Hydrol. Earth Syst. Sci. Discuss., 10, 14041–14093, 2013
www.hydrol-earth-syst-sci-discuss.net/10/14041/2013/
doi:10.5194/hessd-10-14041-2013

Hydrology and Earth System Sciences
Open Access
Discussions

Evaluating topographic wetness indices across central New York agricultural landscapes

B. P. Buchanan[1], M. Fleming[1], R. L. Schneider[2], B. K. Richards[1], J. Archibald[1], Z. Qiu[3], and M. T. Walter[1]

[1]Department of Biological and Environmental Engineering, Cornell University, Ithaca, NY, USA
[2]Department of Natural Resources, Cornell University, Ithaca, NY, USA
[3]Department of Chemistry and Environmental Science, New Jersey Institute of Technology, Newark, NJ, USA

Received: 4 September 2013 – Accepted: 16 October 2013 – Published: 18 November 2013

Correspondence to: B. P. Buchanan (bb386@cornell.edu)

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

Abstract

Accurately predicting soil moisture patterns in the landscape is a persistent challenge. In humid regions, topographic wetness indices (TWI) are widely used to approximate relative soil moisture patterns. However, there are many ways to calculate TWIs and very few field studies have evaluated the different approaches in the US. We calculated TWIs using over 400 unique formulations that considered different: Digital Elevation Model (DEM) resolution (cell size), vertical precision of DEM, flow direction and slope algorithms, smoothing via low-pass filtering, and the inclusion of relevant soil properties. We correlated each TWI with observed patterns of soil moisture at five agricultural fields in central NY, USA; each field was visited 5–8 times between August and November 2012. Using a mixed effects modeling approach, we were able to identify optimal TWI formulations that may provide guidance for practitioners and future studies. Overall, TWIs were moderately well correlated with observed soil moisture patterns; in the best case the relationship between TWI and soil moisture had an average R^2 and Spearman correlation value of 0.61 and 0.78, respectively. In all cases, fine-scale (3 m) LiDAR-derived DEMs worked better than USGS 10 m DEMs and, in general, including soil properties improved the correlations.

1 Introduction

Soil moisture is a key variable controlling a host of important hydrological and biogeochemical processes and, thus, imposes a considerable ecohydrological fingerprint on the landscape. For instance, patterns of soil moisture correlate well with the spatial distribution of: storm runoff, soil properties, nutrient cycling and species composition and richness of plants and wildlife. Many of these processes and attributes have implications for land management, especially in agricultural landscapes where activities to maximize agricultural production should be balanced with decisions that will mitigate nonpoint-source (NPS) pollution. Over the years, numerous researchers have

HESSD
10, 14041–14093, 2013

proposed techniques to better describe and predict the spatial distribution of soil water (e.g., Zhao et al., 1980; Jackson, 1993; Larson et al., 2008; Mallick et al., 2009; Sayde et al., 2010). Perhaps the two most common approaches involve: (i) often complex, distributed watershed models that numerically simulate the physical processes governing soil water dynamics or (ii) more simple terrain-based indices based on topography and sometimes soil properties.

The detailed numerical approach is typically incorporated into distributed hydrologic modeling frameworks and has been shown to provide reasonable simulations of soil moisture patterns (Frankenberger et al., 1999; Motovilov et al., 1999; Mehta et al., 2004; Cuo et al., 2006). However, such models often require extensive data input and calibration, are generally prohibitively complex for conservation planners to use (Lane et al., 2006; White et al., 2010) and frequently suffer from equifinality issues (Beven, 2006).

Terrain indices offer a simpler alternative that, due to their parsimonious formulation and moderate parameterization requirements, can be efficiently applied at larger spatial scales while maintaining a relatively fine spatial resolution. Such indices are typically applied via their cumulative distribution functions, which afford the estimation of total contributing area, as well as the spatial distribution of saturation deficit (or soil moisture) (Western et al., 1999). This facilitates both continuous- and event-based hydrologic predictions as well as targeted environmental management decisions. Although terrain indices can include primary terrain attributes such as curvature, slope or aspect; here we focus on so-called compound terrain derivatives that synthesize several primary indices as they are generally better correlated with observed soil moisture patterns (Moore et al., 1988, 1991; Western et al., 1999).

The most well-known and widely applied compound terrain derivative in hydrology and ecology is the topographic wetness index (TWI) originally proposed by Beven and Kirby (1979). Computed as $\ln(\alpha/\tan\beta)$, where α is the upslope contributing area per unit contour length and $\tan\beta$ is the local slope, the index provides a relative, not absolute, measure of the moisture status of a particular area or pixel. Since its introduction,

HESSD
10, 14041–14093, 2013

the TWI concept has been integrated into many popular hydrologic models (e.g., TOPMODEL, Beven and Kirby, 1979; VSLF, Schneiderman et al., 2007; SWAT-VSA, Easton et al., 2008) and pollution risk indices (Agnew et al., 2006; Reaney et al., 2011; Marjerison et al., 2011; Buchanan et al., 2013). Despite its wide application, large-scale corroboration of TWI-based predictions of landscape-scale soil moisture patterns using actual field observations are the exception rather than the rule. Indeed, most previous empirical validation efforts have focused on collecting high density field observations over very small spatial scales – typically individual hillslopes, fields or plots. Interestingly, these studies have found a wide variety of correlation strengths – with R^2 values ranging from 0–0.89 (Burt and Butcher, 1985; Moore et al., 1988; Ladson and Moore, 1992; Jordan, 1994; Schmidt and Persson, 2003; Western et al., 2004; Tague et al., 2010), and Spearman correlation coefficients between −0.13 and 0.65 (Nyberg, 1996; Hellstrand, 2012). Though some of this variability is undoubtedly attributable to differences in the physiography, geology, climate, and vegetation of the respective study areas, the fundamental reasons for these discrepancies are generally unresolved.

In this study we are focusing on agricultural landscapes in the Northeastern US, where several researchers have concluded that variable source area (VSA) plays a central role in agricultural NPS pollution (e.g., Rossing and Walter, 1995; Frankenberger, 1996; Gburek and Sharpley, 1998; Frankenberger et al., 1999; Walter et al., 2000, 2001; Gburek et al., 2002; Czymmek et al., 2003; Agnew et al., 2006; Qiu et al., 2007; Qiu, 2010; Margerison et al., 2011), which is the leading source of regional freshwater impairment (USEPA, 2009). Risks of VSA storm runoff generation are closely correlated with soil moisture (or soil moisture deficit) (e.g., Walter et al., 2000; Agnew et al., 2006; Lyon et al., 2006a, b; Shaw and Walter, 2009; Cheng et al., 2013). Therefore, identifying effective methods for predicting patterns of soil moisture is important for developing strategies that incorporate VSA hydrology into NPS agricultural pollution mitigation strategies. TWIs are a potentially useful tool for doing this, but it is not clear from previous studies how best to calculate it or which data should be used. Until recently, the latter issue was essentially moot because there were few options.

HESSD
10, 14041–14093, 2013

In recent years, advances in geographic information systems (GIS) and an increase in the availability of high resolution light detection and ranging (LiDAR) data have resulted in detailed and potentially more realistic representations of surface topography, which is the primary data used to calculate a TWI. To some extent, discrepancies in TWI-soil moisture correlations of the previously mentioned studies may be due to variations in the accuracy of the underlying DEM data. Only a few studies have specifically examined the advantages of LiDAR-derived TWIs relative to other less precise DEM sources, such as the standard USGS 10 m DEMs (e.g., Tenenbaum et al., 2006; Murphy and Ogilvie, 2009).

In addition to vertical DEM precision and accuracy, researchers have demonstrated that the TWI is sensitive to many other factors including: DEM cell size, flow direction algorithm, slope algorithm and the inclusion of relevant soil properties. For example, in two boreal forest sites in Sweden, Sørensen et al. (2006) demonstrated that correlations between the TWI and soil moisture, groundwater depth, soil pH and plant species richness varied with the choice of slope and contributing area algorithm. Similarly, Güntner et al. (2004) found that the relationship between TWIs and mapped areas of saturation in two catchments in southwestern Germany were dependent on how slope, contributing area, soil properties and climate were incorporated into the TWI. Remarkably, despite the clear sensitivity of the TWI to these factors, no study has conducted a comprehensive and systematic evaluation in the US.

By correlating TWI maps with observed patterns of soil moisture at numerous agricultural fields in central NY, this study addresses three key research questions: (i) does the TWI provide reasonable estimates of soil water distribution in Northeast US agricultural landscapes, (ii) does that relationship hold across multiple field sites that possess moderately different topographic, land management and soils characteristics and (iii) given the myriad ways of calculating the input variables of the TWI (i.e., slope, contributing area, etc.), is there an optimal TWI formulation?

HESSD
10, 14041–14093, 2013

2 Methods

2.1 Study area

Soil moisture measurements were made at five agricultural field sites located in four different catchments in south-central New York, USA (Fig. 1). The sites are characterized by moderate slopes (4.8–6.6 %) and agricultural land uses (i.e., typically soybean, grass, corn, and fallow; Table 1). Soil types across the field sites were predominantly channery silt loams derived from siltstone, sandstone, shale, and limestone (i.e., Lanford, Eria and Erie-Ellery channery silt loams) and underlain by a shallow frangipan restrictive layer (average depth ∼ 0.4 m). Due to the low permeability of the shallow restrictive layer, soil moisture in the upper soil layer is a key variable influencing runoff generation, which is primarily a saturation-excess process in the study region (Walter et al., 2003; Easton et al., 2008).

2.2 Field data

Volumetric soil moisture readings in the upper 12 cm were collected with Time Domain Reflectometry (TDR) probes across a gradient of TWI values at each site. A minimum of three TDR readings were recorded at each sampling point and used to calculate the average point VWC for each date. All sampling points were located with GPS units (horizontal accuracy ∼ 3 m). Field sites were sampled from mid-August 2012 to the end of November 2012 (Table 2). For storms greater than 6 mm, a minimum of 24 h elapsed before collecting TDR measurements in order to allow for gravity-driven redistribution of soil moisture. All VWC measurements were normalized by the average field soil moisture for each sampling date. Consequently, all soil moisture values represent a relative measure of wetness.

Gravimetric soil moisture measurements, made on soil cores taken from each site were used to calibrate the TDR probe. The soil cores were collected across a range of wetness conditions. A calibration curve, which related TDR period and gravimetric

HESSD
10, 14041–14093, 2013

measurements, was then constructed to correct VWC readings derived from the TDR probe (R^2 = 0.82).

2.3 TWI modifications and analysis methods

We examined how the strength of the correlations between soil moisture and TWI were influenced by different combinations of the following factors: (i) inclusion of soil properties, (ii) vertical accuracy of the DEM source data, (iii) cell size of the base-DEM, (iv) slope algorithm, (v) contributing area algorithm and (vi) smoothing of the final TWI. All TWI values were extracted from the sample point coordinate data using bilinear interpolation from the four nearest grid cells. Bilinear interpolation provided a more representative estimate of the point TWI value given the 3 m horizontal accuracy of our GPS units. The overall analyses resulted in 432 unique TWI formulations. The various parameter combinations used to construct the TWIs are discussed more explicitly below. All terrain analyses were conducted with SAGA-GIS and automated via the RSAGA package in R (Brenning, 2007).

2.3.1 TWI form: STI vs. TI

Two different methods for calculating the TWI were compared: the original topographic index (TI) proposed by Beven and Kirkby (1979) and the soil-topographic wetness index (STI), which extends the purely topography-based TWI by accounting for spatial variation in hydrologically relevant soil properties (Beven, 1986). The standard TI takes the form:

$$\mathrm{TI} = \ln\left(\frac{\alpha}{\tan(\beta)}\right) \quad (1)$$

HESSD
10, 14041–14093, 2013

where α is the upslope contributing area per unit contour length and β is the slope ($m\,m^{-1}$). The STI is expressed as (Walter et al., 2002; Lyon et al., 2004):

$$\mathrm{STI} = \ln\left(\frac{\alpha}{T \cdot \tan(\beta)}\right) \quad (2)$$

where T is the soil transmissivity ($m^2\,d^{-1}$) computed as the product of the average saturated hydraulic conductivity ($m\,day^{-1}$) and the depth to restrictive layer (m); note: this is somewhat different from the STI originally proposed by Beven (1986), which assumed an exponential decrease in hydraulic conductivity with depth and the saturated hydraulic conductivity at the bottom of the soil is approximately zero. However, this way of calculating the STI has been used in several regional modeling studies and been shown to work reasonably well in the Northeast US (e.g., Agnew et al., 2006; Lyon et al., 2006a, b; Schneiderman et al., 2007; Easton et al., 2008). Soil properties were derived from the USDA-NRCS Soil Survey Geographic (SSURGO) database using the Soil Data Viewer application (USDA-NRCS, 2009).

