DOI: 10.1007/s00267-003-5077-9

Delineation and Evaluation of Hydrologic-Landscape Regions in the United States Using Geographic Information System Tools and Multivariate Statistical Analyses

DAVID M. WOLOCK*
U.S. Geological Survey
4821 Quail Crest Place
Lawrence, Kansas 66049 USA

THOMAS C. WINTER
U.S. Geological SurveyMS 413
Denver Federal Center
Lakewood, Colorado 80225 USA

GERARD MCMAHON
U.S. Geological Survey
3916 Sunset Ridge Road
Raleigh, North Carolina 27607 USA

ABSTRACT / Hydrologic-landscape regions in the United States were delineated by using geographic information system (GIS) tools combined with principal components and cluster analyses. The GIS and statistical analyses were applied to land-surface form, geologic texture (permeability of the soil and bedrock), and climate variables that describe the physical and climatic setting of 43,931 small (approximately 200 km^2) watersheds in the United States. (The term "watersheds" is defined in this paper as the drainage areas of tributary streams, headwater streams, and stream segments lying between two confluences.) The analyses grouped the watersheds into 20 noncontiguous regions based on similarities in land-surface form, geologic texture, and climate characteristics. The percentage of explained variance (R-squared value) in an analysis of variance was used to compare the hydrologic-landscape regions to 19 square geometric regions and the 21 U.S. Environmental Protection Agency level-II ecoregions. Hydrologic-landscape regions generally were better than ecoregions at delineating regions of distinct land-surface form and geologic texture. Hydrologic-landscape regions and ecoregions were equally effective at defining regions in terms of climate, land cover, and water-quality characteristics. For about half of the landscape, climate, and water-quality characteristics, the R-squared values of square geometric regions were as high as hydrologic-landscape regions or ecoregions.

The U.S. Geological Survey (USGS) initiated the National Water-Quality Assessment (NAWQA) Program in 1991 to accomplish two primary objectives. The first objective was to assess the status and trends in the quality of the nation's ground-water and surface-water resources, and the second objective was to link the status and trends with an understanding of the natural and human factors that affect water quality (Gilliom and others 1995).

The original NAWQA implementation plan was based on assessing water quality in 60 study areas located throughout the United States. The extensive geographic coverage of the 60 study areas was considered adequate to sample the diversity of hydrologic settings in the nation. The characteristics of hydrologic-cycle components, such as precipitation, evapotranspiration, infiltration, ground-water flow, and overland flow, were used to define the hydrologic setting of each study area. Adequate coverage of the diversity in hydrologic settings is crucial in understanding how natural and human factors affect water quality.

The NAWQA Program design was refined in 2001 for work to be completed in the period 2002–2011. One significant constraint on refining the design of the program was budget; the funding outlook for 2002–2011 was expected to support only about 40 of the original 60 study areas. The need to maintain a wide and representative sample of the hydrologic settings in the United States remained imperative. The challenge, then, was to carefully select study areas and specific data-collection sites that represent the range of hydrologic settings in the United States.

Selecting a group of study areas to represent the range of hydrologic settings in the United States re-

KEY WORDS: Hydrologic landscapes; Hydrologic regions; Cluster analysis; Network design

Published online August 24, 2004.

*Author to whom correspondence should be addressed, *email:* dwolock@usgs.gov

quired a digital map of the factors expected to affect these hydrologic settings. The ideal map needed was one that divided the United States into regions, where areas in the same region have similar hydrologic conditions and areas in different regions have different hydrologic conditions. A region would not necessarily be spatially contiguous; areas similar to each other can be located in different parts of the United States. A grouping of areas into regions on the basis of similarities in landscape and climate characteristics is referred to in this paper as a "regional framework."

Regional frameworks related to hydrologic characteristics have been used in many water-resource and water-quality applications. Most significantly, ecoregions (Omernik 1987, U.S. Environmental Protection Agency 2002) have been used to define areas where conditions affecting stream chemistry are expected to be similar (Omernik and Griffith 1991). Ecoregions, however, were not specifically designed to identify areas throughout the United States with similar hydrologic characteristics. Rather, ecoregions were developed to identify patterns in biotic and abiotic factors thought to generally influence ecological processes at a relatively broad scale.

