Defining a Topographic Index Threshold to Delineate Hydrologically Sensitive Areas for Water Resources Planning and Management

Zeyuan Qiu [1] ⬤ · Steve W. Lyon [2] · Ellen Creveling [3]

Abstract

Hydrologically sensitive areas (HSAs) are runoff-generating areas often targeted for effective water resources planning and management actions. Commonly, HSAs can be mapped as areas in a landscape with a topographic index (TI) greater than a threshold level. This study explored the impact of a gradient of different TI threshold values for delineating HSAs using two popular TIs: a topographic wetness index (TWI) and a soil topographic index (STI). The resultant HSAs for each TI were compared to the Federal Emergency Management Agency (FEMA) 100-year floodplain map in New Jersey and its five water regions. Spatial comparison indicators were used to assess the spatial similarity between the HSAs delineated and the FEMA floodplain map. Such comparisons identified the threshold that delineated HSAs whose spatial distributions were most consistent with the FEMA floodplain at each spatial scale for each TI. For example, the identified threshold for using a TWI to delineate HSAs was 10.5 at the state level; however, this threshold varied by the water region. The HSAs delineated approximate the spatial extent of runoff-contributing areas to the 100-year flood relevant for water resources planning and management actions for flood hazard mitigation.

Keywords Hydrologically sensitive areas · Topographic index · Thresholds · Kappa value · Floodplain · Water resources planning and management

✉ Zeyuan Qiu
zeyuan.qiu@njit.edu

[1] Department of Chemistry and Environmental Science, College of Science and Liberal Arts, New Jersey Institute of Technology, Newark, NJ, USA

[2] School of Environment and Natural Resources, Ohio State University, Columbus, OH, USA

[3] The Nature Conservancy, Southern New Jersey Office, Delmont, NJ, USA

🖄 Springer

1 Introduction

Hydrologically sensitive areas (HSAs) are the parts of a landscape that produce disproportionate amounts of surface runoff. HSAs can be targeted to effectively implement planning and management measures to control nonpoint source pollution (Amin et al. 2017; Qiu et al. 2019) and mitigate negative impacts of stormwater runoff (Qiu et al. 2014, Martin-Mikle et al. 2015; Kaykhosravi et al. 2019). Topographic indices (TIs) have long been used to parsimoniously delineate HSAs based on the concept of variable source area (VSA) hydrology (Walter et al. 2000; Heathwaite et al. 2005; Qiu 2009). Specifically, the TI values mimic the sequential distribution of water storage capacity of a landscape allowing identification of runoff-generating areas (Dahlke et al. 2009). As such, the locations in a landscape with TI values greater than a given threshold delineate HSAs. There are several mechanistic ways to define such a threshold for delineating HSAs in landscapes, including the controlling specific storm event method (Lyon et al. 2006) and the average saturation probability method (Agnew et al. 2006), but these methods require ancillary information about streamflow or runoff production for calibration (e.g. Lyon et al. 2004). Herron and Hairsine (1998) suggested HSAs should be 20% of watershed area with the highest saturation potential. Qiu (2009) used a static TI threshold of 10 for HSA delineation in Raritan River Basin in central New Jersey.

Identification of a proper TI threshold for HSA delineation faces two practical difficulties. The first difficulty is the use of a priori TI derivation, given the variety of different TIs available (Buchanan et al. 2014). TIs vary significantly by the physical attributes being considered and the methods being used to derive those attributes. For example, Buchanan et al. (2014) compiled and compared over 400 unique TI formulations that considered different digital elevation model (DEM) resolutions, vertical precision of the DEMs, flow direction and slope algorithms, and smoothing via low-pass filtering. They found that a soil topographic index (STI) derived using a fine-scale LiDAR DEM with additional of soil information was a better indicator to represent soil moisture conditions in central New York relative to other formulations.

The second practical difficulty is the lack of the empirical validation of the delineated HSAs. Giri et al. (2017) conducted Trellis plot analyses based on a polynomial regression model (order two to four) to analyze the relationship between TIs and the changes in soil moisture measured in two fields. They concluded that the TI threshold for delineating HSAs can range from 9 to 15 when considering a STI in Central New Jersey. Archibald et al. (2014) found good agreement between HSAs mapped using STIs and the observed saturated areas measured through groundwater table positions. However, these empirical studies looking at defining TI thresholds for delineation of HSAs were conducted at field or small-watshed scales. While large-scale patterns of saturated areas or runoff contributing areas can be estimated using vegetation patterns (Kulasova et al. 2014) or satellite remote sensing images (Temimi et al. 2010; Lei et al. 2016), such approaches have thus far not been used to corroborate TI thresholds for HAS delineation from a regional perspective in a manner meaningful to landscape managers.

There is clear need to identify proper TI thresholds for HSA delineation that are scientifically defensible and are more appealing to landscape managers. The objective of this study is thus to identify TI thresholds by comparing the spatial patterns of various delineated HSAs to the spatial pattern of floodplains. While HSAs are the hydrological hotspots in a landscape prone to generate surface runoff, floodplains are the areas that are inundated in a landscape during a storm event. Despite these definitional differences, both are closely related as the

floodplains are often an important part of HSAs and result from runoff accumulation from runoff generation associated with upland areas within the HSAs. As such, the delineated HSAs could approximate the runoff-contributing areas to the floodplains and offer a meaningful spatial extent for landscape managers to direct their planning and management actions designed to mitigate flood and other water-related impacts. To our knowledge, this study is novel as it is the first study that provides a systematic approach for identifying the threshold associated with delineating HSAs for water resources management and planning at large (regional to state) scales. The procedure not only is useful to select the TI threshold, but also helps refine the spatial extent of landscape planning and management to manage flood hazard.