2.3.2 Source data: USGS vs. LIDAR

TWIs were calculated based on publicly available United States Geological Survey (USGS) DEMs as well as high resolution LiDAR data. The USGS DEMs were obtained from the National Elevation Dataset (http://viewer.nationalmap.gov/viewer/) at a 1/3 arc-second (~ 10 m) resolution. Although the National Elevation Dataset does include 1/9 arc-second (~ 3 m) DEMs, they were not available for our study region. The USGS DEMs, typically derived from any of four production methods (i.e., electronic image correlation, manual profiling on stereoplotters, contour-to-grid interpolation or an improved contour-to-grid interpolation known as "LineTrace+"), possess considerably less vertical accuracy than LiDAR DEMs (root mean square error is typically 2.44 m for USGS (USGS, 2013) vs. 0.15 m for our LiDAR dataset). Although it is possible to resample a 10 m USGS DEM to higher cell resolutions, we felt this was not justified as

HESSD
10, 14041–14093, 2013

the resulting grid resolution would exceed the scale at which the original source data were derived – thereby implying an erroneous degree of accuracy in the underlying elevation data. Consequently, all USGS DEMs were evaluated at the original 10 m resolution. The LiDAR DEMs were generated at 3 m and 10 m resolutions from point cloud data of filtered ground shots (average point spacing ~ 0.67 m) via natural neighbor interpolation.

2.3.3 Cell size: 3 m vs. 10 m

Previous studies have demonstrated that terrain derivatives (e.g., slope and contributing area) and, thus, TWIs can be substantially affected by the cell resolution of the base DEM (Hasan et al., 2012; Sørensen and Seibert, 2007). Here we investigate two commonly used DEM resolutions, which are particularly relevant for high resolution distributed hydrologic and water quality modeling: 3 m vs. 10 m. The LiDAR data were interpolated to both 10 m and 3 m DEMs. The overall parameter set for this group includes: (i) 10 m LiDAR TWIs and (ii) 3 m LiDAR TWIs.

2.3.4 Slope calculation: local slope vs. downslope index

Four methods for calculating slope were compared: (i) maximum triangle slope (MTS; Tarboton, 1997), (ii) least squares fitted plane (LSFP; Horn, 1981), (iii) 2nd degree polynomial (SDP; Zevenbergen and Thorne, 1987), and (iv) the downslope index (DSI; Hjerdt et al., 2004). In contrast to MTS, LSFP and SDP, which are considered "local slope" algorithms, because they only consider the cell of interest and its neighbors, DSI is defined as the slope to the closest point that is d meters below the cell of interest. The DSI provides a potential improvement to the local-slope methods because it can account for downslope controls on local soil moisture conditions and thereby relaxes the assumption of parallelism between surface topography and groundwater tables, i.e., the kinematic approximation of water table slope. Importantly, the DSI is controlled by the distance parameter (d) that affects the degree of deviation of the hydraulic and

HESSD
10, 14041–14093, 2013

surface gradients (i.e., large values of d, result in larger downslope influences). For this study, we tested three different values of d: (i) 2 m, (ii) 5 m and (iii) 10 m. Thus, the parameter set for the slope calculation includes a total of six slope types: (i) MTS, (ii) LSFP, (iii) SDP, (iv) DSI with $d = 2$ m, (v) downslope with $d = 5$ m and (vi) downslope with $d = 10$ m. All slopes were computed as straight-line distances as opposed to flowpath distances.

2.3.5 Flow accumulation algorithm

Perhaps the parameter that influences TWI values the most is the contributing area (a) which can be calculated with a variety of different algorithms. Here, we compare six different flow direction algorithms, which, broadly speaking, can be divided into two main categories: single flow direction (SFD) and multiple flow direction (MFD). The principal difference between the single vs. multiple flow direction groups lies in how flow is apportioned to downslope cells. As the name implies, single flow direction algorithms assign all flow to a single downslope cell, whereas MFD algorithms allows flow to be split among multiple cells.

By far, the most commonly used algorithm currently is the D8 form proposed by O'Callaghan and Mark (1984) and coded as the default routine into the hydrologic toolsets of most popular GIS platforms (e.g., ArcGIS, MapWindow, QGIS). The D8 algorithm is simple and computationally efficient. D8 apportions all flow into a single downslope cell determined by the steepest gradient among eight cardinal and intercardinal directions. This may oversimplify actual flow paths, especially in convex terrain and flow-divergence areas, which may lead to incorrect representations of contributing area and flow pathways (overly straight and parallel). D8 is also very sensitive to minor elevation differences in adjacent cells and this can be exacerbated at high DEM resolutions (Park et al., 2009; Erskine et al., 2006).

The second SFD algorithm tested in this study is the randomized single-flow direction method (Rho8) proposed by Fairfield and Leymarie (1991). The method results in stochastic flowpath delineations through the incorporation of a uniformly distributed

HESSD
10, 14041–14093, 2013

random variable into the calculation of slope gradient. This alleviates the overly linear and parallel flow line issues of D8, but flow is still apportioned into a single downslope cell. Moreover, the results are not reproducible and not always physically-based due to the random factor.

The multiple flow direction (MD) approach of Freeman (1991) addresses the shortcomings of SFD approaches by allowing for flow divergence into adjacent downslope cells as a proportion of the slope gradient. The MD method results in smoother, seemingly more physically realistic flow pathways and flow accumulation patterns relative to SFD algorithms – especially in steeper terrain. The biggest drawback is that in valley bottoms and other low-lying areas, flow dispersion can be unrealistic (Costa-Cabral and Burges, 1994; Tarboton, 1997).

The D∞ method of Tarboton (1997) helps reduce the excessive flow dispersion issues of the MD by calculating slope as a function of eight triangular facets, where flow is apportioned to the two downslope cells nearest to the steepest direction weighted as a function of their distance from this direction. Although D∞ does afford multi-cell flow divergence, its restriction to only two downslope cells may become a limitation on convex hillslopes where dispersion is unrealistically confined.

The braunschweiger relief model (BR, Bauer et al., 1985) also allows flow dispersion to multiple, adjacent downslope cells, but restricts dispersion to only three cells – thereby limiting the degree of divergence, but allowing more than D∞. The proportion of flow allotted to each cell is determined by iteratively categorizing the slope direction as defined by an upslope polygon. The upslope polygon is solved for until the source cell is reached. Flow direction is then computed as a function of slope gradient and aspect of the four neighboring pixels (Park et al., 2009).

The final flow direction algorithm we evaluate is multiple triangular flow direction (MD∞). First proposed by Seibert and McGlynn (2007), MD∞ extends the D∞ approach by allowing flow dispersion into more than two downslope cells. MD∞ attempts to strike a balance between the potentially excessive flow dispersion of MD and the restrictive flow dispersion of D∞ – especially on convex slopes.

HESSD
10, 14041–14093, 2013

The parameter set for flow accumulation calculation includes six flow direction algorithms: (i) D8, (ii) Rho8, (iii) BR, (iv) D∞, (v) MD and (vi) MD∞.

2.3.6 Smoothing: filtered vs. unfiltered

High resolution DEMs tend to result in high local variations in TWI values, which may translate to unrealistically irregular predictions of soil moisture and water table depths (Hjerdt et al., 2004; Lanni et al., 2011). Low-pass digital filtering helps to smooth out anomalous local variations by averaging across a user-defined search window. Also, by averaging across non-local grid cells, filtering can potentially incorporate downslope influences, as well as "smear" the results of SFD algorithms, which may result in an intermediate level of flow dispersion. For this study, we applied a 3×3 pixel low-pass mean filter to the TWI maps. Thus, the parameter set for this section include: (i) filtered TWIs and (ii) unfiltered TWIs.

2.4 TWI performance criteria and statistical methods

A mixed effects modeling analysis was used to identify the optimal TWI formulation for the USGS and LiDAR datasets. Subsequently, the optimal models were validated against our observed data by calculating Spearman rank correlation coefficients (r_s) and coefficients of determination (R^2). Both sets of analyses are discussed in detail below.

2.4.1 Mixed effects modeling

To control for the lack of independence among sampling points and field sites (i.e., repeated measures) we applied a linear mixed effect model structure with sampling date and point I.D. as random effects. Fixed effects included the main effect, TWI form, as well as field site and sampling date. The resulting optimal model was validated to verify that the underlying statistical assumptions were not violated; homogeneity of variance was evaluated by plotting residual vs. fitted values, independence was examined by

HESSD
10, 14041–14093, 2013

plotting residuals vs. each explanatory variable, and normality of residuals was evaluated by plotting theoretical quantiles vs. standardized residuals (Q–Q plots). We also evaluated the degree of spatial autocorrelation amongst soil moisture measurements via variogram analysis and no consistent trends were observed (data not shown). We attribute the lack of significant spatial autocorrelation to the fact that our field sampling protocol was conducted using a cluster approach as opposed to linear transects or equal-interval sampling grids.

The relative performance of the 432 different TWIs were evaluated by comparing Akaike Information Criterion (AIC) values derived from the mixed effects models. The AIC is a goodness-fit-index that provides a measure of the relative as opposed to absolute fit. Thus, the AIC is intended to facilitate model comparisons from the same dataset and aids in the selection of optimal models, with lower AIC values indicating a better fitting model (Akaike, 1973, 1974).

The relative performance of the models in the six different TWI parameter-groups listed above (i.e., source data, TI form, cell size, slope algorithm, flow direction algorithm and smoothing) were evaluated by two methods: (i) comparing the mean, median and overall probability distribution of AIC values via violin plots (see Hintze and Nelson, 1998 for a detailed description of these plots) and (ii) by pairwise comparison of TWIs that share the exact same parameter values in all respects except for, of course, the particular TWI parameter in question (see Table 3 for an example).

The following generally accepted guidelines when comparing AICs were adopted for this study (Burnham and Anderson, 2002): models with AIC values within 2 of each other were not considered significantly different; AIC values within only 3–7 units of each other were considered moderately different; AIC values > 10 were considered significantly different from each other.

To facilitate the identification of the overall best fitting TWI from the entire set of models, we also calculated delta AICs (ΔAIC), Akaike weights (AICw$_i$) and evidence ratios (E-ratio). The ΔAIC is simply the difference between the i-th model and the optimal model, calculated as follows:

HESSD
10, 14041–14093, 2013

$$\Delta\mathrm{AIC} = \mathrm{AIC}_i - \mathrm{AIC}_{\mathrm{opt}} \quad (3)$$

where AIC_i is the AIC value for the ith model and $\mathrm{AIC}_{\mathrm{opt}}$ is the AIC value of the best model (minimum AIC value). Akaike weights provide an effective way to interpret the ΔAIC values by comparing the ratio of each model to the best model relative to the entire set of candidate models as follows:

$$\mathrm{AICw}_i = \frac{\exp\left(\frac{-\Delta\mathrm{AIC}_i}{2}\right)}{\sum_{k=1}^{K}\exp\left(\frac{-\Delta\mathrm{AIC}_k}{2}\right)} \quad (4)$$

given a set of K models being evaluated. Evidence ratios provide a more concise way to quantify the weight of evidence in support of one model over another and are calculated simply as the ratio of Akaike weights ($\mathrm{AICw}_{\mathrm{opt}}/\mathrm{AICw}_j$), where $\mathrm{AICw}_{\mathrm{opt}}$ is the estimated best model in the dataset, and j indexes the remaining models in the set. An evidence ratio less than or equal to three relative to another model suggests equivalence between the models (Burnham and Anderson, 2002). All statistical analyses were conducted using the "lme4" package within the R statistical programming environment (Bates et al., 2011; R Core Team, 2011).

2.4.2 Model validation

To validate the optimal models identified via the AIC analysis, we calculated r_s and R^2 values, which were averaged across field sites and sampling dates as a means for controlling for the lack of independence among soil moisture measurements (albeit more crudely than the mixed effects models). The average r_s and R^2 values not only help to evaluate the accuracy of the optimal TWIs, but also facilitate inter-study comparisons as most previous research assessed the strength of correlation between soil moisture patterns (either observed or model generated) and various TWI formulations via these two metrics.

