raw map, a 30 m × 30 m smoothed map, a 50 m × 50 m smoothed map, and a 70 m × 70 m smoothed map. The smoothed maps were obtained by applying a 3 × 3, a 5 × 5 or a 7 × 7 low-pass filter on the 10 m-resolution raw maps. This was done to account for potential non-local influences on saturation areas development as previously done by Lanni *et al*. (2011). The individual wetness indices maps ($n = 1700 \times 4 = 6800$) were then compared with one another with regards to their spatial autocorrelation parameters SAL (shortest autocorrelation length) and STR [surface texture (aspect) ratio]. The SAL parameter, or the fastest decay autocorrelation length, is the shortest distance in which the normalized areal autocorrelation function decays to a threshold value of 0.2 in any possible direction and it identifies the direction in which autocorrelation is minimized. In this study, high values of SAL would indicate the presence of low frequency (long wavelength) components in a given wetness index and they would be associated with wavy spatial patterns. In contrast, lower SAL values would be associated with flatter spatial patterns with a dominance of high frequency (short wavelength) components. The STR parameter was used to characterize the uniformity of texture aspect. It is defined as the ratio of the fastest decay autocorrelation length (SAL) to the slowest decay autocorrelation length and usually takes values in the range of zero to one. Values of STR exceeding 0.5 indicate uniform texture in all directions (no defined lay), while values of STR smaller than 0.3 indicate strong anisotropic phenomena in the wetness indices patterns. It should be noted that the SAL and STR spatial parameters were used for simplicity purposes in this study. Indeed, while these parameters are commonly used in DEM studies, they are unusual in hydrology where geostatistics are commonly relied on to characterize spatial correlation properties of various patterns. The use of geostatistics here would have been challenging because there was no prior knowledge of which theoretical variogram model (exponential, spherical, 'hole effect' or other) should be used to fit the different wetness indices spatial patterns.

Second, the ability of the computed wetness indices to predict the location of connected saturation areas was evaluated. For each date on which connected saturation areas were mapped in the field, binary logistic (or logit) regression was applied to link the values of each wetness index to the spatial presence/absence of surface water connected to the stream. Logistic regression is conceptually similar to linear regression since it aims at evaluating the relationship between one dependent variable and one (or several) independent variable(s). The difference with logistic regression is that the dependent variable is categorical rather than continuous (Hosmer and Lemeshow, 1989). In the binary case, this dependent variable can only take two values and hence, logistic regression returns the posterior probability of a positive binomial outcome. In this study, each connected saturation area map was, in turn, the dependent variable while each wetness index map was, in turn, the independent variable. For each logistic regression conducted, a map was obtained where each grid cell was associated with its probability to be predicted as a saturated one based on the considered wetness index. Each probability produced by logistic regression had to be converted into an actual membership to classify a cell as saturated or non-saturated. Given that probabilities range from zero to one, a cutoff value of 0.75 was used. In the predicted maps, the cells with a probability above the cutoff value were assumed to be saturated whereas the cells with a probability below the cutoff value were assumed to be non-saturated. It is worth mentioning that most studies relying on binary logistic regression often use a 0.5 cutoff probability value and report a range of statistical coefficients to assess the robustness of the logit results (Hosmer and Lemeshow, 1989). Here it was impossible to report such statistical coefficients for 6800 regressions; hence it was decided to focus only on grid cells for which high posterior probabilities (above 0.75) were obtained.

After each logistic regression, a confusion matrix was built to compare the actual connected saturation area classification (field map) and the predicted connected saturation area map (map derived from logit probabilities). This confusion matrix displayed the number of correct and incorrect predictions made by the studied wetness index compared with an actual map of connected saturation areas. A number of performance measures (Table III) were derived from this confusion matrix to quantify the agreement between the actual and the predicted maps; motivation for resorting to several performance measures was that there is no consensus on the most appropriate criteria to compare binary classifications (Güntner *et al*., 2004; Grabs *et al*., 2009). Some measures are very popular because their associated formulas are self-explanatory (sensitivity, specificity, Table III) but those do not take into account the potential effects of randomness. The most commonly used performance measure is Cohen's Kappa because it quantifies the extent to which a model correctly predicts occurrence at rates that are better than chance expectation (Cohen, 1960); however even Cohen's Kappa has been criticized because it is thought to be highly dependent upon sample size and to give a biased appreciation when the size of one class (e.g. the 1 s) far exceeds the size of the other (e.g. the 0 s) (Byrt *et al*., 1993). The Normalized Mutual Information (NMI) statistic has been suggested to solve the latter issue (Cover and Thomas, 1991) but it cannot differentiate the worse-than-random models from the better-than-random models. Hence, in this study all measures listed in Table III were estimated to assess whether the wetness indices selected as the best were highly variable depending on the performance criterion which was used. An objective function was also built by summing up all measures listed in Table III; the assessment of overall best and worst wetness indices was based on the values of this objective function. Our evaluation of the performance of the wetness indices was only based on a cell-to-cell comparison and did not consider the geometric properties (e.g. area, shape) of observed and predicted saturated areas.

Results

The different wetness indices showcased different autocorrelation properties across the whole Girnock catchment as shown in Figure 3. The STR axis was illustrative of a gradient from high

Table III. List of performance criteria and objective function used to evaluate the ability of each wetness index to approximate the different spatial patterns of observed connected saturation areas.

Abbreviated name	Significance	Range (*Optimal*)
Individual performance measures		
Sensitivity	Sensitivity or Recall or True positive rate	0 to 1 (*1*)
Specificity	Specificity or True negative rate	0 to 1 (*1*)
Accuracy	Accuracy or ROC (Receiver Operator Curve) area	0 to 1 (*1*)
PPP	Positive predictive power	0 to 1 (*1*)
NPP	Negative predictive power	0 to 1 (*1*)
DetectionRate	Detection rate	0 to 1 (*1*)
OverallDiagnostic Power	Overall diagnostic power	0 to 1 (*1*)
Kappa	Cohen's kappa	–1 to 1 (*1*)
NMI	Normalized mutual information statistic	0 to 1 (*1*)
Objective function		
ObjF	Sensitivity + Specificity + Accuracy + PPV + NPV + DetectionRate + OverallDiagnosticPower + Kappa + NMI	–1 to 9 (9)

Note: Detailed formulas for computing individual performance measures are reported in Appendix B.