2 Data and Methods

We consider two popular TIs in this study: a topographic wetness index (TWI) based on topography and a soil topographic index (STI) based on both topography and soil. Further, we target New Jersey, USA as a case study given concerns on flooding and water resource management (e.g. Lyon et al. 2018). The entire state of New Jersey can be divided geologically into northern and southern parts defined by what is commonly known as the Fall Line. North of the Fall Line are the Valley and Ridges, the Highlands, and Piedmont Physiographic Provinces. South of the Fall Line is the Coastal Plain Physiographic Province (Watt 2000). The total land mass of New Jersey (excluding the coastal waters and barrier islands) is about 2.02 million hectares. To effectively manage the state's water resources, New Jersey Department of Environmental Protection (NJDEP) divided New Jersey into 20 different watershed management areas (WMAs). NJDEP also defines five different water regions: Lower Delaware, Upper Delaware, Raritan, Atlantic Coast, and Northeast, which represent 24%, 16%, 16%, 31%, and 12% of the land mass in New Jersey, respectively. Each water region is comprised of three to five WMAs (Fig. 1a). Most of the streams and rivers in the Upper Delaware water region and all of the streams and rivers in the Lower Delaware water region drain into the Delaware River, which drains into the Atlantic Ocean. The rivers in the Raritan, Atlantic Coast, and Northeast water regions drain directly into the Atlantic Ocean (Watt 2000).

The following geospatial data were collected and processed for the study: (1) 10-m digital elevation models (DEM) for each of the 20 WMAs in New Jersey from NJDEP; (2) the digital soil survey geographic (SSURGO) databases for each of 21 counties in New Jersey from Natural Resources Conservation Service (NRCS); and (3) the 100-year floodplain map in New Jersey from Federal Emergency Management Agency (FEMA). The data in different coordinate systems were transformed to the New Jersey standard coordinate system for spatial data processing.

The HSAs delineated in this study were compared to the FEMA 100-year flood zone in New Jersey for the state as a whole and for each of the five different water regions (Lower Delaware, Upper Delaware, Raritan, Atlantic Coast and Northeast) with diverse topographic and soil conditions. The FEMA 100-year flood zone is the floodplain area estimated by FEMA to have a 1 % chance of being flooded at least once in a year. Further, the FEMA 100-year floodplain is a standardized product available for the contiguous United States for various applications, making it a viable product for the purpose of comparison and validation.

2.1 Topographic Indices

Topographic indices (TIs) measure the contribution of upslope areas and the ability to transfer water downslope at a given position in a landscape. Under variable source areas assumptions,
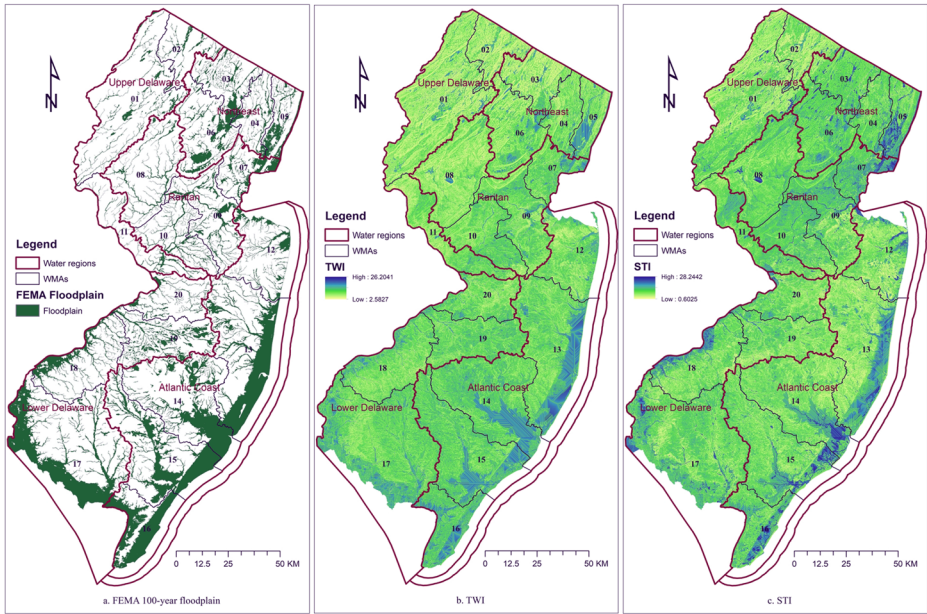
Fig. 1 The distribution of water regions and watershed management areas in New Jersey

TIs also express the likelihood that a location (or point) in a landscape generates saturation excess runoff. This study focuses on two widely used TIs: TWI and STI. The TWI was based on the following equation (Beven and Kirkby 1979):

$$TWI = ln\left(\frac{\alpha}{tan(\beta)}\right) \tag{1}$$

where α is the upslope contributing area per unit contour length (m) and β is the local topographic slope (m/m). Both α and β are topographically dependent and were derived from available DEMs. In addition to topography, a STI considers the spatial variation in hydrologically relevant soil properties (Walter et al. 2002; Anderson et al. 2015) and was calculated using:

$$STI = ln\left(\frac{\alpha}{tan(\beta)K_s D}\right) = ln\left(\frac{\alpha}{tan(\beta)}\right) - ln(K_s D) \tag{2}$$

Where α and β are defined above, K_s is the average saturated hydraulic conductivity (m/day), and D is the depth to the restrictive layer (m). The product of average saturated hydraulic conductivity and depth to restrictive layer terms in Eq. (2) is considered soil transmissivity and measures how water can be transmitted horizontally through soils.