HESSD
10, 14041–14093, 2013

3 Results and discussion

3.1 Source data: USGS vs. LIDAR

A comparison of the means and overall distributions of AIC values reveals that LIDAR-based TWIs consistently provide a better fit to the patterns of observed soil moisture than USGS-based TWIs across the full range of parameter combinations (Fig. 2). Mean AIC values differ by more than 170, with no overlap in distribution. The AIC distribution of the LiDAR dataset is substantially greater than the USGS indicating that the different parameter combinations had a greater influence on the performance of LiDAR TWIs.

The majority of other researchers have evaluated the effect of vertical DEM accuracy on terrain indices by either: (i) comparing TWIs with different vertical information contents to each other via spatial statistics (distribution functions, spatial pattern analysis; Sørensen and Seibert, 2007; Vaze et al., 2010) or (ii) by calculating topographic attributes from DEMs of varying information content and evaluating their effect on hydrologic and water quality model predictions at the basin outlet (Zhang and Montgomery, 1994; Grabs et al., 2009; Kenward et al., 2000). In general, these studies have found that higher quality vertical information results in appreciable improvements in the representation of topographic surfaces, more accurate delineations of hydrologically relevant parameters and more appropriate model outputs especially regarding spatially distributed information. To our knowledge, only two other studies, Tenenbaum et al. (2006) and Murphy and Ogilvie (2009), have used field observations to examine the potential benefits of LiDAR-based DEMs on TWI-soil moisture relationships. Murphy and Ogilvie (2009) compared field-mapped saturated areas in a 193 ha watershed in Canada with both a 1 m LiDAR-TWI and a 10 m photogrammatic-TWI and demonstrated that the LiDAR-TWI yielded better predictions of flow connectivity and overall index distributions. Tenenbaum et al. (2006), on the other hand found more equivocal results. Specifically, they showed that LiDAR-TWIs provided improved predictions of near-surface soil moisture in an urbanizing environment where refined flowpath delin-

HESSD
10, 14041–14093, 2013

eations were necessary, but not in a forested catchment where a coarser, photogrammatic DEM better captured the more generalized soil moisture patterns.

Despite some variability in results, the AIC values in this study indicate that the higher data storage costs and lower computation efficiency associated with higher resolution LiDAR data is justified by the substantial improvement in predictive ability. However, we recognize that such datasets are not always readily obtainable. Consequently, hereafter, we will analyze the LiDAR and USGS datasets separately to facilitate identification of the optimal TWI for both DEM source types.

3.2 TWI form: STI vs. TI

The distribution, mean and median of TWI formulations that incorporated SSURGO soils data correlated with soil moisture patterns better than those that did not include both the LiDAR and USGS datasets (Fig. 3). The average pairwise difference in AIC values was 31 and 23 for the LiDAR and USGS datasets, respectively. This suggests substantial improvement in predictive accuracy due to the inclusion of soils data regardless of DEM source. The pairwise comparison found only four moderate ($\Delta AIC < 10$) exceptions to this rule and only in the LiDAR dataset (Table 4). In other words, the STI fit the empirical dataset better than the TI in 428 out of 432 cases. Also, it should be noted that these four exceptions were well outside the lower quartile range of the STI group and so were not among the better forms of STI. Furthermore, all exceptions used a 10 m cell size and BR flow accumulation, which as we show later, performed relatively poorly and could therefore, be a result of a chance combination of factors.

Few studies have directly examined the benefits of including soils data in the TWI formulation via comparison with empirical data. Güntner et al. (2004) compared the ability of STIs and TIs to predict the aerial extent of saturation as defined by pedological and geobotanical mapping criteria. The incorporation of soils data was found to improve index performance only when transmissivity values were calibrated. However, the authors acknowledge that the soils data used in their study were only "rough estimates" and thus, their results are not necessarily comparable with ours.

HESSD
10, 14041–14093, 2013

Overall, our results suggest that despite the fact that SSURGO soils data are generated at coarse scales, they provide useful, hydrologically relevant information that helps to improve LiDAR- and USGS-based terrain indices at the scale of individual farm fields.

3.3 Cell size: 3 m vs. 10 m

The means and overall distributions of AIC values suggest that higher resolution 3 m TWIs provide a far better fit to observed patterns of soil moisture relative to 10 m LiDAR-based TWIs (Fig. 4). Additionally, the 3 m TWIs outperformed the 10 m across all pairwise comparisons, with an average pairwise AIC difference of over 55. The better correlations to soil moisture with the 3 m TWIs indicate that the added data storage and associated computation costs of the 3 m dataset may be warranted. This is somewhat in contrast with the findings of Zhang and Montgomery (1994) who argue that terrain derivatives computed from 10 m DEMs provide a reasonable compromise between complexity and accuracy. However, their study was conducted almost two decades ago, when data storage limitations and processing rates were more of a concern. Further, their study evaluated the appropriateness of DEM resolution via comparisons of frequency distributions and TOPMODEL predictions, but no field data were used. Sørensen and Seibert (2007), on the other hand, found substantial differences between 5 m and 10 m TWI grids – though they do not necessarily recommend one over the other and instead point out that the appropriate resolution may be dependent on the particular terrain feature or hydrologic characteristic in question.

We attribute the consistently better performance of the 3 m-based TWI-values primarily to more accurate and discrete delineations of flow pathways which, based on our field observations were quite small (often on the order of 1–4 m). According to our field observations, these micro-topographical features exert a considerable influence on downslope soil moisture distribution. However, these features occur at scales much finer than 10 m-based TWIs and are therefore often not captured appropriately. Note, however, that Wolock and Price (1994) found that groundwater table surfaces may be

HESSD
10, 14041–14093, 2013

better represented by coarser DEMs as they result in smoother, more realistic predictions, though it is not obvious that this is strongly comparable to soil moisture.

3.4 Slope calculation: local slope vs. downslope index

3.4.1 LiDAR

The least squares fitted plane (LSFP), maximum triangular slope (MTS) and 2nd degree polynomial (SDP) methods, as well as the downslope index with the distance parameter set to 2 m (DSI-2m) outperformed the downslope index set to 5 and 10 m (DSI-5m, DSI-10m, respectively) across the LiDAR dataset (including both 3 and 10 m cell sizes; Fig. 5). Also, although the means and distributions of the best performing slope algorithms were similar, the maximum triangular slope resulted in the best-fitting TWI and did moderately better than the next best method with a mean pairwise AIC difference of 4. There were, however, 35 cases where the DSI-2m fitted the LiDAR data better (see Buchanan, 2013 for a table of the specific TWIs) and the DSI-2m possessed a lower mean AIC value. This suggests there may be little difference between the two methods. The generally better performance of the local slope algorithms for the LiDAR dataset is consistent with Günter et al. (2004) and Sørensen et al. (2006) who found that the local slope achieved a higher correlation with observed patterns of soil moisture, wetness degree and groundwater depth than the DSI.

3.4.2 USGS

In contrast to the LiDAR dataset, the downslope index with a *d* parameter set to 5 m resulted in the best-fitting USGS TWIs, while the all three local slope algorithms resulted in the worst performing TWIs (Fig. 5). The mean pairwise difference in AIC values for the top two groups was 9, indicating a moderate advantage to using the DSI-5m vs. DSI-2m. Additionally, there were no exceptions to this across the pairwise comparisons.

HESSD
10, 14041–14093, 2013

The likely explanation for the stronger performance of the non-local downslope index in the USGS dataset is that USGS DEMs provide a more generalized representation of the actual topography and therefore emphasize coarser-scale terrain characteristics such as the transition from hilltop to valley bottom. Conversely, LiDAR DEMs capture more subtle micro-topographical features such as small surface depressions and drainages that exert a much smaller influence on upslope drainage conditions. Thus, the DSI is likely be overly sensitive to the highly varied terrain surfaces of LiDAR DEMs and may therefore, lead to erroneous soil moisture predictions upslope of these small features. The DSI is probably more applicable to coarser USGS DEMs that better capture large-scale surface forms which are likely to exert an appreciable effect on upslope drainage. This has important implications for non-local slope algorithms – suggesting the need for an additional parameter that adjusts for the scale of topographic features, such that larger hillslope transitions are emphasized while very small topography is de-emphasized.

3.5 Flow accumulation

3.5.1 LiDAR

The mean and overall distributions of AIC values suggest that the multiple flow direction algorithms fit patterns of observed soil moisture much better than single flow direction when using 3 and 10 m LiDAR-derived TWIs (Fig. 6). The average AIC difference in the SFD vs. MFD groups was roughly 27, which highlights the rather substantial advantage of using MFD. Even so, there was very little difference in the performance level amongst the MFD groups when using LiDAR data (i.e., the means of the top three MFD groups were essentially equal). The only real exception was the BR algorithm of Bauer et al. (1985), which performed considerably worse than the other MFD formulations. The BR method results in similar index values to D∞, MD and MD∞ in upland areas, but much lower values in drainages and low-lying areas (data not shown). These small drainages and subtle convergent zones were important features in the sites used

HESSD
10, 14041–14093, 2013

in this study. Their omission by the BR algorithm is likely the root cause of its poor performance relative to the other MFD algorithms. Overall, the MD∞ of Seibert and McGlynn (2007) achieved the lowest AIC value, suggesting it produced the best-fitting model. However, the mean pairwise difference between MD∞ and the next best group was < 2 AICs and there were over 25 pairwise exceptions (see Buchanan, 2013 for table of the specific TWIs).

Similar to our findings, Güntner et al. (2004), Sørensen et al. (2006), and Park et al. (2009) showed that MFD algorithms resulted in considerably higher correlations with observed soil moisture patterns than SFD. Nevertheless, Park et al. (2009) and Erskine et al. (2006) showed that the relative differences between the SFD and MFD groups were inversely related to cell size, which indicates an interaction between our analysis of cell size and flow contributing algorithms. In particular, Park et al. (2009) found that beyond 20 m cell sizes the performance of single and multiple flow direction algorithms tended to converge. To investigate this potential interaction, we plotted the AIC distributions for each flow accumulation scheme for each cell size in the LiDAR dataset (Fig. 7). It is evident from Fig. 7, that the multi-flow direction algorithms provide better model fits across the range of tested cell sizes. However, the AIC difference between the means of the SFD and MFD groups declines from 35 to 18 when going from a 3 m to 10 m grid size, corroborating the idea that MFD performance declines inversely with cell size. As Erskine et al. (2006) points out, single- and multiple-direction algorithms are most similar in flow convergence zones (e.g., valley bottoms) and as cell size increases, the "percentage of the total drainage area classified in the lower region [convergent areas] increases" – yielding more and more similarity in the represented topography with increasing cell size. Interestingly, Sørensen et al. (2006), Endreny and Wood (2003) and Güntner et al. (2004) used raster DEMs with grid sizes greater than the 20 m similarity threshold identified by Park et al. (2009) and yet still found substantial differences between SFD and MFD algorithms. This discrepancy may be explained by differing vertical accuracies in the base DEMs between studies or differences in the

HESSD
10, 14041–14093, 2013

topography of their unique study sites. Regardless, the lack of inter-study agreement warrants further, more systematic investigation.

3.5.2 USGS

Similarity in the means of the single- and multiple-direction flow accumulation methods suggests the choice of algorithms is not as consequential when using coarser USGS DEMs (Fig. 8). Interestingly, however, when examining the best performing TWIs from each flow direction type, a trend appears that is the reverse of the LiDAR dataset. Namely, that single- as opposed to multiple-direction formulations performed systematically better. It is important to note, however, that the average pairwise AIC difference in the top two groups with the lowest AIC values (i.e., D8 and Rho8) was only 3.5 and, additionally, there were 14 minor to significant exceptions where another flow accumulation algorithm outperformed the D8 in pair-wise comparisons (see Buchanan, 2013 for table of the specific TWIs).

Importantly, the SFD algorithms achieved their best fit to the empirical data only when the TWIs were smoothed via low-pass filtering (Fig. 9). When the TWIs remained un-filtered, MD achieved the best AIC ranking. By smoothing the SFD-based indices, filtering effectively introduces flow dispersion, suggesting that an intermediate level of dispersion may be desirable when using lower resolution USGS DEMs. Indeed, numerous other studies conducted using coarse elevation models, have concluded that an intermediate approach between the SFD and MFD methods achieved the most realistic flow distribution patterns (i.e., Holmgren, 1994; Tarboton, 1997; Endreny and Wood, 2003; Güntner et al., 2004). The implications of the filtering effects are discussed in more detail in the following section.

