EARTH SURFACE PROCESSES AND LANDFORMS
Earth Surf. Process. Landforms (2018)

Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/esp.4431

Modeling wet headwater stream networks across multiple flow conditions in the Appalachian Highlands

Carrie K. Jensen,[1,2]* Kevin J. McGuire,[1,2] Yang Shao[3] and C. Andrew Dolloff[4]

[1] Department of Forest Resources and Environmental Conservation (MC 0324), Cheatham Hall, STE 210, Virginia Tech, 310 West Campus Drive, Blacksburg, VA 24061, USA
[2] Virginia Water Resources Research Center (MC 0444), Cheatham Hall, STE 210, Virginia Tech, 310 West Campus Drive, Blacksburg, VA 24061, USA
[3] Department of Geography (MC 0115), Major Williams Hall, RM 113, Virginia Tech, 220 Stanger Street, Blacksburg, VA 24060, USA
[4] USDA Forest Service, Southern Research Station, Center for Aquatic Technology Transfer, 1710 Research Center Drive, Blacksburg, VA 24060, USA

Received 27 September 2017; Revised 16 May 2018; Accepted 21 May 2018

*Correspondence to: Carrie K. Jensen, Virginia Water Resources Research Center, Cheatham Hall STE 210, Virginia Tech, 310 West Campus Drive, Blacksburg, VA 24061, USA. E-mail: ckjensen@vt.edu

ESPL
Earth Surface Processes and Landforms

ABSTRACT: Despite the advancement of remote sensing and geospatial technology in recent decades, maps of headwater streams continue to have high uncertainty and fail to adequately characterize temporary streams that expand and contract in the wet length. However, watershed management and policy increasingly require information regarding the spatial and temporal variability of flow along streams. We used extensive field data on wet stream length at different flows to create logistic regression models of stream network dynamics for four physiographic provinces of the Appalachian Highlands: New England, Appalachian Plateau, Valley and Ridge, and Blue Ridge. The topographic wetness index (TWI) was the most important parameter in all four models, and the topographic position index (TPI) further improved model performance in the Appalachian Plateau, Valley and Ridge, and Blue Ridge. We included stream runoff at the catchment outlet as a model predictor to represent the wetness state of the catchment, but adjustment of the probability threshold defining wet stream presence/absence to high values for low flows was the primary mechanism for approximating network extent at multiple flow conditions. Classification accuracy was high overall (> 0.90), and McFadden's pseudo R^2 values ranged from 0.69 for the New England model to 0.79 in the Appalachian Plateau. More notable errors included an overestimation of wet stream length in wide valleys and inaccurate reach locations amid boulder deposits and along headwardly eroding tributaries. Logistic regression was generally successful for modeling headwater streams at high and low flows with only a few simple terrain metrics. Modification and application of this modeling approach to other regions or larger areas would be relatively easy and provide a more accurate portrayal of temporary headwaters than existing datasets. © 2018 John Wiley & Sons, Ltd.

KEYWORDS: logistic regression; physiographic province; stream length; temporary streams; geospatial terrain analysis

Introduction

First- and second-order headwater streams account for 70–80% of river network length (Downing *et al.*, 2012) and transport water, sediment, nutrients, organic matter, and contaminants to downstream water bodies (Wohl, 2017). As a result, water quality largely reflects the state of contributing headwaters (Alexander *et al.*, 2007; Dodds and Oakes, 2008). Approximately half of headwaters are temporary ephemeral or intermittent streams that vary in wet length seasonally or between storms (Nadeau and Rains, 2007; Buttle *et al.*, 2012; Datry *et al.*, 2014). Changes in the active or 'wet' stream length can exceed 300% in humid (Jensen *et al.*, 2017) as well as more arid environments (Godsey and Kirchner, 2014). The spatial variability of flow duration along streams impacts solute and sediment loads from catchments, rates and types of biogeochemical reactions, availability of aquatic habitat, and movement of organisms (Larned *et al.*, 2010). However, maps that accurately portray the location of temporary streams across flow conditions are uncommon or nonexistent. Various logistical and technological challenges continue to impede a comprehensive inventory of headwaters and, instead, compel the use of modeling approaches to best represent these low-order stream systems.

The location and length of headwater streams are highly inaccurate on most existing maps (Skoulikidis *et al.*, 2017). The National Hydrography Dataset (NHD), which serves as the standard representation of river networks for many environmental models and watershed management programs in the US can underestimate headwater length by 200% or more

(Elmore *et al.*, 2013; Fritz *et al.*, 2013). Inconsistencies during map production contribute to uncertainty, as source data are not the same for all NHD, topographic, or Natural Resources Conservation Service (NRCS) maps (Colson *et al.*, 2008; Hughes *et al.*, 2011). Aerial photograph interpretation is the most common method for delineating streams, but numerous individuals analyzed photographs over decades using distinct techniques, cartographic standards, and levels of precision (Colson *et al.*, 2008; Hughes *et al.*, 2011). Field surveys provide the source data for some maps, although there is usually no indication of whether the surveyor mapped the actively flowing stream or geomorphic channel as the network, which can result in quite different representations (Adams and Spotila, 2005; Jensen *et al.*, 2017).

Maps rarely show the vast number of temporary headwater channels where networks expand and contract in wet length (Hansen, 2001). The temporary streams that do appear on maps receive categorical flow duration classifications as either intermittent or ephemeral. However, definitions of perennial, intermittent, and ephemeral streams are not standardized; classifications may reflect the frequency of flow (Hedman and Osterkamp, 1982; Hewlett, 1982), water table position (Hansen, 2001; Larned *et al.*, 2007), seasonal position of flow origins at base flow (Paybins, 2003), or geomorphological and biological characteristics (Feminella, 1996; Hansen, 2001; Johnson *et al.*, 2009). Researchers now highlight the need for objective, continuous measures of the frequency and duration of flow that provide more ecologically relevant information than categorical divisions (Leibowitz *et al.*, 2008; Larned *et al.*, 2010, 2011).

Flow duration is difficult to determine from aerial photographs, which capture the stream network at a single moment in time, or with limited field survey data. Perennial-intermittent and intermittent-ephemeral boundaries move between months or years (Fritz *et al.*, 2008). Furthermore, stream networks are often discontinuous, with perennial reaches separated by dry channel segments that may rarely flow (Datry *et al.*, 2014; Godsey and Kirchner, 2014). As a result, flow duration classifications for the NHD and NRCS maps generally have low accuracy, according to the given classification criteria (Svec *et al.*, 2005; Colson *et al.*, 2008, Fritz *et al.*, 2013). Map errors also vary regionally; the NHD tends to overestimate flow duration in arid environments and underestimate flow duration in humid areas (Fritz *et al.*, 2013).

Field surveying continues to be the most accurate and reliable mapping method for streams (Płaczkowska *et al.*, 2015) yet is not feasible over large or remote areas. The increasing availability and advancement of geospatial data and analysis techniques in recent decades offers opportunities as well as unique challenges for characterizing headwaters. High-resolution aerial photographs only permit the identification of streams in areas without heavy vegetative cover (Heine *et al.*, 2004). Satellite imagery and associated wetness or moisture indices can effectively classify larger water bodies and wetlands, but the spatial resolution is usually too coarse to detect headwater streams. One exception is high-resolution near-infrared LiDAR, which Hooshyar *et al.* (2015) employed to extract the wet stream network in a region with exposed channels. Even at sites without dense canopy cover, multiple collections of aerial photographs or LiDAR is necessary during both wet and dry conditions to accurately describe stream length variability. Such efforts are expensive and impractical in many cases, although unmanned aerial vehicles may help to alleviate this burden in the future (Spence and Mengistu, 2016).

Automated stream extraction methods utilize digital elevation models (DEMs), which are often freely available at resolutions finer than 10 m for the continental USA. Channel initiation thresholds based on flow accumulation, upslope area, slope, or stream power are adequate for fourth-order and larger channels (James and Hunt, 2010) but do not consistently locate smaller headwaters (Heine *et al.*, 2004). Alternatively, modeling the stream network with logistic regression outperforms channel initiation thresholds and other extraction techniques (Heine *et al.*, 2004; Sun *et al.*, 2011). Using variables like upslope area, slope, and curvature, logistic regression determines the probability of the presence of a stream channel at each catchment pixel. Despite the success of logistic regression for delineating headwaters, model accuracy remains lower for the temporary channel network (Russell *et al.*, 2015). Alternatively, because flow duration varies widely both along and among non-perennial streams, we believe modeling efforts should shift to continuous predictions of the wet, active network through time rather than a single, stable configuration of perennial or temporary channel length.

The object of this study is to utilize field data on the variability of wet stream length in forested catchments spanning four Appalachian physiographic provinces to create explanatory logistic regression models of the wet stream network as a function of topographic metrics and runoff. Jensen *et al.* (2017) found that expansion and contraction of headwater length with stream flow occurs but also differs considerably in magnitude across the Appalachian region. Flow permanence and, thus, wet stream length largely align with geologic controls such as lithology and the depth and heterogeneity of sediment (Winter, 2007; Jensen *et al.*, 2017), which are not always evident from geologic maps. The current study investigates the efficacy of using topography, which reflects the underlying geology, to predict the distinct observed stream length dynamics of the Appalachians. This project is one of the first known attempts to model headwaters as dynamic networks by including a continuous wetness state variable (stream runoff) as a model parameter (but see the use of accumulated precipitation by Sun *et al.*, 2011), whereas most published studies focus on reproducing the locations of geomorphic channel heads or static representations of channel length. We developed explanatory rather than predictive models, as our goal was to understand how terrain attributes correspond to flow permanence in each province rather than to maximize predictive accuracy (MacNally, 2000; Sainani, 2014). Detailed field data on the location and length of wet stream reaches across flow conditions enabled us to identify additional site characteristics aside from catchment topography that contribute to model inaccuracy and can, therefore, inform future stream delineation efforts.

Study area

We collected field data from headwater catchments in four physiographic provinces of the Appalachian Highlands: the New England (NE), Appalachian Plateau (AP), Valley and Ridge (VR), and Blue Ridge (BR). We selected three forested catchments smaller than 75 ha from each province (Table I), as we discovered that larger watersheds in these regions require more than one day for an individual to map the stream network. The study sites are in the Hubbard Brook Experimental Forest in New Hampshire for NE, Fernow Experimental Forest in West Virginia for AP, Jefferson National Forest at Poverty Creek and the South Fork of Potts Creek in Virginia for VR, and Coweeta Hydrologic Laboratory in North Carolina for BR (Figure 1). We conducted field mapping at Hubbard Brook, Fernow, and Coweeta because these experimental watersheds have weirs recording stream flow and reference areas that are not subject to timber harvests or other experimental manipulation. We chose the VR sites at tributaries to Poverty Creek and the South

Table I. Study area attributes. Modified from Jensen *et al*. (2017)

Physiographic province	Site name[a] (watershed number[b])	Latitude (°N), Longitude (°W)	Drainage area (ha)	Aspect	Mean elevation (m)	Geology	Vegetation
New England	NE13 (WS6)	43.95, 71.74	13.4	SE	690	schist, granulite	northern
	NE25	43.93, 71.77	25.1	NW	740	schist, granulite	hardwoods
	NE42 (WS3)	43.96, 71.72	42.4	S	632	schist, granulite	
Appalachian Plateau	AP14 (WS13)	39.06, 79.70	13.9	NE	773	shale, sandstone	mixed
	AP16 (WS10)	39.05, 79.68	15.7	SW	767	shale, sandstone	hardwoods
	AP37 (WS4)	39.05, 79.69	36.6	SE	822	shale, sandstone	
Valley and Ridge	VR25	37.28, 80.46	25.0	NW	750	shale, sandstone	mixed
	VR35	37.26, 80.48	34.8	N	729	shale, sandstone	hardwoods
	VR70	37.45, 80.49	69.9	S	1029	sandstone	
Blue Ridge	BR12 (WS18)	35.05, 83.44	12.4	NW	823	gneiss, amphibolite	oak-hickory
	BR33 (WS34)	35.06, 83.45	32.7	SE	1019	gneiss, amphibolite	and cove
	BR40 (WS40)	35.05, 83.46	39.6	E	1052	gneiss	hardwoods

[a]Numbers in site name indicate the drainage area in hectares.
[b]If applicable, designated watershed number at the experimental forests.