to low anisotropy while the SAL axis was rather indicative of a gradient from small to large number of low frequency components in the spatial patterns. When looking at DEM derivatives, a clear difference in spatial properties could be observed between flow path maps (i.e. cad8, cadi, tlen) and other wetness indices (Figure 3A); the former were weakly anisotropic while the later showed strong anisotropy. The size of the spatial window (or neighbourhood) had a negligible influence on the computed geomorphic indices as all clustered on Figure 3B, thus illustrating that they have similar autocorrelation properties despite their computation based on different spatial resolutions. Steady-state variants of the TWI lead to very different autocorrelation patterns depending on the transmissivity profiles and the flow direction algorithm used (Figure 3C): for instance, exponential transmissivity profiles tended to create more anisotropy than parabolic profiles and linear profiles, respectively. Differences in STR values (hence anisotropy) could also be seen between indices relying on unidirectional (D8) rather than multi-directional (D∞) flow algorithms. The majority of the steady-state TWI variants were associated with near-zero SAL values, thus indicating wavy spatial patterns. The positioning of steady-state TWI clusters in the SAL-STR space was roughly the same as that of quasi-dynamic TWI clusters (Figures 3C and 3D). Although outside of the scope of this paper, analyses of the quasi-dynamic TWI variants in the SAL-STR space revealed that there was a statistically significant negative correlation (Spearman rank correlation coefficients ranging from -0.98 to -1, $p < 0.05$) between the SAL or STR values and the drainage times: this finding suggests that the

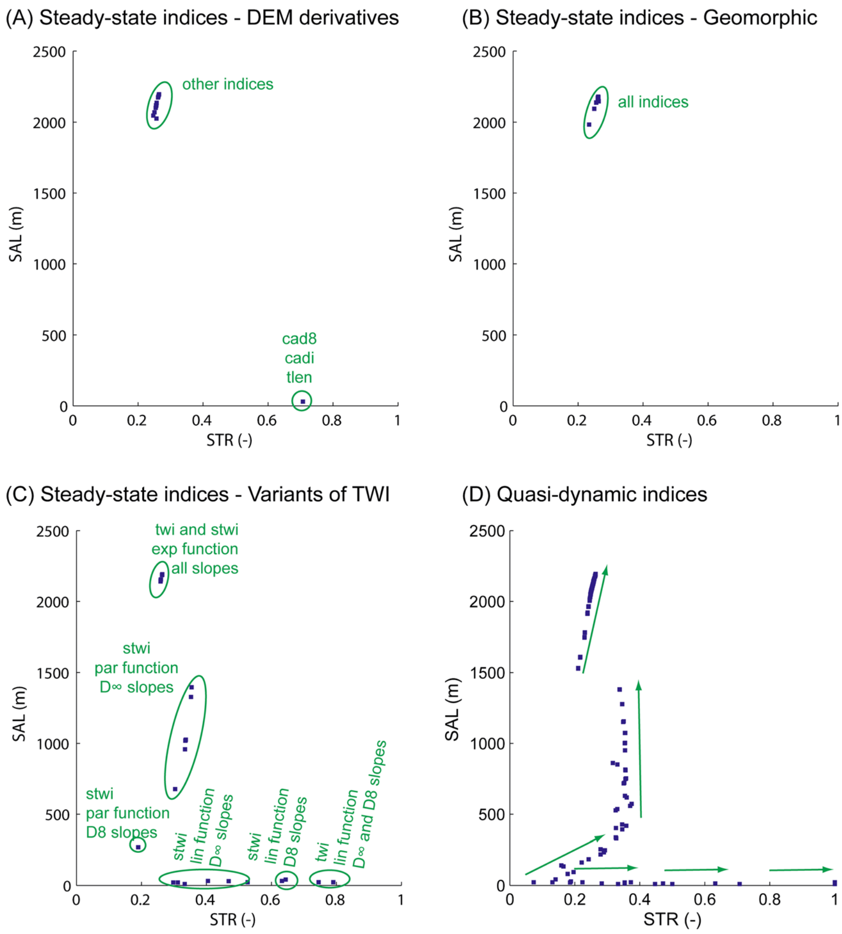

Figure 3. Shortest autocorrelation length (SAL) and surface texture (aspect) ratio (STR) values associated with the raw (non-smoothed) wetness indices maps for the whole Girnock catchment. Note that the limits of the *x*- and *y*-axes differ between the four panels. Clusters of indices are identified to emphasize on map patterns sharing similar autocorrelation properties. In panel D, arrows show the direction of increasing drainage time for each cluster of quasi-dynamic indices. This figure is available in colour online at wileyonlinelibrary.com/journal/espl

longer the drainage period, the less anisotropic the quasi-dynamic wetness indices patterns and the lower the number of low frequency components in those patterns.

For the evaluation of wetness indices with respect to their ability to predict connected saturation areas, we focused on the Bruntland sub-catchment. Objective function values associated with all 1700 raw indices maps for all seven connected saturation area surveys are shown in Figure 4. Quasi-dynamic indices tended to conglomerate in the lower end of the objective function spectrum, which indicate their poor performance with regards to predicting connected saturation areas. Indices that were associated with the higher objective function values were steady-state variants of TWI, followed by geomorphic indices; the former could however be present at both ends of the objective function spectrum. It is worth noting that for the first and last connected saturation area survey dates, which also corresponded to the two wettest survey dates, indices spread over a much larger range of objective function values than for drier surveys (Figure 4). However, objective function values only approached five (over a maximum value of nine) for the two wettest surveys and fell consistently below 4.75 for drier survey dates, which indicates poor to intermediate performances from all wetness indices.

For the Bruntland sub-catchment, comparisons of observed and predicted connected saturation area maps revealed that reasonably good predictions could be achieved for the first and the last survey dates, which also were the wettest ones in terms of connected saturation area extent (more than 30% of overall catchment area) (Figure 5). That was however not the case for the other, drier dates. Two geomorphic indices happened to be selected within the best three indices, namely the topographic position index over a 2000 m neighbourhood and the topographic roughness over a 250 m spatial window. Also, the shape of the predicted connected saturation area was highly variable among survey dates, which means that the algorithms relied on different spatial connectivity rules to define the extent of contributing superficial water. When the best indices were TWI variants, the slope parameter was always the downslope index angle to which a sinus or a tangent function was applied (Figure 5). For the two wettest survey dates, when TWI variants were chosen as the best indices, a parabolic function for the transmissivity profile was generally used. Predicted connected saturation areas also appeared to be more widespread with TWI variants that did not include any soil information. For the whole Girnock catchment, no satisfactory wetness index could be found; predicted connected saturation areas were restricted to permanent stream channels (Figure 6) while the maximum objective function value across the whole range of tested indices never exceeded 3.4.