HESSD
10, 14041–14093, 2013

3.6 Smoothing: filtered vs. unfiltered

3.6.1 LiDAR

The means and overall frequency distributions of the filtered vs. unfiltered TWIs in the LiDAR dataset indicate little advantage to either method (Fig. 10). Even so, unfiltered LiDAR TWIs did achieve the lowest AIC values and the mean pairwise difference between the filtered and unfiltered groups was 7 suggesting a moderate benefit to unsmoothed LiDAR-based TWIs. Although this finding is not strongly supported by our data, it is in direct contrast to Lanni et al. (2011) who demonstrated that their dynamic topographic index, calculated using high resolution (2 m) LiDAR DEMs, performed better when smoothed via a 3×3 low-pass filter. The lack of agreement between our studies may be due to the fact that we employed clustered empirical field data whereas Lanni et al. (2011) evaluated TWI performance via cell-by-cell comparison with a physically based Boussinesq model across a 3.2 km^2 watershed. Additionally, our study sites were characterized by relatively moderate slopes with similar mid-slope topographic positions. In contrast, Lanni et al.'s (2011) study watershed was characterized by varied, high relief terrain including bottom, middle and top of hillslopes. The highly accurate terrain surfaces derived from the unfiltered LiDAR DEMs correctly captured small-scale terrain heterogeneities in our study fields that likely played an important role in determining the direction of runoff into neighboring cells, but may not have exerted a strong upslope influence. The hillslope-scale features examined in Lanni et al.'s (2011) study, which emphasize coarser-scale soil moisture dynamics, were likely better represented by TWIs that incorporated non-local topographic information (i.e. are filtered).

3.6.2 USGS

Unlike the LiDAR dataset, filtering the USGS-TWIs resulted in a substantial improvement in model fit (Fig. 10). The average pairwise difference in AIC values between

HESSD
10, 14041–14093, 2013

filtered vs. unfiltered TWIs was over 17, suggesting substantial improvement in predictive accuracy due to the smoothing of predicted wetness surfaces. Filtering the USGS TWIs likely improves their predictive ability for the same reasons that the downslope index did – because it helps to account for downslope controls on local drainage status which are more appropriately captured by coarser DEMs. In other words, filtering averages out the effects of local anomalies and also incorporates a measure of non-local topographic effects, which results in a smoother, more contiguous, more realistic surface at larger hillslope scales.

3.7 Best overall model

The top ten best performing TWI formulations for both the LIDAR and USGS TWIs are presented in Table 8. Model 1a possessed the lowest AIC value and the highest AIC weight (5917 and 0.39, respectively), indicating that it was the best model among the set of tested models in the LiDAR Dataset (Table 8). However, models 2a–4a all possess evidence ratios of three or less which provides little evidence that model 1a is in fact notably better than models 2a–4a. From this we can conclude that when dealing with LiDAR data, the best TWI formulation will: (i) incorporate soils data, (ii) be interpolated to fine grid resolutions, less than 10 m, (iii) utilize a local slope algorithm such as LSFP, SDP or MTS as opposed to the DSI, (iv) employ a multiple flow direction algorithm such as MD∞ or D∞ and (v) remain unfiltered.

When dealing with coarser USGS-based TWIs, our results suggest that Models 1b and 2b are roughly equal in terms of their fit to observed moisture patterns (i.e., evidence ratios ≤ 3) (Table 8). Thus, the optimal parameter set when using USGS TWIs will: (i) incorporate soils data, (ii) utilize a slope algorithm that accounts for downslope controls such as DSI set to an intermediate d parameter (e.g., 5 m), (iii) employ a single flow direction algorithm and (iv) importantly, be smoothed via low-pass filtering.

For comparative purposes we have also included a plot of the best performing LiDAR and USGS TWIs vs. observed soil moisture for each sample date, along with the associated R^2 value (Figs. 11 and 12, respectively). The mean R^2 across all sampling

HESSD
10, 14041–14093, 2013

dates for the LiDAR dataset was roughly 0.61, i.e., approximately 61 % of the variation in soil moisture was explained by this TWI. However, the best-performing USGS TWI only explained roughly 32 % of the soil moisture variation on average. Higher R^2 values for both the USGS and LiDAR datasets were achieved when the soil moisture readings were binned, according to their TWI values, into integer categories or wetness classes similar to Schneiderman et al. (2007) and Easton et al. (2008). Binning the TDR readings has the effect of averaging over larger spatial scales, which helps to reduce the effect of anomalous TDR readings, thus improving soil moisture-TWI correlations. For example, by binning the TDR readings into TWI-integer classes, the mean R^2 values increased substantially from 0.61 to 0.79 for the LiDAR TWIs and from 0.32 to 0.72 for the USGS TWIs.

The majority of studies conducted prior to the year 2000 found coefficients of determination that seldom exceeded 0.5, and typically ranged from 0–0.4 (e.g., Burt and Butcher, 1985; Moore et al., 1988; Ladson and Moore, 1992; Jordan, 1994; Western et al., 1999, 2004; Tague et al., 2010). Notably, much like the USGS DEMs used in this research, these older studies generally used DEMs derived from lower-quality elevation data. The fact that the range of R^2 values from our USGS TWI is consistent with those of the older studies implies that elevation accuracy may have played a strong role in limiting predictive ability.

Taking advantage of the availability of higher quality elevation data, several more recent studies have reported improved soil moisture-TWI correlations. For instance, using a 5 m LiDAR-derived TWI, Tague et al. (2010) demonstrated an average R^2 value of 0.74 across two experimental plots in MD, USA. Likewise, Tenenbaum et al. (2006), Sulebak et al. (2000), and Schmidt and Persson (2003) used high resolution TWIs (< 10 m) derived from high resolution elevation source data (not necessarily LiDAR), and found R^2 values ranging from 0.51 to 0.87.

The average Spearman coefficients corroborate the R^2 and AIC analyses (Fig. 13a). The upper range of Spearman values observed in this study (i.e., 0.7–0.78) were comparable with those of Sørensen et al. (2006), but are considerably higher than Tromp-

HESSD
10, 14041–14093, 2013

van Meerveld and McDonnell (2006) and Cantón et al. (2004). Note, both the Spearman vs. AIC analyses generally indicated the similar performance for the TWIs, e.g., high Spearman coefficients were strongly correlated with low AICs (Fig. 13b). Accordance of Spearman and AIC values lends credence to our findings and statistical methods and helps to facilitate comparisons with other studies that employed the Spearman metric.

Despite the fact that we were able to demonstrate good correlations between LiDAR-derived TWIs and observed soil water patterns, which were consistent with those of other more recent research, on average 40 % of the variation remained unexplained. This is, perhaps, unsurprising considering that these simple indices are overlooking several other well-proven factors that influence the spatial distribution of soil water. The effect of ET on soil moisture is particularly influential and varies based on vegetation type, aspect, and solar radiation; to name a few factors that are not included in the TWI indices. Another factor that may account for the discrepancies between TWIs and measured soil moisture is the inherently different scales between the base data used to generate TWIs and the scale at which the TDR probe measures soil moisture, even with multiple measurements to characterize a sampling point.

Moreover, soil moisture dynamics are known to change not only through space, but also through time. Nevertheless, a core assumption of the TWIs examined here, and in most other research, is that of steady-state, wherein time-dependent storage terms are neglected. As pointed out by Barling et al. (1994), rainstorms will rarely be of sufficient depth or duration to achieve steady-state subsurface flow. To address this issue, several researchers have explored more dynamic topographic indices which relax the steady-state assumption (e.g., Barling et al., 1994; Wilson et al., 2005). Although these may offer some improvements in terms of physical realism over the standard TWI, the dynamic and quasi-dynamic indices have yet to be widely adopted or well-tested beyond their original papers. Additionally, these more advanced conceptualizations require considerably more input data and are sufficiently complex to start blurring the line between what constitutes a distributed hydrological model and a wetness index.

HESSD
10, 14041–14093, 2013

The simplicity of the standard TWI is really at the heart of its popularity and yet this simplicity also leads to variable results when trying to represent dynamic processes via a static index.

4 Conclusions

We identified some notable differences among different formulations of TWIs and their correlation to spatial patterns of soil moisture in agricultural settings in central, NY. Most importantly, we found that some TWI-forms correlate relatively well with soil moisture. Our principal findings include:

- LiDAR-derived TWIs achieved good correlations with observed patterns of soil moisture in fields in northeastern US.
- LiDAR-derived TWIs achieved appreciably better correlations than USGS-based TWIs. Thus, when given a choice between using LiDAR or USGS DEMs for constructing TWI maps, we recommend the former.
- TWIs that include soil transmissivity (STI) work better than the simpler TI (we used the SSURGO dataset and calculated transmissivity as the product of the soil depth to a restrictive layer and average saturated hydraulic conductivity of that soil).
- The optimal formulation for a LiDAR TWI will:
  - Use a fine resolution DEM (we used 3 m).
  - Use the Maximum Triangular Slope algorithm to compute slope (Tarboton, 1997).
  - Use the Multiple Triangular Flow Direction algorithm (Seibert and McGlynn, 2007) to compute flow accumulation values.

HESSD
10, 14041–14093, 2013

- Not apply a low-pass smoothing filter (we used a 3 × 3 low-pass filter).

- The optimal formulation for a USGS TWI will:
  - Use the Downslope Index (Hjerdt et al., 2004) with a *d* parameter set to 5 m to compute slope (Tarboton, 1997).
  - Use the D8 Flow Direction algorithm (O'Callaghan and Mark, 1984) to compute flow accumulation values.
  - Smooth via a 3 × 3 low-pass filter.

Despite the encouraging LiDAR-based TWI-soil moisture correlations observed in this study, on average, roughly 40 % of the variation in soil moisture remained unexplained by the TWI. This is perhaps unsurprising considering we were attempting to describe an inherently dynamic process with a static index. Future studies may want to evaluate the cost-effectiveness (in terms of complexity and computational efficiency) of other TWI formulations, which either relax steady-state assumptions (e.g., Lanni et al., 2011), incorporate a measure of spatio-temporal variations in evapotranspiration (Ludwig and Mauser, 2000) or account for other terrain attributes such as aspect or time-variable channel initiation thresholds (Xiande et al., 2002; Gomez-Plaze et al., 2001; Kim and Lee, 2004; respectively).

References

Agnew, L. J., Lyon, S., Gérard-Marchant, P., Collins, V. B., Lembo, A. J., Steenhuis, T. S., and Walter, M. T.: Identifying hydrologically sensitive areas: bridging the gap between science and application, J. Environ. Manage., 78, 63–76, 2006.

Akaike, H.: Information theory and an extension of the maximum likelihood principle, in: Second International Symposium on Information Theory, vol. 1, edited by: Petrov, B. N. and Csaki, F., Akademiai Kiado, Budapest, 267–281, 1973.

Akaike, H.: A new look at the statistical model identification, IEEE T. Automat. Contr., 19, 716–723, doi:10.1109/TAC.1974.1100705, 1974.