TWI for the entire State of New Jersey was derived from the 20 WMA DEMs at a 10-m resolution using the System for Automated Geoscientific Analyses (SAGA) geographic information systems (GIS), which is in a public domain (Conrad et al. 2015). The DEMs at a WMA level were merged to form five larger DEMs for five water regions following the delineation by NJDEP. These five water regions were then merged to form the DEM for New Jersey. To ensure a continuous flow in each grid of the DEM, depressions were filled using the GIS function *fill sinks* (Planchon and Darboux 2001). The least square fitted plane method

(Horn 1981; Costa-Cabral and Burgess 1996) was used to calculate the slope of each grid cell in the DEMs while the multiple triangular flow direction method (Seibert and McGlynn 2007) was applied to obtain the catchment area. The resulting slope and catchment area layers were used to obtain the TWI for the state. Since the grid size in the DEMs was measured in feet, a correction factor of −1.191387, i.e. $ln(0.3038)$, was added to the resulting TWI to get the TWI measured in meters as defined by Eq. (1).

Soil transmissivity for each county in New Jersey was calculated using the soil saturated hydraulic conductivity and topsoil depth. Both were extracted from the SSURGO database using a soil data viewer, which is an add-on ArcGIS tool developed and maintained by NRCS. Using the aggregation methods within the ArcGIS soil data viewer, the topsoil depth to a restrictive layer and the geometrically weighted average soil saturated hydraulic conductivity layers were created. Both layers were multiplied to calculate the soil transmissivity value. Since the topsoil depth was given in cm and the soil saturated hydraulic conductivity in μm/s, a correction factor of 0.000864 was multiplied to the soil transmissivity values to convert the unit into m^2/day as defined in Eq. (2). The second term in Eq. (2) was calculated by taking the natural logarithm of the resulting soil transmissivity values. Some polygons in the soil map were classified as water, rough broken land, shale, quarry, and pits, sand and gravel, for which there was no attribute value on soil saturated hydraulic conductivity in the soil database. Since these soil classifications did not have any soil transmissivity, the grids with these soil classifications were assigned a value of −2, which approximated the lowest value of $ln(K_sD)$ calculated in New Jersey.

The resulting soil transmissivity layers for all counties were merged into a single transmissivity layer for New Jersey. The soil transmissivity layer was added to the statewide TWI to form the STI layer for New Jersey based on Eq. (2). The TWIs and STIs for the five water regions in New Jersey were clipped from the statewide TWI and STI layers based on the boundaries of water regions to ensure consistency when comparing across the regions.

2.2 Hydrologically Sensitive Areas Delineation

Hydrologically sensitive areas (HSAs) can be defined as the areas in a landscape with TI values greater than a given threshold. To explore the impact of threshold selection, this study adopted a wide range of thresholds from 8 to 14.5 with interval of 0.5 resulting in 14 different thresholds for delineating HSAs. Given two TIs (TWI and STI) and six regional scales (five water regions plus the state), this led to 168 maps of potential HSAs delineated for comparison with the FEMA map. The FEMA 100-year floodplain map was first rasterized using the statewide DEM as a snap layer so that the resulting floodplain raster grids were aligned with the TWI and STI grids. The statewide FEMA floodplain layer was then clipped into regional FEMA floodplain layers based on the boundaries of water regions. There were six FEMA floodplain layers (five water regions plus the state). Each FEMA floodplain layer was compared to the different delineated HSAs to determine the suitability of the TIs and the various thresholds for delineating HSAs representative of the 100-year floodplain.

2.3 Spatial Pattern Comparison Indicators

An error matrix can be used to conceptualize the comparison between a delineated HSAs (Map 1) and a FEMA floodplain map (Map 2) with two categories: in the HSAs or FEMA floodplain as indicated by "yes" (positive or 1) and outside as indicated by "no" (negative or 0)

(Fig. 2). In such an error matrix, the notation n_{ij} indicates the number of cells with category i in Map 1 and category j in Map 2 and i and j are either "1" or "0"; N_1 and N_0 are the numbers of cells within and outside the HSAs in a region; M_1 and M_0 are the numbers of cells within and outside the FEMA floodplains in a region; and N is the total number of cells in a region. Kuhnert et al. (2005) assess several spatial pattern comparison indicators based on such error matrix values. Four of these pattern comparison indicators were used here to assess the consistency between the delineated HSAs and the FEMA floodplain.