Lastly, the performance of wetness indices was not systematically improved nor worsened when considering smoothed maps rather than raw maps. Minimum and maximum objective functions values did not vary in a consistent manner with map resolution (Table IV); in fact, the minimum objective function

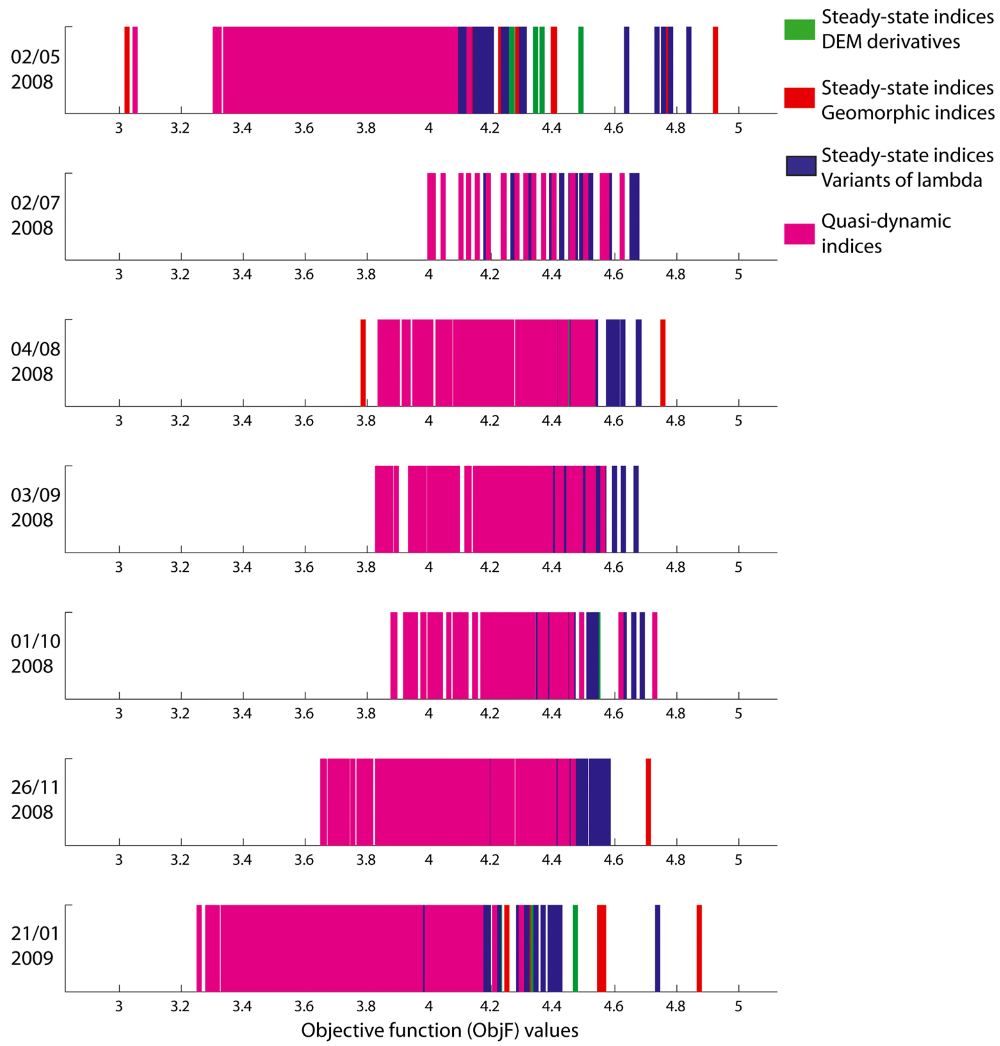

Figure 4. Objective function values for wetness indices pertaining to four different categories. Results are associated with raw (non-smoothed) maps for the Bruntland sub-catchment. This figure is available in colour online at wileyonlinelibrary.com/journal/espl

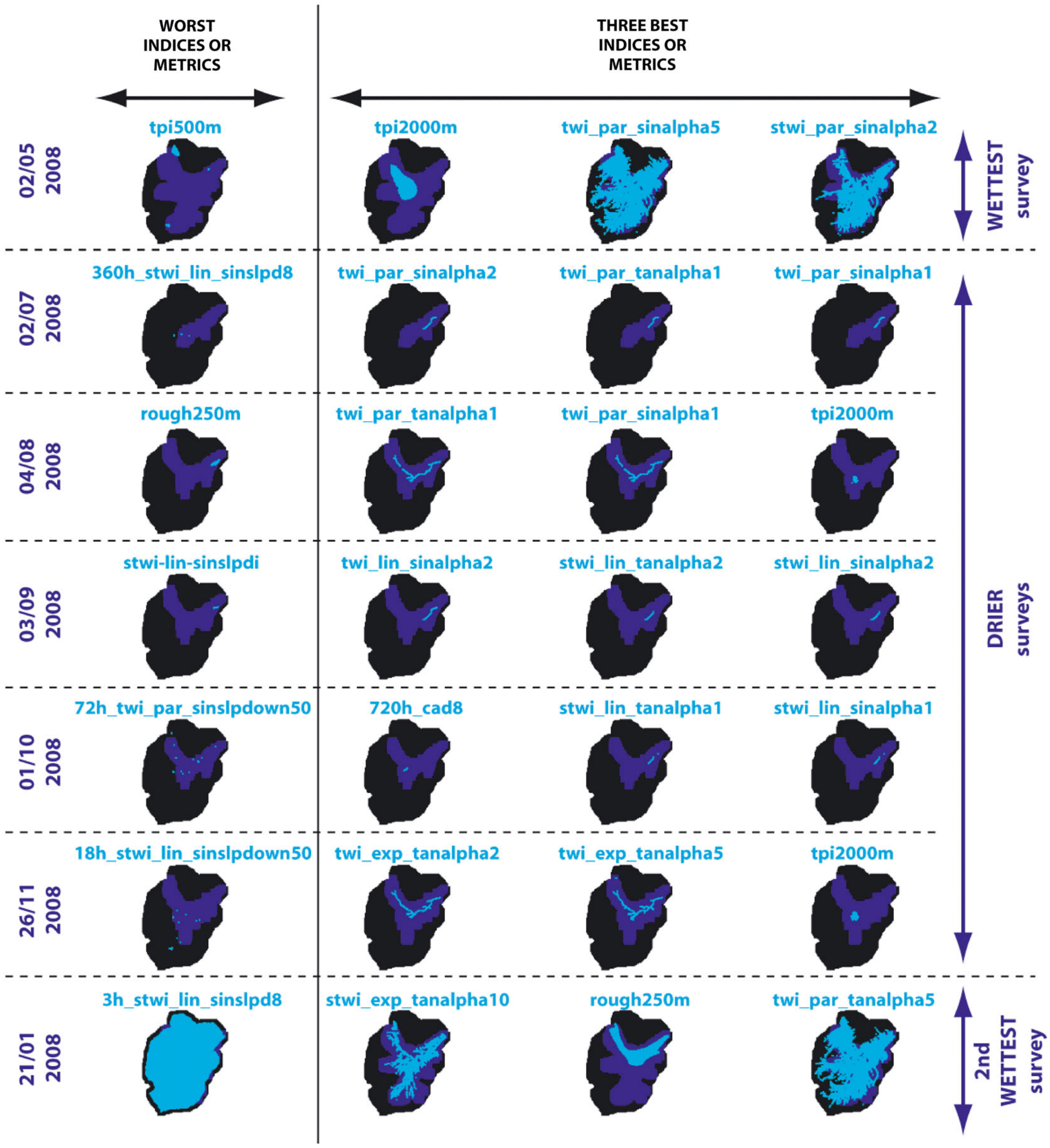

Figure 5. Worst and best wetness indices identified for the prediction of connected saturation areas in the Bruntland sub-catchment for each survey date. Observed connected saturation areas, predicted connected saturation areas and non-saturated or non-contributing areas are illustrated in blue, cyan and black, respectively. This figure is available in colour online at wileyonlinelibrary.com/journal/espl