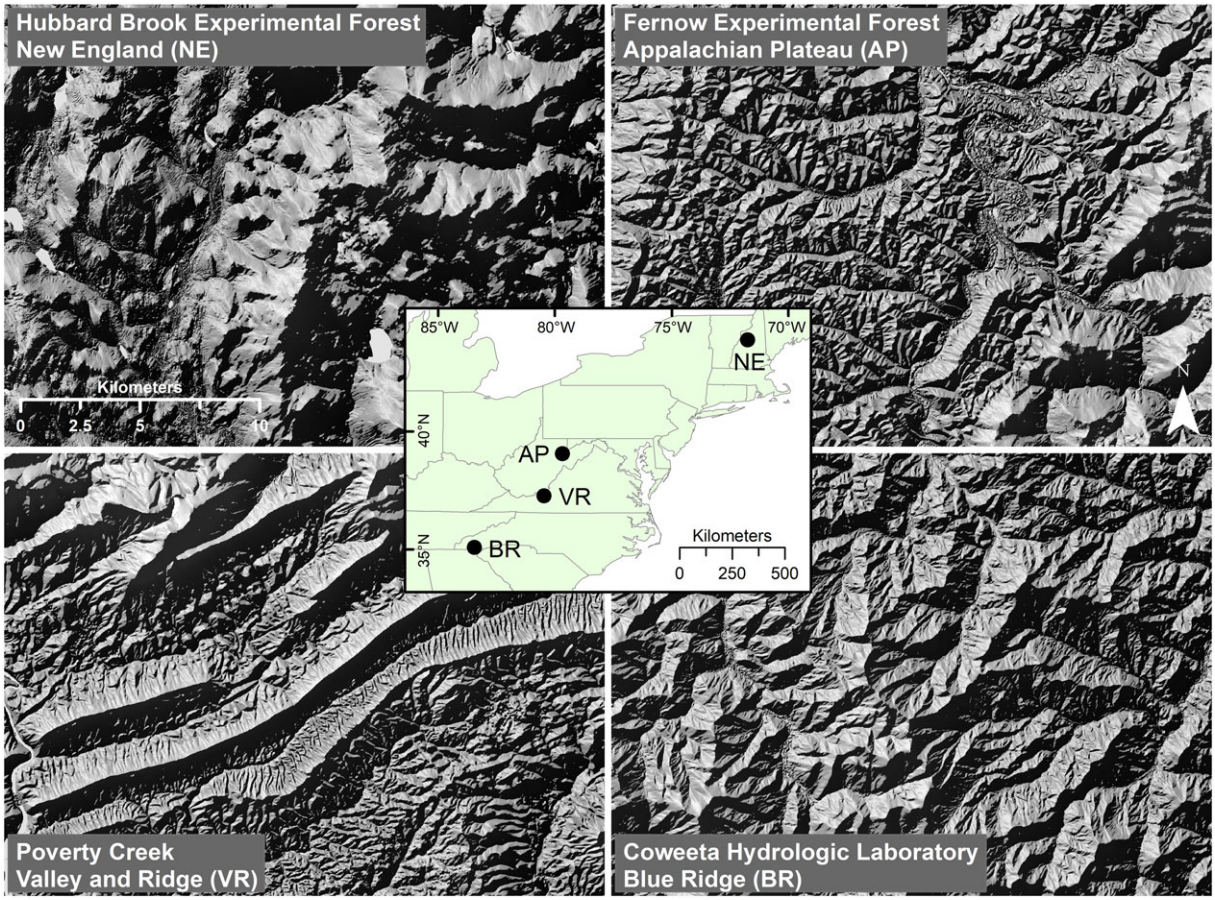

Figure 1. Hillshade views of regional topography in the four study areas. [Colour figure can be viewed at wileyonlinelibrary.com]

Fork of Potts Creek based on the availability of 3 m DEM data, mature forest coverage, road access, and the location of the catchments on National Forest.

The climate, geology, topography, and vegetation vary across the four Appalachian provinces (Table I; see Jensen *et al*. (2017) for further details). NE is the coldest study area and receives the most snow, so peak stream flow usually occurs after snowmelt in the late spring. The VR catchments are in a rain shadow and have the lowest precipitation totals, while the BR sites are the wettest. Snow is rare in BR due to warm winters. Substrate at the NE sites consists of glacial till and drift deposits that vary in thickness up to 8 m (Benettin *et al*., 2015) and are underlain by Lower Silurian pelitic schist and granulite of the Upper and Lower Rangeley Formation (Barton, 1997). The glacial landscape is rounded and hummocky with little channel incision (Figure 1). Soils are predominantly well-drained sandy loam Spodosols (Likens, 2013). The AP catchments are located in the gently folded Allegheny Mountains, which are characterized by steep slopes and broad, flat uplands (Morisawa, 1962). Devonian shales and sandstones of the Hampshire Formation form the underlying bedrock (Cardwell *et al*., 1968), and soils are shallow (< 1 m) silt and sandy loam Inceptisols

(Losche and Beverage, 1967). Folded and thrust faulted sedimentary rock with differing erodibilities creates the parallel ridges and associated trellis drainage pattern of the VR. The VR sites at Poverty Creek are underlain by Devonian shales, sandstones, and siltstones of the Brallier and Chemung Formations (Virginia Division of Mineral Resources, 1993), and bedrock at the South Fork of Potts Creek catchment includes Devonian Oriskany and Silurian Keefer and Rose Hill sandstones (Schultz *et al.*, 1986). Soils consist of shallow, well-drained stony and clay loam Inceptisols (Adams and Stephenson, 1983). Middle to Late Proterozoic biotite gneiss and amphibolite of the Coweeta Group and Tallulah Falls Formation underlie thick saprolite at the BR catchments (Hatcher, 1988). Dominant soils include sandy and gravelly loam Inceptisols and Ultisols (Velbel, 1988). All of the study areas have second-growth forests that range from northern hardwoods in NE to oak–hickory associations in BR.

Methods

Field data collection

We mapped the wet stream network of each study catchment seven times throughout 2015 and 2016 with a Bad Elf GNSS Surveyor Global Positioning System (GPS) unit, as detailed in Jensen *et al.* (2017). During each mapping, we walked along the stream from the catchment outlet until we found the flow origin of every tributary, marking disconnections in the wet network between flowing reaches or non-flowing pools. We included all surface water greater than 1 m in length as part of the wet network. The reported accuracy of the GPS unit is 1 m, but measured accuracy varied from 3 to 10 m depending on weather conditions and tree cover. To compensate for the lower accuracy, we used pin flags and field notes in addition to the GPS points to compare the location of the wet stream across mapping dates. The same individual also completed all mapping with the same GPS unit.

Eight of the twelve study catchments have weirs that gauge stream discharge at 5-min intervals. We performed salt dilution gauging (Calkins and Dunne, 1970) to measure stream flow at the remaining sites (NE25, VR25, VR35, and VR70) during each mapping. Our goal was to map the wet network across a range of flows between at least the 25 and 75% exceedance probabilities of mean daily discharge, rather than major storms or droughts. We only mapped during non-storm conditions or on the recession limb of storm events at least several hours after the hydrograph peak to minimize changes in discharge during an individual survey. Precipitation data were available for all sites, but some of the gauges were not in or immediately adjacent to the study catchments, which is especially problematic for isolated summer thunderstorms. Antecedent precipitation indices showed lower correlations with stream length than runoff (discharge normalized by catchment area), so we limited our wetness state variable to runoff as the more representative measurement of moisture conditions.

Terrain analysis

Following the procedure in Jensen *et al.* (2017), we brought the GPS points into ArcGIS (ArcMap version 10.3.1, ESRI 2015, Redlands, CA) to digitize the wet stream network along lines of high flow accumulation according to the multiple triangular flow direction algorithm (Seibert and McGlynn, 2007) applied to 3 m DEMs. DEMs with resolutions between 1 and 5 m tend to create more accurate stream networks (Li and Wong, 2010) and better predict channel head locations (Tarolli and Dalla Fontana, 2009) than coarser resolutions. Similarly, Gillin *et al.*, 2015a) found that terrain metrics calculated from 3 and 5 m DEMs are most comparable with field-derived values for one of the NE catchments used in this study. We completed all DEM processing and terrain analysis in ArcGIS and the System for Automated Geoscientific Analyses software (SAGA version 2.3.1). The 1/9 arc-second DEM for AP, BR, and VR70 are from the 2003 West Virginia and North Carolina National Elevation Datasets. LiDAR data collections occurred during leaf-off and snow-free conditions at NE in 2012 for the White Mountain National Forest and at VR25 and VR35 in 2011 for the Virginia Geographic Information Network. We re-sampled the bare earth DEMs classified from the LiDAR data to 1 m and then coarsened the DEMs to 3 m by mean cell aggregation. We used a low-pass (3 ×3) filter and sink-filling algorithm (Wang and Liu, 2006) for hydrological correction of all DEMs.

We moved GPS points located in low flow accumulation pixels due to positional error to the nearest cell of comparatively high flow accumulation, although the maximum displacement we allowed was 3 pixels (9 m) in accordance with the average GPS accuracy. Because we mapped each catchment seven times, we were able to compare GPS points from all of the mappings to delineate the most accurate network possible. We assigned points from each mapping date to the flow lines of this 'master' network to ensure consistency as we determined changes in wet stream length. We consulted field notes during this process to verify that point displacements between mappings were the result of variability in the wet network rather than GPS error.

We derived terrain metrics from the hydrologically-corrected 3 m DEMs. Metrics included: upslope accumulated area, calculated from the multiple triangular flow direction algorithm (Seibert and McGlynn, 2007); maximum slope (Travis *et al.*, 1975); topographic wetness index (Beven and Kirkby, 1979), based on the multiple triangular flow direction and maximum slope algorithms; downslope index (Hjerdt *et al.*, 2004); stream power index (Moore *et al.*, 1991); profile, planform, tangential, longitudinal, and cross-sectional curvature (Evans, 1979; Wood, 1996; Wilson and Gallant, 2000); topographic position index (Guisan *et al.*, 1999); and local lateral contributing area, calculated as the sum of the left- and right-side contributions (Grabs *et al.*, 2010). We used the DEM Surface Tools for ArcGIS (version 2.1.375) (Jenness, 2013) to produce the curvature metrics. We created an additional raster grid for each metric by calculating the mean of all pixel values in the upslope accumulated area contributing to a given cell. Thus, we developed 'local' as well as 'mean upslope' rasters for all metrics except for upslope accumulated area and local lateral contributing area.

Logistic regression modeling

We used logistic regression to model the probability of each catchment pixel being a 'wet' stream as a function of terrain metrics and runoff. In logistic regression, the relationship between the probability (p) of a binary response variable (in this case, the presence or absence of a wet stream at a pixel) and model predictors (x_i) can be expressed as

$$\hat{p} = \exp(b_0 + b_1x_1 + b_2x_2 + \ldots b_nx_n) / (1 + \exp(b_0 + b_1x_1 + b_2x_2 + \ldots b_nx_n)), \quad (1)$$

where $\hat{p}$ is the regression estimate of logit (p), which is

$$\text{logit}\,(p) = \log\,(p/(1-p)), \quad (2)$$

and b_0 is the model intercept, b_i (i = 1, 2, ..., n) are the regression coefficients, and x_i (i = 1, 2, ..., n) are the independent variables.

We transformed the delineated stream network from each mapping into a raster denoting the presence (1) or absence (0) of a wet stream. We utilized 100% of the wet stream cells with a value of 1 and randomly sampled 10% of the non-stream pixels with a value of 0 from across the entire catchment for modeling. Preliminary models using varying proportions of non-stream pixels showed a decrease in classification accuracy with more sample points but no apparent impact on the selection of model explanatory variables. We limited the sample of non-stream pixels to 10% owing to the comparatively low number of wet stream points and, additionally, because predictive accuracy was not the primary goal. We split the sampled stream and non-stream pixels into training (70%) and testing (30%) data for model building and validation, respectively.