2.3.1 Hit Rate

The hit rate measures the percentage of cells in the FEMA floodplain map that are also identified as a part of HSAs. The hit rate indicates how successfully the HSAs would predict the floodplain. The hit rate simply indicates how well the model (the HSAs here) replicates the benchmark data (the FEMA floodplain here) without penalizing for overprediction (Sampson et al. 2015). Given the error matrix in Fig. 2, the hit rate is calculated as:

$$R_{hit} = {}^{n_{11}}/{M_1}. \tag{3}$$

2.3.2 Agreement Rate

A cell-by-cell comparison takes a cell in the HSA map and matches it with the corresponding cell in the FEMA floodplain map. In such a comparison, a match (either a positive or negative match) is counted as 1, and a mismatch is 0. The agreement rate is the number of matched cells divided by the total number of cells being compared (Kuhnert et al. 2005). Based on the error matrix in Fig. 2, the agreement rate based on the cell-by-cell comparison is calculated as follows:

$$R_{cbc} = (n_{11} + n_{00})/N. \tag{4}$$

2.3.3 Overall Comparison Rate

The overall comparison is an improvement over the cell-by-cell comparison by measuring the similarity after taking into consideration the disagreement measured by the differences in the number of cells in each category (Pontius 2002; Kuhnert et al. 2005). Given the error matrix and notations defined in Fig. 2, the overall agreement rate based on the overall comparison is calculated as follows:

		FEMA 100-Year Floodplain		Row Total
		Yes (1)	No (0)	
Hydrologically Sensitive Areas (HSAs)	Yes (1)	n_{11}	n_{10}	$N_1=n_{11}+n_{10}$
	No (0)	n_{01}	n_{00}	$N_0=n_{01}+n_{00}$
Column Total		$M_1=n_{11}+n_{01}$	$M_0=n_{10}+n_{00}$	$N=n_{11}+n_{10}+n_{01}+n_{00}$

Fig. 2 The error matrix for comparing the FEMA 100-year floodplain and the delineated hydrologically sensitive areas

$$R_{oall} = 1 - \frac{1}{N} \sum_{i=0}^{1} |N_i - M_i| = 1 - 2|n_{10} - n_{01}|/N. \tag{5}$$

2.3.4 Kappa Value

The kappa value measures the agreement of two spatial patterns by taking away some of the agreement that likely occurred by chance (Cohen 1960) and has been widely used for spatial pattern comparison. Given the error matrix in Fig. 2, the chance agreement can be measured as

$$R_c = \sum_{i=0}^{1} (N_i \cdot M_i)/N^2. \tag{6}$$

The kappa value is then measured as the ratio of the difference between the observed cell-by-cell agreement and the chance agreement to the difference between the maximum agreement and the chance agreement calculated as follows:

$$R_{kappa} = (R_{cbc} - R_c)/(1 - R_c). \tag{7}$$

It should be noted that a higher TI threshold would delineate fewer areas in a region as HSAs (i.e. a smaller HSA) and would consequentially result in fewer positive matches between the HSA and floodplain maps (i.e. n_{11}). Therefore, a lower hit rate (R_{hit}) would be expected for a higher TI threshold. A higher TI threshold might decrease the positive matches (i.e. n_{11}) but would increase the negative matches (i.e. n_{00}). The change in the agreement rate based on the cell-by-cell comparison (R_{cbc}) depends on how much the increase in the negative matches offsets the decrease in the positive matches. The agreement rate would increase when the increase in the negative matches is faster than the decrease in the positive matches and decrease when the relationship is slower. The overall comparison rate (R_{oall}) is very similar to the agreement rate with the goal of minimizing the mismatches ($n_{10} - n_{01}$). The kappa value (R_{kappa}) is similar to the overall comparison rate with additional consideration of taking away the matches by chance. It is expected that both the overall comparison rate and the kappa value would behave similarly to the agreement rate as the TI threshold increases. The change in these rates (R_{cbc}, R_{oall}, and R_{kappa}) from increasing to decreasing indicate that there would be peak levels in these indicators; in other words, these indicators tend to converge to their own levels across the gradient of TI thresholds.

2.4 Selecting TI Thresholds for HSA Delineation

HSAs and floodplains are different but interconnected features in a landscape. They overlap because the floodplain is the area in a landscape with sufficient accumulated runoff. The runoff-contributing areas of the HSAs vary by different storm events and consequently would result in different extents of flooded areas. Given their interconnection, a plausible explanation can be made using the occurrence of peak values in the spatial pattern comparison indicators. Namely, the spatial extents of the HSAs delineated by the specific TI threshold corresponding to the highest comparison indicator value represent the runoff-contributing areas to the FEMA 100-year floodplain.

In other words, the TI thresholds corresponding to these peak levels would delineate the HSAs that have the most consistent spatial patterns with the FEMA floodplain according to the respective spatial comparison indicators. Therefore, we can use these indicator values as

objective functions such that identification of their peak levels allows for a systematic approach for selecting "best" TI thresholds and subsequently delineating management-relevant HSAs from regional to state scales. We present results of one possible approach to do that leveraging our process-based understanding of hydrology at the regional and state scales. Further, we consider potential strengths and weaknesses of this approach.

3 Results

The distribution of the FEMA 100-year floodplain, as expected, was closely associated with waterbodies, such as streams, rivers and oceans (Fig. 1a). The darker areas (Fig. 1b and c) indicate higher TWI and STI values and therefore represent HSAs more prone to generate runoff. The pattern of these darker areas not only captures the distribution of the floodplain, but also includes some HSAs outside the floodplain that likely contribute to flood waters. The STI pattern was similar to that of the TWI with some subtle variability across the state caused by the inclusion of soil transmissivity. The TWI values ranged from 2.60 to 26.20 with a mean of 9.04 and a standard deviation of 2.43. The STI values ranged from 0.92 to 28.23 with a mean of 7.94 and a standard deviation of 2.77. The deep topsoil in the Atlantic Coast region generally made STI values lower than TWI values (i.e. more dark areas in the region as shown in Fig. 1b) while the shallow topsoil and urbanized landscape in the Northeast region made STI values higher than TWI values (i.e. more dark areas in the region as shown in Fig. 1c).