For the purpose of selecting areas to be studied in the second decade of NAWQA, a map of hydrologic-landscape regions (HLRs) that was based on the hydrologic-landscapes concept of Winter (2001) was produced. The hydrologic-landscapes concept reflects fundamental hydrologic processes that are expected to affect water quality and other environmental characteristics. A map of HLRs was generated using readily available spatial data layers, Geographic Information System (GIS) tools, and statistical analyses to produce a consistent and reproducible hydrologic characterization of the United States. The methods used to delineate HLRs are similar to procedures reported in Hargrove and Hoffman (1999), Hargrove and others (2003), and Preston (2000).

The purpose of this paper is to present the process of delineating and evaluating HLRs. The specific objectives are

- to describe how the HLRs were derived;
- to quantify the statistical power, as measured by the R-squared value in an analysis of variance, of the HLRs to delineate similar areas throughout the United States; and
- to compare the statistical power of different regional frameworks. The statistical power of HLRs is compared to that of level-II ecoregions, a widely used regional framework, and to that of square geometric regions, a regional framework with no physical meaning. The square geometric regions serve as a baseline with which to compare the HLRs and ecoregions.

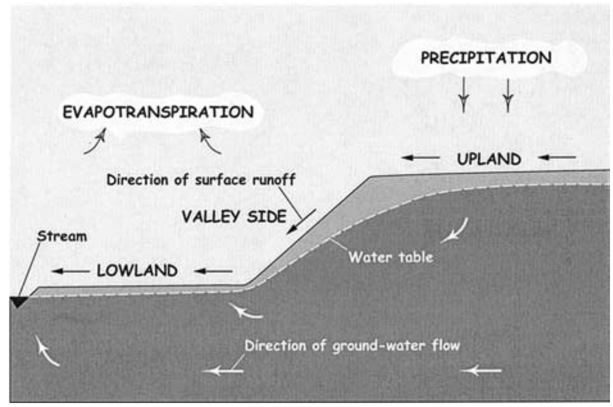

Figure 1. Conceptual diagram of a fundamental hydrologic landscape unit.

The HLRs presented in this paper were derived specifically to select study areas for the NAWQA Program. If the application were for a different purpose, covered a different spatial extent, or was based on a different dataset, then the final product (that is, the regions) would be different. An important goal of this paper is to present a regional delineation approach that can be customized for different applications and to illustrate an evaluation approach that can be used for any regional framework.

The Hydrologic-Landscapes Concept

The characteristics of the Earth that affect the location, movement, and chemistry of water are extremely complex. With respect to the movement of water, however, many seemingly diverse landscapes have some features in common. It is these commonalties that need to be identified in developing a conceptual framework. Only by evaluating landscapes from a common conceptual hydrologic framework can processes common to some or all landscapes be distinguished from processes unique to particular landscapes.

The concept of hydrologic landscapes is based on the idea that a single, simple physical feature is the basic building block of all landscapes. This feature is termed a fundamental hydrologic landscape unit and is defined as upland adjacent to lowland separated by a valley side (Figure 1). The hydrologic system of a fundamental hydrologic landscape unit consists of (1) the movement of surface water, which is controlled by the slopes and permeability of the landscape; (2) the movement of ground water, which is controlled by the hydraulic characteristics of the geologic framework; and (3) atmospheric-water exchange, which is controlled by climate (Winter 2001).

All hydrologic landscapes can be conceived of as variations and multiples of fundamental hydrologic landscape units, and these can be used to define general landscape types that describe major physiographic features of the Earth. Some examples are (1) a landscape consisting of narrow lowlands and uplands separated by high and steep valley sides, characteristic of mountainous terrain; (2) a landscape with wide lowlands separated from much narrower uplands by steep valley sides, characteristic of basin-and-range physiography and basins of interior drainages; and (3) a landscape having narrow lowlands separated from very broad uplands by valley sides of various slopes and heights, characteristic of plateaus and high plains.

The movement of water over the surface and through the subsurface of generalized landscapes is controlled by common physical principles regardless of the geographic location of the landscapes. For example, in a landscape with low permeability soils, surface runoff will be extensive, and recharge to ground water will be limited. In contrast, in a landscape with highly permeable soils, surface runoff will be limited, and ground-water recharge will be significant. In landscapes that have a shallow water table, transpiration directly from ground water may have a substantial effect on the volume of ground water in storage and on the movement of ground water to and from surface water.

