ORNL/TM-2018/898

Site Classification for Standard Modular Hydropower Development: Characterizing Stream Reaches by Module Need

Mark Bevelhimer
Chris DeRolph
Adam Witt

July 2018

Approved for public release. Distribution is unlimited.

OAK RIDGE NATIONAL LABORATORY
MANAGED BY UT-BATTELLE FOR THE US DEPARTMENT OF ENERGY

ORNL/TM-2018/898

Environmental Sciences Division

STANDARD MODULAR HYDROPOWER: SITE CLASSIFICATION

Mark Bevelhimer
Chris DeRolph
Adam Witt

Date Published: July 2018

Prepared by
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, TN 37831-6283
managed by
UT-BATTELLE, LLC
for the
US DEPARTMENT OF ENERGY
under contract DE-AC05-00OR22725

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS, ACRONYMS, AND INITIALISMS

cfs	cubic feet per second
DB	database
EPA	Environmental Protection Agency
HUC	Hydrologic Unit Code
MAF	mean annual flow
MSU	Michigan State University
NERC	North American Electric Reliability Corporation
NHAAP	National Hydropower Asset Assessment Program
NHD	National Hydrographic Dataset
NLCD	National Land Cover Database
NPD	non-powered dam
NSD	new stream-reach development
NSDI	National Spatial Data Infrastructure
ORNL	Oak Ridge National Laboratory
RPS	Renewable Portfolio Standard
SMH	standard modular hydropower
USGS	US Geological Service
UTK	University of Tennessee–Knoxville
WRD	Water Resources Division

ACKNOWLEDGMENTS

The authors would like to acknowledge and express their appreciation to all the following individuals and programs for their review, comments, and support of this report and support of Standard Modular Hydropower research efforts.

US Department of Energy

- Water Power Technologies Office (project sponsor)

Oak Ridge National Laboratory

- Charlie Horak, Technical writer/editor
- Deborah Counce, Technical writer/editor
- Kevin Stewart, technical reviewer
- Missy Miller, administrative assistant
- Priscilla Henson, ORNL Publications manager
- Scott DeNeale, project manager

University of Tennessee–Knoxville

- Dr. Thanos Papanicolaou, professor and director of Hydraulics and Sedimentation Laboratory
- Dr. Benjamin Abban, post-doctoral research associate
- Sarah Pierce, MS student in water resource engineering

1. INTRODUCTION

Oak Ridge National Laboratory (ORNL) is leading a research effort to redefine small hydropower development through a new research effort called "Standard Modular Hydropower" (SMH), which seeks to establish the site characteristics, design envelope specifications, and technology characteristics of next-generation small, low-head hydropower plants. The SMH philosophy is that a limited number of standardized passage, generation, and foundation modules can be deployed at a single site and across several sites to assemble a fully functional, environmentally compatible hydroelectric facility. The modules together must reduce the environmental impacts of development compared with conventional approaches, and the facility must deploy at a cost competitive with that of comparable renewable energy resources.

Site classification is a key standardizing concept and component of SMH research, addressing the perception that hydropower development is predominantly site-specific. Site classification consists of analyses and tools that identify similarities in stream-reach, landscape, and biological characteristics across river systems. These similarities are classified into a finite number of clusters such that differences among characteristics within a cluster are limited. In theory, stream reaches within a given cluster share enough common characteristics that, should development be pursued on any given stream reach, design requirements would not vary significantly across sites. A combination of standard generation, passage, and foundation modules could be deployed at a large group of sites within a given cluster with few to no changes in major design features.

Site classification has two main objectives: (1) to identify classes of module needs (e.g., fish passage or sediment transport) and (2) to align need classes with functional design requirements (e.g., achieving injury-free downstream passage for fish). For example, sites or river reaches that have similar stream gradient, hydrology, and migratory fish species would be expected to have similar fish passage needs and design requirements. Site classification will include information on and incorporate issues related to the presence of migratory fish species, landscape characteristics, population density, sediment characteristics, existing water quality issues, and recreational services provided by the river.

To achieve a standard site classification, it is necessary to develop tools that classify or group potential sites or development regions into similar classes so that development is less site-specific. Site classification will leverage previous and ongoing research into stream classification, mitigation prediction, and environmental metrics to maximize the efficiencies that can result from systematically applying knowledge and rubrics for how environmental and ecological systems respond to disturbances. It is impossible to eliminate all site specificity from hydropower development, but future development efforts can include judicious application of validated site classification principles to select technology modules that are most appropriate for a site class, providing greater transparency, clarity, and predictability of outcomes for stakeholders.

A specific goal of site classification is to develop a framework for classifying potential SMH sites in terms useful for informing SMH development, module need, and module design requirements using existing and new classification schemes. While the current vision for SMH has focused on applications for new stream-reach development (NSD), many of the site classifications developed to date are equally useful for other applications, such as modular development at non-powered dams (NPDs) and other hydropower co-development opportunities. This report summarizes the progress made in site classification through June 2018 and presents the future direction of the site classification task.

2. CLASSIFICATION APPROACH

Site classification incorporates multivariate statistical methods to classify sites (i.e., stream reaches) into clusters—i.e., groups of sites with similar attributes—based on a simultaneous consideration of physical and biological attributes of those sites. Site classification is conducted separately for six module types:

- Hydroelectric generation (generation)
- Water quality
- Sediment passage
- Downstream fish passage/upstream fish passage (fish passage)
- Structural foundation support (foundation)
- Recreation passage

This first iteration of site classification was carried out to inform module need. We classified US stream reaches based on national-scale data sets that provide physical and biological information that a hydropower developer would generally be required to know to proceed through a pre-feasibility analysis of a site. Our goal was not to rank or prioritize sites for development, or to determine optimum designs for each cluster, but rather to use dozens of variables per stream reach to establish clusters that would likely require similar passage, generation, or foundation technologies to sustain an important river function. In the future, we will conduct a second round of classifications for many of the modules to inform the development of specific module design requirements. We will also evaluate whether an amalgamation of the individual module classifications can be used to inform SMH consideration as a whole, or whether that might need to be a separate classification altogether.

Classification assessment unit

We limit classification to stream reaches with mean annual flows of between 50 and 25,000 cubic feet per second (cfs), as defined in the SMH exemplary design report (Witt et al. 2017)[1] as the range most likely to be amenable to SMH development (i.e., low-head and <10 MW of installed capacity). Initially, we needed to decide what to classify, e.g., watersheds, catchments, individual rivers, stream reaches, or specific sites. It was additionally necessary that the object for classification be identifiable on a national level. We considered the following four options as the object of classification:

- Hydrologic Unit Codes (HUCs), 12-digit: HUCs are watersheds of various sizes that include stream reaches of different lengths. There are ~87,700 12-digit HUCs nationwide. Classifying by HUC would run the risk of averaging conditions that represent many stream reaches and would make it more difficult to limit the analysis to streams of certain sizes.

- National Hydrographic Dataset (NHD) stream reach: River reaches defined as part of the NHD have been used in past National Hydropower Asset Assessment Program (NHAAP) work. NHD stream segments are typically defined from confluence to confluence (stream intersections) and are thus not consistent in length or associated watershed size. Across the United States, there are 2,600,000 NHD reaches, 363,000 of which have flows >50 cfs and <25,000 cfs.

- Specified reach distance: An alternative to using NHD-defined reaches is to build our own stream network of a target reach length (e.g., 20 km). This however, would require significant work and not provide commensurate benefit.

---

[1] Witt et al. 2017. Exemplary Design Envelope Specification for Standard Modular Hydropower Technology. ORNL/TM-2016/298/R1.

- NSD site: NSD sites generated in a previous Department of Energy–funded project at ORNL already identify 11,041 potential hydropower development sites; however, because of the objectives of that analysis, the list probably doesn't include all possible SMH sites.

After consideration of these options, we chose to use the NHD stream reaches as the object of Site Classification. NHD reaches include complete national coverage and are a common reference used for a variety of academic and industry purposes. One drawback to using the NHD reaches is that each has a unique length, which means that care needs to be taken with regard to some descriptive characteristics that might be functions of or related to reach length.

Variable selection

Before performing the statistical classification analysis for a module or issue, we addressed three questions that identified the specific data coverages needed for the classification.

1. **What are the classification objectives for a particular module?**

   It is necessary to define what high-level questions need to be answered to determine whether a module is needed and what kind of functionality would be required. For example, in the case of sediment transport, the primary question is

- Is there a need or desire to maintain sediment transport at the project site?

2. **If a module proves to be needed, then what information is needed to make a decision about module need and functionality?**

   In the case of the sediment transport module, the second set of questions might be

   - How much sediment and what types of sediment (e.g., cobble, sand, silt) need to be passed through the site?
   - When and how often does sediment need to be passed?

3. **What data sets are needed (and available) to inform the statistical clustering exercise?**

   The information/data needed for classification will typically be of three types:

   - Instream biotic and abiotic (e.g., hydrology, fisheries, water quality)
   - Watershed (e.g., land use, land cover, soil type)
   - External (e.g., meteorology, human dimensions)

The precise information needed, as identified in the previous steps, will often not be available from a single data source. And in some cases, the specific data might not be available and, therefore, other variables might need to be identified that are correlated with or informative regarding the desired information.

In the case of the sediment transport example, specific variables that could be used to characterize similar groups of stream reaches might be

- stream flow
- runoff
- water velocity

- stream slope
- percentage of impervious surfaces in the watershed
- percentage of agriculture in watershed
- measure of suspended sediment concentrations

The set of variables used to inform module need will likely be different from those used to inform module design requirements, although with some overlap. Table A.1 in Appendix A provides an example of how variables might differ between module need and module design analyses.

2.1 STATISTICAL METHODS

We used the K-means clustering analysis from the grouping analysis tool set in ArcGIS to perform the site classification. In general, this approach creates groups by iteratively placing stream reaches in groups (hereinafter referred to as "clusters") in such a way that the overall differences among reaches (as defined by the selected set of input variables) within a cluster are minimized.

For this first round of clustering analysis, we chose to set the number of clusters at ten. We thought that this number was large enough to provide some clusters of meaningful sizes from the 300,000+ stream reaches, yet not so large that we couldn't describe the main characteristics of each of the clusters based on output information that describes which variables are most important to differentiate each cluster from the others. Because the clustering was not constricted to produce clusters of equal numbers of reaches, some clusters may be very small (i.e., fewer than 100 reaches); these clusters are generally ignored in our summary analysis. In addition, no spatial constraints were applied to the analysis, so reaches within the same cluster might be at quite a distance from one another.

The number of input variables for the clustering analysis differed among the different modules and varied from 5 to 12 (Table 1). We chose variables for each clustering exercise that best addressed the data needs defined in our variable selection as described above and avoided variables that were highly correlated.

The results of the clustering will not in themselves identify which stream reaches are the best candidates for SMH development. But the characteristics that are most correlated with each cluster (e.g., high flow or the presence of migratory salmon), and that produce differentiation among clusters, can inform various decisions regarding project development. To provide greater interpretation of clusters relative to their favorability for SMH development, we compared the distribution of NHD stream reaches in the clusters with the distribution of potential NSD sites and existing hydropower sites as identified in Oak Ridge National Laboratory's (ORNL's) NHAAP database. Clusters with a higher relative distribution of NSD sites or existing hydropower dams than the NHD reaches in the clusters might be considered as clusters that are more favorable for development, since their potential has been previously demonstrated.