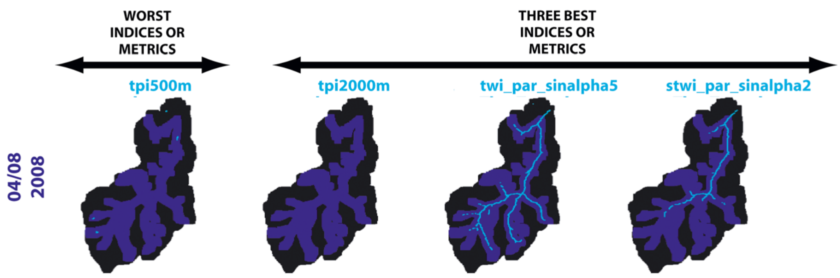

Figure 6. Worst and best wetness indices identified for the prediction of connected saturation areas in the whole Girnock catchment for the single survey date. Observed connected saturation areas, predicted connected saturation areas and non-saturated or non-contributing areas are illustrated in blue, cyan and black, respectively. This figure is available in colour online at wileyonlinelibrary.com/journal/espl

value was almost the same for all four smoothening options but the 50 m × 50 m one. As for the maximum objective function value, it tended to increase with the size of the smoothening window even though it was higher for the 50 m × 50 m maps in comparison to the 70 m × 70 m maps. It could also be observed that the differences between the performance of raw and smoothed maps varied with the survey date (Table V). On the first and the last survey dates, which were the wettest ones, the use of a bigger smoothening window (30 m × 30 m rather than raw, 70 m × 70 m rather than 50 m × 50 m) consistently led to improved objective function values (as illustrated by moderate to high positive percentage change differences). In contrast, in drier conditions, the use of a bigger smoothening window sometimes led to worse objective function values as indicated by negative percentage change differences in Table V.

Table IV. Summary statistics of the objective function values computed over all raw and smoothed wetness indices maps.

	Minimum	Maximum	Coefficient of variation
Raw maps	3.02	4.92	0.43
30 m × 30 m smoothed maps	3.02	4.92	0.66
50 m × 50 m smoothed maps	3.01	5.24	0.80
70 m × 70 m smoothed maps	3.02	5.07	0.88

Discussion

Predictive power of wetness indices across space and time

In spite of the exhaustive investigation conducted here, very few wetness indices were useful to model the connectivity of saturation areas to the stream in the Bruntland and Girnock catchments; in fact, prediction results were fair only in the wettest conditions. These poor results are however not unusual. Indeed, the TWI has previously been shown to have a significant explanatory value only in wet conditions (Western *et al.*, 1999; Güntner *et al.*, 2004). Some authors have even discussed the limited predictive power of deterministic wetness indices which are unable to capture the natural, and sometimes random variability of landscape characteristics (Güntner *et al.*, 2004). Other studies have also showed how the water table may or may not closely follow the topography (Haitjema and Mitchell-Bruker, 2005). The typical wide, low-gradient valley bottoms in the Bruntland and Girnock catchments likely gave rise to some computational errors since the ground surface cannot be used to describe subsurface flow pathways in such areas. In flat landscapes or in areas dominated by wetlands, the actual hydraulic gradient is often smaller than the surface slope, thus leading to an underestimation of the state of wetness (Grabs *et al.*, 2009). This explains why TWI variants relying on the surface slope were rarely chosen as ‘best performing’ indices in contrast to TWI variants relying on the downslope gradient (Figure 5). The fact that the best TWI variants in wet conditions relied on a parabolic soil transmissivity profile suggests that catchment soils have a maximum soil storage deficit that cannot be exceeded, in contrast to soils associated with an exponential transmissivity profile and for which there is no upper limit of soil storage deficit (Ambroise *et al.*, 1996a). The ability of geomorphic indices to identify saturation-prone areas was also evaluated but prediction results were inconsistent, mainly because of issues regarding the size of the spatial neighbourhood and the static definition of the morphological units. It was however interesting to find that the use of smoothed wetness indices maps over large spatial neighbourhoods led to better connected saturation area predictions in wet conditions but not in drier conditions (Table V); this is consistent with the preferential states hypothesis (Grayson *et al.*, 1997) that opposes the influence of non-local topographic controls on shallow soil moisture during wet conditions to that of local, small-scale and even random topographic controls on shallow soil moisture in dry conditions. Fine-tuning this type of analysis by testing out different sizes of smoothening windows could therefore lead to the identification of different characteristic scales for hydrological processes in wet versus dry conditions.

Added value of soil information for the predictive power of wetness indices

While soil type has proven to be a good predictor of global catchment hydrological dynamics in Scotland (water transit times, see Hrachowitz *et al.*, 2009, 2010), the HOST classification has not been used towards the specific prediction of surface saturation areas. The indices tested in the current study that incorporated information from the HOST classification did not lead to any significant improvement over indices which did not include any soil information. These results strongly contrast with the fairly good approximation of connected saturation areas previously obtained based on hydroclimatic variables only (Birkel *et al.*, 2010) while ignoring both topographic and soil characteristics. Indeed, within the framework of a lumped dynamic saturation area model (SAMdyn), the extent of the connected saturation area, expressed in proportion of catchment area, was accurately modelled in the whole Girnock catchment based on a seven-day antecedent precipitation exponential decay function. Hence, coupling fine scale hydroclimatic data with physical wetness indices might be needed to better take into account the temporal dynamics of soil hydrological processes. It is also worth discussing a decision that was made at the beginning of this study, namely the decision to focus only on saturation areas connected to the hydrographic network rather than trying to predict the location of all saturation areas, including those which likely cannot route surface water to the streams. This methodological choice was made while hypothesizing that soils in the valley bottom hollows should have a better transmissivity than hillslope soils, hence their potential to enhance hydrologic connectivity to the stream (Ambroise *et al.*, 1996b). The coarse resolution of the HOST maps did not however make it possible to make that subtle distinction between morphological units. The importance of spatial connectivity when studying saturation areas has been argued before, notably by recommending that performance criteria including explicit information about cell

Table V. Percentage change differences in the values of the objective function between raw and smoothed wetness indices maps.