3.1 Comparison at the State Scale

As expected, both HSAs and hit rate (R_{hit}) decreased as the selected threshold increases (Fig. 3). As shown in Fig. 3a, when using the threshold value of 8.0, for example, the TWI delineated 64% of the state as HSAs with the hit rate of 0.90 such that 90% of the FEMA 100-year floodplain was within the HSAs. As the threshold level was increased to 14.5, only 1% of the state was delineated as HSAs with a corresponding hit rate of 4%. The same pattern also held when using STI to delineate HSAs (Fig. 3b). The higher hit rate indicates a larger overlap between the FEMA floodplain and the delineated HSA. Although hit rate cannot be used to identify the TI threshold due to its monotonic change in response to the changes in threshold value, it can be used to assess which comparison indicators better identify the TI threshold that delineates a HSA that would be most consistent with the FEMA floodplain.

The other three spatial comparison indicators converged toward peak values at various TI thresholds. The cell-by-cell agreement rate (R_{cbc}) converged to a peak for a TWI of 11.5 and a STI of 11.0, which delineated 15% and 12% of the state as HSAs, respectively, and had corresponding hit rates of 45% and 34%, respectively. The overall comparison rate (R_{oall}) was the highest for a TWI of 10.5 and a STI of 9.5, which delineated 23% and 23% of the state as HSAs, respectively, and had corresponding hit rates of 59% and 52%, respectively. The highest kappa values (R_{kappa}) were achieved for a TWI of 10.5 and a STI of 10.0, which delineated 23% and 18% of the state as HSAs, respectively, with hit rates of 59% and 46%, respectively. However, the convergence pattern (e.g. the peak of the objective function) was much clearer for R_{oall} and R_{kappa} than for R_{cbc}. After its peak value, R_{cbc} only decreases slightly while R_{oall} and R_{kappa} had clear optima.

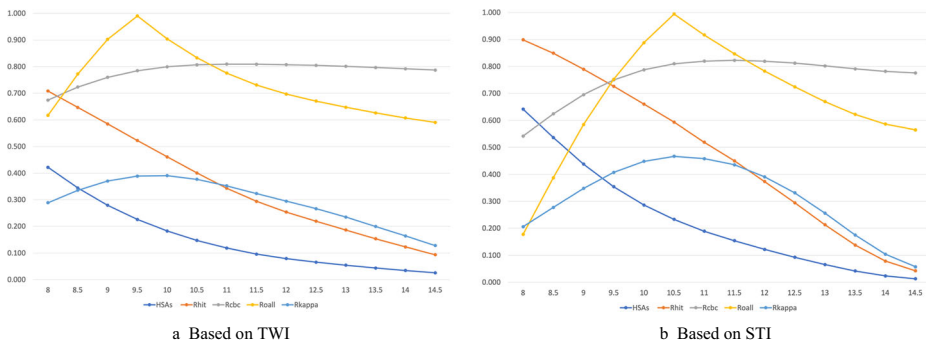a Based on TWI   b Based on STI

Fig. 3 Spatial comparison indicators under different thresholds in New Jersey

Such convergence implies that the HSAs delineated using a certain threshold have the most consistent spatial pattern in comparison with the FEMA 100-year floodplain. The HSAs delineated for these given thresholds would closely approximate the runoff generating areas that contribute water to the 100-year flood. Specifically, the convergence patterns derived from the spatial comparison have helped narrow the possible values to be considered as threshold candidates from 28 to 5 and begin to outline a systematic approach to identifying management areas based on HSAs.

3.2 Comparison at the Regional Scale

In the Atlantic Coast water region (Table 1), R_{cbc} and R_{kappa} converged at a TWI of 11.0, which delineated 30% of the region as HSAs with a hit rate of 61%. R_{oall} was at its highest level for a TWI of 10.5, delineating 36% of the Atlantic Coast region as HSAs with a hit rate of 68%. R_{cbc} converged to a peak for a STI of 9.5 which delineated 26% of the region as HSA with a hit rate of 54%. The highest R_{oall} value occurred at a STI value of 8.5, which delineated 35% of the region as HSAs with the hit rate of 66%. R_{kappa} was the highest for a STI of 9.0 delineating 30% of the region as HSAs with a hit rate of 60%.

In the Lower Delaware water region, R_{cbc} converged for a TWI of 11.5, delineating 15% of the region as HSAs with a hit rate of 37% while both R_{oall} and R_{kappa} peaked at a TWI of 10.5 delineated 25% of the region as HSAs with the hit rate of 54%. The R_{cbc} was the highest for a STI of 11.0 delineating 11% of the region as HSAs with a hit rate of 29%. R_{oall} and R_{kappa} converged to peaks at STIs of 9.0 and 9.5, respectively, which delineated 27% of the region as HSAs with a hit rate of 53% and 21% of the region as HSAs with a hit rate of 46%, respectively.

In the Northeast water region, R_{cbc} converged at a TWI of 11.0 and a STI of 11.5, which delineated 15% of the region as HSAs with a hit rate of 50% and 13% of the region as HSAs with a hit rate of 41%, respectively. With a TWI of 10.0 there was 21% of the region delineated as HSA with a hit rate of 62% corresponding to the highest values in both R_{oall} and R_{kappa}. The R_{oall} was the highest for a STI of 10.0, which delineated 21% of the region as HSAs with a hit rate of 57%. The R_{kappa} was the highest at a STI of 10.5 that delineated 18% of the region as HSAs with a hit rate of 52%.