Table 1. Variables used in clustering analyses. Lowercase x's identify those variables with R^2s <0.5, indicating those with the least influence on cluster determination.

Metric	Description (units)	Module					
		Generation	**Water quality**	**Sediment passage**	**Fish passage**	**Foundation**	**Recreation**
QA_MA	Mean annual flow (MAF) (cfs)	X	X	X	X	X	X
QA_CV	Coefficient of variation for flow based on monthly averages and annual mean	x		x			
SLOPE	Slope of stream segment	X		x		X	x
VA_MA	MAF velocity (m/s)	x		x		x	
RunoffWs	Mean runoff in watershed	X					
IEOFCAT	Mean infiltration-excess overland flow in catchment	X					
BFICAT	Base-flow index	X					
ElevDiffWS	Difference between maximum and minimum elevation in watershed	X					
PctAgWs	Percent agricultural land cover in watershed		X				
PctImp2006Ws	Percent imperviousness from 2006 in watershed		X				
PctImp2006Cat	Percent imperviousness from 2006 in catchment		X				
PctAgCat	Percent agricultural land cover in catchment		X				
FarmNCat	Sum total of nitrogen from farm areas in catchment		X				
KffactWs	The Kffactor—relative index of susceptibility of bare, cultivated soil to particle detachment and transport by rainfall in watershed		x				
KffactCat	The Kffactor—relative index of susceptibility of bare, cultivated soil to particle detachment and transport by rainfall in catchment		x				
PctForWetWs	Percent forest or wetland land cover in watershed		X				
PctForWetCat	Percent forest or wetland land cover in catchment		X				
IEOFCat	Mean infiltration-excess overland flow in catchment		X				
PopDns10Cat	Population density from 2010 census in catchment		x				
RckDepCat	Mean depth to bedrock in catchment (cm)			X		X	
PctImp2006Ws	Percent imperviousness from 2006 in watershed			X			
PctAgWs	Percent agricultural land cover in watershed			X			
PctForRipWs	Percent riparian forest land cover in watershed			X			
RunoffWs	Mean runoff in watershed			X			
KffactWs	The Kffactor—relative index of susceptibility of bare, cultivated soil to particle detachment and transport by rainfall in watershed			X			
PctClayWs	Percent clay content of soils in watershed			x			
PctSandWs	Percent sand content of soils in watershed			X			
DamUNDR	Upstream network dam density per unit stream network length (#/100 km)				X		
DamDMD	Downstream mainstem dam density per unit downstream mainstem length (#/100 km)				X		
HUC2PctFP	Percent of mitigation sites in the mitigation database within the HUC2 that had Tier 1 fish passage mitigation required				X		

Table 1. Variables used in clustering analyses (continued).

Metric	Description (units)	Module					
		Generation	Water quality	Sediment passage	Fish passage	Foundation	Recreation
AnadAcip	Number of ocean-run sturgeon species (Acipenseriformes) within the reach's HUC8 (count)				x		
PotAcip	Number of inland sturgeon/paddlefish species (Acipenseriformes) within the reach's HUC8 (count)				X		
AnadClup	Number of ocean-run clupeid species within the reach's HUC8 (count)				X		
EelsLamp	Number of ocean-run eel/lamprey species within the reach's HUC8 (count)				X		
AnadSalm	Number of ocean-run salmonid species within the reach's HUC8 (count)				X		
PotSalm	Number of inland salmonid species within the reach's HUC8 (count)				x		
PotOthr	Number of other inland migratory species within the reach's HUC8				X		
PowerQS	Measure of stream power (mean annual flow x slope)					X	
UCSLow	Low value for range of unconsolidated rock score					X	
UCSHigh	High value for range of unconsolidated rock score					X	
GrndAccel	Earthquake susceptibility					X	
Fishing_Salt	Saltwater fishing locations in HUC6 (count)						X
NRIRecHUC6KM	Length of stream identified as having outstanding recreational value in the National Rivers Inventory (m) [1]						X
AWHUC6KM	Length of stream identified as American Whitewater paddling runs (m)						X
Boat_Ramp_Un	Number of undeveloped boat ramps in HUC6 (count)						X
Boat_Ramp	Number of developed boat ramps in HUC6 (count)						X
PopDns10HUC6	Mean population density in HUC6						X
PopDns10CAT	Mean population density in stream reach catchment						X
Fishing_Cold	Number of cold-water fishing locations in HUC6 (count)						X

2.2 MODULE-SPECIFIC CLASSIFICATION CATEGORIES

2.2.1 Generation

The primary objective of classifying sites based on generation potential relies on understanding the ability of the stream to supply adequate flow, velocity, and head to produce a viable source of hydropower. This analysis is not intended to be a detailed resource assessment but only to provide a very broad classification. For the clustering analysis, we included variables that were related to or that characterized (1) the amount of water in a reach (QA_MA, RunoffWs, IEOFCAT), (2) temporal aspects of the hydrograph (QA_CV, BFICAT), and (3) the potential energy associated with the stream (SLOPE, VA_MA, ElevDiffWS) (Table 1). Fine-scale information (i.e., hourly or daily) on flow variability is not readily available by NHD reach, so we calculated a flow coefficient of variation based on mean monthly averages to at least include seasonal variations in flow in the analysis.

2.2.2 Water Quality

The primary objective of classifying sites to inform the need for a water quality module is to group sites of similar existing or future water quality and water quality risk. We included variables that were related to or characterized (1) indicators of existing compromised water quality (i.e., PctAgWs, PctImp2006Ws, PctImp2006Cat, PctAgCat, FarmNCat, PopDns10Cat), (2) land cover characteristics that affect water quality (KffactWs, KffactCat, PctForWetWs, PctForWetCat), and (3) instream water quantity (i.e., QA_MA, IEOFCat) (Table 1).

2.2.3 Sediment Passage

The primary objective of classifying sites to inform the need for a sediment transport module was to understand if there will likely be a need or desire to pass sediment and, if so, what kind of sediment, how much and when. We therefore included variables that were related to or characterized (1) the quality and quantity of sediment in the stream (i.e., RckDepCat, PctClayWs, PctSandWs), (2) the ability of the stream to transport sediment (i.e., QA_MA, Slope, VA_MA), (3) watershed impacts on sediment transport (PctImp2006Ws, PctAgWs, PctForRipWs, RunoffWs, KffactWs), and (4) the temporal aspects of sediment transport dynamics (i.e., QA_CV) (Table 1).

Because direct sediment transport data are generally not available at the level of national coverage, we are also pursuing an alternate approach with the University of Tennessee’s Water Resources Research Center that includes a more mechanistic understanding of the hydrodynamics of stream geomorphology and sediment transport. If successful, this effort will result in a suite of variables that can be used in a clustering analysis, which are more closely related to specific sediment type and transport hydrology.

2.2.4 Fish Passage

To determine if an upstream or downstream fish passage module is needed, it is important to know whether fish are present that would normally pass through the site and, if so, what their life stages are and when, how frequently, and how they can pass (for example by spill, through a turbine, or by way of ladder or bypass). The primary objective of classifying sites based on fish passage need was to include variables that characterized or were related to (1) existing barriers in the stream network (DamUNDR, DamDMD), (2) the existence of passage mitigation at nearby hydropower dams (i.e., HUC2PctFP), and (3) the presence and absence of various migratory fish species (i.e., AnadAcip, PotAcip, AnadClup, EelsLamp, AnadSalm, PotSalm, PotOthr) (Table 1).

2.2.5 Foundation

To better understand what type of foundation module might be needed, we need to know something about the stability of the river geomorphology and the water flow forces that are present. The primary objective of classifying sites based on foundation need was to include variables that were related to or useful for characterizing (1) streambed morphology (RckDepCat), (2) erodibility of local soil and rock (UCSLow, UCSHigh), (3) water force that would be experienced by an in-river structure (QA_MA, Slope, VA_MA, PowerQS), and (4) a measure of earthquake instability or risk (GrndAccel) (Table 1).

2.2.6 Recreation

To better understand the possible need for and purpose of a recreation support module, we need to know something about the types of river-related recreation that are presently supported in the region and the potential for new recreational opportunities—including the number of people that might take advantage of them. The primary objective of classifying sites based on recreation need was to include variables that were related to or characterized (1) the present status of recreation in the region (Fishing_Salt, NRIRecHUC6KM, AWHUC6KM, Boat_Ramp_Un, Boat_Ramp, Fishing_Cold), (2) the numbers of active or potential recreation users (PopDns10HUC6, PopDns10CAT), and (3) river size (QA_MA, Slope) (Table 1).

3. RESULTS

3.1 GENERATION

Of the initial 304,035 NHD stream reaches, 264,180 were grouped into 10 clusters based on similarities in the amount of water in a reach, temporal aspects of the hydrograph, and potential energy associated with the stream. Table 2 shows how the NHD reaches are distributed among the ten clusters, what the defining characteristics of each cluster are, and where the reaches within a cluster are distributed geographically. Figure 1 illustrates the geographic distribution of the ten clusters. Six of the clusters (numbers 1, 2, 3, 7, 8, and 9) contain attributes that are correlated with viable hydropower development (e.g., some combination of high velocity, high baseflow, or steep gradient), and four of the clusters (numbers 4, 5, 6, and 10) are dominated by characteristics that are not favorable to development (e.g., low baseflow, low gradient, low velocity).

Most of the sites (4,841 of 8,489; 57%) identified in ORNL's NSD analysis (NSD sites) are in reaches in clusters 3 and 8 (Figure 2). Similarly, 993 (61%) of the 1,636 existing hydropower sites in the United States are in reaches contained in clusters 1, 3, 7, and 8. Clusters with a higher relative distribution of NSD sites or existing dams than the NHD reaches in the clusters might be considered more favorable to development, since their potential has been previously demonstrated. In the case of the generation analysis, using a criterion of a >5% difference from the NHD distribution, we found that NSD sites were disproportionately distributed in clusters 2, 3, and 10; and existing dams were disproportionately distributed in reaches in clusters 3 and 8. About 39% of NSD sites and 25% of existing dams are in cluster 3 stream reaches, which contain only 8% of the NHD stream reaches.

Table 2. Number and defining characteristics of stream reaches classified into ten different clusters as a result of K-means clustering analysis for generation potential.