Wet stream length dynamics and the mobility of flow origins reflect geologic attributes including lithology, structure, transmissivity, and depth to bedrock that vary at both local and regional scales (Winter, 2007; Whiting and Godsey, 2016; Jensen *et al.*, 2017). We created a separate model for each physiographic province in our study to examine how regional, rather than site-specific, topography can help explain flow permanence in headwaters. We extracted terrain metric values from the raster grids at the sample points and assigned the same runoff value as a constant representing overall catchment wetness to all points from a single mapping. We log-transformed variables with non-normal distributions and performed all statistical analysis and modeling in R 3.2.1 (R Core Team, 2015). Variable selection for each of the four final models was based on *P*-values

Table II. Model parameters. The *P*-value of all model parameters is << 0.001

Model	McFadden's pseudo R^2	Parameters	Coefficient	Standard error	t-statistic
NE	0.69	TWI (ln)	15.57	0.16	94.50
		Flow	0.12	0.007	18.57
AP	0.79	TWI (ln)	8.11	0.22	37.61
		TPI (mean[a])	–4.75	0.15	–31.91
		Cross-sectional Curv.	–0.28	0.01	–19.39
		Flow	0.13	0.02	7.13
VR	0.77	TWI (ln)	9.56	0.12	80.12
		TPI (mean[a])	–3.26	0.08	–42.18
		Longitudinal Curv.	0.17	0.006	30.93
		Flow	0.18	0.009	18.59
BR	0.76	TWI (ln)	7.69	0.10	73.95
		TPI	–1.75	0.04	–40.23
		Profile (mean[a]) Curv.	2.14	0.08	26.89
		Flow	0.02	0.004	4.21

[a]Variable denoted as 'mean' refers to the mean value of upslope pixels as opposed to the local value.

Table III. Probability threshold values selected by the optimization procedure and associated accuracy statistics for high and low flows. Omission and commission errors ≥ 0.20 in bold

Site	Flow	Probability threshold	Classification accuracy	Balanced accuracy	Omission error	Commission error
NE42	High	0.50	0.95	0.92	0.11	0.12
	Low	0.57	0.98	0.96	0.06	0.12
NE25	High	0.50	0.90	0.85	**0.24**	**0.20**
	Low	0.80	0.96	0.86	**0.27**	**0.21**
NE13	High	0.50	0.94	0.94	0.06	0.11
	Low	0.63	0.95	0.90	0.17	0.09
AP37	High	0.50	0.98	0.95	0.10	0.09
	Low	0.57	0.99	0.91	0.17	0.15
AP16	High	0.53	0.99	0.95	0.11	0.00
	Low	0.65	0.99	0.96	0.08	**0.21**
AP14	High	0.53	0.97	0.88	**0.22**	0.09
	Low	0.56	0.98	0.91	0.18	0.05
VR70	High	0.58	0.98	0.90	0.18	0.10
	Low	0.62	0.99	0.90	**0.20**	0.09
VR35	High	0.56	0.96	0.90	0.19	0.04
	Low	0.94	0.99	0.83	**0.33**	**0.43**
VR25	High	0.50	0.96	0.92	0.16	0.03
	Low	0.80	0.99	0.96	0.07	0.14
BR40	High	0.50	0.97	0.93	0.12	0.09
	Low	0.52	0.97	0.93	0.13	0.09
BR33	High	0.54	0.97	0.93	0.13	0.14
	Low	0.70	0.97	0.91	0.16	0.16
BR12	High	0.50	0.95	0.89	**0.20**	0.10
	Low	0.56	0.96	0.88	**0.24**	0.04

and step-wise improvements in McFadden's pseudo R^2 values, although we also sought to eliminate collinear predictors with a Pearson's correlation greater than 0.70 (Kuhn, 2008). We evaluated model performance with McFadden's pseudo R^2 values, classification accuracy, and errors of omission and commission of stream pixels. We repeated this process for each physiographic province to produce four models.

Selection of any probability threshold is possible with logistic regression to signify the presence of the response variable. We created an accuracy optimization procedure to choose threshold values for each catchment that produced the maximum classification accuracy for the highest and lowest stream flows of the testing dataset. We applied the models to the three study catchments in each of the four provinces to examine the spatial distribution of omission and commission errors for the selected probability thresholds.

Results

Model parameters

The natural logarithm of the topographic wetness index (TWI) was the most important parameter in all of the final models (Table II). We tested single-variable models using TWI to predict wet stream presence/absence, which resulted in McFadden's pseudo R^2 values of 0.67–0.68. TWI reflects both the upslope accumulated area and local slope at a pixel (Seibert and McGlynn, 2007). High TWI values indicating large contributing areas or low slopes increased the likelihood of stream presence. The AP, VR, and BR models all included either the local or mean upslope topographic position index (TPI), which increased McFadden's pseudo R^2 values by an additional 0.06–0.09 after the inclusion of TWI. TPI compares the elevation of a cell with the mean elevation in a neighborhood defined by a 100 m radius around the pixel: positive values correspond to ridges, and negative values indicate valleys (Guisan *et al.*, 1999). TPI model coefficients were all negative. We also selected curvature metrics for the AP, VR, and BR models, but these parameters only increased McFadden's pseudo R^2 values by 0.02–0.03. The AP model included cross-sectional curvature, which is the same as 'planform curvature' in the ArcGIS Spatial Analyst Toolbox (Zevenbergen and Thorne, 1987). Positive values of cross-sectional curvature correspond to convexities where water diverges, and negative values show concavities where flow converges (Jenness, 2013). Longitudinal curvature appeared in the VR model and is the same as 'profile curvature' in ArcGIS (Zevenbergen and Thorne, 1987). The BR model incorporated mean upslope profile curvature. Both longitudinal and profile curvature indicate whether water accelerates

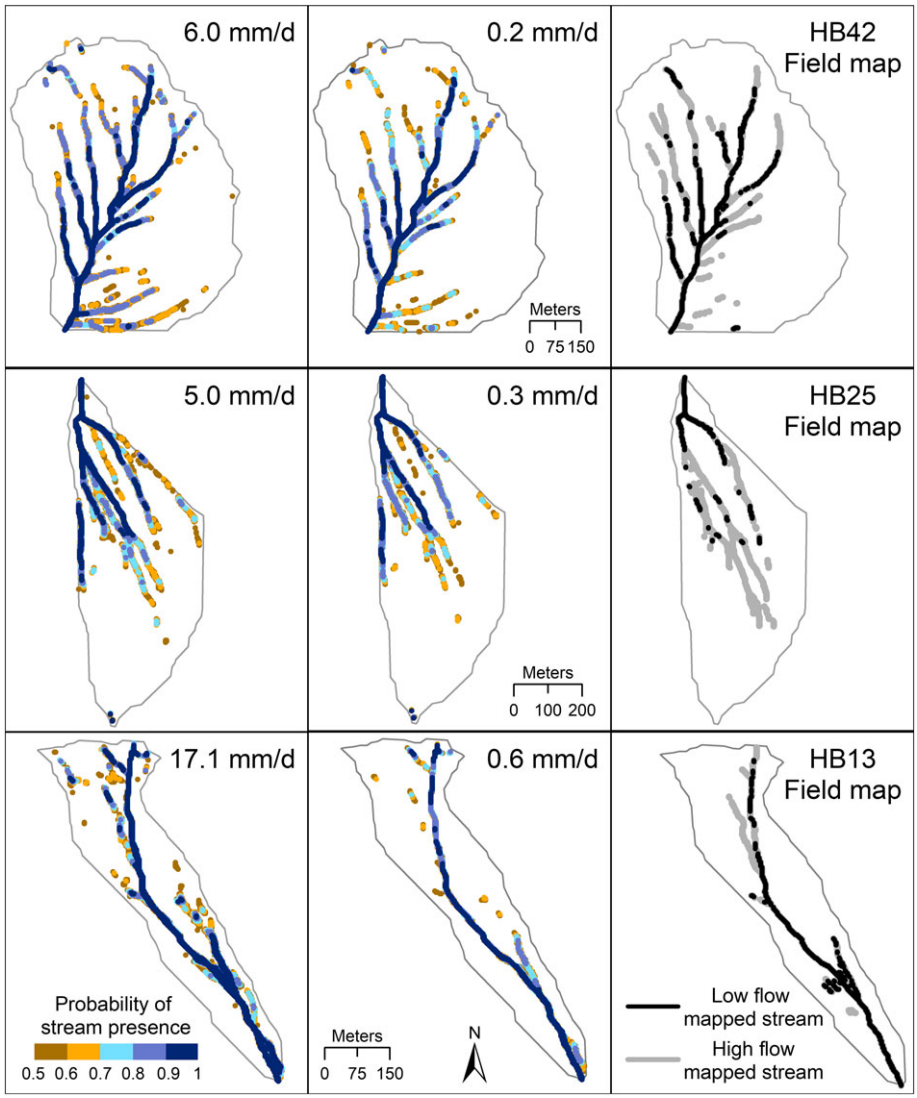

Figure 2. Modeled stream networks for high (left) and low (center) flows in NE. Right columns show field-surveyed streams for the modeled flows. Only probability values > 0.50 are shown. [Colour figure can be viewed at wileyonlinelibrary.com]

or decelerates when flowing over a point, with positive values showing deceleration at concavities and negative values suggesting acceleration at convex parts of the landscape (Jenness, 2013). Runoff also appeared in all four models yet increased McFadden's pseudo R^2 values only slightly (< 0.02). Owing to the large number of sample points, all of the final model predictors were highly significant ($P << 0.001$) in the models.

Model performance

McFadden's pseudo R^2 values were highest for AP (0.79) but were similar for VR (0.77) and BR (0.76) (Table II). NE had the lowest McFadden's pseudo R^2 value of 0.69. Classification accuracy was high overall, ranging from 0.90–0.99 across the study catchments (Table III). The optimum probability threshold to maximize classification accuracy for each catchment varied from 0.50–0.94, with higher threshold values corresponding to low flows. Omission and commission errors did not display systematic trends but were greatest for NE25, VR35, and BR12 at low flows.

The NE model was able to reproduce the complex pattern of tributaries in the three Hubbard Brook catchments (Figure 2). One of the more prominent errors was the presence of high probability values for a wet stream along the western border of NE25. We did not map a stream in the field at the western edge of NE25, but we believe the catchment boundary delineated from the DEM is slightly inaccurate, as the stream appears to flow into the neighboring catchment. Thus, a wet stream may have actually been present at this location on some of the mapping dates. Although we optimized the probability threshold for each site, we found a slightly higher threshold value of approximately 0.75 yielded a more visually accurate stream network pattern at high and low flows (Figure 3), particularly for NE13 and NE42.

The AP model effectively portrayed the simple network pattern of the Fernow catchments yet tended to overestimate the presence of wet streams at higher elevations (Figures 4 and 5). For AP16, the modeled network precisely matched the field map for the high flow but did not reflect the top-down contraction in stream length that occurred during the low flow, even at probability thresholds greater than 0.90. Similarly, we never observed surface water in the southwest tributary of AP37 as far upslope as the model indicated, and flow duration along the upper reaches of the main stem was much lower than the model predicted.

Model performance for the VR differed between VR70 at the South Fork of Potts Creek and the two catchments draining to Poverty Creek, VR25 and VR35. The VR model produced realistic stream networks in VR25 and VR35 at high flows but greatly overestimated wet stream length for the low flows (Figures 6 and 7). Stream length changed considerably between high and low flow conditions at these catchments, and we found that threshold values of 0.95–0.99 were necessary for a closer visual match between the model and field map at low flows. Nevertheless, the modeled network contraction patterns with decreasing runoff and increasing threshold values were consistent with field observations, with many of the same reaches drying first. Overestimation of stream length at low flows was less of an issue for VR70, which displayed much less network expansion and contraction. However, the modeled network continued too far upslope for both flow conditions, so increasing the probability thresholds to 0.80–0.90 visually improved the