Test maps	Reference maps	02/05 2008	02/07 2008	04/08 2008	03/09 2008	01/10 2008	26/11 2008	21/01 2009
Raw	Smoothed 30 m × 30 m	2.35%	–1.21%	–2.32%	–1.35%	–0.11%	–0.33%	2.25%
Raw	Smoothed 50 m × 50 m	4.06%	3.13%	–0.55%	–1.15%	0.35%	1.56%	3.45%
Raw	Smoothed 70 m × 70 m	6.13%	0.12%	–0.98%	–0.99%	0.40%	1.27%	5.68%
Smoothed 30 m × 30 m	Smoothed 50 m × 50 m	2.23%	4.63%	2.12%	0.35%	0.20%	2.16%	2.10%
Smoothed 30 m × 30 m	Smoothed 70 m × 70 m	4.80%	1.85%	1.34%	1.01%	0.38%	1.93%	5.01%
Smoothed 50 m × 50 m	Smoothed 70 m × 70 m	1.98%	–2.61%	–0.66%	0.16%	0.08%	0.00%	2.21%

Note: The percentage difference is computed as follows: (Value of predicted map – Value of actual map) × 100/Value of actual map.

neighbourhood be used when evaluating wetness indices (Güntner *et al.*, 2004). When the aim is to identify connected saturation areas for rainfall – runoff modelling, future applications might therefore consider the use of connectivity metrics (Western *et al.*, 2001; Ali and Roy, 2010) so as to facilitate the identification of the best wetness indices under contrasted conditions.

Influence of spatial data resolution on the predictive power of wetness indices

The results obtained in this study suggest that the 1 km^2 resolution of the HOST classification was too coarse to achieve sensible connected saturation area predictions in small to mesoscale catchments. In fact, acceptable index performances only occurred when 30% of the Bruntland sub-catchment was saturated (Table II, Figure 5); 30% of the Bruntland sub-catchment area equates to ~1.02 km^2, which is consistent with the 1 km^2 resolution of the HOST maps. Studies evaluating the performance of soil-topographic indices often rely on soil parameters averaged for whole catchments rather than spatially distributed ones, and here it was expected that resorting to spatially distributed parameters would significantly enhance the explanatory power of the tested wetness indices. The generic soil parameter values used in this study were however too coarse for yielding high quality results. A similar issue has been reported before (Western and Grayson, 2001; Güntner *et al.*, 2004), as to the difference in availability and detail of soil types and parameter values in comparison to standard topographic data. One of the implicit objectives of this study was to evaluate whether the combination of high resolution topographic data (10 m) and low resolution soil data (1 km) could significantly improve our prediction of connected saturation areas but results reveal that such was not the case. In a study focusing on a German catchment, Güntner *et al.* (2004) showed that saturation area predictions could be improved if soil transmissivity values were optimized. Such was not possible here as the only saturated hydraulic conductivity and soil depth values were generic ones obtained from a UK-wide dataset and not specific to the Bruntland or Girnock catchments. The inadequate resolution of the HOST maps might also explain why quasi-dynamic indices did not outperform steady-state ones, a result that is contrary to what has been shown previously (Barling *et al.*, 1994; Borga *et al.*, 2002; Tarolli *et al.*, 2008; Grabs *et al.*, 2009). Digital soil mapping and modelling (DSM, McBratney *et al.*, 2003) at a better resolution than 1 km^2 is therefore required towards getting more accurate connected saturation area predictions.

Conclusions

The main contribution of this study was the evaluation of a wide range of different wetness indices for predicting the connectivity of saturation areas to the stream network. The focus was on two Scottish catchments for which multiple actual maps of connected saturation areas were available so that the performance of wetness indices could be evaluated in contrasted conditions. Several computational options for terrain slope and flow directions were combined, and steady-state geomorphic indices which had not been exhaustively tested in hydrological studies before were considered. The HOST classification was also coupled with terrain data towards the improved simulation of surface saturation areas connected to the hydrographic network. Results show that connected saturation area predictions were very poor in dry conditions and at best fair in wet conditions, thus confirming what was reported in previous studies. However, somehow unexpectedly, soil-topographic indices did not bear more explanatory power than topographic-only indices, this even though spatially distributed soil parameter values were used rather than catchment averaged ones. The 1 km^2 resolution HOST classification was likely too coarse for predicting the spatiotemporal variability of connected saturation areas in both study catchments where neatly dissected morphological units showcase contrasting shallow water storage dynamics.

Despite our negative results, the work presented in this paper bears important implications onto how the discipline of Hydrology should approach hydrologic connectivity investigation. There is an ongoing debate around using either wetness indices (steady-state variants of the TWI) or proper ‘connectivity metrics’ (*sensu* Western *et al.*, 2001; Ali and Roy, 2010) for the accurate prediction of areas where hydrologic coupling of hillslopes, riparian zones and streams occur. The development of dynamic models (quasi-dynamic variants of the TWI) was also done in an attempt to predict how lateral flow processes responsible for hillslope–stream coupling turn on or shut off under certain conditions. While the discipline has not yet converged on a preferred index of hydrological connectivity, variants of the TWI remain the most used in the connectivity-focused literature because they are easy to derive at any resolution, and researchers often neglect the fact that the TWI was not specifically designed to capture connected saturated areas and does not account for important connectivity-prone conditions such as transient saturation. Our comparison exercise hints that ignoring such conditions is not an acceptable assumption, even in a catchment where topography and soil play a critical role and should, theoretically, lead to a good performance of topographic indices (TWI) and soil-topographic indices (STWI) in the prediction of connected saturated areas.

Our results also reveal that while topography is an important driver of hydrologic connectivity, it can be considered at different spatial scales through different wetness indices, with variable success. Indeed, while both variants of the TWI and geomorphic indices are topography-driven, the former are defined using cell-to-cell topography and assumptions about upslope contributing areas whereas the latter are defined using ‘regional’ topography to derive automated landform classifications. The better performance of geomorphic indices found in this paper highlights the importance of catchment architecture and suggests that the influence of topography on hydrologic connectivity should be conceptualized from the top-down (from the catchment to the morphological or hydrological response unit) rather than from the bottom-up (from the cell to the contributing area). Although conceptualizing water movement based on indices of topographic position alone might seem simplistic, it is certainly a reasonable, non-parametric approach that enables hydrological connectivity to develop as an emergent, landscape-dependent property and as such, we hypothesize that geomorphic indices would lead to acceptable predictions of connected saturated areas in a variety of environments. Since topographic position is an inherently scale-dependent property, geomorphic indices have an undeniable advantage over TWI variants in that they can allow us to examine the different scales over which surface hydrological connections are made, from small surface depressions in pothole-dominated systems to well-defined hillslope–riparian–stream transitions in mountainous terrain. We can therefore foresee that geomorphic indices would perform reasonably well even in regions where flow generation does not follow the network of topographic lows but rather depends on hummocky surface topography and the temporal cascade of ‘filling up and spilling over' events in surface depressions. Besides, whilst the inability of the indices to predict non-wet conditions may be problematic from a hydrological point of view, this is less the case from a geomorphic perspective. As most sediment transfer will occur in the wettest, most connected conditions, the indices may still be fit for purpose for geomorphic applications (e.g. sediment transfer processes) in landscape evolution models.