#	# Reaches	Defining characteristics	Locale
1	37,500	Low Q, high baseflow, high seasonal variability (snow melt)	Rockies, Sierras
2	10,000	Med Q, steep grade, high baseflow, low seasonal variability	Cascades
3	21,500	Med Q, high velocity	National
4	9,000	Med Q, low grade, low baseflow	Great Plains, Texas
5	18,000	Low Q, low grade, low baseflow, low velocity, high seasonal variability	National, valleys and plains
6	69,000	Low Q, low grade, low baseflow, low velocity	Mississippi Valley, Midwest and Midsouth
7	2,700	Low Q, steep grade, high baseflow	Rockies, Sierras, Cascades
8	86,500	High baseflow, low Q, moderate runoff, low velocity	Great Lakes, Appalachians, Atlantic Coast, West Coast foothills
9	20	Med Q, very steep gradient, high velocity	–
10	9,500	High Q, low grade	National

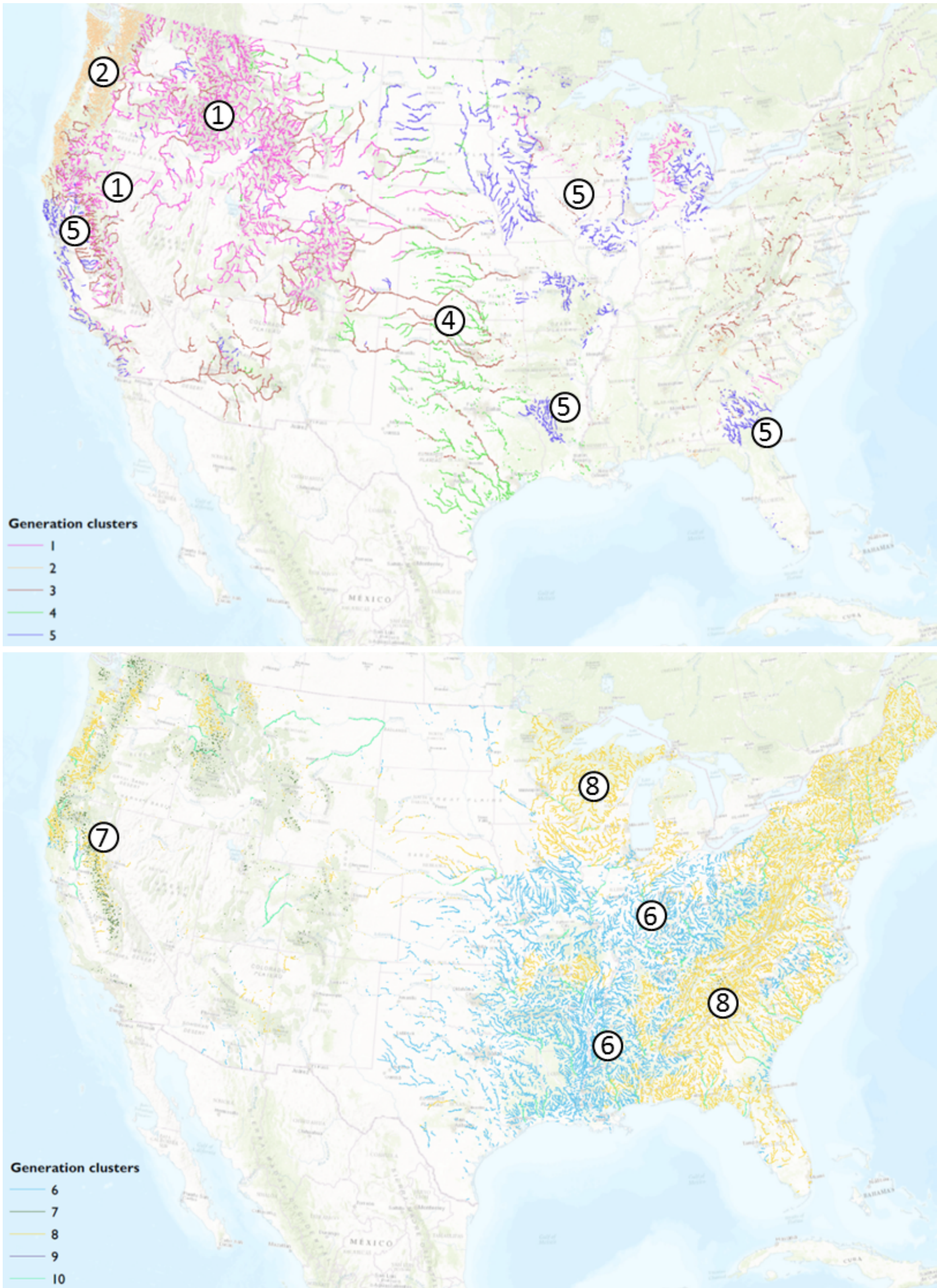

Figure 1. Maps of the k-means clusters 1 through 5 (top panel) and 6 through 10 (bottom panel) based on characteristics related to hydropower generation. Clusters not identified in the figures are either few in number or widely distributed.

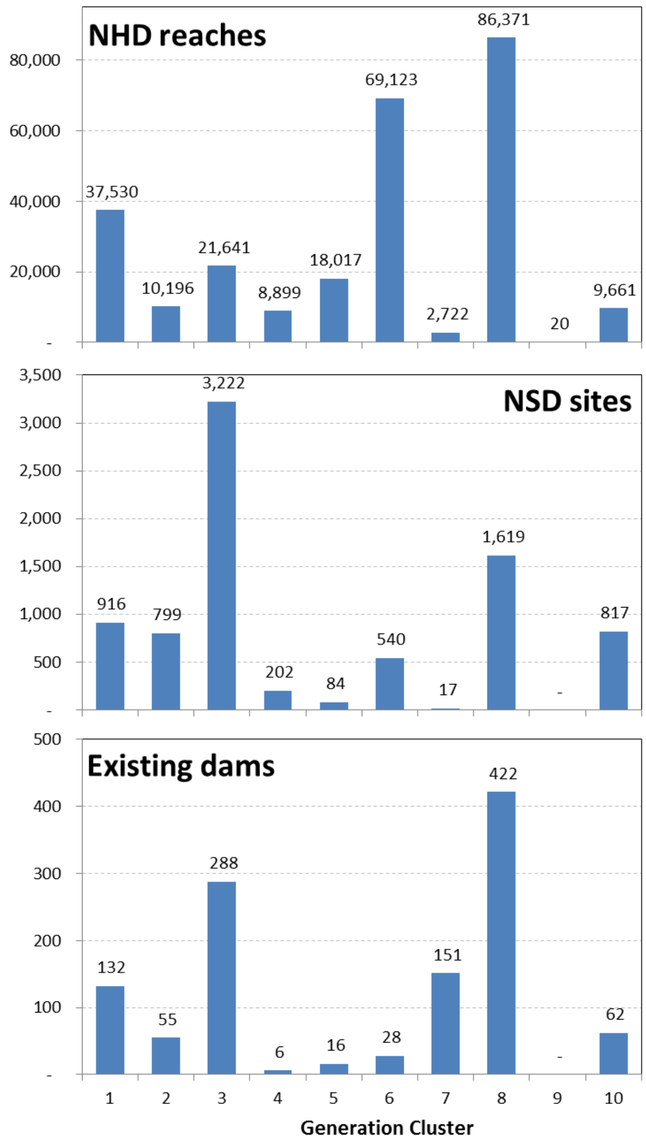

Figure 2. Count of US NHD stream reaches (1) in each of ten K-means generation clusters (top panel), (2) associated with identified NSD sites (middle panel), and 3) associated with existing hydropower dams (bottom panel).

3.2 WATER QUALITY

Of the initial 304,035 NHD stream reaches, 252,709 were grouped into 10 clusters based on similarities in the amount of water in a reach, indicators of existing compromised water quality, and land cover characteristics that affect water quality (Table 3, Figure 3). Within this analysis, it is difficult to pre-select desirable or viable hydropower sites based on water quality. For example, high-gradient streams with high generation potential are often located in remote mountainous areas where water quality is generally very good. On the other hand, development in streams with already degraded water quality could result in fewer water quality concerns related to hydropower development

Most of the sites (5,357 of 8,489; 63%) identified in ORNL's NSD analysis (NSD sites) are in reaches in clusters 1, 5, and 7 (Figure 4). Similarly, 1,219 (75%) of the 1,636 existing hydropower sites in the United States are in reaches contained in clusters 1, 5, 7, and 9. Clusters with a higher relative distribution of NSD sites or existing dams than the NHD reaches in the clusters might be considered more favorable to development, since their potential has been previously demonstrated. In the case of the water quality analysis, using a criterion of a >5% difference from the NHD distribution, we found that NSD sites were disproportionately distributed in clusters 1 and 4; and existing dams were disproportionately distributed in clusters 5 and 9. About 43% of existing dams are in cluster 5, which contains only 22% of the NHD stream reaches.

Table 3. Number and defining characteristics of stream reaches classified into ten different clusters as a result of K-means clustering analysis for water quality.

#	# Reaches	Defining characteristics	Locale
1	38,238	Unforested, low agriculture, low erodibility	Plains
2	30,675	Agricultural, high erodibility	Mississippi and Ohio River valleys
3	114	Agricultural, nitrogen runoff	–
4	10,144	Large rivers,	National
5	54,936	Forested, low erodibility	National
6	40,053	Agricultural, high erodibility, nitrogen runoff	Midwest, Ohio and Mississippi River valleys
7	59,040	Forested, low agriculture, moderate erodibility	National
8	1,793	Small streams, urban, impervious surfaces	National
9	7,509	Suburban, impervious	
10	10,207	Agricultural, unforested, nitrogen runoff	Great Plains

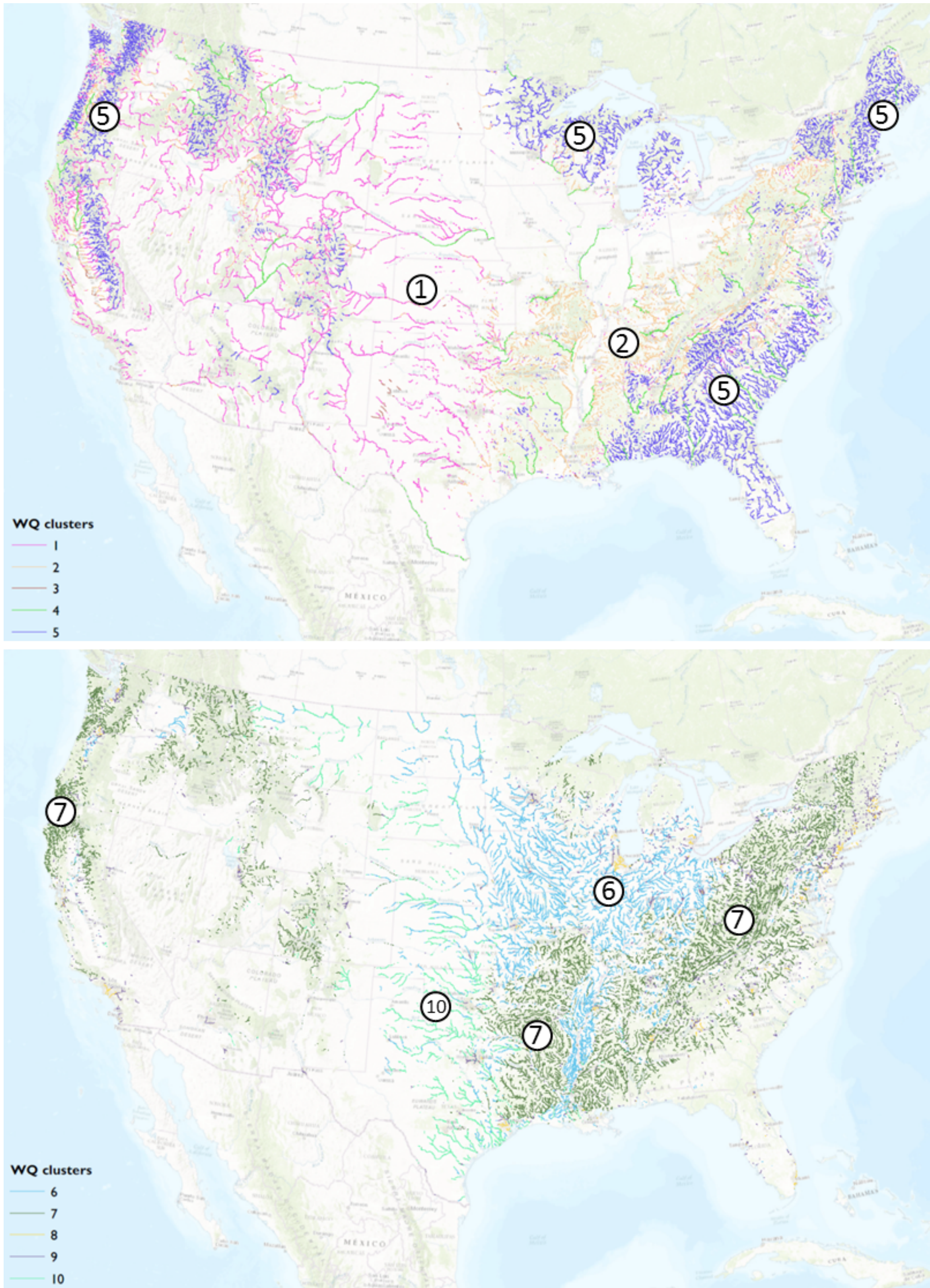

Figure 3. Maps of the k-means clusters 1 through 5 (top panel) and 6 through 10 (bottom panel) based on characteristics related to water quality. Clusters not identified in the figures are either few in number or widely distributed.