HESSD
10, 14041–14093, 2013

Barling, R. D., Moore, I. D., and Grayson, R. B.: A quasi-dynamic wetness index for characterising the spatial distribution of zones of surface saturation and soil water content, Water Resour. Res., 30, 1029–1044, 1994.
Bates, D., Maechler, M., and Bolker, B.: lme4: linear mixed-effects models using S4 classes, edited by: Bates, D., Maechler, M., and Bolker, B., Comprehensive R Archive Network, available at: http://cran.r-project.org/package=lme4 (last access: May 2013), 2011.
Bauer, J., Rohdenburg, H., and Bork, H. R.: Ein digitales reliefmodell als vorraussetzung fuer ein deterministisches modell der wasser- und stofffluesse, in: Parameteraufbereitung fuer deterministische Gebietswassermodelle, Grundlagenarbeiten zu Analyse von Agrar-Oekosystemen, edited by: Bork, H. R. and Rohdenburg, H., Technische Universitat Braunschweig, Braunschweig, Germany, 1–15, 1985.
Beven, K. J.: A manifesto for the equifinality thesis, J. Hydrol., 320, 18–36, doi:10.1016/j.jhydrol.2005.07.007, 2006.
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.
Brenning, A.: RSAGA: SAGA geoprocessing and terrain analysis in R, available at: http://cran.r-project.org/web/packages/RSAGA/index.html (last access: May 2013), 2007.
Buchanan, B. P.: Variable source area hydrology, artificial drainage and non-point source pollution in northeastern US agricultural landscapes, Ph.D. thesis, Department of Natural Resources, Cornell University, Ithaca, NY, 2013.
Burnham, K. P. and Anderson, D. R.: Model Selection and Multimodel Inference: a practical information-theoretic approach, 2nd Edn., Springer, New York, 2002.
Burt, T. P. and Butcher, D. P.: Topographic controls of soil moisture distributions, J. Soil Sci., 36, 469–486, 1985.
Cantón, Y., Solé-Benet, A., and Domingo, F.: Temporal and spatial patterns of soil moisture in semiarid badlands of SE Spain, J. Hydrol., 285, 199–214, doi:10.1016/j.jhydrol.2003.08.018, 2004.
Cheng, X., Dahlke, H. S., Shaw, S., Marjerison, R., Yearick, C., and Walter, M. T.: Improving risk estimates of runoff producing areas: formulating variable source areas as a bivariate process, Hydrol. Process., in review, 2013.
Costa-Cabral, M. C. and Burges, S. J.: Digital elevation model networks (DEMON): a model of flow over hillslopes for computation of contributing and dispersal areas, Water Resour. Res., 30, 1681–1692, doi:10.1029/93WR03512, 1994.

HESSD
10, 14041–14093, 2013

Cuo, L., Giambelluca, T. W., Ziegler, A. D., and Nullet, M. A.: Use of the distributed hydrology soil vegetation model to study road effects on hydrological processes in Pang Khum Experimental Watershed, northern Thailand, Forest Ecol. Manage., 224, 81–94, doi:10.1016/j.foreco.2005.12.009, 2006.

Czymmek, K. J., Ketterings, Q. M., Geohring, L. D., and Albrecht, G. L.: The New York Phosphorus Runoff Index User's Manual and Documentation, CSS Extension Publication E03-13, available at: http://nmsp.cals.cornell.edu/publications/extension/PI_User_Manual.pdf, last access: 19 July 2013, 64 pp., 2003.

Easton, Z. M., Fuka, D. R., Walter, M. T., Cowan, D. M., Schneiderman, E. M., and Steenhuis, T. S.: Re-conceptualizing the soil and water assessment tool (SWAT) model to predict runoff from variable source areas, J. Hydrol., 348, 279–291, doi:10.1016/j.jhydrol.2007.10.008, 2008.

Endreny, T. A. and Wood, E. F.: International Journal of Geographical Information Science maximizing spatial congruence of observed and DEM- delineated overland flow networks, Int. J. Geogr. Inf. Sci., 17, 699–713, 2003.

Erskine, R. H., Green, T. R., Ramirez, J. A., and MacDonald, L. H.: Comparison of grid-based algorithms for computing upslope contributing area, Water Resour. Res., 42, W09416, doi:10.1029/2005WR004648, 2006.

Fairfield, J. and Leymarie, P.: Drainage networks from grid digital elevation models, Water Resour. Res., 27, 709–717, doi:10.1029/90WR02658, 1991.

Frankenberger, J. R.: Identification of critical runoff generating areas using a variable source area model, Ph.D. thesis, Department of Agricultural and Biological Engineering, Cornell University, Ithaca, NY, 207 pp., 1996.

Frankenberger, J. R., Brooks, E. S., Walter, M. T., Walter, M. F., and Steenhuis, T. S.: A GIS-based variable source area hydrology model, Hydrol. Process., 13, 805–822, doi:10.1002/(SICI)1099-1085(19990430)13:6<805::AID-HYP754>3.0.CO;2-M, 1999.

Freeman, T. G.: Calculating catchment area with divergent flow based on a regular grid, Comput. Geosci., 17, 413–422, doi:10.1016/0098-3004(91)90048-I, 1991.

Gburek, W. J. and Sharpley, A. N.: Hydrologic controls on phosphorus loss from upland agricultural watersheds, J. Environ. Qual., 27, 267–277, doi:10.2134/jeq1998.00472425002700020005x, 1998.

HESSD
10, 14041–14093, 2013

Gburek, W. J., Drungil, C. C., Srinivasan, M. S., Needelman, B. A., and Woodward, D. E.: Variable-source-area controls on phosphorus transport: bridging the gap between research and design, J. Soil Water Conserv., 57, 534–543, 2002.

Grabs, T., Seibert, J., Bishop, K., and Laudon, H.: Modeling spatial patterns of saturated areas: a comparison of the topographic wetness index and a dynamic distributed model, J. Hydrol., 373, 15–23, doi:10.1016/j.jhydrol.2009.03.031, 2009.

Güntner, A., Seibert, J., and Uhlenbrook, S.: Modeling spatial patterns of saturated areas: an evaluation of different terrain indices, Water Resour. Res., 40, W05114, doi:10.1029/2003WR002864, 2004.

Hasan, A., Pilesjö, P., and Persson, A.: On Generating Digital Elevation Models from LiDAR Data: Resolution vs. Accuracy and Topographic Wetness Index Indices in Northern Peatlands, Taylor and Francis, London, UK, 2012.

Hellstrand, E.: Spatial and temporal mapping of shallow groundwater tables in the riparian zone of a Swedish headwater catchment, Department of Earth Sciences, Uppsala University, Uppsala, Sweden, 2012.

Hintze, J. L. and Nelson, R. D.: Violin plots: a box plot-density trace synergism, Am. Stat., 52, 181–184, doi:10.2307/2685478, 1998.

Hjerdt, K. N., McDonnell, J. J., Seibert, J., and Rodhe, A.: A new topographic index to quantify downslope controls on local drainage, Water Resour. Res., 40, W05602, doi:10.1029/2004WR003130, 2004.

Holmgren, P.: Multiple-flow direction algorithms for runoff modelling in grid based elevation models: an empirical evaluation, Hydrol. Process., 8, 327–334, 1994.

Horn, B. K. P.: Hill shading and the reflectance map, P. IEEE, 69, 14–47, doi:10.1109/PROC.1981.11918, 1981.

Jackson, T. J.: Measuring surface soil moisture using passive microwave remote sensing, Hydrol. Process., 7, 139–152, doi:10.1002/hyp.3360070205, 1993.

Jordan, J. P.: Spatial and temporal variability of stormflow generation processes on a Swiss catchment, J. Hydrol., 153, 357–382, 1994.

Kenward, T., Lettenmaier, D. P., Wood, E. F., and Fielding, E.: Effects of digital elevation model accuracy on hydrologic predictions, Remote Sens. Environ., 74, 432–444, 2000.

Ladson, A. R. and Moore, I. D.: Soil water prediction on the Konza Prairie by microwave remote sensing and topographic attributes, J. Hydrol., 138, 385–407, doi:10.1016/0022-1694(92)90127-H, 1992.

HESSD
10, 14041–14093, 2013

Lane, S. N., Brookes, C. J., Heathwaite, A. L., and Reaney, S.: Surveillant science: challenges for the management of rural environments emerging from the new generation diffuse pollution models, J. Agr. Econ., 57, 239–257, doi:10.1111/j.1477-9552.2006.00050.x, 2006.

Lanni, C., McDonnell, J. J., and Rigon, R.: On the relative role of upslope and downslope topography for describing water flow path and storage dynamics: a theoretical analysis, Hydrol. Process., 25, 3909–3923, doi:10.1002/hyp.8263, 2011.

Larson, K. M., Small, E. E., Gutmann, E. D., Bilich, A. L., Braun, J. J., and Zavorotny, V. U.: Use of GPS receivers as a soil moisture network for water cycle studies, Geophys. Res. Lett., 35, 1–5, doi:10.1029/2008GL036013, 2008.

Lyon, S. W., Walter, M. T., Grard-Marchant, P., and Steenhuis, T. S.: Using a topographic index to distribute variable source area runoff predicted with the SCS curve-number equation, Hydrol. Process., 18, 2757–2771, doi:10.1002/hyp.1494, 2004.

Lyon, S. W., Lembo, A. J., Walter, M. T., and Steenhuis, T. S.: Defining probability of saturation with indicator kriging on hard and soft data, Adv. Water Resour., 29, 181–193, doi:10.1016/j.advwatres.2005.02.012, 2006a.

Lyon, S. W., McHale, M., Walter, M. T., Steenhuis, T. S.: Effect of runoff generation mechanism on estimating land use control of P concentrations, J. Am. Water Resour. As., 42, 793–804, 2006b.

Mallick, K., Bhattacharya, B. K., and Patel, N. K.: Estimating volumetric surface moisture content for cropped soils using a soil wetness index based on surface temperature and NDVI, Agr. Forest Meteorol., 149, 1327–1342, doi:10.1016/j.agrformet.2009.03.004, 2009.

Marjerison, R. D., Dahlke, H., Easton, Z. M., Seifert, S., and Walter, M. T.: A phosphorus index transport factor based on variable source area hydrology for New York State, J. Soil Water Conserv., 66, 149–157, doi:10.2489/jswc.66.3.149, 2011.

Mehta, V. K., Walter, M. T., Brooks, E. S., Steenhuis, T. S., Walter, M. F., Johnson, M., Boll, J., and Thongs, D. J.: Application of SMR to modeling watersheds in the Catskill Mountains, Environ. Model. Assess., 9, 77–89, doi:10.1023/B:ENMO.0000032096.13649.92, 2004.

Moore, I. D., Burch, G. J., and Mackenzie, D. H.: Topographic effects on the distribution of surface soil water and the location of ephemeral gullies, T. ASAE, 31, 1098–1107, 1988.

Moore, I. D., Grayson, R. B., and Ladson, A. R.: Digital terrain modelling: a review of hydrological, geomorphological, and biological applications, Hydrol. Process., 5, 3–30, doi:10.1002/hyp.3360050103, 1991.

HESSD
10, 14041–14093, 2013

Motovilov, Y. G., Gottschalk, L., Engeland, K., and Rodhe, A.: Validation of a distributed hydrological model against spatial observations, Agr. Forest Meteorol., 98–99, 257–277, doi:10.1016/S0168-1923(99)00102-1, 1999.
Murphy, P. N. C., Ogilvie, J., and Arp, P.: Topographic modelling of soil moisture conditions: a comparison and verification of two models, Eur. J. Soil Sci., 60, 94–109, doi:10.1111/j.1365-2389.2008.01094.x, 2009.
Nyberg, L.: Spatial variability of soil water content in the covered catchment at Gårdsjön, Sweden, Hydrol. Process., 10, 89–103, 1996.
O'Callaghan, J. F. and Mark, D. M.: The extraction of drainage networks from digital elevation data, Comput. Vision Graph., 28, 323–344, doi:10.1016/S0734-189X(84)80011-0, 1984.
Park, S. J., Ruecker, G. R., Agyare, W. R., Akramhanov, A., Kim, M., and Vlek, P. L. G.: Influence of grid cell size and flow routing algorithm on soil-landform modeling, J. Korean Geogr. Soc., 44, 122–145, 2009.
Qiu, Z.: Variable source pollution: turning science into action to manage and protect critical source areas in landscapes, J. Soil Water Conserv., 65, 137A–141A, 2010.
Qiu, Z., Walter, M. T., and Hall, C.: Managing variable source pollution in agricultural watersheds, J. Soil Water Conserv., 62, 115–122, 2007.
R Core Team: R: A language and environment for statistical computing, http://www.R-project.org/, R Foundation for Statistical Computing, Vienna, Austria, 2011.
Reaney, S. M., Lane, S. N., Heathwaite, A. L., and Dugdale, L. J.: Risk-based modelling of diffuse land use impacts from rural landscapes upon salmonid fry abundance, Ecol. Model., 222, 1016–1029, doi:10.1016/j.ecolmodel.2010.08.022, 2011.
Rossing, J. M. and Walter, M. F.: Hydrologically-based identification of critical areas for water quality protection in the New York City watershed, ASAE Paper No. 952580, Presented at ASAE Annual International Meeting, Chicago, IL, 10 pp., 1995.
Sayde, C., Gregory, C., Gil-Rodriguez, M., Tufillaro, N., Tyler, S., van de Giesen, N., English, M., Cuenca, R., and Selker, J. S.: Feasibility of soil moisture monitoring with heated fiber optics, Water Resour. Res., 46, W06201, doi:10.1029/2009WR007846, 2010.
Schmidt, F. and Persson, A.: Comparison of DEM data capture and topographic wetness indices, Precis. Agric., 4, 179–192, 2003.
Schneiderman, E. M., Steenhuis, T. S., Thongs, D. J., Easton, Z. M., Zion, M. S., Neal, A. L., Mendoza, G. F., and Walter, M. T.: Incorporating variable source area hydrology into a curve-