For the Raritan water region, R_{cbc} identified the optimal TWI of 12.5 and STI of 13, which delineated 5% of the region as HSAs with a hit rate of 22% and 4% of the region as HSAs with a hit rate of 21%, respectively. R_{oall} converged at a TWI of 10.5 and a STI of 10.5 that

Table 1 The list of TI threshold candidates identified by spatial comparison in New Jersey and its five water regions

Water Region	TI	Indicator	Threshold (a)	HSAs (%) (b)	Hit Rate (%) (c)	Effectiveness (d*)
State	**TWI**	R_{cbc}	11.5	15.39	45.01	2.92
		R_{oall}/R_{kappa}	**10.5**	**23.31**	**59.40**	**2.55**
	STI	R_{cbc}	11	11.84	34.32	2.90
		R_{oall}	9.5	22.57	52.26	2.32
		R_{kappa}	10	18.25	46.11	2.53
Atlantic Coast	**TWI**	**R_{cbc}/R_{kappa}**	**11**	**30.2**	**61.1**	**2.03**
		R_{oall}	10.5	35.8	67.7	1.89
	STI	R_{cbc}	9.5	25.9	54.1	2.09
		R_{oall}	8.5	35.4	65.7	1.85
		R_{kappa}	9	30.4	60.0	1.98
Lower Delaware	**TWI**	R_{cbc}	11.5	14.6	36.7	2.52
		R_{oall}/R_{kappa}	**10.5**	**24.7**	**54.1**	**2.19**
	STI	R_{cbc}	11	10.8	28.9	2.69
		R_{oall}	9	26.9	52.9	1.97
		R_{kappa}	9.5	21.4	46.1	2.16
Northwest	TWI	R_{cbc}	11	15.2	50.4	3.31
		R_{oall}/R_{kappa}	10	21.4	62.3	2.92
	STI	R_{cbc}	11.5	12.5	41.3	3.32
		R_{oall}	10	21.2	57.3	2.71
		R_{kappa}	**10.5**	**17.6**	**51.8**	**2.95**
Raritan	**TWI**	R_{cbc}	12.5	4.8	21.5	4.45
		R_{oall}	10.5	14.6	45.9	3.15
		R_{kappa}	**11**	**11.2**	**39.7**	**3.55**
	STI	R_{cbc}	13	4.3	20.9	4.88
		R_{oall}	10.5	14.0	41.4	2.95
		R_{kappa}	11	10.8	36.2	3.36
Upper Delaware	TWI	R_{oall}	11	7.7	37.6	4.92
		R_{kappa}	10.5	9.5	43.1	4.53
	STI	R_{oall}	10.5	8.5	38.0	4.48
		R_{kappa}	**11**	**6.7**	**33.6**	**4.99**

*Column d is the ratio of Column c to Column b

delineated 15% and 14% of the region as HSAs, respectively, with hit rates of 46% and 41%, respectively. The highest R_{kappa} values were achieved at a TWI of 11.0 and a STI of 11.0, which both delineated 11% of the region as HSAs with hit rates of 40% and 36%, respectively.

In the Upper Delaware water region, R_{cbc} did not identify any threshold, as it did not converge. The R_{oall} values reached to their peak levels at a TWI of 11.0 and a STI of 10.5, which delineated roughly 8% and 9% of the region as HSAs, respectively, and both had a hit rate of 38%. The highest R_{kappa} identified a TWI of 10.5 and a STI of 12.0 delineating 10% and 7% of the region as HSAs, respectively, with hit rates of 43% and 34%, respectively.

3.3 TI Threshold Selection

Using three spatial comparison indicators, namely agreement rate (R_{cbc}), overall comparison rate (R_{oall}) and kappa value (R_{kappa}), complimented with the hit rate (R_{hit}), the threshold candidates can be quickly narrowed from 28 to something between 4 to 6 candidates across the water regions in New Jersey as presented above. While this signifies a much smaller range of TI values to consider, it is still a daunting task to select one TI threshold for delineating

HSAs for water resources management. To address this, we propose a possible two-step process to finalize the selection.

First, the agreement rate is not a reliable indicator to be used to identify the TI threshold, as it fails to identify a clear threshold in some cases. Further, even if the agreement rate does converge, it always identifies the TI thresholds that delineate the most restrictive sets of HSAs. For example, the STI threshold of 11.0, which was identified by the agreement rate, delineated 12% of the state as HSAs while STI thresholds of 10.0 and 9.5, which were identified by the kappa value and overall comparison rate, respectively, delineated 18% and 23% of the state as HSAs, respectively. Such an observation holds true for the TWI thresholds at the state scale as well as for both TIs across the five water regions in New Jersey.

Second, the effectiveness of identification can be used to further refine the selection of the TI threshold. The identification effectiveness measures how successfully the identified HSAs are with regard to their overlap with the FEMA floodplain. The identification effectiveness is defined as the ratio of the hit rate to the percentage of HSA (see last column of Table 1). The best TI threshold, thus, would be the one with the highest effectiveness. For example, the TWI threshold of 11.0 would be selected as the optimal threshold because it has the highest effectiveness ratio (2.03) among the TI threshold candidates at the state scale.

Using the two-step process, the final selection of the TI thresholds in New Jersey and its five water regions are highlighted in bold in Table 1. It is also worthy to note that all recommended TIs and their corresponding thresholds (after other considerations) are consistent with the recommendations based on the kappa value. Such congruence between the effectiveness ratio and kappa value indicates that the kappa value is an efficient spatial comparison indicator for identifying the TI threshold. Of course, this holds for delineating HSAs that approximate runoff contributing areas to the floodplain. The efficiency of the kappa value comes from considering both positive and negative matches and controlling for matches by chance.