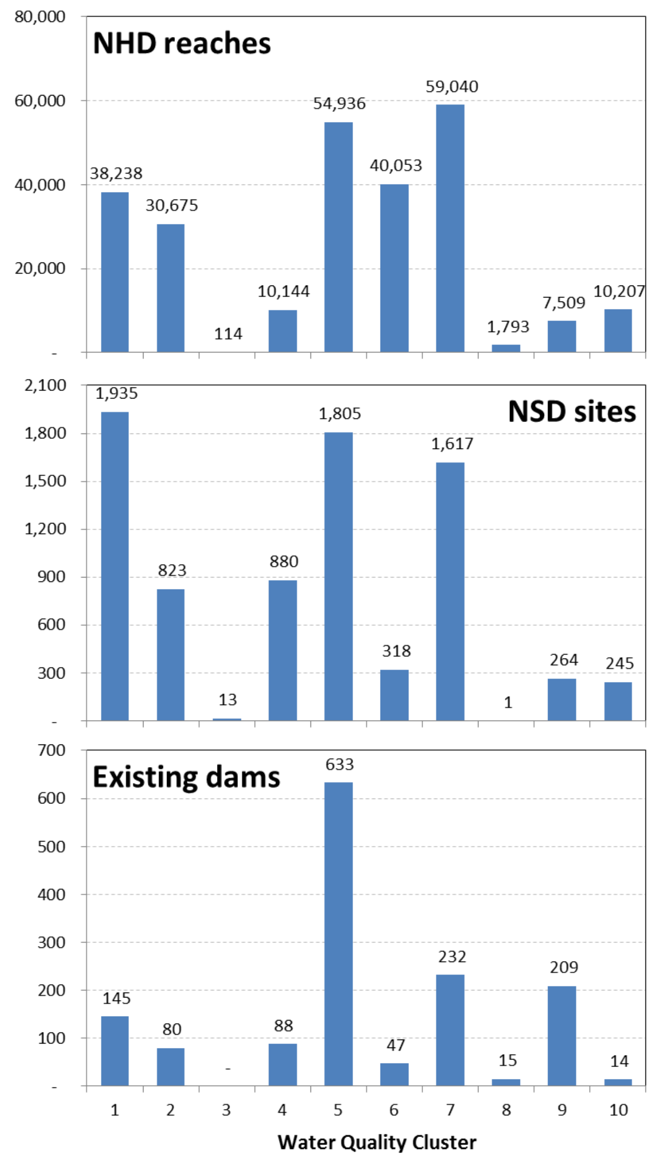

Figure 4. Count of US NHD stream reaches (1) in each of ten K-means water quality clusters (top panel), (2) associated with identified NSD sites (middle panel), and (3) associated with existing hydropower dams (bottom panel).

3.3 SEDIMENT PASSAGE

Of the initial 304,035 NHD stream reaches, 271,170 were grouped into 10 clusters based on similarities in the quality and quantity of sediment in the stream, the ability of the stream to transport sediment, and watershed impacts on sediment load and transport (Table 4, Figure 5). Within this analysis, it is difficult to pre-select desirable or viable hydropower sites based on sediment passage. However, developers might prefer sites with little need for sediment transport—such as streams in heavily forested watersheds, high-velocity streams, and those with low agricultural input (e.g., clusters 2, 6, 7, and 10) or where sediment transport is easier to achieve, such as streams with fine suspended sediment or sand (e.g., clusters 3, 4, and 8) as opposed to coarse gravel and cobble.

Most of the sites (5,014 of 8,489; 59%) identified in ORNL's NSD analysis (NSD sites) are in reaches in clusters 1, 6, and 7 (Figure 6). Similarly, 986 (60%) of the 1,636 existing hydropower sites in the United States are in reaches contained in clusters 1, 6, 7, 8, and 9. Clusters with a higher relative distribution of NSD sites or existing dams than the NHD reaches in the clusters might be considered more favorable to development, since their potential has been previously demonstrated. In the case of the sediment passage analysis, using a criterion of a >5% difference from the NHD distribution, we found that NSD sites were disproportionately distributed in clusters 1, 9 and 10; and existing dams were disproportionately distributed in clusters 6 and 9. About 41% of existing dams are in cluster 6, which contains only 15% of the NHD stream reaches.

Table 4. Number and defining characteristics of stream reaches classified into ten different clusters as a result of K-means clustering analysis for sediment passage.

#	# Reaches	Defining characteristics	Locale
1	34,841	Low runoff, variable flow	Mountain West and plains
2	81	Moderate steady flow, low ag, high runoff, high velocity	
3	44,804	Low velocity, clay	Upper Midwest
4	62,824	Agricultural, slow, high erodibility, clay	Midsouth
5	3,592	Small streams, slow, urban	National
6	40,154	Forested, low ag	Northeast, Northwest, Appalachians
7	32,939	Rocky streams, forested	National
8	30,742	Sandy, low erosion, slow	Southeast, Great Lakes
9	11,471	Large rivers, high velocity	National
10	9,785	Forested, low ag, high runoff, steady flow	Pacific Northwest

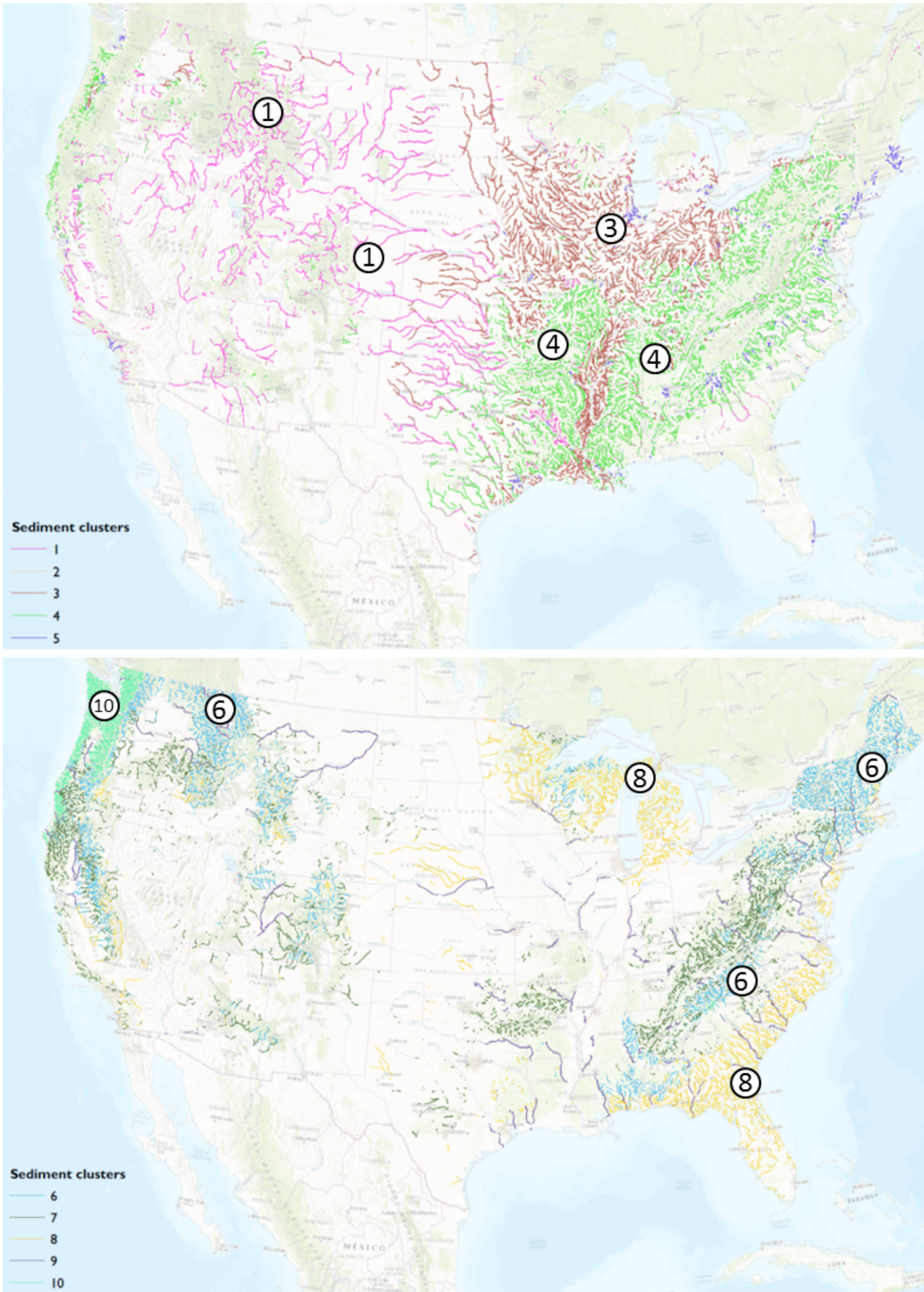

Figure 5. Maps of the k-means clusters 1 through 5 (top panel) and 6 through 10 (bottom panel) based on characteristics related to sediment passage. Clusters not identified in the figures are either few in number or widely distributed.