One issue that still lacks understanding is the prediction of hydrologic connectivity over time, and the results presented in this paper reveal that both process-based indices (TWI variants) and non-process-based indices (geomorphic indices) fail to predict the disconnection sequence of saturated areas in transient and dry conditions. Bracken *et al*. (2013) recently suggested that one way to better capture the temporal variability in hydrologic connectivity would be to move away from the use of topographic and soil–moisture and rather investigate 'how storage of water occurs in different catchments and how these stores fill up (or down) and link (or not) to produce (dis) connected flow' (p. 31). The better success of geomorphic indices found in this paper suggests that water storages could also be examined at the sub-catchment scale, from one morphological unit to the other, provided that the dominant flow process is either saturation-excess overland flow or a fill-and-spill-like mechanism. Future research on these aspects is definitely needed: we believe that a consensus amongst researchers is required around a range of wetness indices that are both sensitive to dry-transient-wet conditions and scale-dependent, thus making them applicable to various catchments dominated by surface flow processes.

Acknowledgements—The authors are grateful to the Scottish Environment Protection Agency (SEPA), the Macaulay Land Use Research Institute (MLURI) and the Fisheries Research Services (FRS) for access to their data and use of experimental facilities. The discharge data were provided by Derek Fraser, SEPA.

Appendix A

Abbreviations list and computation details for some wetness indices evaluated in this study.

DEM	Digital elevation model
aspect	Aspect
plancurv	Plan curvature
profilecurv	Profile curvature
tanslpd8	Tangent (tan) of slope (slpd8) estimated with the D8 algorithm
cad8	D8 contributing area
tanslpdown20 tanslpdown30 tanslpdown40 tanslpdown50	Tangent of average slope over a 20, 30, 40 or 50 m distance down D8 flow directions
plen	Length of the longest upslope flow path terminating at each grid cell
Tlen	Total length of all upslope flow paths terminating at each grid cell
d8stream	D8 distance to the stream
d8sar	D8 slope on contributing area ratio
tanslpdi	Tangent (tan) of slope (slpdi) estimated with the D∞ algorithm
cadi	D∞ contributing area
tanalpha1 tanalpha2 tanalpha5 tanalpha10	Downslope index computed given a 1, 2, 5 or 10 m drop in elevation
ddaverage	Weighted average of the D∞ distance to the stream
ddminimum	Minimum D∞ distance to the stream
ddmaximum	Maximum D∞ distance to the stream
disar	D∞ slope on contributing area ratio
rough50m rough100m rough250m	Topographic roughness (focal standard deviation of elevation values) computed over a spatial window (neighborhood) of 50, 100 or 250 m
tpi50m tpi100m tpi250m tpi500m tpi2000m	Topographic position index; difference between a central grid cell and the mean of its surrounding cells over a given spatial window (neighborhood) of 50, 100, 250, 500 or 2000 m
slopepos50m slopepos100m slopepos250m slopepos500m slopepos2000m	Slope position classification (1 = Valley; 2 = Toe slope; 3 = Flat; 4 = Midslope; 5 = Upper slope; 6 = Ridge) determined over a spatial window (neighborhood) of 50, 100, 250, 500 or 2000 m
landforms100–250 m landforms 100–2000 m	Landscape classification (1 = Canyons, deeply incised streams; 2 = Midslope drainages, shallow valleys; 3 = Upland drainages, headwaters; 4 = U-shaped valleys; 5 = Plains; 6 = Open slopes; 7 = Upper slopes, mesas; 8 = Local ridges, hills in valleys; 9 = Midslope ridges, small hills in plains; 10 = Mountain tops, high ridges) determined over spatial windows of 100–250 m and 100–2000 m, respectively
trmi50m trmi100m trmi250m trmi500m trmi2000m	Topographic relative moisture index, which classifies slope position, shape, steepness and aspect over a spatial window (neighborhood) of 50, 100, 250, 500 or 2000 m
strmi50m	

Continues

Table A1. Continued

strmi100m strmi250m strmi500m strmi2000m	Soil-topographic relative moisture index, which classifies slope position, shape, steepness, aspect, topsoil texture and depth over a spatial window (neighborhood) of 50, 100, 250, 500 or 2000 m
Spi	Stream power index [cad8 · slpd8]
Sai	Slope aspect index [aspect · slpd8]
Swc	Soil water content [(cad8/tanslpd8) · plancurv]
Tc	Transport capacity [cad8 · (tanslpd8)^2]
srvgd	Surface runoff velocity and gully development [cad8 · tanslpd8 · plancurv]
host	United Kingdom Hydrology of Soil Types classification
lambda	Generic denomination of a topographic index computed as follows: [function(contributing area/(slope · transmissivity))] The transmissivity can be spatially variable or constant (equal to 1). The contributing and slope terms can be computed using different flow direction algorithms. Different functions can be applied to the ratio contributing area/(slope · transmissivity)
Twi	Topographic wetness index, or topographic version of lambda (unit constant transmissivity)
stwi	Soil-topographic wetness index, or soil-topographic version of lambda (spatially variable transmissivity)
lin par exp	Linear, parabolic or exponential functions used to compute variants of the topographic index lambda
tan sin	Tangent or sinus of slope angle
Qd	Quasi-dynamic

Appendix B.

Table BI. Formulas for the computation of classification performance measures from a confusion matrix.

Abbreviated name	Description and formulas	Range (*Optimal*)
Sensitivity	Proportion of observed 'saturated' areas that are predicted as 'saturated' $\frac{TP}{(TP+FN)}$	0 to 1 (*1*)
Specificity	Proportion of observed 'non saturated' areas that are predicted as 'non saturated' $\frac{TN}{(TN+FP)}$	0 to 1 (*1*)
Accuracy	Proportion of predictions ('saturated' and 'non saturated') that are correct $\frac{(TP+TN)}{N}$	0 to 1 (*1*)
PPP	Proportion of 'saturated' cells that are correctly predicted $\frac{TP}{(TP+FP)}$ or $\frac{(\text{Sensitivity}\cdot\text{Prevalence})}{((\text{Sensitivity}\cdot\text{Prevalence})+((1-\text{Specificity})\cdot(1-\text{Prevalence})))}$	0 to 1 (*1*)
NPP	Proportion of 'non saturated' cells that are correctly predicted $\frac{TN}{(TN+FN)}$ or $\frac{(\text{Specificity}\cdot(1-\text{Prevalence}))}{(((1-\text{Sensitivity})\cdot\text{Prevalence})+((\text{Specificity})\cdot(1-\text{Prevalence})))}$	0 to 1 (*1*)
DetectionRate	$\frac{TP}{N}$	0 to 1 (*1*)
OverallDiagnosticPower	$\frac{(FP+TN)}{N}$	0 to 1 (*1*)
Kappa	$\frac{\left[(TP+TN)-\left(\frac{((TP+FN)\cdot(TP+FP)+(FP+TN)\cdot(FN+TN))}{N}\right)\right]}{\left[N-\left(\frac{((TP+FN)\cdot(TP+FP)+(FP+TN)\cdot(FN+TN))}{N}\right)\right]}$	–1 to 1 (*1*)
NMI	$1-\frac{[-TP\cdot\ln(TP)-FP\cdot\ln(FP)-FN\cdot\ln(FN)-TN\cdot\ln(TN)+(TP+FP)\cdot\ln(TP+FP)+(FN+TN)\cdot\ln(FN+TN)]}{[N\cdot\ln(N)-((TP+FN)\cdot\ln(TP+FN)+(FP+TN)\cdot\ln(FP+TN))]}$	0 to 1 (*1*)

Note: TP is the number of cells correctly predicted as saturated (true positive), FP is the number of cells incorrectly predicted as saturated (false positive), FN is the number of cells incorrectly predicted as non-saturated (false negative), and TN is the number of cells correctly predicted as non-saturated (true negative). Values of TP, FP, FN and TN were obtained after logit regression using a high cutoff point of 0.75. The range of expected values and the optimal one associated with each performance criterion is reported in the last column. Note that N = TP + TN + FP + FN. Also, the prevalence is the proportion of truly or falsely predicted connected saturation areas: Prevalence = (TP + FN)/N.