HESSD
10, 14041–14093, 2013

number-based watershed model, Hydrol. Process., 21, 3420–3430, doi:10.1002/hyp.6556, 2007.
Seibert, J. and McGlynn, B. L.: A new triangular multiple flow direction algorithm for computing upslope areas from gridded digital elevation models, Water Resour. Res., 43, 1–8, doi:10.1029/2006WR005128, 2007.
Shaw, S. B. and Walter, M. T.: Improving runoff risk estimates: formulating runoff as a bivariate process using the SCS-Curve Number Method, Water Resour. Res., 45, W030404, doi:10.1029/2008WR006900, 2009.
Sørensen, R. and Seibert, J.: Effects of DEM resolution on the calculation of topographical indices: TWI and its components, J. Hydrol., 347, 79–89, doi:10.1016/j.jhydrol.2007.09.001, 2007.
Sørensen, R., Zinko, U., and Seibert, J.: On the calculation of the topographic wetness index: evaluation of different methods based on field observations, Hydrol. Earth Syst. Sci., 10, 101–112, doi:10.5194/hess-10-101-2006,2006.
Sulebak, J. R., Tallaksen, L. M., and Erichsen, B.: Estimation of areal soil moisture by use of terrain data, Geogr. Anal., 82A, 89–105, 2000.
Tague, C., Band, L., Kenworthy, S., and Tenebaum, D.: Plot- and watershed-scale soil moisture variability in a humid Piedmont watershed, Water Resour. Res., 46, W12541, doi:10.1029/2009WR008078, 2010.
Tarboton, D. G.: A new method for the determination of flow directions and upslope areas in grid digital elevation models, Water Resour. Res., 33, 309–319, doi:10.1029/96WR03137, 1997.
Tenenbaum, D. E., Band, L. E., Kenworthy, S. T., and Tague, C. L.: Analysis of soil moisture patterns in forested and suburban catchments in Baltimore, Maryland, using high-resolution photogrammetric and LIDAR digital elevation datasets, Hydrol. Process., 20, 219–240, doi:10.1002/hyp.5895, 2006.
Tromp-van Meerveld, H. J. and McDonnell, J. J.: On the interrelations between topography, soil depth, soil moisture, transpiration rates and species distribution at the hillslope scale, Adv. Water Resour., 29, 293–310, doi:10.1016/j.advwatres.2005.02.016, 2006.
USDA-NRCS: Soil Data Viewer, avalaible at: http://soils.usda.gov/sdv/ (last access: 15 May 2013), 2009.
USEPA: National Water Quality Inventory: Report to Congress: 2004 reporting cycle, United States Environmental Protection Agency, Office of Water, Washington, D.C., 52 pp., 2004.

HESSD
10, 14041–14093, 2013

USGS – United States Geologic Survey, Report: Vertical Accuracy of the Nation Elevation Dataset, availabe at: http://ned.usgs.gov/downloads/documents/NED_Accuracy.pdf, last access: May 2013.

Walter, M. T., Walter, M. F., Brooks, E. S., Steenhuis, T. S., Boll, J., and Weiler, K. R.: Hydrologically sensitive areas: variable source area hydrology implications for water quality risk assessment, J. Soil Water Conserv., 55, 277–284, 2000.

Walter, M. T., Brooks, E. S., Walter, M. F., Steenhuis, T. S., Scott, C. A., and Boll, J.: Evaluation of soluble phosphorus loading from manure-applied fields under various spreading strategies, J. Soil Water Conserv., 56, 329–335, 2001.

Walter, M. T., Steenhuis, T. S., Mehta, V. K., Thongs, D., Zion, M., and Schneiderman, E.: Refined conceptualization of TOPMODEL for shallow subsurface flows, Hydrol. Process., 16, 2041–2046, doi:10.1002/hyp.5030, 2002.

Walter, M. T., Mehta, V. K., Marrone, A. M., Boll, J., Gérard-Merchant, P., Steenhuis, T. S., and Walter, M. F.: A simple estimation of the prevalence of Hortonian flow in New York City's watersheds, ASCE J. Hydrol. Eng., 8, 214–218, 2003.

Western, A. W., Grayson, R. B., Blöschl, G., Willgoose, G. R., and McMahon, T. A.: Observed spatial organization of soil moisture and its relation to terrain indices, Water Resour. Res., 35, 797–810, doi:10.1029/1998WR900065, 1999.

Western, A. W., Zhou, S.-L., Grayson, R. B., McMahon, T. A., Blöschl, G., and Wilson, D. J.: Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes, J. Hydrol., 286, 113–134, doi:10.1016/j.jhydrol.2003.09.014, 2004.

White, M. J., Storm, D. E., Busteed, P. R., Smolen, M. D., Zhang, H., and Fox, G. A.: A quantitative phosphorus loss assessment tool for agricultural fields, Environ. Modell. Softw., 25, 1121–1129, doi:10.1016/j.envsoft.2010.03.017, 2010.

Wilson, D. J., Western, A. W., and Grayson, R. B.: A terrain and data-based method for generating the spatial distribution of soil moisture, Adv. Water Resour., 28, 43–54, doi:10.1016/j.advwatres.2004.09.007, 2005.

Wolock, D. M. and Price, C. V.: Effects of digital elevation model map scale and data resolution on a topography-based watershed model, Water Resour. Res., 30, 3041–3052, 1994.

Zevenbergen, L. W. and Thorne, C. R.: Quantitative analysis of land surface topography, Earth Surf. Proc. Land., 12, 47–56, doi:10.1002/esp.3290120107, 1987.

HESSD
10, 14041–14093, 2013

Zhang, W. and Montgomery, D. R.: Digital elevation model grid size, landscape representation, and hydrologic simulations, Water Resour. Res., 30, 1019–1028, doi:10.1029/93WR03553, 1994.
Zhao, R. J., Zuang, Y., Fang, L. R., Lin, X. R., and Zhang, Q. S.: The Xinanjiang model, IAHS-AISH P., 129, 351–356, 1980.

HESSD
10, 14041–14093, 2013

Table 1. Soil, topographic and land use characteristics of each study site. Average transmissivity values were derived from SSURGO soil data (USDA-NRCS, 2009).

Site	Average (range) Slope ($m\,m^{-1}$)	Area (ha)	Average Transmissivity ($\mu m\,d^{-1}$)	Primary Land Use
1	5.29 (0.004–71.31)	5.6	0.9	Corn and Soybean
2	6.34 (0.016–46.35)	2.3	0.4	Orchard and Grass
3	4.81 (0.001–105.10)	10.2	1.6	Corn, Soybean and Grass
4	6.63 (0.014–44.16)	6.0	2.8	Corn, Soybean and Grass
5	6.56 (0.305–22.018)	2.3	0.6	Grass/Fallow

HESSD
10, 14041–14093, 2013

Table 2. Summary of TDR measurements at each field site. VWC represents volumetric water content. Site numbers correspond with Fig. 1.

Site	Sampling Date	7-Day Antecedent Rainfall (mm)	Mean VWC (%)	SD (%)	Sample Size
1	12 Sept 2012	1.23	25.3	7	22
	21 Sep 2012	1.61	32.7	7.4	22
	26 Oct 2012	1.68	39.1	7.9	22
	2 Nov 2012	1.6	43.9	7.1	22
	8 Nov 2012	0.47	41.7	7.4	27
	15 Nov 2012	0.69	44.9	7	27
	30 Nov 2012	0	39.6	7.8	27
2	3 Oct 2012	0.66	36.6	9.7	13
	2 Nov 2012	1.6	45.3	8.6	20
	12 Nov 2012	0.04	38.2	9	20
	20 Nov 2012	0.68	40.2	9.1	26
	28 Nov 2012	0	39.3	7.6	23
3	21 Aug 2012	1.64	23.1	7.3	25
	5 Sep 2012	0	31.8	9.6	31
	14 Sep 2012	0.74	20.7	8.8	19
	17 Oct 2012	0.24	34.2	9.9	36
	31 Oct 2012	2.14	45	11.3	38
	7 Jul 2012	0.74	43.5	10.5	40
	15 Nov 2012	0.69	43.9	8.9	50
	28 Nov 2012	0	40.8	9	46
4	23 Aug 2012	0.1	26.2	11.6	27
	6 Sep 2012	0.45	24.2	13.6	25
	19 Sep 2012	1.34	36.4	10.1	35
	24 Oct 2012	1.76	41.6	7.9	16
	2 Nov 2012	1.5	40.5	7.9	36
	9 Nov 2012	0.13	38.3	8.9	54
	16 Nov 2012	0.78	40.8	9.8	51
	30 Nov 2012	0.04	39.1	11	50
5	19 Oct 2012	0.24	38	6.2	35
	31 Oct 2012	2.14	42.4	7.7	38
	7 Nov 2012	0.74	42.7	9.3	22
	14 Nov 2012	0.69	44.6	8.6	44
	30 Nov 2012	0	39.8	8.3	46

HESSD
10, 14041–14093, 2013

Table 3. Example of TWI differences in pairwise comparisons.

TWI	Source Type	Cell Size	Slope	Flow Accum	Smoothing
1	LIDAR	3 m	LSFP	D8	Filtered
2	LIDAR	3 m	LSFP	D8	Un-Filtered

HESSD
10, 14041–14093, 2013

Table 4. Summary of the cases where the TI performed better than the STI in the LiDAR dataset. In the filtered column, "Y" stands for yes and "N" stands for no.

SourceData	CellSize	Slope	FlowAccum	Filtered	AIC_{STI}	AIC_{TI}	ΔAIC
LIDAR	10	MTS	BR	Y	6048.2	6045.6	−2.6
LIDAR	10	DSI-2m	BR	Y	6050.4	6041.7	−8.7
LIDAR	10	SDP	BR	Y	6062.9	6057.5	−5.4
LIDAR	10	LSFP	BR	Y	6064.0	6057.3	−6.6

HESSD
10, 14041–14093, 2013

Table 5. Top ten best performing TWIs from the mixed effects analysis for the LiDAR and USGS datasets.