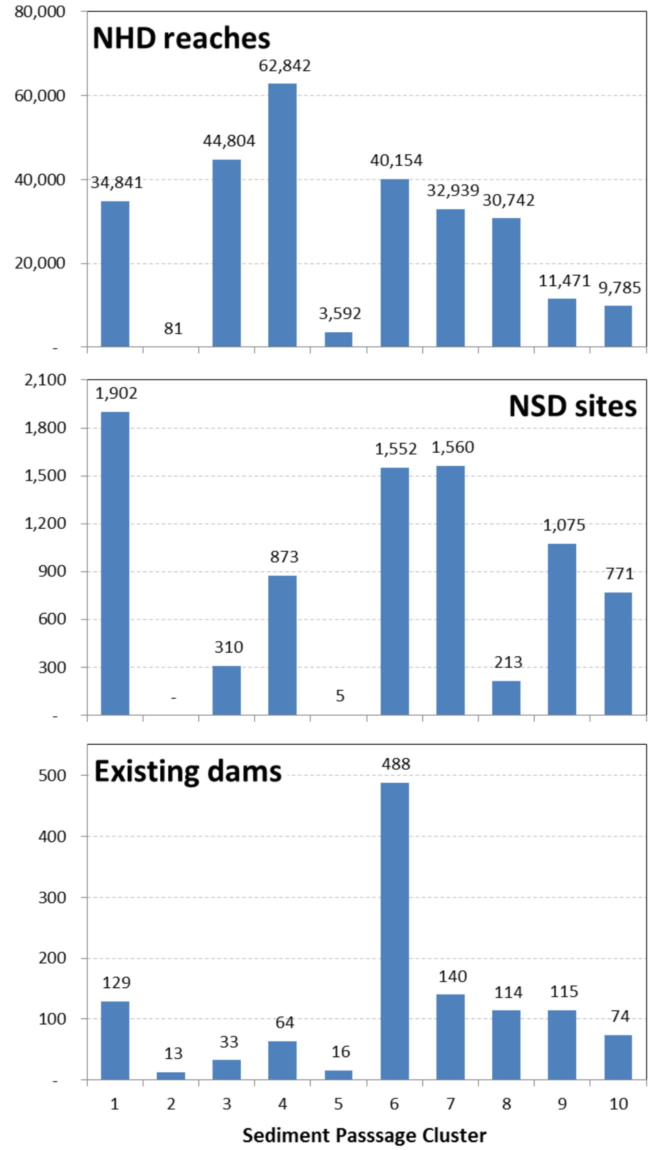

Figure 6. Count of US NHD stream reaches (1) in each of ten K-means sediment passage clusters (top panel), (2) associated with identified NSD sites (middle panel), and (3) associated with existing hydropower dams (bottom panel).

3.4 FISH PASSAGE

Of the initial 304,035 NHD stream reaches, 261,134 were grouped into ten clusters based on similarities in the number of existing barriers in the stream network, the existence of passage mitigation at nearby hydropower dams, and the presence or absence of various migratory fish species (Table 5, Figure 7). Developers that prefer to avoid fish passage mitigation might favor clusters with low numbers of migratory fish species present, with many downstream dams that already prevent anadromous fish migration, and/or near existing licensed dams with minimal existing fish passage requirements, e.g., clusters 5, 7, and 9).

Most of the sites (6,864 of 8,489; 81%) identified in ORNL's NSD analysis (NSD sites) are in reaches in clusters 1, 2, 3, and 7 (Figure 8). Similarly, 885 (54%) of the 1,636 existing hydropower sites in the United States are in reaches contained in clusters 1, 2, 7, and 10. Clusters with a higher relative distribution of NSD sites or existing dams than the NHD reaches in the clusters might be considered more favorable to development, since their potential has been previously demonstrated. In the case of the fish passage analysis, using a criterion of a >5% difference than the NHD distribution, we found that NSD sites were disproportionately distributed in clusters 1, 3 and 6; and existing dams were disproportionately distributed in clusters 1, 2, and 10.

Table 5. Number and defining characteristics of stream reaches classified into ten different clusters as a result of K-means clustering analysis for fish passage.

#	# Reaches	Defining characteristics	Locale
1	41,002	Potamodromous salmonids, high downstream dam count, high existing passage mitigation	Appalachia, Texas, Northwest
2	48,378	High other potamodromous species, low anadromous species	Great Lakes, upper Midwest, upper Ohio River, Gulf Coast
3	17,218	Anadromous salmonids, potamodromous salmonids, low upstream and downstream dam count, high existing passage mitigation, anadromous lampreys	Pacific Northwest
4	10,446	Some anadromous clupeids, high upstream and downstream dam count, low MAF	South central
5	19,626	Low existing passage mitigation, low or absent salmonid presence, eels, low downstream dam count	Lower Mississippi River drainage
6	7,731	High MAF, inland sturgeon, and other inland species	Scattered nationally
7	94,507	Very low numbers of all major migratory species, low existing passage mitigation	Scattered nationally
8	2	–	–
9	8,467	Inland sturgeons and other inland potamodromous species, low downstream dam count, low existing passage mitigation, low anadromous species	Upper Mississippi River drainage
10	13,757	Anadromous clupeids, ocean-run sturgeons, eels, high upstream and downstream dam count,	Atlantic Coast

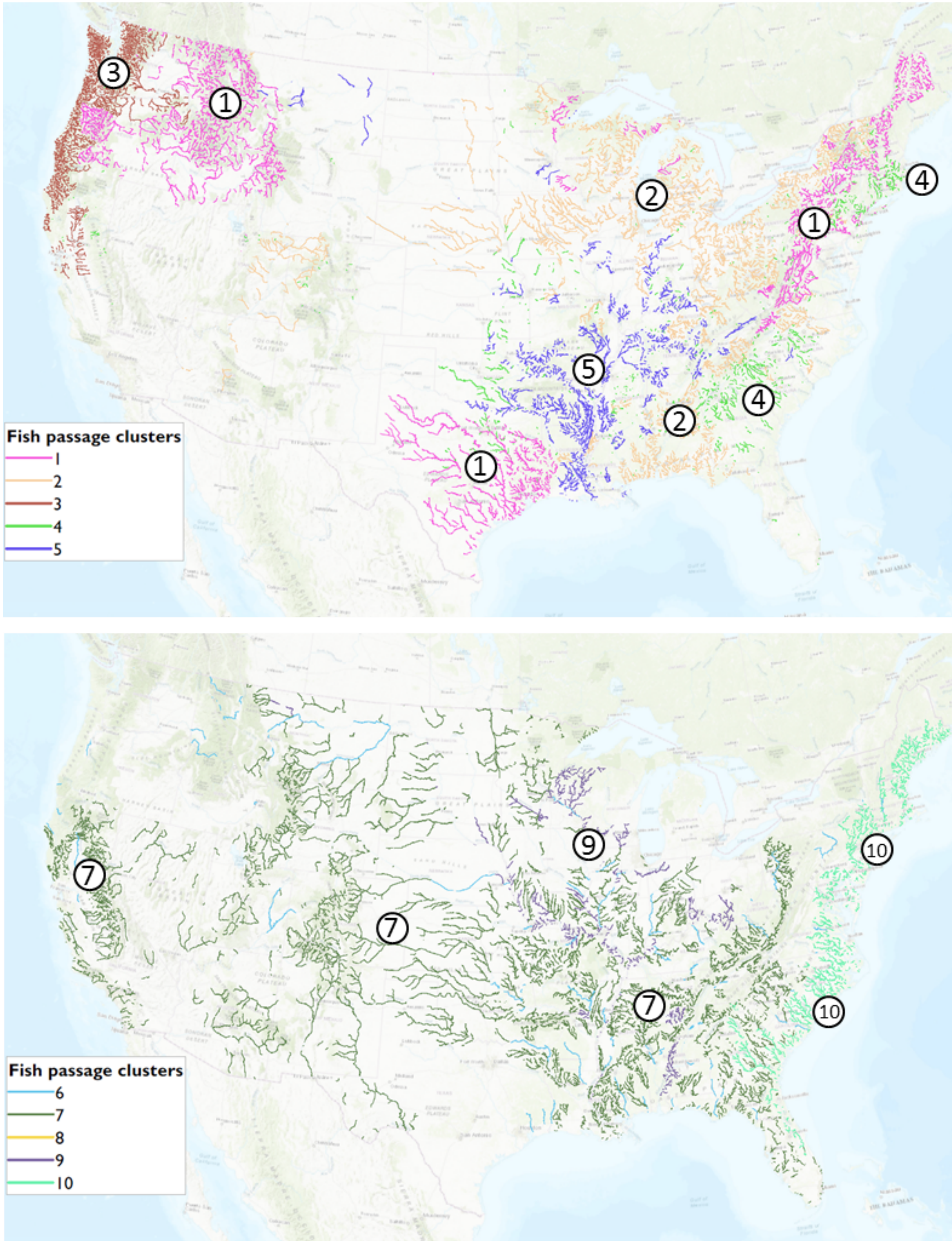

Figure 7. Maps of the k-means clusters 1 through 5 (top panel) and 6 through 10 (bottom panel) based on characteristics related to fish passage. Clusters not identified in the figures are either few in number or widely distributed.

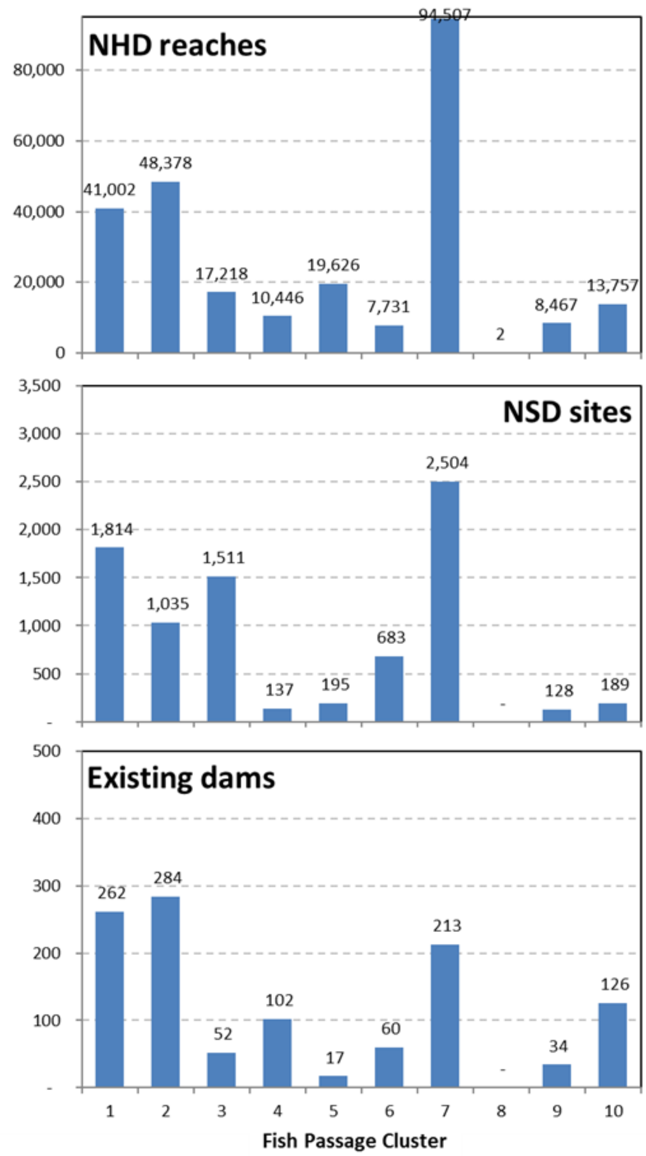

Figure 8. Count of US NHD stream reaches (1) in each of ten K-means fish passage clusters (top panel), (2) associated with identified NSD sites (middle panel), and (3) associated with existing hydropower dams (bottom panel).

3.5 FOUNDATION

Of the initial 304,035 NHD stream reaches, 271,271 were grouped into 10 clusters based on similarities in the hydrology (flow and velocity), depth to bedrock, risk of earthquake disturbance, and stream power (Table 6, Figure 9). Sites in the cluster defined by potential earthquake severity (i.e., cluster 3) would likely require additional consideration of dam safety in assessing foundational support, as would those with higher flows and higher velocities (i.e., clusters 5, 7, 8, and 10). Those reaches where depth to bedrock is shallow (i.e., clusters 1 and 7) would likely require different foundation anchoring from those where the bedrock is relatively deep (i.e., clusters 2, 4, 6, and 8).