References

Ali GA, Roy AG. 2010. Shopping for hydrologically representative connectivity metrics in a humid temperate forested catchment. *Water Resources Research* **46**: W12544. DOI. 10.1029/2010WR009442

Ambroise B. 2004. Variable, 'active' versus 'contributing' areas or periods: a necessary distinction. *Hydrological Processes* **18**: 1149–1155.

Ambroise B, Beven K, Freer J. 1996a. Toward a generalization of the topmodel concepts: topographic indices of hydrological similarity. *Water Resources Research* **32**: 2135–2145.

Ambroise B, Freer J, Beven K. 1996b. Application of a generalized topmodel to the small ringelbach catchment, Vosges, France. *Water Resources Research* **32**: 2147–2159.

Aryal SK, Mein RG, O'Loughlin EM. 2003. The concept of effective length in hillslopes: assessing the influence of climate and topography on the contributing areas of catchments. *Hydrological Processes* **17**: 131–151.

Barling RD, Moore ID, Grayson RB. 1994. A quasi-dynamic wetness index for characterizing the spatial-distribution of zones of surface saturation and soil–water content. *Water Resources Research* **30**: 1029–1044.

Beven KJ. 1986. Hillslope runoff processes and flood frequency characteristics. In Hillslope Processes, Abrahams AD (ed.). Allen and Unwin: Concord, MA; 187–202.

Beven KJ, Kirkby MJ. 1979. A physically based variable contributive area model of basin hydrology. *Hydrological Sciences Bulletin* **24**: 43–69.

Birkel C, Tetzlaff D, Dunn SM, Soulsby C. 2010. Towards a simple dynamic process conceptualization in rainfall–runoff models using multi-criteria calibration and tracers in temperate, upland catchments. *Hydrological Processes* **24**: 260–275.

Birkel C, Tetzlaff D, Dunn SM, Soulsby C. 2011. Using time domain and geographic source tracers to conceptualize streamflow generation processes in lumped rainfall–runoff models. *Water Resources Research* **47**: W02515. DOI. 10.1029/2010WR009547

Boorman DB, Hollis JM, Lilly A. 1995. Hydrology of Soil Types: A Hydrological Classification of the Soils of the United Kingdom, Institute of Hydrology Report 126. Institute of Hydrology: Wallingford.

Borga M, Fontana GD, Cazorzi F. 2002. Analysis of topographic and climatic control on rainfall-triggered shallow landsliding using a quasi-dynamic wetness index. *Journal of Hydrology* **268**: 56–71.

Bracken LJ, Wainwright J, Ali G, Roy AG, Smith MW, Tetzlaff D, Reaney S. 2013. Concepts of hydrological connectivity: research approaches, pathways and future agendas. *Earth-Science Reviews* **119**: 17–34.

Byrt T, Bishop J, Carlin JB. 1993. Bias, prevalence and kappa. *Journal of Clinical Epidemiology* **46**: 423–429.

Cappus P. 1960. Etude des lois de l'écoulement. Application au calcul et à la prévision des débits. *La Houille Blanche* **A**: 493–518.

Cohen J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* **20**: 37–46.

Cover TM, Thomas JA. 1991. Elements of Information Theory. John Wiley & Sons: New York.

Duan JF, Miller NL. 1997. A generalized power function for the subsurface transmissivity profile in topmodel. *Water Resources Research* **33**: 2559–2562.

Dunn SM, Lilly A. 2001. Investigating the relationship between a soils classification and the spatial parameters of a conceptual catchment-scale hydrological model. *Journal of Hydrology* **252**: 157–173.

Dunne T, Black RD. 1970a. An experimental investigation of runoff production in permeable soils. *Water Resources Research* **6**: 478–490.

Dunne T, Black RD. 1970b. Partial area contributions to storm runoff in a small New England watershed. *Water Resources Research* **6**: 1296–1311.

ESRI 2008. ArcGIS 9.3. Environmental Systems Research Institute (ESRI): Redlands, CA.

Franks SW, Gineste P, Beven KJ, Merot P. 1998. On constraining the predictions of a distributed model: the incorporation of fuzzy estimates of saturated areas into the calibration process. *Water Resources Research* **34**: 787–797.

Gómez-Plaza A, Martínez-Mena A, Albaladejo J, Castillo VM. 2001. Factors regulating spatial distribution of soil water content in small semiarid catchments. *Journal of Hydrology* **253**: 211–226.

Grabs T, Seibert J, Bishop K, Laudon H. 2009. Modeling spatial patterns of saturated areas: a comparison of the topographic wetness index and a dynamic distributed model. *Journal of Hydrology* **373**: 15–23.

Grayson RB, Western AW, Chiew FHS, Blöschl G. 1997. Preferred states in spatial soil moisture patterns: local and nonlocal controls. *Water Resources Research* **33**: 2897–2908.

Güntner A, Seibert J, Uhlenbrook S. 2004. Modeling spatial patterns of saturated areas: an evaluation of different terrain indices. *Water Resources Research* **40**: W05114. DOI. 10.1029/2003WR002864

Haitjema HM, Mitchell-Bruker S. 2005. Are Water tables a subdued replica of the topography? *Ground Water* **43**: 781–786.

Hjerdt KN, McDonnell JJ, Seibert J, Rodhe A. 2004. A new topographic index to quantify downslope controls on local drainage. *Water Resources Research* **40**: W05602.

Hosmer DW, Lemeshow S. 1989. Applied Logistic Regression. John Wiley & Sons: New York.

Hrachowitz M, Soulsby C, Tetzlaff D, Dawson JJC, Malcolm IA. 2009. Regionalization of transit time estimates in montane catchments by integrating landscape controls. *Water Resources Research* **45**: W05421. DOI. 10.1029/2008WR007496

Hrachowitz M, Soulsby C, Tetzlaff D, Speed M. 2010. Catchment transit times and landscape controls – does scale matter? *Hydrological Processes* **24**: 117–125.