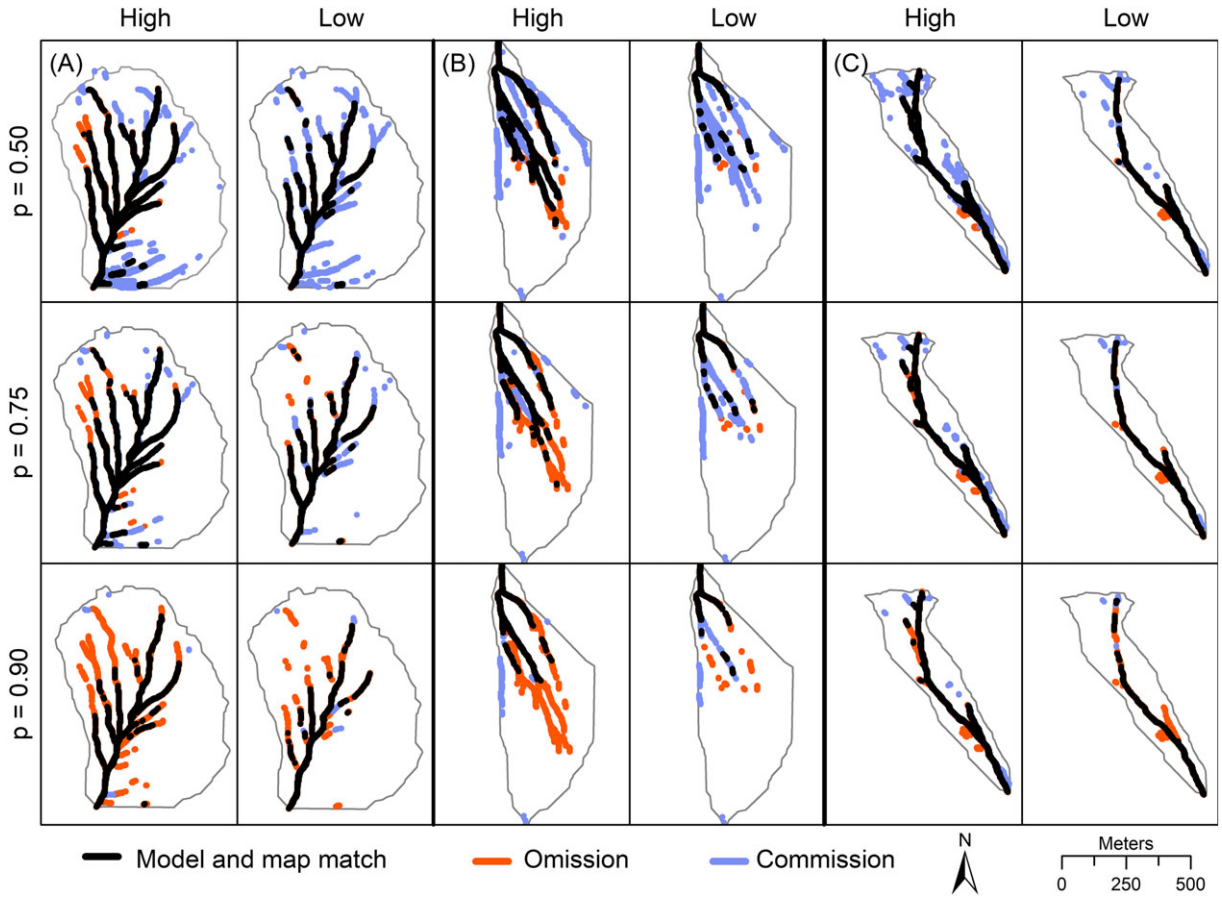

Figure 3. Omission and commission errors at NE42 (A), NE25 (B), and NE13 (C) at the highest and lowest mapped flows for probability thresholds of 0.50, 0.75, and 0.90. See Figure 2 for high and low runoff values. [Colour figure can be viewed at wileyonlinelibrary.com]

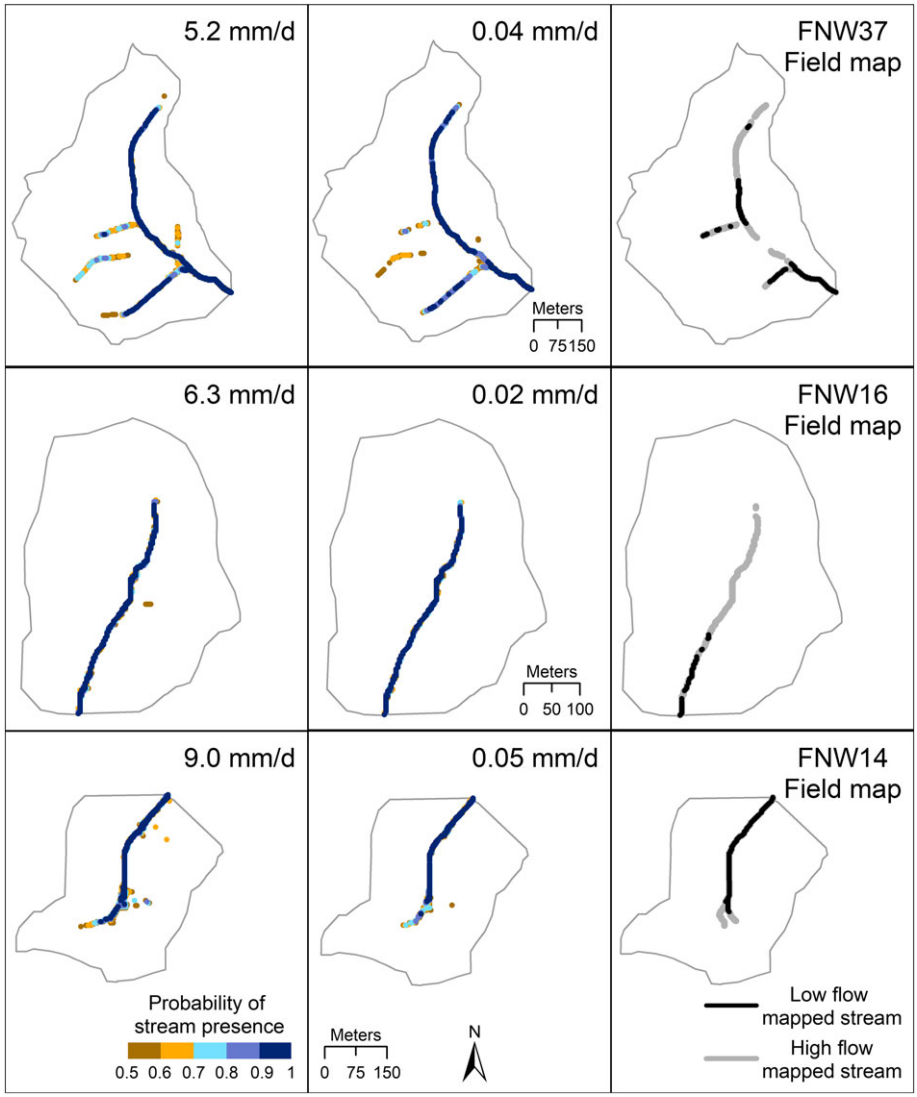

Figure 4. Modeled stream networks for high (left) and low (center) flows in AP. Right columns shows field-surveyed streams for the modeled flows. Only probability values > 0.50 are shown. [Colour figure can be viewed at wileyonlinelibrary.com]

delineation. The model also omitted a number of short, disconnected reaches that activated during high flows at VR70, although the total missing stream length was relatively minor.

The BR model accurately delineated wet stream length along the main channel stems in each Coweeta catchment but overestimated the length of several small tributaries in BR33 and BR40 (Figures 8 and 9), which often corresponded to bedrock springs. As a result, probability thresholds of 0.75–0.90 yielded the most realistic stream networks. As was the case for the other study areas, these threshold values were higher than the threshold optimization procedure suggested (Table III). The BR field maps showed almost negligible network expansion and contraction between high and low flows, but the model was not able to reproduce the more notable changes in stream length that did occur. In BR12, the model did not show contraction of the wet stream along the eastern tributary that we observed in the field and, instead, erroneously indicated drying further downstream.

Discussion

Terrain metrics

TWI is the most critical topographic metric for modeling wet stream length in our Appalachian catchments (Table II). The TWI calculation incorporates the upslope area and slope at a pixel, which are generally among the best predictors of channel head locations (James and Hunt, 2010; Julian *et al.*, 2012; Avcioglu *et al.*, 2017) and most significant variables in stream network models (Sun *et al.*, 2011; Elmore *et al.*, 2013; Russell *et al.*, 2015; González-Ferreras and Barquín, 2017). Substituting the upslope accumulated area (UAA) for TWI lowers model accuracy only slightly. We calculate upslope area for UAA and TWI according to the multiple triangular flow direction algorithm (Seibert and McGlynn, 2007) and use 3 m DEMs, so variable importance and model performance may differ for other flow direction algorithms or DEM resolutions.

The effect of TWI on the probability of stream presence differs most notably between NE and the other three study areas (Table II). For a given TWI value, the log odds of a wet stream in NE are 50–100% greater than for AP, VR, and BR. In other words, a stream is more likely to be wet in NE for a given TWI value. Likewise, a higher TWI owing to a larger area or lower slope would be necessary in the other provinces to have the same probability of stream presence as NE. TWI is the only terrain metric in the NE model yet is able to predict the complex arrangement of tributaries at the Hubbard Brook catchments surprisingly well. A lack of stream incision into the glacial till results in planar hillslopes and few defined valleys (Figure 1), which complicate channel delineation both from DEMs and aerial photographs as well as in the field. The

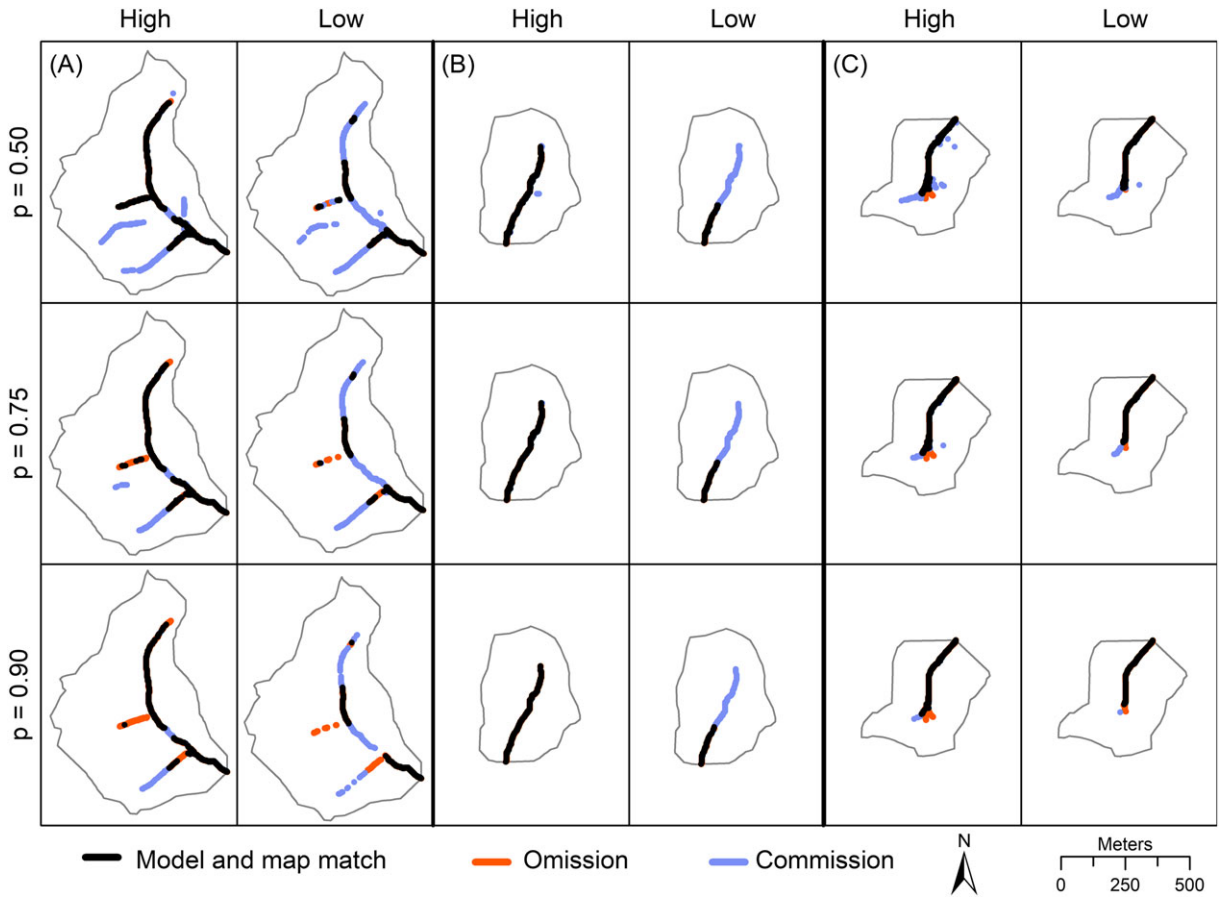

Figure 5. Omission and commission errors at AP37 (A), AP16 (B), and AP14 (C) at the highest and lowest mapped flows for probability thresholds of 0.50, 0.75, and 0.90. See Figure 4 for high and low runoff values. [Colour figure can be viewed at wileyonlinelibrary.com]

success of the model suggests wet stream length at NE is mostly a function of upslope area and, to a lesser extent, slope in the unincised catchments.