Most of the sites (5,960 of 8,489; 70%) identified in ORNL's NSD analysis (NSD sites) are in reaches in clusters 1, 5, 7, and 8 (Figure 10). Similarly, 872 (53%) of the 1,636 existing hydropower sites in the United States are in reaches contained in clusters 4, 7, 8, and 9. Clusters with a higher relative distribution of NSD sites or existing dams than the NHD reaches in the clusters might be considered more favorable to development, since their potential has been previously demonstrated. In the case of the foundation analysis, using a criterion of a >5% difference from the NHD distribution, we found that NSD sites were disproportionately distributed in clusters 5 and 8; and existing dams were disproportionately distributed in clusters 7, 8, and 9.

Table 6. Number and defining characteristics of stream reaches classified into ten different clusters as a result of K-means clustering analysis for project foundation.

#	# Reaches	Defining characteristics	Locale
1	34,133	Low flow, shallow bedrock,	Low mountain streams: foothills of Appalachia, Ozark, Sierras, etc.
2	43,474	Low power, low flow, low gradient, deep bedrock	Lowlands: Glaciated Great Lakes, Upper Mississippi
3	12,912	Highest earthquake hazard, high erodibility	Pacific Coast, New Madrid fault, South Carolina coast
4	42,554	Low power, low erodibility, low flow, low gradient, deep bedrock	Lowlands: Midwest, Northeast, and Southeast
5	10,110	High flow, high power, high velocity	Large rivers: national
6	75,378	Low power, high erodibility, low flow, low gradient, deep bedrock	Lowlands: Southeast and Gulf Coast, Central Valley California
7	20,837	Moderately high power, high velocity, low erodibility, shallow bedrock	Foothill streams: Pacific Northwest, Rockies, Appalachians, Maine
8	29,151	High erodibility, high flow, moderately high velocity, deep bedrock	National
9	2,644	Low erodibility, very high power, shallow bedrock, high gradient, moderately high velocity	–
10	78	Low erodibility, high flow, very high gradient, very high velocity	–

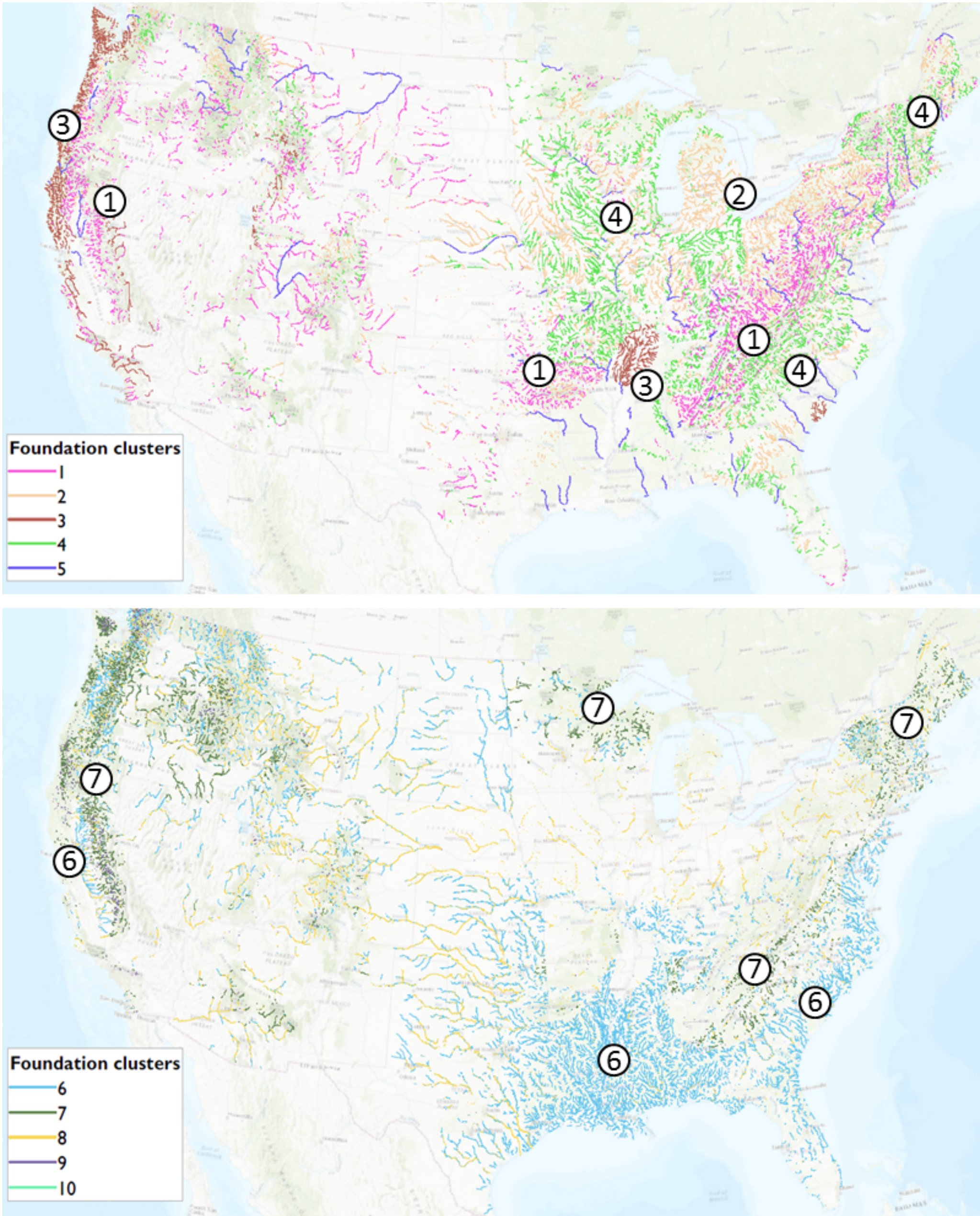

Figure 9. Maps of the k-means clusters 1 through 5 (top panel) and 6 through 10 (bottom panel) based on characteristics related to foundational support. Clusters not identified in the figures are either few in number or widely distributed.

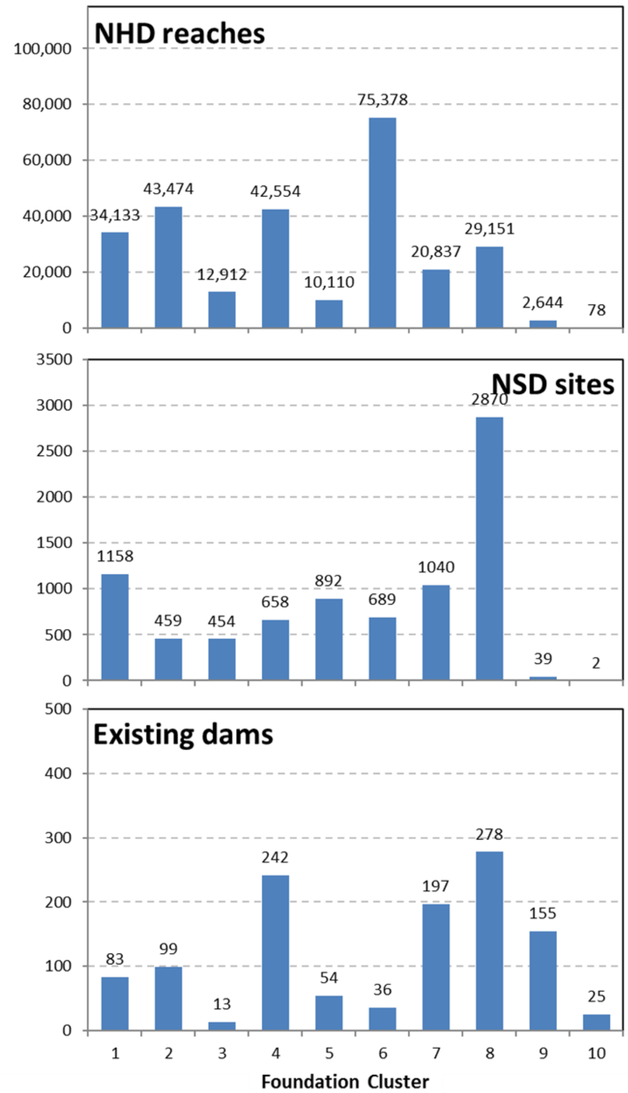

Figure 10. Count of US NHD stream reaches (1) in each of ten K-means foundation clusters (top panel), (2) associated with identified NSD sites (middle panel), and (3) associated with existing hydropower dams (bottom panel).

3.6 RECREATION

Of the initial 304,035 NHD stream reaches, 301,472 were grouped into 10 clusters based on similarities in indicators of existing recreational resources (i.e., boat ramps, recognized whitewater sections, designated recreation and fishing reaches), proximity to user populations, river flow, and stream gradient (Table 7, Figure 11). One might expect that clusters defined by high levels of existing river recreation would likely require a module to maintain such services (i.e., clusters 4, 6, 7, and 8), whereas those with low present recreation might have less need (i.e., clusters 2, 9, and 10). On the other hand, clusters low in recreation but near population centers (i.e., clusters 2 and 3) might be perfect candidates for the creation of new recreational opportunities (e.g., whitewater parks) as part of "co-development" with hydropower.

Most of the sites (7,023 of 8,489; 83%) identified in ORNL's NSD analysis (NSD sites) are in reaches in clusters 4, 9, and 10 (Figure 12). Similarly, 1,163 (71%) of the 1,636 existing hydropower sites in the United States are in reaches contained in clusters 4, 7, and 10. Clusters with a higher relative distribution of NSD sites or existing dams than the NHD reaches in the clusters might be considered more favorable to development, since their potential has been previously demonstrated. In the case of the recreation analysis, using a criterion of a >5% difference from the NHD distribution, we found that NSD sites were disproportionately distributed in clusters 4 and 9; and existing dams were disproportionately distributed in clusters 3 and 7.

Table 7. Number and defining characteristics of stream reaches classified into ten different clusters as a result of K-means clustering analysis for recreation.