4 Discussions

This study presents a feasible approach to identify a TI threshold for delineating HSAs at a regional to state scale in a landscape where HSAs approximate the runoff-contributing areas to the FEMA 100-year floodplain. Spatial comparison indicators allowed us to effectively narrow a wide range of potential thresholds to smaller ranges of TI values; however, these indicators alone were not sufficient to identify a specific TI threshold without comparison for validation. Clearly, additional discretion and knowledge are needed to select the proper TI threshold to delineate HSAs. Specifically, for this research, we determined that the agreement rate alone was not reliable enough for selecting the TI threshold. Rather, based on our experience in this region and on the process-based connections between HSAs and floodplains, the concept of identification effectiveness needed to be introduced to finalize the selection of one "best" TI threshold for each region (Table 1).

While our two-step approach is not the only way to finalize the TI selection, we feel it does achieve the goal of bringing in knowledge of the hydrological processes in the region relevant for water resource management. It is interesting to note that the two-step process allowed TWI thresholds to be selected for the Atlantic Coast and Lower Delaware water regions and STI thresholds to be selected for the Northeast and Upper Delaware water regions (Table 1), which is consistent with our local knowledge of hydrological processes (Qiu 2009; Lyon et al. 2018).

The Atlantic Coast and Lower Delaware water regions, for example, are located within the physiographic province of the Coastal Plain dominated by shallow water tables. These regions also have been experiencing extensive floods in recent years, which likely make TWI better to delineate HSAs in these regions as topographical convergence dominates flow pathway distributions. The Northeast and Upper Delaware water regions are located in the physiographic provinces of the Valley and Ridges, Highlands and Piedmont. Those regions have less extensive floods relative to the Coastal Plain province and also have various restrictive soil layers that dictate the interflow movement of water. These natural conditions likely make STI better to delineate HSAs in those regions as soil transmissivity is a limiting factor for flow pathway distributions. The only exception here is the selection of TWI for delineating HSAs in the Raritan water region. Our prior knowledge identified STI is an appropriate index to delineate HSAs in the region dominated by the Highlands and Piedmont physiographic provinces (Qiu 2009; Qiu et al. 2014). This discrepancy highlights the challenge of selecting simple techniques for mapping out complex hydrological interactions across regional scales and helps reiterates the need for interpretation based on the understanding of underlying hydrological process.

Furthermore, the two-step approach presented here was developed to eventually be use by landscape managers and/or regulatory agencies to address local water resource concerns allowing for the involvement of diverse stakeholders. Stakeholder engagement would eventually introduce other factors such as local stakeholder preferences, implementation feasibility, and availability of resources for conservation into the delineation of HSAs (Qiu 2009). Engagement of stakeholders through a participatory process is a powerful technique to implement such an approach successfully as it would allow for feedback and criterion-based voting systems to ensure buy-in on water resources planning and management actions (Lyon et al. 2018). Our approach presented here can provide a short list of TI threshold candidates that would be ideal for such a stakeholder engagement process to produce a meaningful threshold from which realistic and consensual goals can be met. As such, our approach provides a dynamic rather static basis to landscape managers for selecting a TI threshold and delineating a meaningful spatial extent to direct their planning and management actions at regional to state scales.

5 Conclusions

This study used four spatial comparison indicators to assess the consistency between the FEMA 100-year floodplain and HSAs delineated using a gradient of TI thresholds in New Jersey at the state scale and across five water regions. Convergence of indicators varied by regions. Such information was combined with other local knowledge, such as flood extent and local topographic and soil conditions, to identify the "best" TI and the corresponding optimal threshold for delineating HSAs to approximate the runoff-contributing areas to a 100-year flood. Altogether, this study provides a practical approach for HSA delineation to support efficient and effective landscape and water resources management. The assessments presented in this study offer a first-cut analysis of the complicated spatial hydrological responses of a landscape at regional to state scales. The main strength in a threshold-based approach for HSA delineation such as that developed here are the simplicity and the lack of requirement of ancillary data for training models. Depending on the landscape problems to be addressed, such as flood mitigation, nonpoint source pollution prevention or soil erosion control, the approach

developed here offers transparency and flexibility in selecting slightly different TI thresholds to delineate larger or smaller HSAs for targeting specific best management practices.

References

Agnew LJ, Walter MT, Lembo A, Gérard-Marchant P, Steenhuis TS (2006) Identifying hydrologically sensitive areas: bridging science and application. J Environ Manag 78:64–76

Amin MM, Veith TL, Collick AS, Karsten HD, Buda AR (2017) Simulating hydrological and nonpoint source pollution processes in a karst watershed: a variable source area hydrology model evaluation. Agric Water Manag 180:212–223

Anderson TR, Groffman PM, Walter MT (2015) Using a soil topographic index to distribute denitrification fluxes across a northeastern headwater catchment. J Hydrol 522:123–134

Archibald J, Buchanan B, Fuka DR, Georgakakos CB, Lyon SW, Walter MT (2014) A simple, regionally parameterized model for predicting nonpoint source areas in the northeastern US. Journal of Hydrology: Region Studies 1:74–91. https://doi.org/10.1016/j.ejrh.2014.06.003

Beven KJ, Kirkby MJ (1979) A physically based, variable contributing area model of basin hydrology. Hydrol Sci Bull 24:43–69