TPI considerably improves model performance for AP, VR, and BR. As a measure of the relative elevation of a pixel, TPI distinguishes ridges and valleys and, accordingly, the degree of stream incision. More deeply incised valleys have a lower depth to bedrock, increasing the likelihood that channels lie below the seasonally shifting saturated zone to provide stream flow (Whiting and Godsey, 2016). Likewise, we observed in the field that highly confined stream valleys tend to have a higher flow duration than wide, aggraded valleys, which studies have also noted elsewhere (Stanley *et al.*, 1997). The lack of defined stream valleys at Hubbard Brook (Figure 1) may explain why neither the local nor mean upslope TPI appears in the final NE model.

We are not certain why the AP and VR models incorporate the mean upslope TPI parameter while the BR model includes the local TPI metric. Elmore *et al.* (2013) determined that slope and curvature variables averaged over the upslope contributing area were always more significant than local values in models of the geomorphic channel network for the Eastern USA. Flow origins in the AP and VR provinces tend to occur at upslope areas that are approximately an order of magnitude larger than those in NE and BR (Jensen *et al.*, 2017). However, substitution of the local and mean TPI variables in each model reduces performance only marginally. Thus, we do not believe the specification of a mean or local value is integral to achieve reasonable representations of headwaters. We are not aware of any studies that examine TPI for stream network modeling, but Gillin *et al.*, 2015b) did select TPI for a logistic regression model of hydropedologic units in NE42. Our results suggest this metric may be useful for future headwater modeling studies.

Curvature metrics improve model performance only slightly for AP, VR, and BR after the addition of TWI and TPI (Table II). We tested a global model for all of the catchments using a factor variable to differentiate the physiographic provinces; TWI and TPI are the only two terrain metrics in the final model, reinforcing the comparatively minor role of curvature. In addition, the different types of curvature are highly correlated with each other and all generally indicate the degree of land surface convergence or divergence, so the selection of cross-sectional or profile versus longitudinal curvature in our models is not extremely meaningful. Even though curvature does not greatly increase model accuracy, Russell *et al.* (2015) found that curvature variables can help differentiate perennial and intermittent streams. Longitudinal and cross-sectional curvature (often known as 'profile' and 'planform' curvature, respectively, in ArcGIS) commonly correlate with the location of channel heads (Tarolli and Dalla Fontana, 2009; Julian *et al.*, 2012) and flow origins (Whiting and Godsey, 2016) and are often of secondary importance in stream network models following upslope area and slope (Sun *et al.*, 2011; Elmore *et al.*, 2013; Russell *et al.*, 2015). However, TWI and TPI seem to largely capture the topographic information that curvature metrics provide while producing higher model accuracy.

Characterizing network dynamics

Stream runoff is significant in all four models but provides only small improvements in model accuracy. As a constant value for each mapping rather than a metric that differs for every pixel,

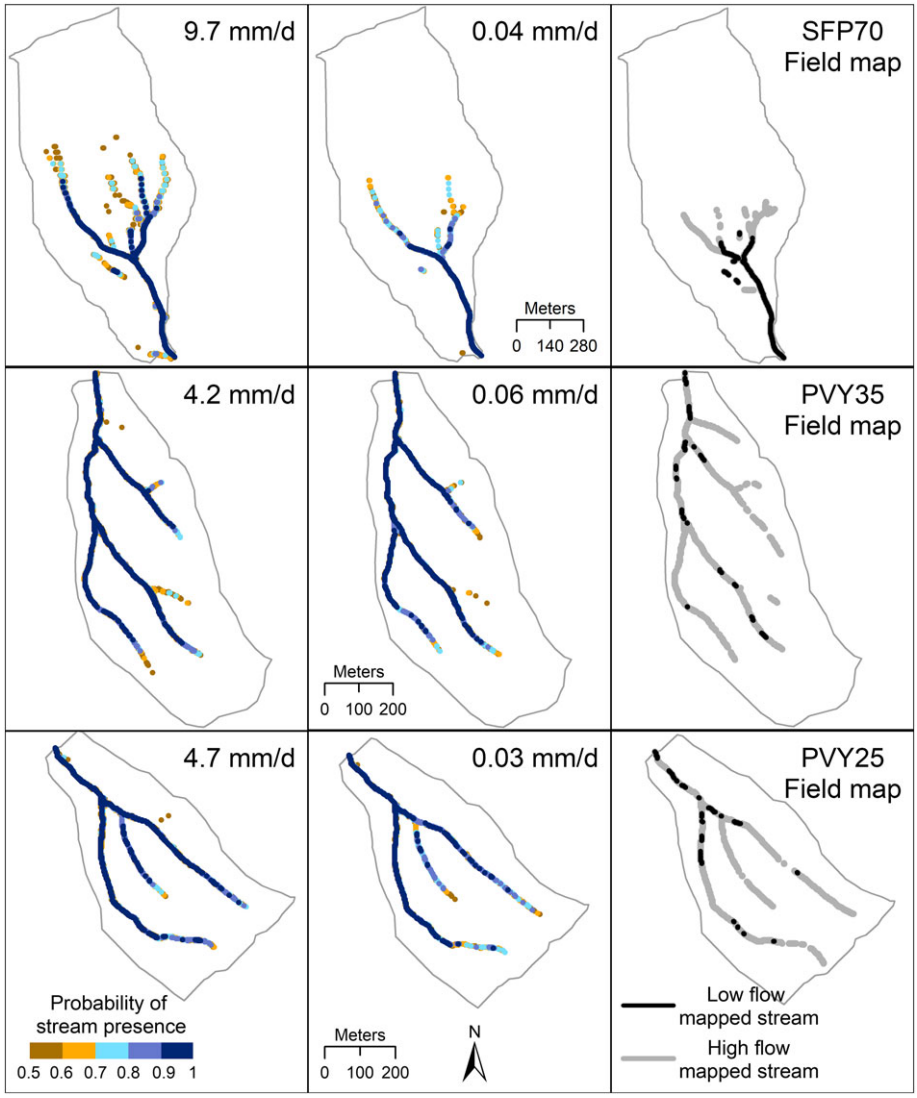

Figure 6. Modeled stream networks for high (left) and low (center) flows in VR. Right columns shows field-surveyed streams for the modeled flows. Only probability values > 0.50 are shown. [Colour figure can be viewed at wileyonlinelibrary.com]

the purpose of runoff in the models is to represent different wetness states of the catchment by increasing the probability of stream presence for higher flow conditions. For this reason, runoff has a greater model coefficient and t-statistic in the provinces where stream length changes more with flow (Table II; Figure 10). Jensen *et al.* (2017) explain stream length dynamics in the study catchments in terms of the permeability and water storage capacity of the underlying sediment and bedrock. Network expansion and contraction is minimal at BR (Figure 10), which is probably due to deep, permeable soils that can store and transport perennial flow to streams (Hewlett and Hibbert, 1967; Hatcher Jr., 1988). Conversely, VR25 and VR35 at Poverty Creek display extreme wet network variability, as the underlying shale has lower permeability and, thus, produces less base flow (Carlston, 1963) than sandstones and other more permeable geology. At a finer spatial scale, wet stream length and flow duration tend to be lower amid transmissive boulder deposits and in wide, sediment-filled valleys, highlighting the additional need for valley incision to the saturated zone for greater flow permanence (Whiting and Godsey, 2016).

The minor contributions of runoff to the models suggest this variable may not be useful for predicting network dynamics. Because stream length is quite stable at Coweeta, we can easily remove runoff from the BR model. For the other three provinces, we created a separate model of the wet stream network at high and low flows, excluding runoff as a potential explanatory variable, to determine if distinct topographic metrics correlate with stream length during wet and dry conditions. The final parameters for the high and low flow models (Table IV) are nearly the same as those in the original models (Table III). For NE, UAA is the single best metric at low flow instead of TWI (Table IV), although TWI and UAA are highly correlated and result in similar McFadden's pseudo R^2 values. The mean upslope downslope index (Hjerdt *et al.*, 2004) replaces cross-sectional curvature for the low flow AP model, but each of these variables increases McFadden's pseudo R^2 values by just 0.03 after TWI and TPI. Longitudinal curvature is absent from the low flow VR model but provides only a small contribution to the high flow and original models. Overall, TWI and TPI remain the most important parameters at both flows. The primary difference between the high and low flow models is the model intercepts and coefficients, which serve to modify the probability of wet stream presence. We can produce a similar result with fewer models by simply varying the probability threshold value of the original model outputs. For this reason, we prefer the original modeling approach that retains runoff for the catchments in our study.

The optimum probability threshold values we calculate for each catchment are greater for low flows than high flows, since increasing the threshold excludes more pixels from

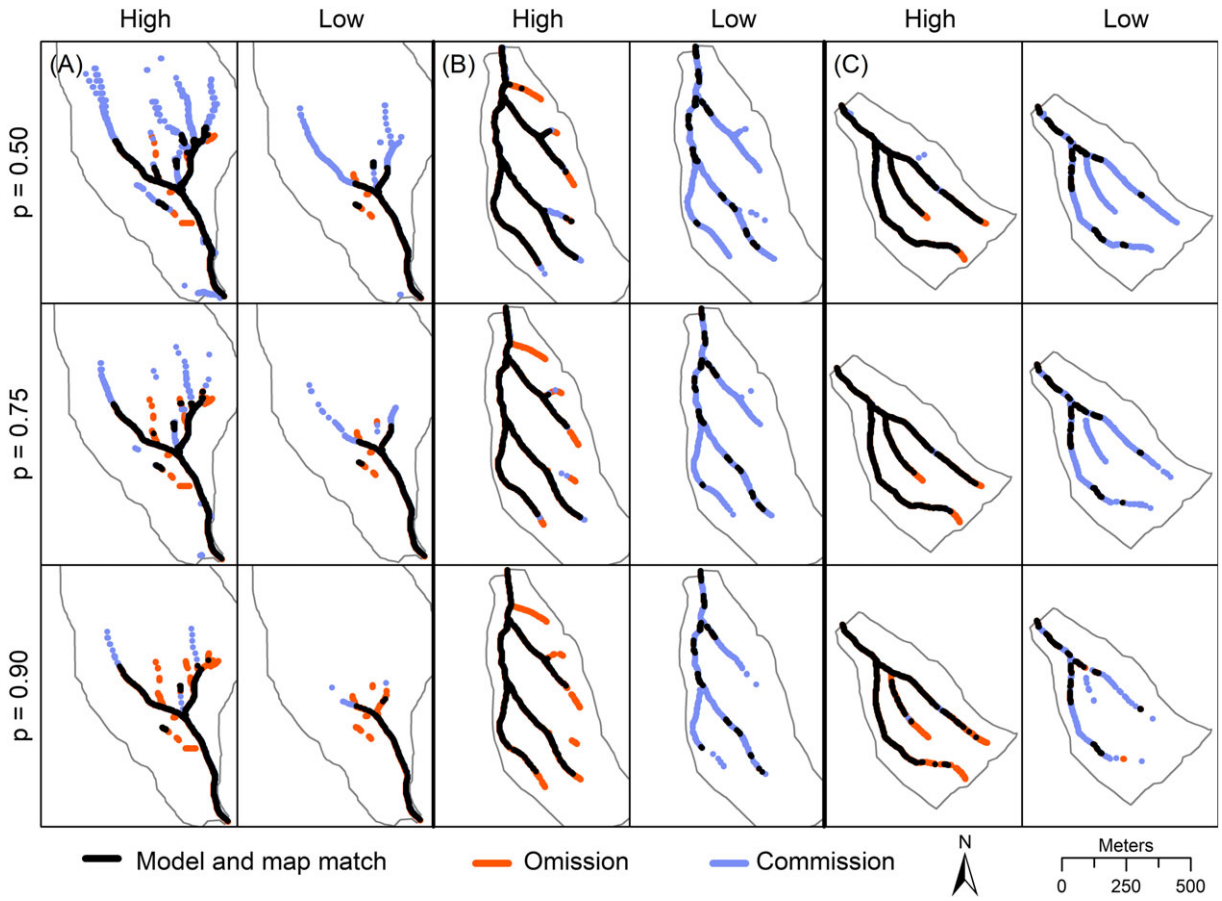

Figure 7. Omission and commission errors at VR70 (A), VR35 (B), and VR25 (C) at the highest and lowest mapped flows for probability thresholds of 0.50, 0.75, and 0.90. See Figure 6 for high and low runoff values. [Colour figure can be viewed at wileyonlinelibrary.com]

classification as a wet stream. Catchments with highly dynamic networks, such as VR35, require a substantial change in threshold values between high and low flows (Table III), indicating the lack of model sensitivity to the runoff variable. Nearly identical probability thresholds are suitable for both flow conditions at sites with more constant stream length, like AP14. We tried incorporating interaction effects between runoff and the topographic metrics as well as the individual slopes of the log–log power relationship between stream length and runoff (Figure 10) as model parameters to help account for the inconsistency in threshold values and eliminate the need for threshold optimization. None of these attempts corrected the issue. In addition, further adjustment of the threshold is still necessary for most of the catchments to create a modeled network that is consistent with field observations. The optimized threshold values are almost always too low. We attempted to remedy this issue by doubling the number of non-stream pixels for model-building and validation to 20%. However, increasing the proportion of non-stream pixels heightens the imbalance with the number of wet stream sample points and results in optimized threshold values that are even further from the visually-selected thresholds than when we only use 10%.