#	# Reaches	Defining characteristics	Locale
1	10,788	Marine species fishing, small streams, limited boat access	Atlantic, Pacific, and Gulf coasts
2	1,669	Urban streams	National
3	7,464	Suburban, small streams, cold-water fishing	National
4	42,225	High whitewater use, cold-water fishing	Appalachians, Sierras, and Rocky Mountains
5	6	High gradient, low population density, limited boat access, cold-water fishing	–
6	39,195	High recreational preservation value, low gradient	Mid-central, Southeast, far Northwest
7	24,667	Low gradient, high boat access, some whitewater	Maine, Wisconsin, Minnesota, and Arkansas
8	3,454	Marine species fishing, high recreational preservation value, high whitewater use, high boat access, cold-water fishing	Puget Sound
9	12,598	Large rivers, low gradient	National
10	159,406	Rural, limited boat access, low gradient	Ohio and Mississippi River valleys, eastern Great Lakes, Great Plains

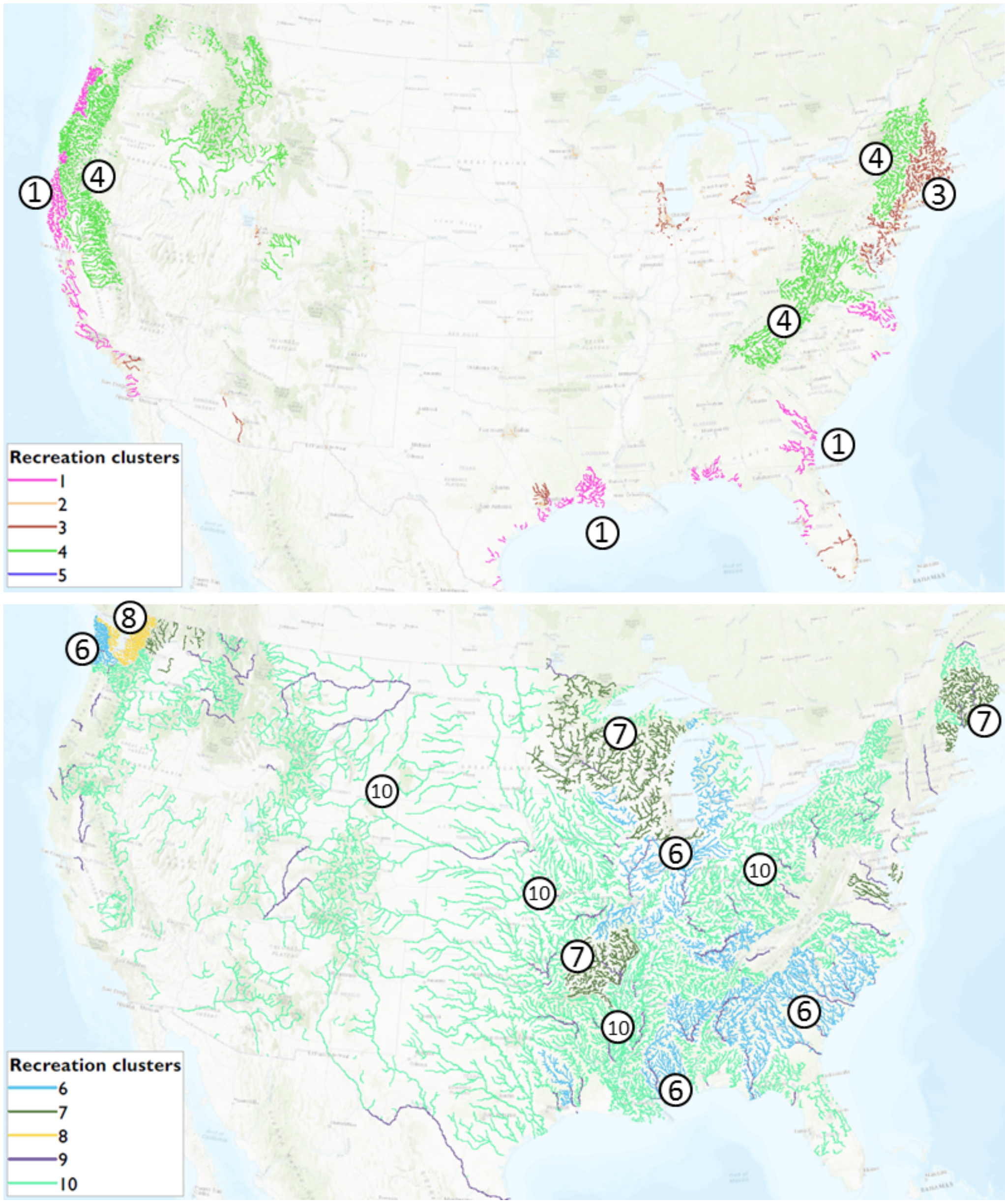

Figure 11. Maps of the k-means clusters 1 through 5 (top panel) and 6 through 10 (bottom panel) based on characteristics related to supporting recreation uses. Clusters not identified in the figures are either small in number or widely distributed.

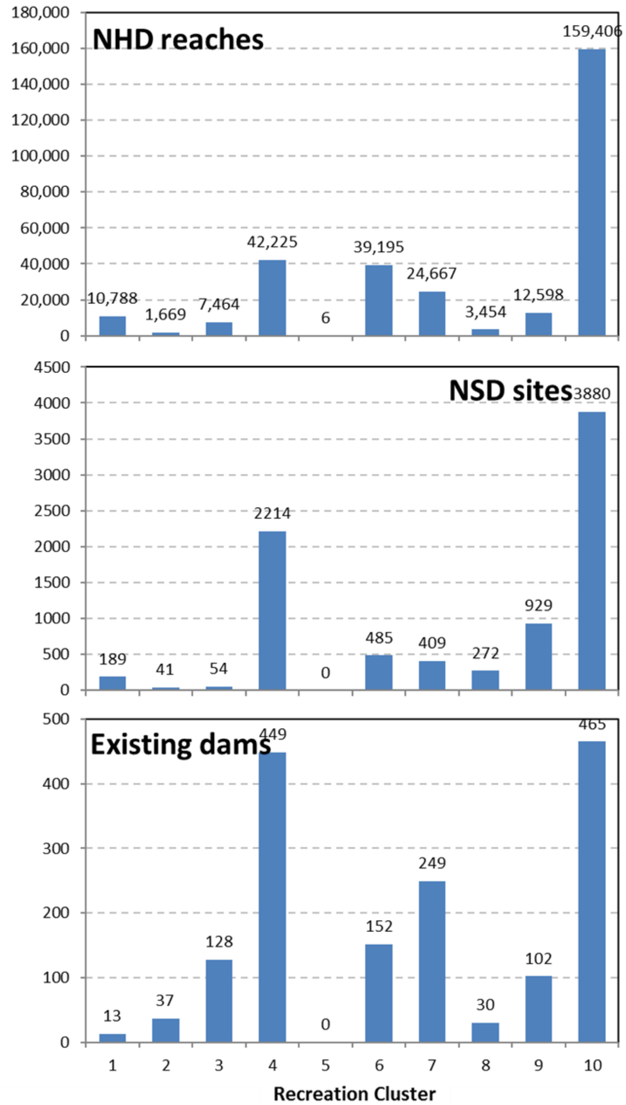

Figure 12. Count of US NHD stream reaches (1) in each of ten K-means recreation clusters (top panel), (2) associated with identified NSD sites (middle panel), and (3) associated with existing hydropower dams (bottom panel).

4. DISCUSSION

The results of the site classification analyses are summarized in a searchable Excel spreadsheet and in a soon-to-be-released, geospatial, web-based tool called SMH Explorer. Both formats will include for each of the 300,000+ NHD reaches the following information:

- the cluster number (1 to 10) for each of the six module classifications completed to date
- the value of the key variables that went into the clustering analysis (e.g., mean annual flow, percentage of impervious surfaces in the watershed, presence of anadromous salmon)
- the length of each reach
- presence of existing dams in the reach (and generation capacity)
- presence of NSD sites in the reach

The online SMH Explorer, which is presently undergoing beta testing, is intended for SMH module developers, SMH project developers, and environmental regulators. The intended uses of the tool might vary depending on user group.

Module developers

- Application space—A module developer might want to know how much opportunity there is for a particular module design. This could be useful for pre-development decision-making and post-development marketing. Example: How many sites or how big an area might benefit from a fish passage module that passes a particular species group?
- Module design questions—A module developer might want to know the range of conditions within a certain type of project site. Example: What is the underlying geomorphology (e.g., sand or bedrock) for a particular region of particular stream size?

Project developers

- Module need—A project developer will want to know what modules (and functions) are likely to be needed for a site of interest. Example: Will sediment transport be necessary and, if so, what type and amount of sediment and how often?
- Issue avoidance—A developer might want to find sites where they can avoid certain issues. Example: For whatever reason, a developer might want to avoid the need to include a fish passage module.
- Issue attraction—A developer might have a reason to develop at sites with existing environmental issues. Example: A stream with low dissolved oxygen can present an opportunity to couple power generation with environmental improvement.
- Location centered—A developer might have a reason to develop within a specified region and want to know what opportunities and conditions are present.
- Similar development opportunity—A developer might have successfully developed a site and want to know where there are similar sites that could be approached in the same way.

Regulators

- Environmental issue identification—A regulator might want to quickly find information on other development within an area.
- Environmental conditions—A regulator might want to get a quick overview of the range of conditions and stream types relative to potential SMH development.

4.1 FUTURE GOALS

1. We will continue to collaborate with expert faculty and associates in hydraulic design and sedimentation in the University of Tennessee–Knoxville (UTK) Department of Civil and Environmental Engineering to develop a list of variables for sediment transport classification that are more mechanistically related to sediment transport dynamics.

2. We will showcase research outcomes through the public release of SMH Explorer, a web-based environmental design platform for SMH facilities. The database that underlies this tool will be geo-spatially located so that all data query results can be produced as data tables for qualifying reaches and their attributes, or as maps of qualifying stream reaches. This tool will allow users to

   a. drill down into individual clusters or combinations of clusters to better understand the application space (e.g., number of potential sites) of particular modular technologies

   b. filter reaches to identify reaches for possible new development that have similar characteristics to already developed or developing sites

3. We will begin selecting variables and performing cluster analyses that are more specifically aimed at understanding specific design requirements. For example, a fish passage module, data, and analysis will be developed that can be used to group potential sites (i.e., stream reaches) based on migratory species' needs for specific fish ladder velocities and gradients. Table 8 provides an example of how some variables are useful for informing module need, others for design requirements, and some for both need and design.

Table 8. Examples of classification data to be used to inform module need and module design requirements.

Classification variable	Module need	Design requirements
Downstream fish passage		
Anadromous/catadromous migratory species presence • Timing of migrations/movements • Size/life stage of migrants	X	X
Resident migratory species present • Timing of migrations/movements • Size/life stage of migrants	X	X
Stream gradient		X
Flow metrics		X
Upstream fish passage		
Migratory species presence • Spatial and temporal	X	
Resident species passage needs • Spatial and temporal	X	
Stream gradient		X
Species' swimming abilities • Prolonged velocity and duration • Burst velocity and duration		X
Species' jumping abilities		X
Flow metrics		X
Sediment transport		
Bedload characterization	X	
Bedload movement	X	
Stream gradient		X
Watershed contribution to sediment load • Soil type • Soil erodibility	X	
Watershed land use • Percent forested • Percent agricultural (row crops)	X	
Flow • Mean annual • Flood frequency	X	X
Recreational boating		
Existing use • Public access / launches	X	
Local recreational boating vendors	X	
Flow metrics • Seasonality (monthly quartiles)	X	X
Gradient		X
Local population density	X	
Water quality		
Existing water quality issues • 303d listings	X	
Water quality constituents • Nitrogen runoff • Phosphorus runoff • Turbidity	X	X
Watershed characteristics • Percent agricultural (row crops)	X	X

APPENDIX A. DATA VARIABLES AND SOURCES

Table A1. Descriptors of variables (and their sources) used in site classification clustering analysis.