It is important to recognize that the hydrologic-landscapes concept considers only certain land-surface form, geologic, and climate characteristics of the landscape. Other factors, such as land use and land cover, will affect hydrologic processes in important ways but were not used to define regions in this paper. The hydrologic-landscape regions can be viewed as a starting point, based on a simple set of landscape factors, which can be used to help understand hydrologic diversity.

Delineation of Hydrologic-Landscape Regions

The overall approach for delineation of HLRs for the United States involved three steps:

- Delineate a set of watersheds that covers all 50 States. In this paper, the term "watersheds" denotes the drainage areas of tributary streams, headwater streams, and stream segments lying between two confluences.
- Determine metrics to quantify the land-surface form, geologic texture (herein defined as the permeability of the soil and bedrock), and climate characteristics that define hydrologic landscapes, and then average the metrics for each watershed.
- Use principal components and cluster analyses to assign the watersheds to groups according to their similarity in land-surface form, geologic texture, and climate characteristics. Each group of similar watersheds comprises an HLR.

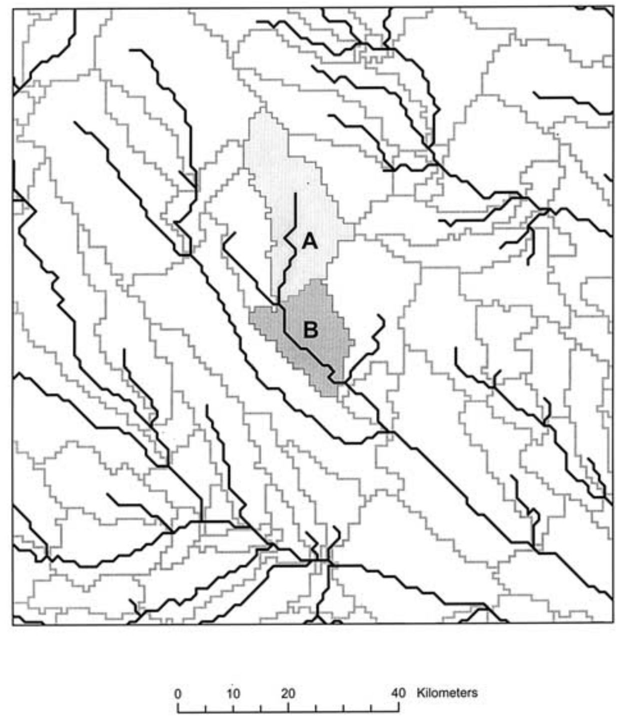

Figure 2. Synthetic stream network and watersheds for a section of central Iowa.

ARC/INFO GIS was used in the first two steps just listed. (The use of brand, trade, or firm names in this paper is for identification purposes only and does not constitute endorsement by the U.S. Geological Survey.) All the GIS data were analyzed at 1-km resolution in the Lambert Azimuthal Equal Area projection system (central meridian, –100°; latitude of origin, 45°; sphere of influence radius, 6,370,997 m).

Delineation of Watersheds

A set of small watersheds (each about 200 km^2 in area) completely covering the 50 States was derived from 1-km-resolution digital elevation model (DEM) data (Verdin and Greenlee 1996). This watershed size was selected because it is compatible with the spatial resolution of nationally available datasets on land-surface form, geologic texture, and climate characteristics