We initially intended to use stream runoff as a continuous variable to delineate the wet network for a specific flow condition. While the models do not successfully predict wet stream length for a precise runoff value, characterizing the wet and dry extremes of network extent is possible by choosing the appropriate lower (0.50–0.75) or higher (0.75–1.0) threshold value, respectively, even if runoff data are not available. Any prior knowledge of the magnitude of stream length dynamics, which we can estimate from geologic characteristics or measurements like the base flow index (Jensen *et al.*, 2017), can further inform the appropriate threshold for more realistic model outputs.

Model strengths and shortcomings for stream network delineation

The models are able to locate several disconnected wet reaches that often either escape detection or appear as continuous tributaries on maps. Coarse surficial deposits cover the southeastern portion of NE42 and coincide with short, disconnected tributaries, which is distinct from the network pattern in the remainder of the catchment. Terrain metrics do not necessarily reflect the surficial geology, so overestimation of wet stream length is common in this area. However, the NE model accurately predicts a lower probability of wet stream pixels in the vicinity of the boulder deposits of NE42 (Figures 2 and 3). The BR model also indicates a disconnected reach in the northern part of BR40 (Figures 8 and 9). This wet reach occurs in a valley with sediment fill from an old landslide deposit. The modeled location and length of the high probability pixels are quite similar to field observations of the disconnected tributary. The VR model does not identify all of the disconnected reaches in VR70 but, surprisingly, is able to find two isolated wetlands without a tributary inlet or outlet that persist during both high and low flows (Figures 6 and 7): one wetland lies west of the main stem, and the other is just above the junction of the two primary tributaries.

Whereas some wet stream reaches and wetlands never have a surface water connection to the catchment outlet across the flow conditions observed in the field, discontinuous wet reaches also

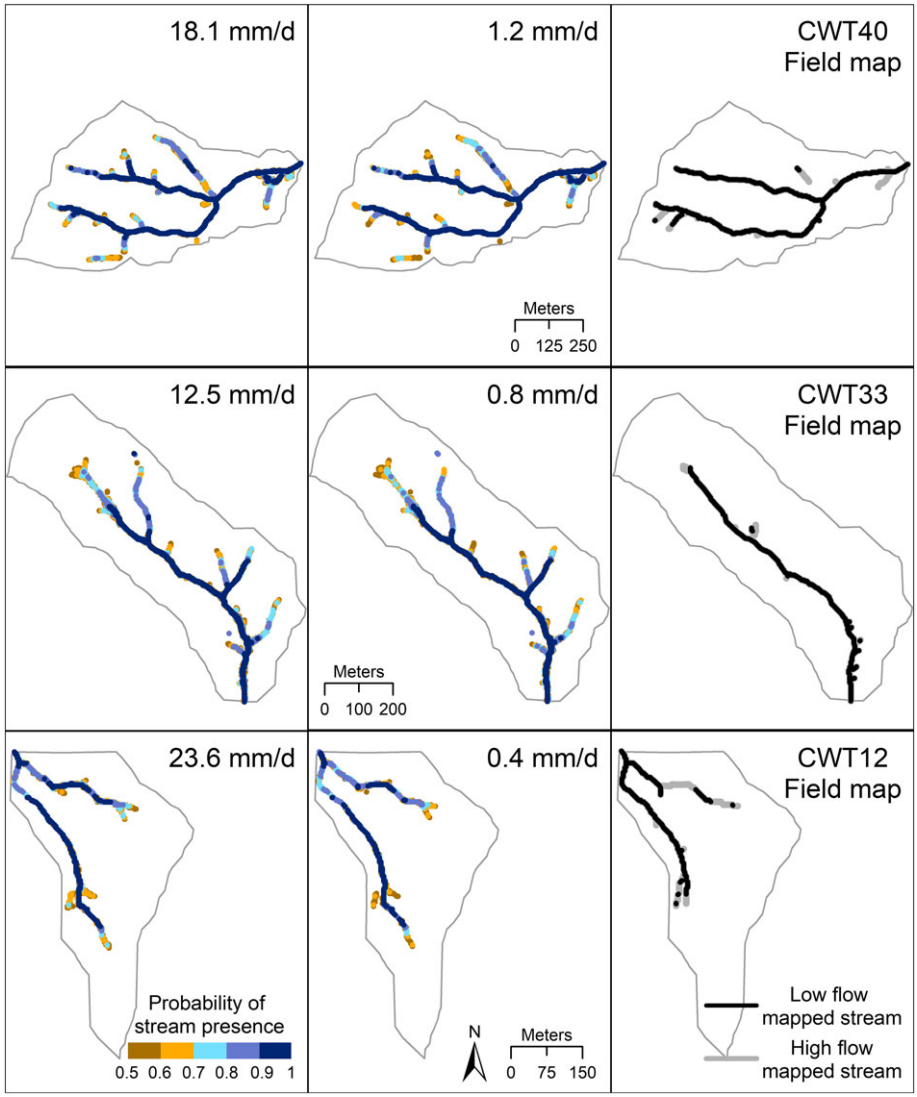

Figure 8. Modeled stream networks for high (left) and low (center) flows in BR. Right columns shows field-surveyed streams for the modeled flows. Only probability values > 0.50 are shown. [Colour figure can be viewed at wileyonlinelibrary.com]

develop during channel drying. The modeled network contraction patterns evident by increasing the probability threshold values show the formation of discontinuous reaches in addition to more sequential top-down drying (Figures 2–9). Disconnection of streams into separate wet reaches is common during dry periods (Godsey and Kirchner, 2014), but most maps and stream network models cannot recreate the fine-scale spatial variability in flow duration that actually exists in headwaters. Researchers increasingly recognize the significance of isolated wetlands (Cohen *et al.*, 2016) and temporary or discontinuous stream reaches (Stanley *et al.*, 1997; Larned *et al.*, 2010; Datry *et al.*, 2014) for ecological services including biodiversity, nutrient cycling, and downstream water quality. Disconnected water bodies frequently maintain a subsurface connection to perennial streams through hyporheic exchange (Boulton *et al.*, 1998) and also serve as habitat refugia and storage sites for sediment, organic matter, and contaminants moving through the catchment (Larned *et al.*, 2010). As a result, locating isolated wetlands and disconnected wet reaches is a high priority yet also frequently a challenge for scientists and policy-makers.

The models fail to adequately represent other aspects of the stream network that correspond to discontinuities in the surficial geology or watershed evolution processes. The main eastern tributary in BR12 has a major disconnection at a valley fill from a landslide deposit (Figures 8 and 9). A wide, sediment-filled valley floor also aligns with the disconnection along the main stem of AP37 (Figures 4 and 5). In both cases, the models overestimate stream length in wider valley sections where flow duration is lower. In VR70, channels with perennial flow originate at the base of boulder-filled hollows. Upslope of the origins, water emerges on top of the boulders for only short distances at high flows before again infiltrating into the subsurface. The VR model does not identify most of these short reaches and, additionally, extends the streams too far up the catchment, as the model cannot account for the transmissive boulder deposits with the available terrain metrics (Figures 6 and 7). Finally, the AP model dramatically overestimates stream length and flow duration at higher elevations, especially for the southwest tributary of AP37 (Figures 4 and 5). Field observations suggest that the tributaries in this catchment are headwardly eroding into a relict upland in the same manner as a hanging valley. Thus, there will probably be streams at the predicted locations higher in the catchment in the future, but, meanwhile, the model does not accurately portray current conditions. In all cases, the overprediction of stream length occurs in areas where the channel has not incised into a valley fill, coarse boulder deposits, or a relict upland. Despite the lack of incision, these locations all coincide with well-defined valleys that have strongly negative TPI values to increase the modeled probability of wet stream presence. In contrast, locations that accurately show wet stream disconnections in NE42, BR40, and portions of VR70 are not in highly confined

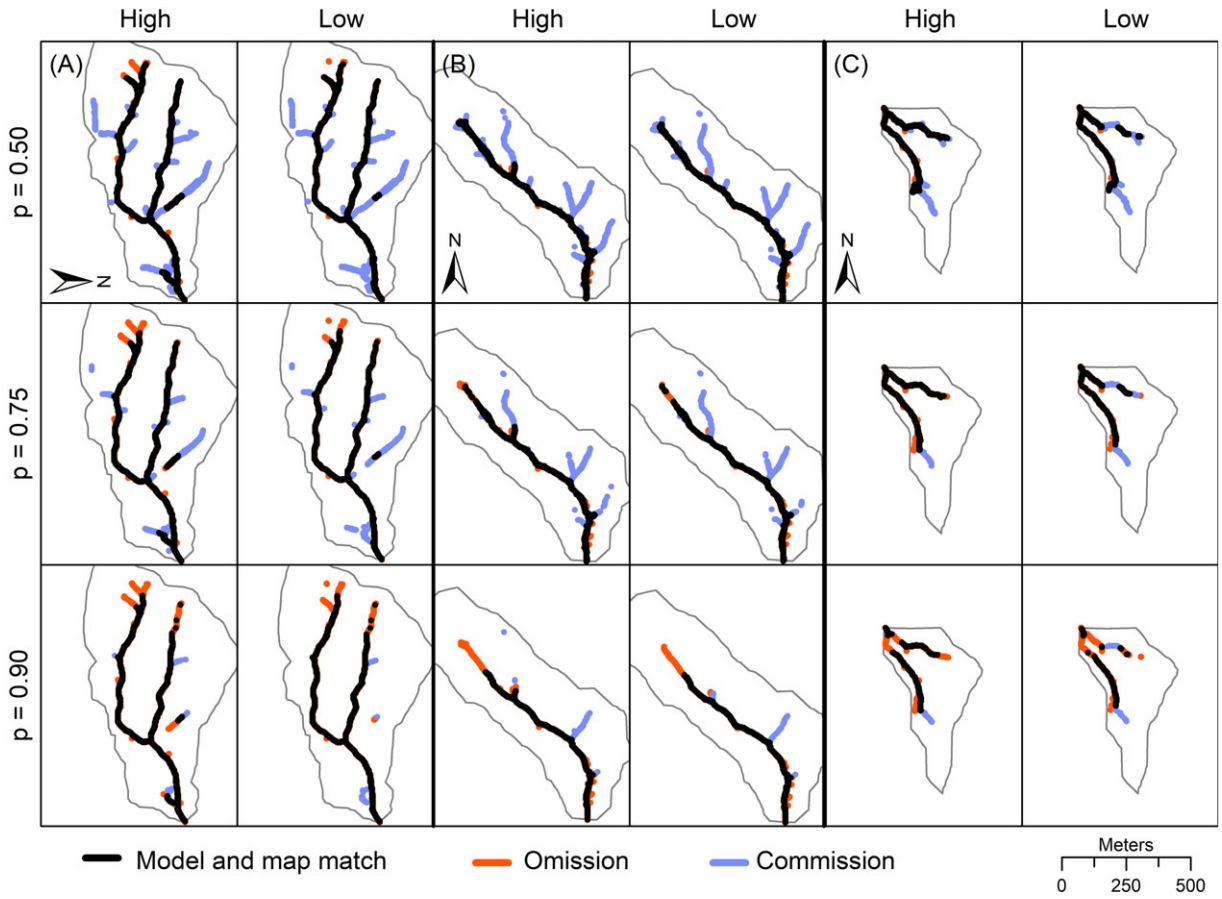

Figure 9. Omission and commission errors at BR40 (A), BR33 (B), and BR12 (C) at the highest and lowest mapped flows for probability thresholds of 0.50, 0.75, and 0.90. See Figure 8 for high and low runoff values. Note north orientation. [Colour figure can be viewed at wileyonlinelibrary.com]