Field alias	Field name	Description	Source	Data scale	Units
DB_ID	objectid	Unique database identifier			
COMID	comid	Unique stream reach identifier	NHDPlusV2	Reach	NA
Stream name	gnis_name	Stream name	NHDPlusV2	Reach	NA
State	state	State in which the reach lies	NHDPlusV2	State	NA
Mean annual flow	qa_ma	Mean annual flow from runoff	ORNL SMH	Reach	cfs
Generation cluster number	gen10clstr	Generation clusters	NHDPlusV2	Reach	NA
Generation cluster description	genclusterdef	Short text description of generation clusters	ORNL SMH	Reach	NA
Water quality cluster number	wq10clstrs	Water quality clusters	ORNL SMH	Reach	NA
Water quality cluster description	wqclusterdef	Short text description of water quality clusters	ORNL SMH	Reach	NA
Sediment cluster number	sed10clstr	Sediment clusters	ORNL SMH	Reach	NA
Sediment cluster description	sedclusterdef	Short text description of sediment clusters	ORNL SMH	Reach	NA
Fish passage cluster number	fpsg10clst	Fish passage clusters	ORNL SMH	Reach	NA
Fish passage cluster description	fshclusterdef	Short text description of fish passage clusters	ORNL SMH	Reach	NA
Foundation cluster number	fnd10clstrs	Foundation clusters	ORNL SMH	Reach	NA
Foundation cluster description	fndclusterdef	Short text description of foundation clusters	ORNL SMH	Reach	NA
Recreation cluster number	rec10clstrs	Recreation clusters	ORNL SMH	Reach	NA
Recreation cluster description	recclusterdef	Short text description of recreation clusters	ORNL SMH	Reach	NA

Field alias	Field name	Description	Source	Data scale	Units
Distance to nearest substation	dist2sub	Distance to nearest substation from reach midpoint	ORNL SMH	Reach	m
NPD count	npd_count	Number of NPDs in reach	ORNL NPD	Reach	Count
NPD MW	npd_mw	Total potential MW from NPDs in reach	ORNL NPD	Reach	MW
NSD count	nsd_count	Number of NSD sites in reach	ORNL NSD	Reach	Count
NSD MW	nsd_mw	Total potential MW from NSD sites in reach	ORNL NSD	Reach	MW
Support RPS	supportrps	Percent of residents within county that support renewable energy portfolio standards	Yale Climate Opinion Maps	County	%
NERC subregion	subregid	NERC subregion ID	EIA	NERC Subregion	NA
Subregion future population	popchngsub	Projected population increase by 2050 in NERC subregion	ORNL LandCast	NERC Subregion	Millions of individuals
Population density	popdns10cat	Population density from 2010 census in catchment	StreamCat	Catchment	People per square km
Fish group 1	grp1	Number of ocean-run sturgeon species (Acipenseriformes) within the reach's HUC8	NatureServe	HUC8	Count
Fish group 2	grp2	Number of inland sturgeon/paddlefish species (Acipenseriformes) within the reach's HUC8	NatureServe	HUC8	Count
Fish group 3	grp3	Number of ocean-run clupeid species within the reach's HUC8	NatureServe	HUC8	Count
Fish group 4	grp4	Number of ocean-run eel/lamprey species within the reach's HUC8	NatureServe	HUC8	Count
Fish group 5	grp5	Number of ocean-run salmonid species within the reach's HUC8	NatureServe	HUC8	Count
Fish group 6	grp6	Number of inland salmonid species within the reach's HUC8	NatureServe	HUC8	Count
Fish group 7	grp7	Number of other inland migratory species within the reach's HUC8	NatureServe	HUC8	Count
Fish passage mitigation	huc2prcntfp	Percent of mitigation sites in the mitigation database within the HUC2 that had Tier 1 fish passage mitigation required	ORNL Environmental Mitigation	HUC2	%
303d listed for temperature	d303_temp	Stream listed as impaired for temperature on EPA 303d list	US EPA	Reach	NA

Field alias	Field name	Description	Source	Data scale	Units
303d listed for sediment	d303_sdmnt	Stream listed as impaired for sediment on EPA 303d list	US EPA	Reach	NA
303d listed for pH	d303_ph	Stream listed as impaired for pH on EPA 303d list	US EPA	Reach	NA
303d listed for nutrients	d303_ntrnt	Stream listed as impaired for nutrients on EPA 303d list	US EPA	Reach	NA
303d listed for turbidity	d303_trbdt	Stream listed as impaired for turbidity on EPA 303d list	US EPA	Reach	NA
303d listed for dissolved oxygen	d303_do	Stream listed as impaired for dissolved oxygen on EPA 303d list	US EPA	Reach	NA
Upstream mainstem dams	umct	Upstream mainstem dam count	MSU Dam metrics DB	Mainstem	Count
Upstream dams	unct	Total upstream dam count	MSU Dam metrics DB	Watershed	Count
Downstream mainstem dams	dmct	Downstream mainstem dam count	MSU Dam metrics DB	Downstream flowpath	Count
Total mainstem dam count	tmct	Total mainstem dam count	MSU Dam metrics DB	Mainstem	Count
Distance to upstream mainstem dam	um2d	Distance to upstream mainstem dam	MSU Dam metrics DB	Mainstem	Count
Distance to downstream mainstem dam	dm2d	Distance to downstream mainstem dam	MSU Dam metrics DB	Mainstem	Count
Mainstem dist. between up/down dams	tm2d	Total mainstem distance between upstream and/or downstream mainstem dams	MSU Dam metrics DB	Mainstem	Count
Percentage discharge stored in reservoirs	udor	Percentage of estimated annual discharge stored in upstream reservoirs	MSU Dam metrics DB	Watershed	Count
Upstream dam density	damundr	Upstream network dam density per unit stream network length	MSU Dam metrics DB	Watershed	(#/100 km)
Downstream dam density	damdmd	Downstream mainstem dam density per unit downstream mainstem length	MSU Dam metrics DB	Downstream flowpath	(#/100 km)
Depth to bedrock	rckdepcat	Mean depth to bedrock in catchment	USGS WRD NSDI	Catchment	cm
Percent clay in soils	pctclayws	Percent clay content of soils in watershed	StreamCat	Watershed	%
Percent sand in soils	pctsandws	Percent sand content of soils in watershed	StreamCat	Watershed	%

Field alias	Field name	Description	Source	Data scale	Units
Ground acceleration	grndaccel	Earthquake susceptibility from national seismic hazard map	USGS seismic hazard maps	Reach	NA
K-factor in catchment	kffactcat	The Kffactor—relative index of susceptibility of bare, cultivated soil to particle detachment and transport by rainfall in catchment	StreamCat	Catchment	NA
K-factor in watershed	kffactws	The Kffactor—relative index of susceptibility of bare, cultivated soil to particle detachment and transport by rainfall in watershed	StreamCat	Watershed	NA
UCS low	ucsLow	Unconfined compressive strength of primary lithology—low end of range	UTK Hydraulics and Sedimentation Lab	Geologic unit	NA
UCS high	ucsHigh	Unconfined compressive strength of primary lithology—high end of range	UTK Hydraulics and Sedimentation Lab	Geologic unit	NA
Primary lithology	lithlgyprmry	Primary lithology	USGS geologic maps	Geologic unit	NA
Rock type	rocktype	Type of underlying bedrock	UTK Hydraulics and Sedimentation Lab	Geologic unit	NA
Stream order	streamorde	Strahler stream order	NHDPlusV2	Reach	NA
Elevation	minelevsmo	Elevation at downstream end of reach	NHDPlusV2	Reach	cm
Slope	slope	Slope of stream reach	NHDPlusV2	Reach	m/m
Stream power	powerQS	SLOPE * QA_MA	ORNL SMH	Reach	NA
Mean annual velocity	va_ma	Mean annual velocity for QA	NHDPlusV2	Reach	fps
Infiltration-excess overland flow	ieofcat	Mean infiltration-excess overland flow in catchment	USGS WRD NSDI	Catchment	% total stormflow
Runoff in watershed	runoffws	Mean runoff in watershed	StreamCat	Watershed	mm
Flow variation	qa_cv	Coefficient of variation for flow based on monthly averages and annual mean	NHDPlusV2/ORNL	Reach	NA
Elevation change in watershed	elevdiffws	Difference between maximum and minimum elevation in watershed	NHDPlusV2/ORNL	Watershed	cm
Base flow index	bficat	Base flow index is the ratio of base flow to total flow, expressed as a percentage	USGS WRD NSDI	Catchment	%
Agriculture in catchment	pctagcat	Percent agricultural land cover in catchment	StreamCat	Catchment	%
Agriculture in watershed	pctagws	Percent agricultural land cover in watershed	StreamCat	Watershed	%

Field alias	Field name	Description	Source	Data scale	Units
Nitrogen from farms	farmncat	Sum total of nitrogen from farm areas in catchment	USGS WRD NSDI	Catchment	%
Forest/wetlands in catchment	pctforwetcat	Percent forest and wetland land cover in watershed	NLCD 2011	Catchment	%
Forest/wetlands in watershed	pctforwetws	Percent forest and wetland land cover in watershed	NLCD 2011	Watershed	%
Imperviousness in catchment	pctimprv06cat	Percent imperviousness from 2006 in catchment	StreamCat	Catchment	%
Imperviousness in watershed	pctimprv06ws	Percent imperviousness from 2006 in watershed	StreamCat	Watershed	%
Riparian forest in watershed	pctforripws	Percent riparian forest land cover in watershed	StreamCat	Watershed	%
HUC6	huc6	6-digit HUC watershed	USGS WBD	HUC6	NA
Population density in HUC6	popdns10huc6	Mean population density in HUC6	StreamCat/ORNL	HUC6	People per sq km
Boat ramps in HUC6	boat_ramp	Number of developed boat ramps in HUC6	Delorme/ORNL	HUC6	Count
Undeveloped boat ramps in HUC6	boat_ramp_undeveloped	Number of undeveloped boat ramps in HUC6	Delorme/ORNL	HUC6	Count
Fishing cold water	fishing_coldwater	Number of cold-water fishing locations in HUC6	Delorme/ORNL	HUC6	Count
Fishing saltwater	fishing_saltwater	Saltwater fishing locations in HUC6	Delorme/ORNL	HUC6	Count
Whitewater paddling in HUC6	awhuc6km	Length of stream identified as whitewater paddling runs in HUC6	American Whitewater/ORNL	HUC6	m
Outstanding rivers in HUC6	nrirechuc6km	Length of stream identified as having outstanding recreational value in HUC6	National Rivers Inventory	HUC6	m
Reach length	st_length(shape)	Reach length	NHDPlusV2		m

Table A2. Data sources with links.

Source	URL
NHDPlusV2	http://www.horizon-systems.com/nhdplus/NHDPlusV2_home.php
ORNL SMH	https://hydropower.ornl.gov/smh/
ORNL NPD	https://nhaap.ornl.gov/content/non-powered-dam-potential
ORNL NSD	https://nhaap.ornl.gov/nsd
Yale Climate Opinion Maps	http://climatecommunication.yale.edu/visualizations-data/ycom-us-2016/
EIA	https://www.eia.gov/maps/layer_info-m.php
ORNL LandCast	http://www.pnas.org/content/112/5/1344
StreamCat	https://www.epa.gov/national-aquatic-resource-surveys/streamcat
NatureServe	http://www.natureserve.org/conservation-tools/data-maps-tools/digital-distribution-native-us-fishes-watershed
ORNL Environmental Mitigation	https://nhaap.ornl.gov/environmental-mitigation
US EPA	https://www.epa.gov/waterdata/waters-geospatial-data-downloads
MSU Dam metrics DB	https://www.sciencedirect.com/science/article/pii/S004896971730308X?via%3Dihub
USGS WRD NSDI	https://water.usgs.gov/lookup/getgislist
USGS seismic hazard maps	https://earthquake.usgs.gov/hazards/hazmaps/
UTK Hydraulics and Sedimentation Lab	http://hsl.engr.utk.edu/
USGS geologic maps	https://mrdata.usgs.gov/geology/state/
NLCD 2011	https://www.mrlc.gov/nlcd2011.php
USGS WBD	https://nhd.usgs.gov/wbd.html
Delorme/ORNL	https://developer.garmin.com/datasets/overview
American Whitewater/ORNL	https://www.americanwhitewater.org/
National Rivers Inventory	https://www.nps.gov/subjects/rivers/data.htm