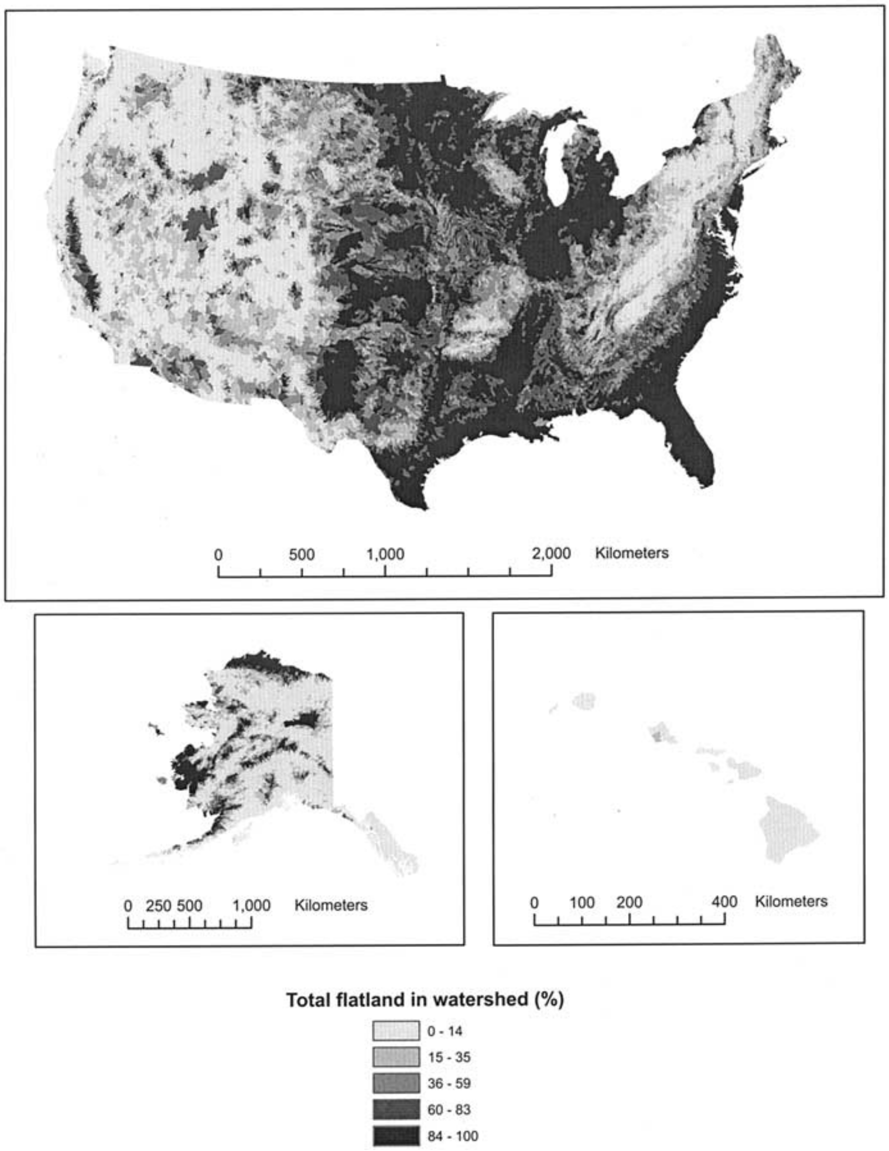

Figure 3. Total percentage of flatland in small watersheds for the United States.

(see the next section). This watershed size is large enough to include lowlands and uplands, but small enough to provide spatial detail among watersheds.

First, a synthetic stream network was extracted from the DEM by specifying a minimum drainage area of 100 km^2 to initiate a stream channel. Each stream segment in the network was assigned a unique identifier value such that all grid cells in the same segment had the same unique identifier value. The entire set of stream segments then was used to delineate the watersheds from the DEM. The result was a set of 43,931 watersheds with an average area of 212 km^2.

The synthetic stream network and delineated watersheds are illustrated in Figure 2, which shows a $100 \times 100\text{-km}^2$ area in central Iowa. The stream segment labeled A is an example of a headwater stream. It contains multiple grid cells that all have the same value; the watershed draining to stream segment A is shaded light gray. The stream segment labeled B is an example of a stream that lies between two confluences; its watershed is a darker shade of gray.

Quantification of the Land-Surface Form, Geologic Texture, and Climate Characteristics

The critical features used to describe Winter's hydrologic landscapes are (1) land-surface form, to quantify the effects of gravity on the movement of water through the landscape; (2) geologic texture, to estimate permeability of soil and bedrock materials that affect surface runoff, infiltration, and ground-water

Table 1. Lithologic groups of principal aquifers and bedrock permeability classes: 1 is the lowest permeability and 7 is the highest permeability

Lithologic group	Bedrock permeability class)
Not a principal aquifer	1
Sandstone	2
Semiconsolidated sand	3
Basalt and other volcanic rocks	4
Sandstone and carbonate rocks	5
Unconsolidated sand and gravel	6
Carbonate rock	7

flow; and (3) climate characteristics, to approximate water available to surface- and ground-water systems.