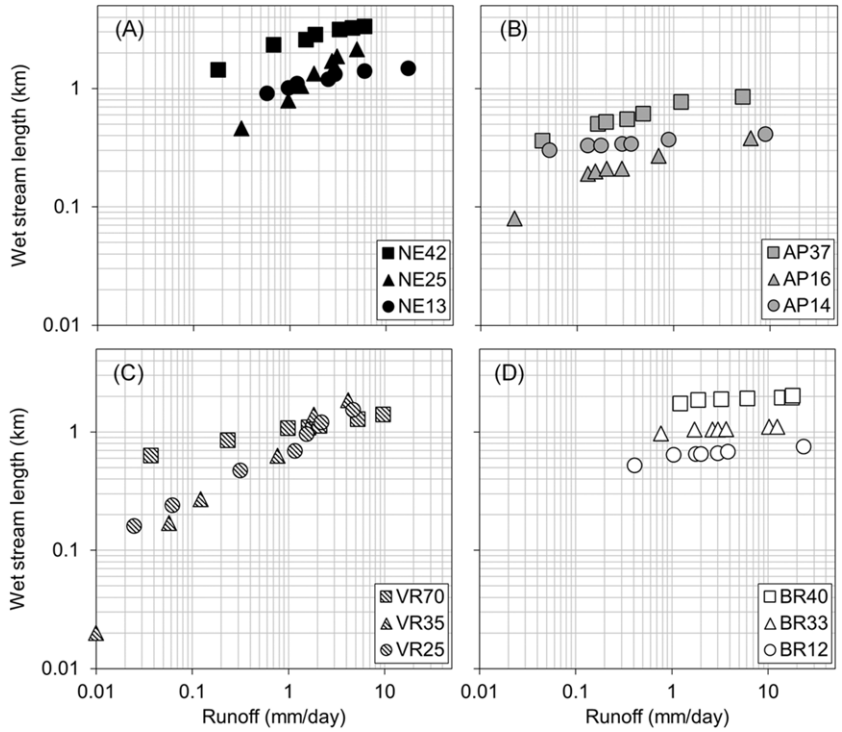

Figure 10. Stream length versus runoff in NE (A), AP (B), VR (C), and BR (D). Modified from Jensen *et al.* (2017).

Table IV. High and low flow model parameters for NE, AP, and VR. The *P*-value of all model parameters is << 0.001

Province	Model	McFadden's pseudo R^2	Parameters	Coefficient	t-statistic
NE	High	0.68	Intercept	–34.97	–39.08
			TWI (ln)	15.74	38.89
	Low	0.71	Intercept	–35.57	–30.32
			UAA (ln)	1.71	31.25
AP	High	0.81	Intercept	–19.68	–17.41
			TWI (ln)	7.97	14.90
			TPI (mean[a])	–4.80	–13.47
			Cross-sectional Curv.	–0.37	–9.67
	Low	0.79	Intercept	–33.12	–15.69
			TWI (ln)	12.11	16.77
			TPI (mean[a])	–4.05	–7.72
			Downslope index (mean[a])	16.90	7.04
VR	High	0.79	Intercept	–24.79	–33.62
			TWI (ln)	10.97	31.99
			TPI (mean[a])	–2.71	–15.65
			Longitudinal Curv.	0.24	18.46
	Low	0.71	Intercept	–22.39	–24.89
			TWI (ln)	9.12	24.00
			TPI (mean[a])	–3.47	–12.88

[a]Variable denoted as'mean' refers to the mean value of upslope pixels as opposed to the local value.

valleys or have clear topographic breaks to delimit the wet reach. This finding again emphasizes the importance of valley incision to the zone of saturation for flow permanence but also indicates a need for better metrics of relative incision and aggradation to improve future models.

Model application

Our study constitutes an exploratory analysis of the terrain variables that correlate with the location of the wet stream network in four Appalachian physiographic provinces. The overall agreement between the modeled probability of wet stream pixels and field maps indicates logistic regression is an effective way to characterize headwaters across flow conditions. Similar studies successfully apply logistic regression to delineate headwater channels (Russell *et al.*, 2015), and our results demonstrate that this fairly simple approach is also suitable for modeling network expansion, contraction, and disconnection at a finer reach scale by varying the probability threshold values of model outputs. Traditionally, physically mapping wet stream length multiple times is the only method for measuring wet network variability. A key improvement in our models is the prediction of disconnected reaches along streams for both wet and dry conditions, which provides a more realistic map of flow permanence than assigning a single flow duration classification to an entire tributary (Russell *et al.*, 2015; Williamson *et al.*, 2015).

Our models rely on terrain metrics, which are easy to derive from DEMs with programs like ArcGIS and SAGA. We can estimate the high and low extremes of wet stream length by simply altering the probability threshold of the logistic regression model output. We also include stream runoff as a wetness state variable in the models because of visual and statistical improvements in the predicted probability of stream pixels, particularly for NE, AP, and VR. More training data of stream networks at different runoff conditions would undoubtedly increase the contribution of this variable to the models. Other wetness state variables such as an antecedent precipitation index would be an appropriate substitute if stream discharge measurements are not possible. However, neither runoff nor precipitation is essential to approximate the magnitude of stream length dynamics.

The twelve catchments in this study represent a limited sample of headwaters, despite the extensive time and effort associated with field data collection. The climate, geology, vegetation, and land use are consistent across the selected catchments in each province. Thus, our models may not be applicable to all headwaters in a given physiographic region. Incorporation of additional variables such as annual precipitation, evapotranspiration, lithology, and land cover would likely be necessary for a comprehensive stream extraction over larger areas.

Conclusion

Logistic regression is a suitable and straightforward method for modeling wet stream length dynamics in Appalachian headwaters. The most significant terrain metrics correlating with the probability of stream presence at a pixel include the topographic wetness index (TWI) followed by the topographic position index (TPI). In effect, catchment locations with larger upslope areas, lower slopes, and more valley incision have a higher flow duration. TPI is not a common parameter for stream network modeling, but our results show that this variable can greatly improve channel delineations. The contributions of curvature and runoff to the models are comparatively minor for the selected study areas.

The four models have high overall accuracy and predict high and low probability values of wet stream pixels that match the network patterns we observed in the field. In general, greater threshold values (> 0.75) correspond to dry conditions, and lower values (0.50–0.75) represent high flows. Catchments with highly variable lengths require more adjustment of the probability threshold. The models are able to distinguish several disconnected stream reaches and wetlands but do not always portray the correct stream location and length in areas with less channel incision amid boulder deposits, wide valley segments, or along headwardly eroding tributaries. Future modeling efforts should further investigate these sources of error to improve the reliability of network predictions.

Accurate maps of headwaters are rare yet increasingly essential for ecological modeling and the enforcement of water policy for effective watershed conservation. Because many headwater streams have temporary flow, data describing not only the length but the associated flow duration of tributaries, disconnected reaches, and wetlands are critical for informing the research and management of river systems. Our logistic regression models approximate the wet stream network of catchments at high and low flows rather than provide a single prediction of static channel length. The modeling methodology is relatively simple, employing basic terrain metrics, which permits the application of this approach in other regions to improve the characterization of temporary headwaters.

Acknowledgements—This study was supported by the Cunningham Graduate Fellowship and Graduate Student Association Graduate Research Development Fund Award at Virginia Tech, the Consortium of Universities for the Advancement of Hydrologic Science Pathfinder Fellowship, and the National Science Foundation under grant LTER DEB 1114804. Coweeta Hydrologic Laboratory and Fernow and Hubbard Brook Experimental Forests are operated and maintained by the US Forest Service. We would like to thank the George Washington and Jefferson National Forests for granting access to field sites and participating in the project. We are also grateful to Jake Diamond, Dave Jensen, Joe Famularo, and Morgan Schulte for assistance with field work and two anonymous reviewers for helpful comments.

References

Adams HS, Stephenson SL. 1983. A description of the vegetation on the south slopesof Peters Mountain. *southwestern Virginia. Bulletin of the Torrey Botanical Club* **110**: 18–22. https://doi.org/10.2307/2996512.

Adams RK, Spotila JA. 2005. The form and function of headwater streams based on field and modeling investigations in the Southern Appalachian Mountains. *Earth Surface Processes and Landforms* **30**: 1521–1546. https://doi.org/10.1002/esp. 1211.

Alexander RB, Boyer EW, Smith RA, Schwarz GE, Moore RB. 2007. The role of headwater streams in downstream water quality. *Journal of the American Water Resources Association* **43**: 41–59. https://doi.org/10.1111/j.1752-1688.2007.00005.x.

Avcioglu B, Anderson CJ, Kalin L. 2017. Evaluating the slope-area method to accurately identify stream channel heads in three physiographic regions. *Journal of the American Water Resources Association* **53**: 562–575. https://doi.org/10.1111/1752-1688.12512.

Barton CC. 1997. Bedrock geologic map of Hubbard Brook Experimental Forest and maps of fractures and geology in roadcuts along Interstate 93, Grafton County, New Hampshire, Scale 1:2000. US Geological Survey Miscellaneous Investigation Series I–2562.

Benettin P, Bailey SW, Campbell JL, Green MB, Rinaldo A, Likens GE, Botter G. 2015. Linking water age and solute dynamics in streamflow at the Hubbard Brook Experimental Forest, NH. *USA. Water Resources Research* **51**: 9256–9272. https://doi.org/10.1002/2015WR017552.

Beven KJ, Kirkby MJ. 1979. A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Journal* **24**: 43–69. https://doi.org/10.1080/02626667909491834.

Boulton AJ, Findlay S, Marmonier P, Stanley EH, Valett HM. 1998. The functional significance of the hyporheic zone in streams and rivers. *Annual Review of Ecology and Systematics* **29**: 59–81. https://doi.org/10.1146/annurev.ecolsys.29.1.59.

Buttle JM, Boon S, Peters DL, Spence C, van Meerveld HJ, Whitfield PH. 2012. An overview of temporary stream hydrology in Canada. *Canadian Water Resources Journal* **37**: 279–310. https://doi.org/10.4296/cwrj2011-903.

Calkins D, Dunne T. 1970. A salt tracing method for measuring channel velocities in small mountain streams. *Journal of Hydrology* **11**: 379–392. https://doi.org/10.1016/0022-1694(70)90003-X.

Cardwell DH, Erwin RB, Woodward HP. 1968 Geologic map of West Virginia, 2 sheets, Scale 1:250 000. West Virginia Geological and Economic Survey.

Carlston CW. 1963. Drainage density and streamflow. US Geological Survey Professional Paper 422–C.

Cohen MJ, Creed IF, Alexander L, Basu NB, Calhoun AJ, Craft C, Jawitz JW. 2016. Do geographically isolated wetlands influence landscape functions? *Proceedings of the National Academy of Sciences* **113**: 1978–1986. https://doi.org/10.1073/pnas. 1512650113.

Colson T, Gregory J, Dorney J, Russell P. 2008. Topographic and soil maps do not accurately depict headwater stream networks. *National Wetlands Newsletter* **30**: 25–28.

Datry T, Larned ST, Tockner K. 2014. Intermittent rivers: a challenge for freshwater ecology. *BioScience* **64**: 229–235. https://doi.org/10.1093/biosci/bit027.

Dodds WK, Oakes RM. 2008. Headwater influences on downstream water quality. *Environmental Management* **41**: 367–377. https://doi.org/10.1007/s00267-007-9033-y.

Downing JA, Cole JJ, Duarte CM, Middelburg JJ, Melack JM, Prairie YT, Tranvik LJ. 2012. Global abundance and size distribution of streams and rivers. *Inland Waters* **2**: 229–236. https://doi.org/10.5268/IW-2.4.502.

Elmore AJ, Julian JP, Guinn SM, Fitzpatrick MC. 2013. Potential stream density in Mid- Atlantic US watersheds. *PlOS ONE* **8** e74819. DOI: https://doi.org/10.1371/journal.pone.0074819.

Evans S. 1979. An integrated system of terrain analysis and slope mapping. Final report on grant DA-ERO-591-73-G0040. University of Durham, England.

Feminella JW. 1996. Comparison of benthic macroinvertebrate assemblages in small streams along a gradient of flow permanence. *Journal of the North American Benthological Society* **15**: 651–669. https://doi.org/10.2307/1467814.