#	TI form	Cell Size	Slope	Flow Accum	Filtered	AIC	ΔAIC	AIC_w	E-Ratio
LiDAR									
1a	STI	3 m	MTS	MD∞	N	5916.8	0	0.39	1
2a	STI	3 m	MTS	D∞	N	5918	1.2	0.22	1.78
3a	STI	3 m	LSFP	MD∞	N	5918.7	1.9	0.15	2.57
4a	STI	3 m	SDP	MD∞	N	5918.9	2.1	0.14	2.83
5a	STI	3 m	SDP	D∞	N	5921.8	5	0.03	12.28
6a	STI	3 m	LSFP	D∞	N	5921.9	5	0.03	12.34
7a	STI	3 m	MTS	D∞	Y	5924.3	7.4	0.01	41.45
8a	STI	3 m	DSI-2m	MD∞	N	5924.9	8.1	0.01	57.42
9a	STI	3 m	SDP	MD	N	5925.5	8.7	0.01	75.8
10a	STI	3 m	LSFP	MD	N	5925.5	8.7	0.01	75.9
USGS									
1b	STI	10 m	DSI-5m	D8	Y	6141.3	0	0.59	1
2b	STI	10 m	DSI-5m	Rho8	Y	6142.7	1.4	0.3	1.97
3b	STI	10 m	DSI-5m	BR	Y	6147	5.7	0.03	19.67
4b	STI	10 m	DSI-5m	MD	Y	6148.3	7	0.02	29.5
5b	STI	10 m	DSI-2m	D8	Y	6148.6	7.3	0.02	29.5
6b	STI	10 m	DSI-5m	MD∞	Y	6150.2	8.8	0.01	59
7b	STI	10 m	DSI-5m	D∞	Y	6150.3	9	0.01	59
8b	STI	10 m	DSI-2m	Rho8	Y	6150.6	9.3	0.01	59
9b	STI	10 m	DSI-2m	BR	Y	6150.7	9.4	0.01	59
10b	STI	10 m	DSI-10m	D8	Y	6151.6	10.2	0	–

HESSD
10, 14041–14093, 2013

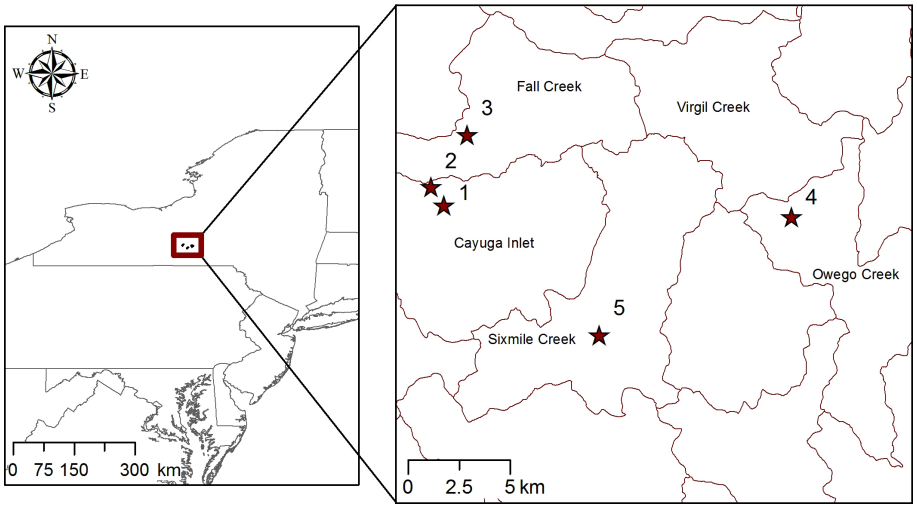

Fig. 1. Study site locations (red stars). Watershed boundaries are depicted by red polygons (USGS, 12-digit hydrologic unit codes).

HESSD
10, 14041–14093, 2013

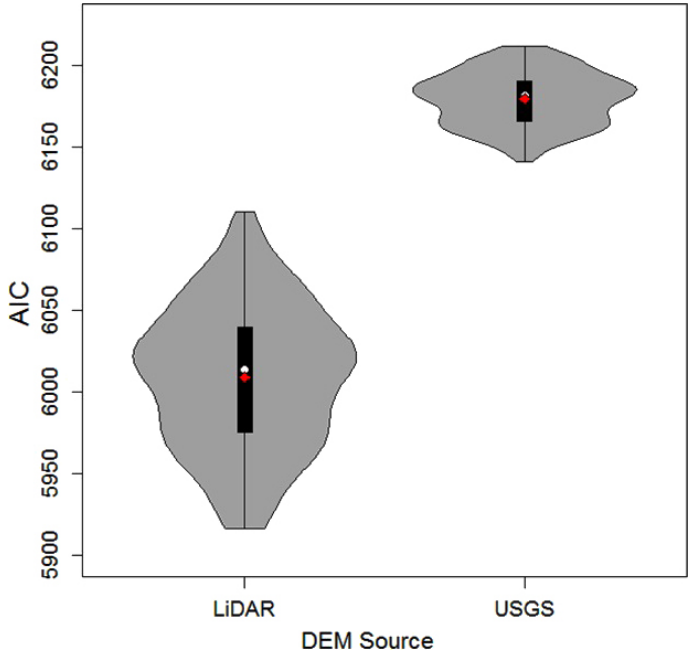

Fig. 2. Violin plot of AIC values of the LIDAR and USGS DEM source groups. The median and mean are indicated by the white dots and red diamonds, respectively; the interquartile range is indicated by the thick vertical bar, the density distribution is shown as the symmetrical (mirrored) grey area and highlights the peaks and valleys of the AIC distribution.

HESSD
10, 14041–14093, 2013

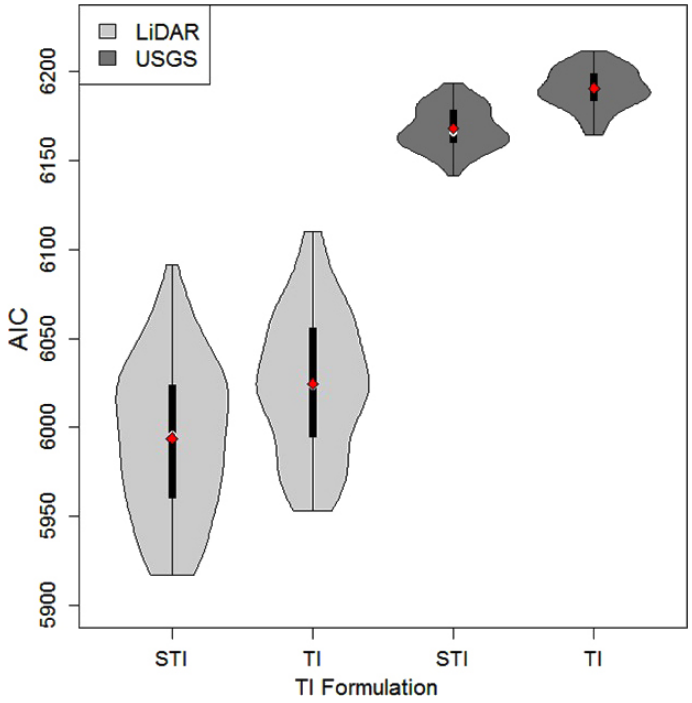

Fig. 3. Violin plot of AIC values vs. TI form for the LIDAR and USGS datasets. Mean and median values are depicted as red diamonds and white dots, respectively.

HESSD
10, 14041–14093, 2013

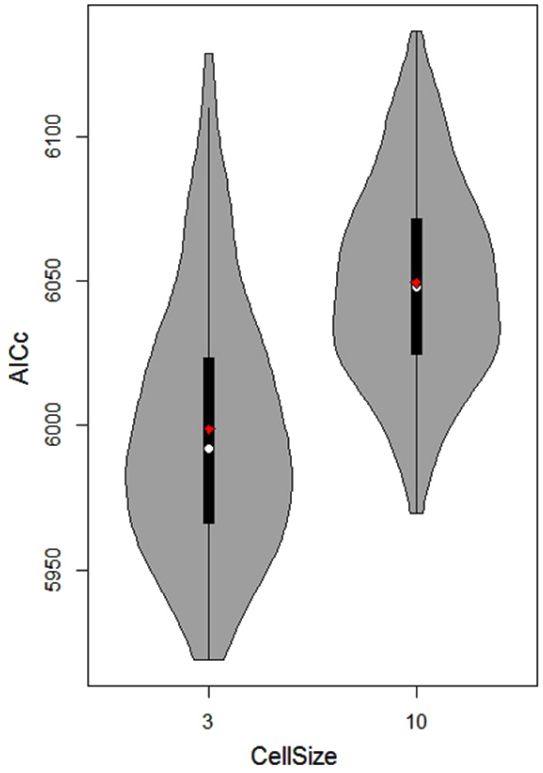

Fig. 4. AIC values vs. grid cell size for the LiDAR dataset. Mean and median values are depicted as red diamonds and white dots, respectively.

HESSD
10, 14041–14093, 2013

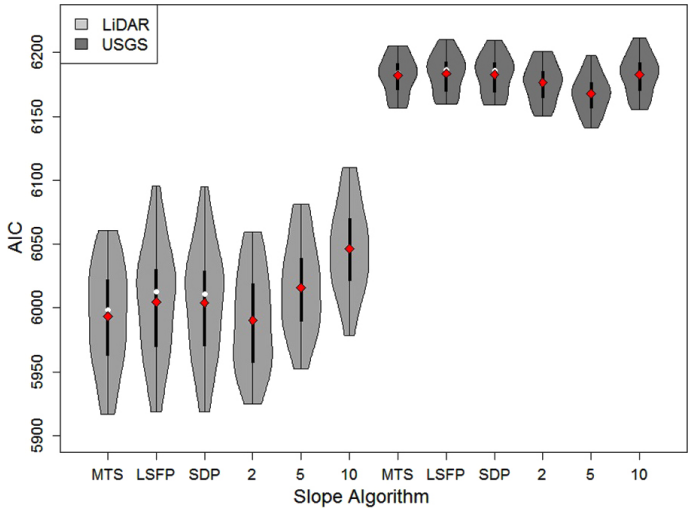

Fig. 5. AIC values vs. flow direction algorithm for the LiDAR (light grey; 3 and 10 m) and USGS (dark grey) datasets. Mean and median values are depicted as red diamonds and white dots, respectively.

HESSD
10, 14041–14093, 2013

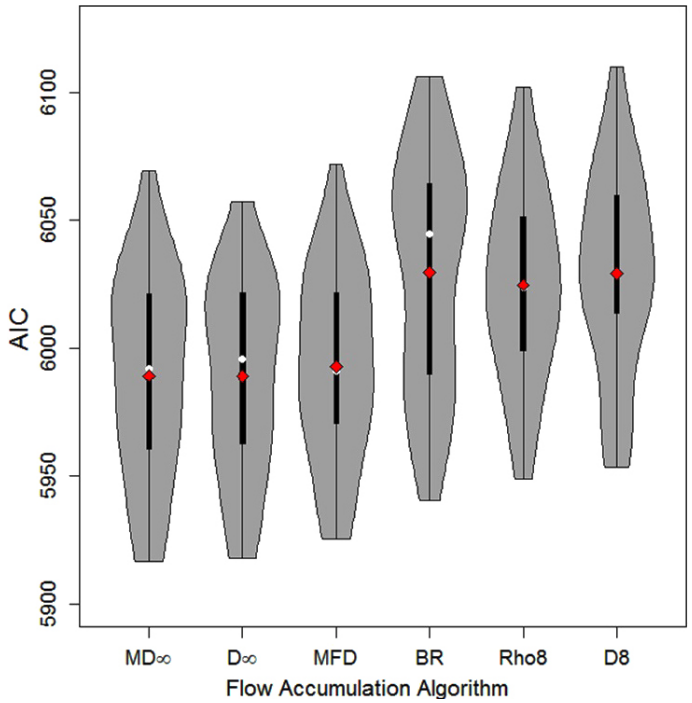

Fig. 6. AIC distributions for each of the six test flow accumulation algorithms using LiDAR-derived DEMs. Mean and median values are depicted as red diamonds and white dots, respectively.

HESSD
10, 14041–14093, 2013

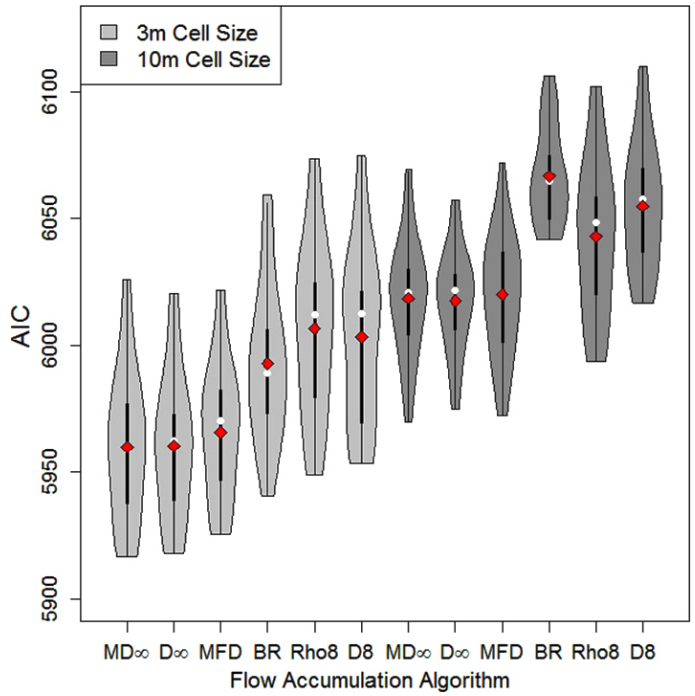

Fig. 7. AIC values vs. flow accumulation algorithm for both the 3 m (light grey) and 10 m (dark grey) LiDAR dataset.