Buchanan BP, Fleming M, Schneider RL, Richards BK, Archibald J, Qiu Z, Walter MT (2014) Evaluating topographic wetness indices across Central New York agricultural landscapes. Hydrol Earth Syst Sci 18: 3279–3299

Cohen J (1960) A coefficient of agreement for nominal scales. Educational and Psychological Measurements 20(1):37–46

Conrad O, Bechtel B, Bock M, Dietrich H, Fischer E, Gerlitz L, Wehberg J, Wichmann V, Boehner J (2015) System for automated Geoscientific analyses (SAGA) v. 2.1.4. Geosci. Model Dev 8:1991–2007. https://doi.org/10.5194/gmd-8-1991-2015

Costa-Cabral MC, Burges SJ (1996) Digital elevation model networks (DEMON): a model of flow over hillslopes for computation of contributing and dispersal areas. Water Resour Res 30:1681–1692

Dahlke HE, Easton ZM, Fuka DR, Lyon SW, Steenhuis TS (2009) Modelling variable source area dynamics in a CEAP watershed. Ecohydrology 2:337–249. https://doi.org/10.1002/eco.58

Giri S, Qiu Z, Zhang Z (2017) A novel technique for establishing a soil topographic index threshold in defining hydrologically sensitive areas in landscapes. J Environ Manag 200:391–399

Heathwaite L, Quinn PF, Hewett C (2005) Modelling and managing critical source areas of diffuse pollution from agricultural land using flow connectivity simulation. J Hydrol 304(1–4):446–461

Herron NF, Hairsine PB (1998) A scheme for evaluating the effectiveness of riparian zones in reducing overland flow to streams. Aust J Soil Res 36:683–698

Horn BK (1981) Hill shading and the reflectance map. Proc IEEE 69:14–47

Kaykhosravi S, Abogadil K, Khan UT, Jadidi MA (2019) The low-impact development demand index: a new approach to identifying locations for LID. Water 11(11):2341

Kuhnert M, Voinov A, Seppelt R (2005) Comparing raster map comparison algorithms for spatial modeling and analysis. Photogramm Eng Remote Sens 71(8):975–984

Kulasova A, Blazkova S, Beven K, Rezacova D, Cajthaml J (2014) Vegetation pattern as an indicator of saturated areas in a Czech headwater catchment. Hydrol Process 28(20):5297–5308

Lei S, Chen H, Bian Z, Liu Z (2016) Evaluation of integrating topographic wetness index with backscattering coefficient of TerraSAR-X image for soil moisture estimation in a mountainous region. Ecol Indic 61(2): 624–633

Lyon SW, Cunha L, DiBlasio M, Creveling E (2018) Leveraging a participatory process for restoration return on investment: the nature Conservancy's floodplain investment tool. J Am Water Resour Assoc 54(6):1285–1301. https://doi.org/10.1111/1752-1688.12687

Lyon SW, Gérard-Marcant P, Walter MT, Steenhuis TS (2004) Using a topographic index to distribute variable source area runoff predicted with the SCS-curve number equation. Hydrol Process 18(15):2757–2771

Lyon SW, Lembo AJ, Walter MT, Steenhuis TS (2006) Defining probability of saturation with indicator kriging on hard and soft data. Adv Water Resour 29(2):181–193

Martin-Mikle CJ, de Beurs KM, Julian JP, Mayer PM (2015) Identifying priority sites for low impact development (LID) in a mixed-use watershed. Lands and urban planning 140:29–41

Planchon O, Darboux F (2001) A fast, simple and versatile algorithm to fill the depressions of digital elevation models. Catena 46:159–176

Pontius RG (2002) Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions. Photogramm Eng Remote Sens 68(10):1041–1049

Qiu Z (2009) Assessing critical source areas in watersheds for conservation buffer planning and riparian restoration. Environ Manag 44:968–980

Qiu Z, Hall C, Drewes D, Messinger G, Prato T, Hale K, Van Abs D (2014) Hydrologically sensitive areas, land use controls, and protection of healthy watersheds. J. Water Resour. Plann. Manage 140(7):04014011-1-10. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000376

Qiu Z, Kennen JG, Giri S, Walter MT, Kang Y, Zhang Z (2019) Reassessing the relationship between landscape alteration and aquatic ecosystem degradation from a hydrologically sensitive area perspective. Sci Total Environ 650:2850–2862

Sampson CC, Smith AM, Bates PD, Neal JC, Alfieri L, Freer JE (2015) A high-resolution global flood hazard model. Water Resour Res 51(9):7358–7381

Seibert J, McGlynn BL (2007) A new triangular multiple flow direction algorithm for computing upslope areas from gridded digital elevation models. Water Resour. Res. 43(4):W04501. https://doi.org/10.1029/2006WR005128

Temimi M, Leconte R, Chaouch N, Sukumal P, Khanbilvardi R, Brissette F (2010) A combination of remote sensing data and topographic attributes for the spatial and temporal monitoring of soil wetness. J Hydrol 388(1–2):28–40

Walter MT, Steenhuis TS, Mehta VK, Thongs D, Zion M, Schneiderman E (2002) A refined conceptualization of TOPMODEL for shallow-subsurface flows. Hydrol Process 16:2041–2046

Walter MT, Walter MF, Brooks ES, Steenhuis TS, Boll J, Weiler K (2000) Hydrologically sensitive areas: variable source area hydrology implications for water quality risk assessment. J Soil Water Conserv 55(3): 277–284

Watt MK (2000) A hydrologic primer for New Jersey watershed management. USGS report no. 2000-4140. US geological survey, West Trenton, New Jersey