Fritz KM, Hagenbuch E, D'Amico E, Reif M, Wigington PJ, Leibowitz SG, Comeleo RL, Ebersole JL, Nadeau TL. 2013. Comparing the extent and permanence of headwater streams from two field surveys to values from hydrographic databases and maps. *Journal of the American Water Resources Association* **49**: 867–882. https://doi.org/10.1111/jawr. 12040.

Fritz KM, Johnson BR, Walters DM. 2008. Physical indicators of hydrologic permanence in forested headwater streams. *Journal of the North American Benthological Society* **27**: 690–704. https://doi.org/10.1899/07-117.1.

Gillin CP, Bailey SW, McGuire KJ, Prisley SP. 2015a. Evaluation of Lidar-derived DEMs through terrain analysis and field comparison. *Photogrammetric Engineering and Remote Sensing* **81**: 387–396. https://doi.org/10.14358/PERS.81.5.387.

Gillin CP, Bailey SW, McGuire KJ, Gannon JP. 2015b. Mapping of hydropedologic spatial patterns in a steep headwater catchment. *Soil Science Society of America Journal* **79**: 440–453. https://doi.org/10.2136/sssaj2014.05.0189.

Godsey SE, Kirchner JW. 2014. Dynamic, discontinuous stream networks: hydrologically driven variations in active drainage density. *flowing channels and stream order. Hydrological Processes* **28**: 5791–5803. https://doi.org/10.1002/hyp. 10310.

González-Ferreras AM, Barquín J. 2017. Mapping the temporary and perennial character of whole river networks. *Water Resources Research* **53**: 6709–6724. https://doi.org/10.1002/2017WR020390.

Grabs TJ, Jencso KG, McGlynn BL, Seibert J. 2010. Calculating terrain indices along streams: a new method for separating stream sides. *Water Resources Research* **46**. https://doi.org/10.1029/2010WR009296.

Guisan A, Weiss SB, Weiss AD. 1999. GLM versus CCA spatial modeling of plant species distribution. *Plant Ecology* **143**: 107–122. https://doi.org/10.1023/A:1009841519580.

Hansen WF. 2001. Identifying stream types and management implications. *Forest Ecology and Management* **143**: 39–46. https://doi.org/10.1016/S0378-1127(00)00503-X.

Hatcher RD, Jr. 1988. Bedrock geology and regional geologic setting of Coweeta Hydrologic Laboratory in the eastern Blue Ridge. In *Ecological Studies: Forest Hydrology and Ecology at Coweeta*, Swank WT, Crossley DA, Jr (eds). Springer-Verlag: New York; 81–92.

Hedman ER, Osterkamp WR. 1982. Streamflow characteristics related to channel geometry of streams in western United States. US Geological Survey Water Supply Paper no. 2193.

Heine RA, Lant CL, Sengupta RR. 2004. Development and comparison of approaches for automated mapping of stream channel networks. *Annals of the Association of American Geographers* **94**: 477–490. https://doi.org/10.1111/j.1467-8306.2004.00409.x.

Hewlett JD. 1982. *Principles of Forest Hydrology*. University of Georgia Press.

Hewlett JD, Hibbert AR. 1967. Factors affecting the response of small watersheds to precipitation in humid areas. *Forest Hydrology* **1**: 275–290.

Hjerdt KN, McDonnell JJ, Seibert J, Rodhe A. 2004. A new topographic index to quantify downslope controls on local drainage. *Water Resources Research* **40**. https://doi.org/10.1029/2004WR003130.

Hooshyar M, Kim S, Wang D, Medeiros SC. 2015. Wet channel network extraction by integrating LiDAR intensity and elevation data. *Water Resources Research* **51**: 10029–10046. https://doi.org/10.1002/2015WR018021.

Hughes RM, Kaufmann PR, Weber MH. 2011. National and regional comparisons between Strahler order and stream size. *Journal of the North American Benthological Society* **30**: 103–121. https://doi.org/10.1899/09-174.1.

James LA, Hunt KJ. 2010. The LiDAR-side of headwater streams: mapping channel networks with high-resolution topographic data. *Southeastern Geographer* **50**: 523–539. https://doi.org/10.1353/sgo.2010.0009.

Jenness J. 2013. DEM Surface Tools for ArcGIS (surface_area.exe). Jenness Enterprises. Available at http://www.jennessent.com/arcgis/surface_area.htm.

Jensen CK, McGuire KJ, Prince PS. 2017. Headwater stream length dynamics across four physiographic provinces of the Appalachian Highlands. *Hydrological Processes* **31**: 3350–3363. https://doi.org/10.1002/hyp.%2011259

Johnson BR, Fritz KM, Blocksom KA, Walters DM. 2009. Larval salamanders and channel geomorphology are indicators of hydrologic permanence in forested headwater streams. *Ecological Indicators* **9**: 150–159. https://doi.org/10.1016/j.ecolind.2008.03.001.

Julian JP, Elmore AJ, Guinn SM. 2012. Channel head locations in forested watersheds across the mid-Atlantic United States: a physiographic analysis. *Geomorphology*. **177**: 194–203. https://doi.org/10.1016/j.geomorph.%202012.07.029

Kuhn M. 2008. Caret package. *Journal of Statistical Software* **28**: 1–26.

Larned ST, Datry T, Arscott DB, Tockner K. 2010. Emerging concepts in temporary-river ecology. *Freshwater Biology* **55**: 717–738. https://doi.org/10.1111/j.1365-2427.2009.02322.x.

Larned ST, Datry T, Robinson CT. 2007. Invertebrate and microbial responses to. inundation in an ephemeral river reach in New Zealand: effects of preceding dry periods. *Aquatic Sciences* **69**: 554–567. https://doi.org/10.1007/s00027-007-0930-1.

Larned ST, Schmidt J, Datry T, Konrad CP, Dumas JK, Diettrich JC. 2011. Longitudinal river ecohydrology: flow variation down the lengths of alluvial rivers. *Ecohydrology* **4**: 532–548. https://doi.org/10.1002/eco.%20126

Leibowitz SG, Wigington PJ, Jr, Rains MC, Downing DM. 2008. Non-navigable streams and adjacent wetlands: addressing science needs following the Supreme Court's Rapanos decision. *Frontiers in Ecology and the Environment* **6**: 364–371. https://doi.org/10.1890/070068.

Li J, Wong DW. 2010. Effects of DEM sources on hydrologic applications. *Computers, Environment and Urban Systems* **34**: 251–261. https://doi.org/10.1016/j.compenvurbsys.2009.11.002.

Likens GE. 2013. *Biogeochemistry of a Forested Ecosystem*. Springer Science and Business Media: New York.

Losche CK, Beverage WW. 1967. Soil survey of Tucker County and part of Northern Randolph County, West Virginia. Soil Survey Reports. US Department of Agriculture.

MacNally R. 2000. Regression and model-building in conservation biology, biogeography and ecology: the distinction between – and reconciliation of –'predictive' and 'explanatory' models. *Biodiversity and Conservation* **9**: 655–671. https://doi.org/10.1023/A:1008985925162.

Moore ID, Grayson RB, Ladson AR. 1991. Digital terrain modelling: a review of hydrological, geomorphological. *and biological applications*. *Hydrological Processes* **5**: 3–30. https://doi.org/10.1002/hyp.%203360050103

Morisawa ME. 1962. Quantitative geomorphology of some watersheds in the Appalachian Plateau. *Geological Society of America Bulletin* **73**: 1025–1046. https://doi.org/10.1130/0016-7606(1962)73[1025:QGOSWI]2.0.CO;2.

Nadeau TL, Rains MC. 2007. Hydrological connectivity between headwater streams and downstream waters: how science can inform policy. *Journal of the American Water Resources Association* **43**: 118–133. https://doi.org/10.1111/j.1752-1688.2007.00010.x.

Paybins KS. 2003. Flow origin, drainage area, and hydrologic characteristics for headwater streams in the mountaintop coal-mining region of southern West Virginia, 2000-2001. US Department of the Interior, US Geological Survey.

Płaczkowska E, Górnik M, Mocior E, Peek B, Potoniec P, Rzonca B, Siwek J. 2015. Spatial distribution of channel heads in the Polish Flysch Carpathians. *Catena* **127**: 240–249. https://doi.org/10.1016/j.catena.2014.12.033.

R Core Team. 2015. R: A language and environment for statistical computing. In *R Foundation for Statistical Computing, Vienna, Austria*. http://www.r-project.org.

Russell PP, Gale SM, Muñoz B, Dorney JR, Rubino MJ. 2015. A spatially explicit model for mapping headwater streams. *Journal of the American Water Resources Association* **51**: 226–239. https://doi.org/10.1111/jawr.%2012250

Sainani KL. 2014. Explanatory versus predictive modeling. PM and R. **6**: 841–844. https://doi.org/10.1016/j.pmrj.2014.08.941.

Schultz AP, Stanley CB, Gathright II TM, Rader EK, Bartholomew MJ, Lewis SE, Evans NH. 1986. Geologic map of Giles County, Virginia, Scale 1:50 000. Virginia Division of Mineral Resources.

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

Skoulikidis NT, Sabater S, Datry T, Morais MM, Buffagni A, Dörflinger G, Tockner K. 2017. Non-perennial Mediterranean rivers in Europe: status, pressures, and challenges for research and management. *Science of the Total Environment* **577**: 1–18. https://doi.org/10.1016/j.scitotenv.2016.10.147.

Spence C, Mengistu S. 2016. Deployment of an unmanned aerial system to assist in mapping an intermittent stream. *Hydrological Processes* **30**: 493–500. https://doi.org/10.1002/hyp. 10597.

Stanley EH, Fisher SG, Grimm NB. 1997. Ecosystem expansion and contraction in streams. *Bioscience* **47**: 427–435. https://doi.org/10.2307/1313058.

Sun X, Thompson CJ, Croke B. 2011. Using a logistic regression model to delineate channel network in southeast Australia. In *19th International Congress on Modelling and Simulation (MODSIM 2011)*. Modelling and Simulation Society of Australia and New Zealand Inc: Perth, Australia; 1916–1922.

Svec JR, Kolka RK, Stringer JW. 2005. Defining perennial, intermittent, and ephemeral channels in eastern Kentucky: application to forestry best management practices. *Forest Ecology and Management* **214**: 170–182. https://doi.org/10.1016/j.foreco.2005.04.008.

Tarolli P, Dalla Fontana G. 2009. Hillslope-to-valley transition morphology:new opportunities from high resolution DTMs. *Geomorphology* **113**: 47–56. https://doi.org/10.1016/j.geomorph.2009.02.006.

Travis MR, Elsner GH, Iverson WD. 1975. VIEWIT: computation of seen areas, slope and aspect for land-use planning. US Forest Service General Technical Report PSW-11. US Department of Agriculture.

Velbel MA. 1988. Weathering and soil-forming processes. In *Ecological Studies: Forest Hydrology and Ecology at Coweeta*, Swank WT, Crossley DA, Jr (eds). Springer-Verlag: New York; 93–102.

Virginia Division of Mineral Resources. 1993. Geologic Map of Virginia, Scale 1:500 000. Virginia Division of Mineral Resources.

Wang L, Liu H. 2006. An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling. *International Journal of Geographical Information Science* **20**: 193–213. https://doi.org/10.1080/13658810500433453

Whiting JA, Godsey SE. 2016. Discontinuous headwater stream networks with stable flowheads, Salmon River basin. *Idaho. Hydrological Processes* **30**: 2305–2316. https://doi.org/10.1002/hyp.%2010790

Williamson TN, Agouridis CT, Barton CD, Villines JA, Lant JG. 2015. Classification of ephemeral, intermittent, and perennial stream reaches using a TOPMODEL-based approach. *Journal of the*

American Water Resources Association **51**: 1739–1759. https://doi.org/10.1111/1752-1688.12352.

Wilson JP, Gallant JC. 2000. *Terrain Analysis: Principles and Applications*. Wiley: New York.

Winter TC. 2007. The role of ground water in generating streamflow in headwater areas and in maintaining base flow. *Journal of the American Water Resources Association* **43**: 15–25. https://doi.org/10.1111/j.1752-1688.2007.00003.x.

Wohl E. 2017. The significance of small streams. *Frontiers of Earth Science* **11**: 447–456. https://doi.org/10.1007/s11707-017-0647-y.

Wood J. 1996. The Geomorphological Characterisation of Digital Elevation Models. Doctoral Dissertation, University of Leicester.

Zevenbergen LW, Thorne CR. 1987. Quantitative analysis of land surface topography. *Earth Surface Processes and Landforms* **12**: 47–56. https://doi.org/10.1002/esp.%203290120107