HESSD
10, 14041–14093, 2013

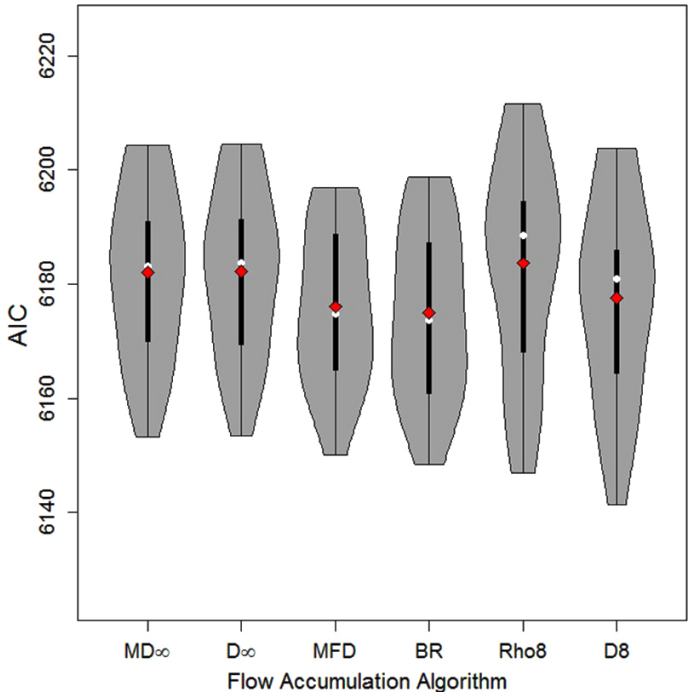

Fig. 8. AIC values vs. flow accumulation algorithm for the USGS dataset.

HESSD
10, 14041–14093, 2013

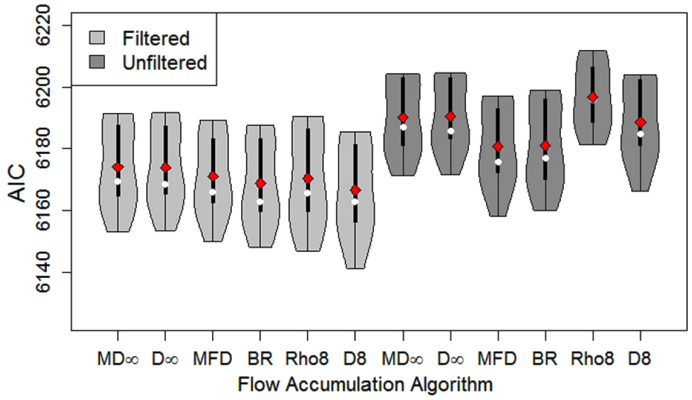

Fig. 9. AIC values vs. flow accumulation algorithm for the filtered (light grey) and unfiltered (dark grey) TWIs.

HESSD
10, 14041–14093, 2013

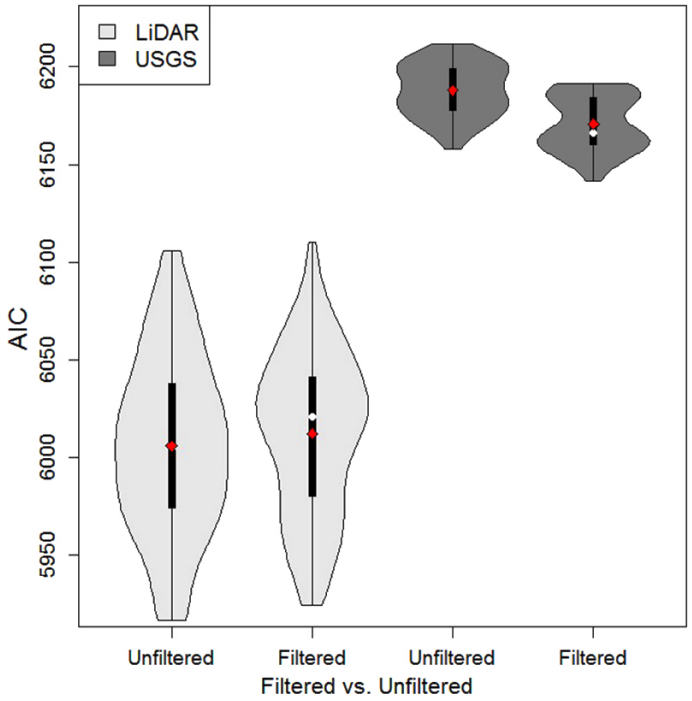

Fig. 10. AIC values vs. smoothing for the LiDAR (light grey) and USGS (dark grey) TWIs.

HESSD
10, 14041–14093, 2013

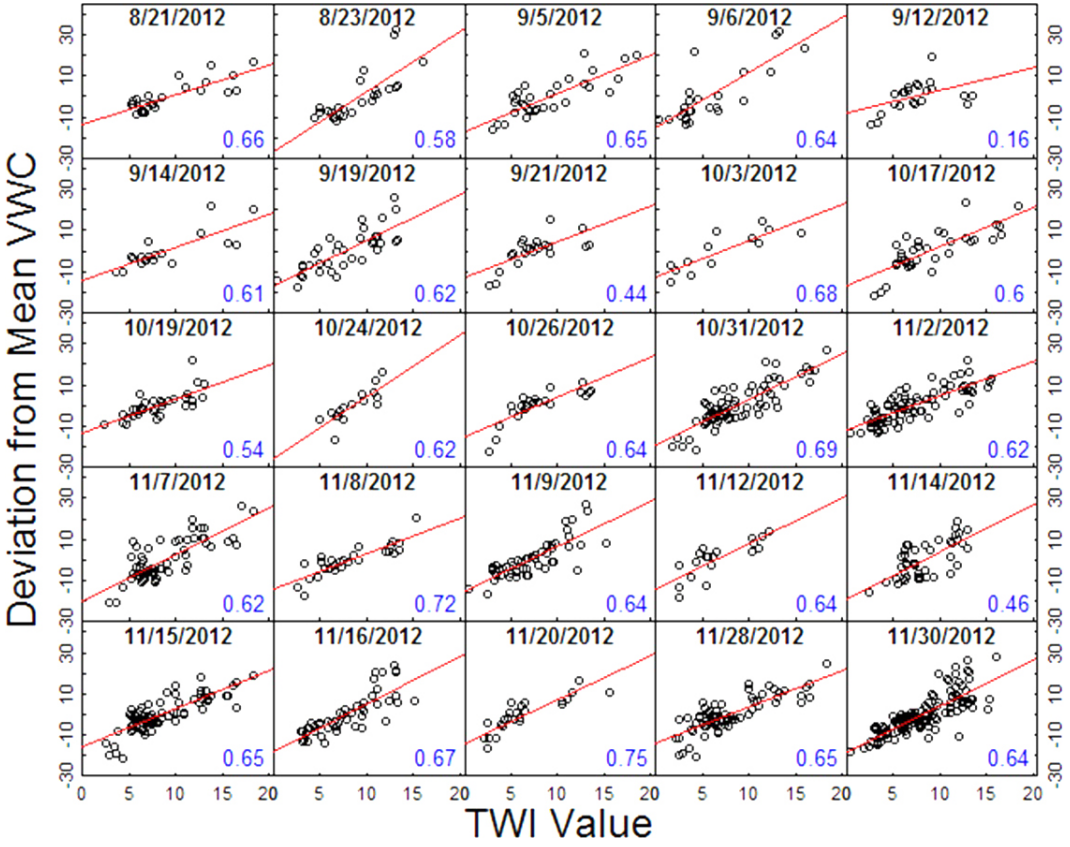

Fig. 11. Mean-normalized volumetric water content (%) vs. index value of the optimal TWI for the LiDAR dataset. R^2 values are shown in the lower right corner of each graph.

HESSD
10, 14041–14093, 2013

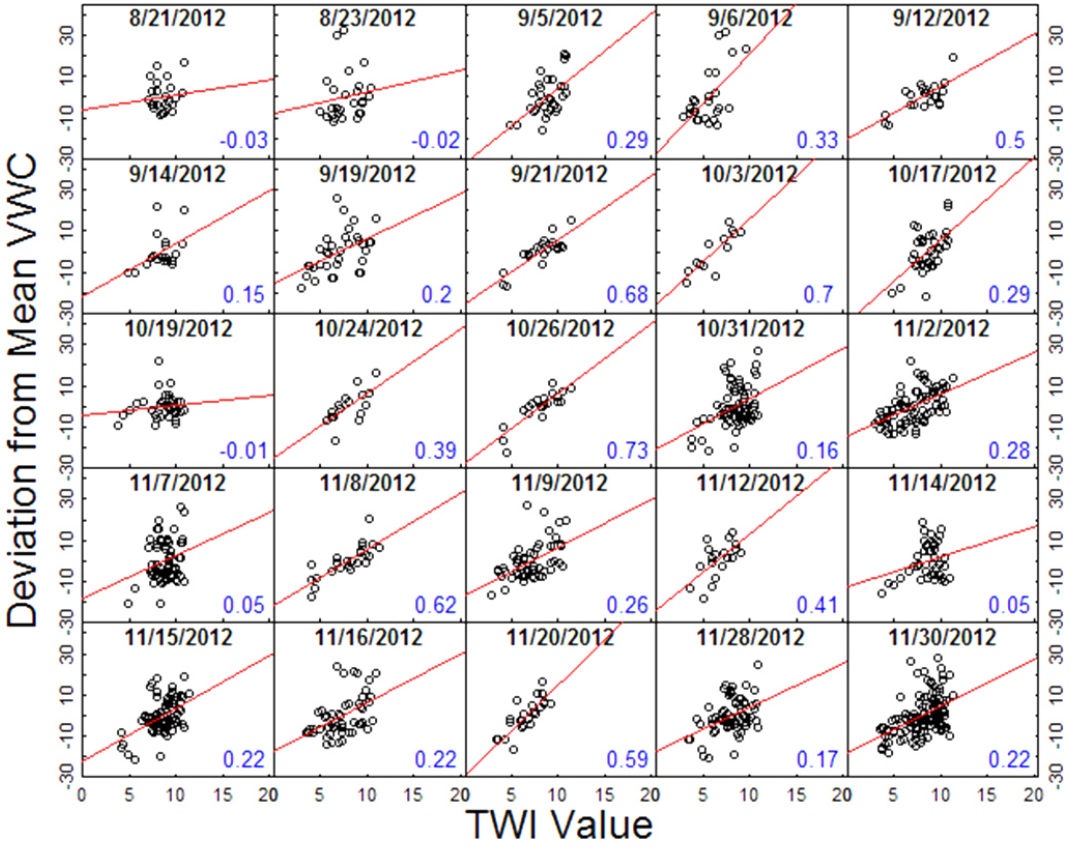

Fig. 12. Mean-normalized volumetric water content (%) vs. index value of the optimal TWI for the USGS dataset. R^2 values are shown in the lower right corner of each graph.

HESSD
10, 14041–14093, 2013

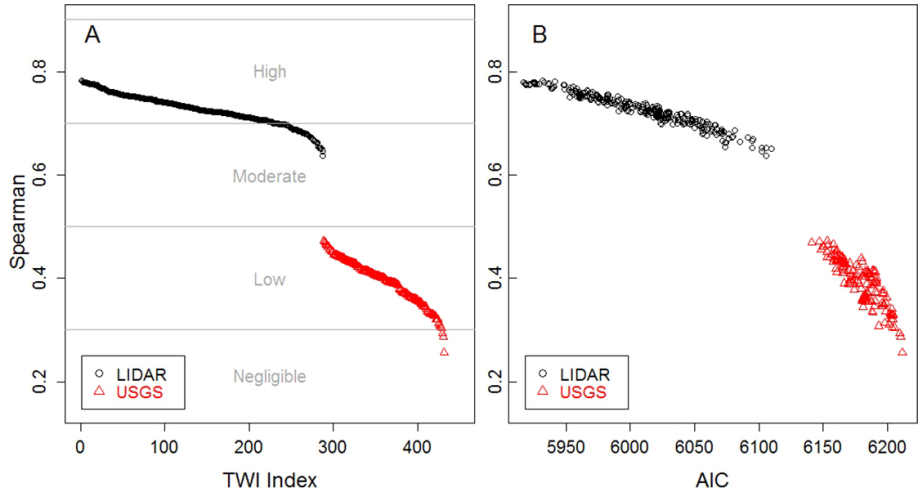

Fig. 13. (A) Spearman correlation coefficients vs. TWI. **(B)** Spearman correlation coefficients vs. AIC values from all 432 mixed effect models.

HESSD
10, 14041–14093, 2013