James AL, Roulet NT. 2007. Investigating hydrologic connectivity and its association with threshold change in runoff response in a temperate forested watershed. *Hydrological Processes* **21**: 3391–3408.

Jenness J. 2006. Topographic Position Index (tpi_jen.avx) Extension for ArcView 3.x. Jenness Enterprises. http://www.jennessent.com

Kirkby M. 1975. Hydrograph modelling strategies. In Processes in Physical and Human Geography, Peel R, Chisholm M, Haggett P (eds). Butterworth-Heinemann: Woburn, MA; 69–90.

Lanni C, Rigon R, McDonnell JJ. 2011. On the relative role of upslope and downslope topography for describing water flow path and storage dynamics. *Hydrological Processes* **25**: 3909–3923. DOI. 10.1002/hyp.8263

Malcolm IA, Soulsby C, Youngson AF. 2006. High frequency logging technologies reveal state dependence of hyporheic process dynamics: implications for hydroecology. *Hydrological Processes* **20**: 615–622.

McBratney AB, Mendonça Santos ML, Minasny B. 2003. On digital soil mapping. *Geoderma* **117**: 3–52.

McNamara JP, Chandler D, Seyfried M, Achet S. 2005. Soil moisture states, lateral flow, and streamflow generation in a semi-arid, snowmelt-driven catchment. *Hydrological Processes* **19**: 4023–4038.

Montgomery DR, Dietrich WE. 2002. Runoff generation in a steep, soil-mantled landscape. *Water Resources Research* **38**(9): 1168. DOI. 10.1029/2001WR000822.

Moore ID, Grayson RB, Ladson AR. 1991. Digital terrain modeling – a review of hydrological, geomorphological, and biological applications. *Hydrological Processes* **5**: 3–30.

Moore ID, Norton TW, Williams JE. 1993. Modelling environmental heterogeneity in forested landscapes. *Journal of Hydrology* **150**: 717–747.

Moore RJ, Bell VA, Cole SJ, Jones DA. 2007. Rainfall–Runoff and Other Modelling for Ungauged/Low-benefit Locations. Research Contractor: CEH Wallingford, Environment Agency: Bristol.

O'Loughlin EM. 1986. Prediction of surface saturation zones in natural catchments by topographic analysis. *Water Resources Research* **22**: 794–804.

Ogden FL, Watts BA. 2000. Saturated area formation on nonconvergent hillslope topography with shallow soils: a numerical investigation. *Water Resources Research* **36**: 1795–1804.

Parker AJ. 1982. The topographic relative moisture index: an approach to soil-moisture assessment in mountain terrain. *Physical Geography* **3**: 160–168.

Quinn P, Beven K, Chevallier P, Planchon O. 1991. The prediction of hillslope flow paths for distributed hydrological modeling using digital terrain models. *Hydrological Processes* **5**: 59–79.

Rinderer M, Kollegger A, Fischer BMC, Stähli M, Seibert J. 2012. Sensing with boots and trousers – qualitative field observations of shallow soil moisture patterns. *Hydrological Processes* **26**: 4112–4120.

Robson A, Beven K, Neal C. 1992. Towards identifying sources of subsurface flow – a comparison of components identified by a physically based runoff model and those determined by chemical mixing techniques. *Hydrological Processes* **6**: 199–214.

Seibert J, McGlynn B. 2007. A new triangular multiple flow direction algorithm for computing upslope areas from gridded digital elevation models. *Water Resources Research* **43**: W04501. DOI. 10.1029/2006WR005128

Seibert J, Stendahl J, Sorensen R. 2007. Topographical Influences on soil properties in Boreal Forests. *Geoderma* **141**: 139–148.

Shand P, Haria AH, Neal C, Griffiths KJ, Gooddy DC, Dixon AJ, Hill T, Buckley DK, Cunningham JE. 2006. Hydrochemical heterogeneity in an upland catchment: further characterisation of the spatial, temporal and depth variations in soils, streams and groundwaters of the Plynlimon forested catchment. *Hydrology Earth System Sciences* **9**: 621–644.

Sivapalan M, Wood EF, Beven KJ. 1990. On hydrologic similarity: 3. A dimensionless flood frequency model using a generalized geomorphologic unit hydrograph and partial area runoff generation. *Water Resources Research* **26**: 43–58.

Sorensen R, Zinko U, Seibert J. 2006. On the calculation of the topographic wetness index: evaluation of different methods based on field observations. *Hydrology and Earth System Sciences* **10**: 101–112.

Soulsby C, Malcolm IA, Youngson AF, Tetzlaff D, Gibbins CN, Hannah DM. 2005. Groundwater–surface water interactions in upland scottish rivers: hydrological, hydrochemical and ecological implications. *Scottish Journal of Geology* **41**: 39–49.

Tarboton DG. 1997. A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research* **33**: 309–319.

Tarboton DG. 2010. Terrain Analysis Using Digital Elevation Models (TauDEM) 5.0. http://hydrology.neng.usu.edu/taudem/

Tarolli P, Borga M, Fontana GD. 2008. Analysing the influence of upslope bedrock outcrops on shallow landsliding. *Geomorphology* **93**: 186–200.

Tarolli P, Borga M, Chang KT, Chiang SH. 2011. Modeling shallow landsliding susceptibility by incorporating heavy rainfall statistical properties. *Geomorphology* **133**: 199–211.

Tetzlaff D, Soulsby C, Waldron S, Malcolm IA, Bacon PJ, Dunn SM, Lilly A, Youngson AF. 2007. Conceptualization of Runoff Processes Using a Geographical Information System and Tracers in a Nested Mesoscale Catchment. *Hydrological Processes* **21**: 1289–1307.

Tromp-Van Meerveld HJ, McDonnell JJ. 2006. Threshold relations in subsurface stormflow: 2. The fill and spill hypothesis. *Water Resources Research* **42**: W02411. DOI. 10.1029/2004WR003800

Van de Grift J. 2006. The Relative Site Moisture Index: An Expansion of the Topographic Relative Moisture Index. Dissertation, University of Wisconsin-Madison, Madison, WI.

Weiss AD. 2001. Topographic Position and Landforms Analysis. Dissertation, ESRI User Conference, Poster presentation. San Diego, CA.

Western A, Grayson R. 2001. Soil moisture and runoff processes at Tarrawarra. In Spatial Patterns in Catchment Hydrology, Grayson R, Blöschl G (eds). Cambridge University Press: Cambridge; 209–246.

Western AW, Blöschl G, Grayson RB. 2001. Toward capturing hydrologically significant connectivity in spatial patterns. *Water Resources Research* **37**: 83–97.

Western AW, Grayson RB, Blöschl G, Willgoose GR, McMahon TA. 1999. Observed spatial organization of soil moisture and its relation to terrain indices. *Water Resources Research* **35**: 797–810.

Yair A, Raz-Yassif N. 2004. Hydrological processes in a small and catchment: scale effects of rainfall and slope length. *Geomorphology* **61**: 155–169.