Land-surface form is described by terrain characteristics that are computed from 1-km resolution DEM data for the small watersheds. The characteristics are as follows: relief (maximum elevation minus minimum elevation in the watershed); total percentage of flatland (areas with less than 1% slope) in the watershed; the percentage of flatland located in upland areas of the watershed (flat areas with elevation greater than the midpoint elevation); and the percentage of flatland in lowland areas of the watershed (flat areas with elevation less than or equal to the midpoint elevation). Figure 3, for example, shows a map of total percentage of flatland in each watershed. These terrain characteristics

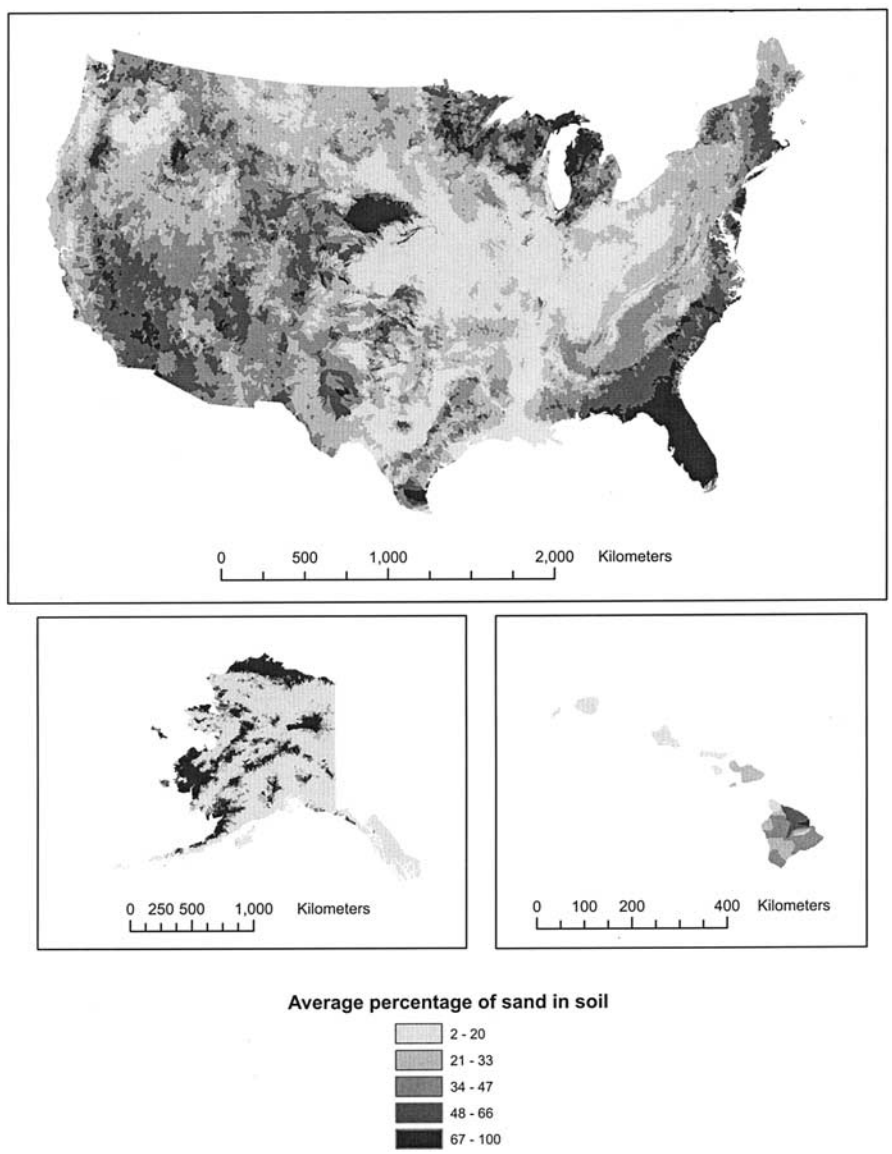

Figure 4. Average percentage of sand in the soil for small watersheds in the United States.

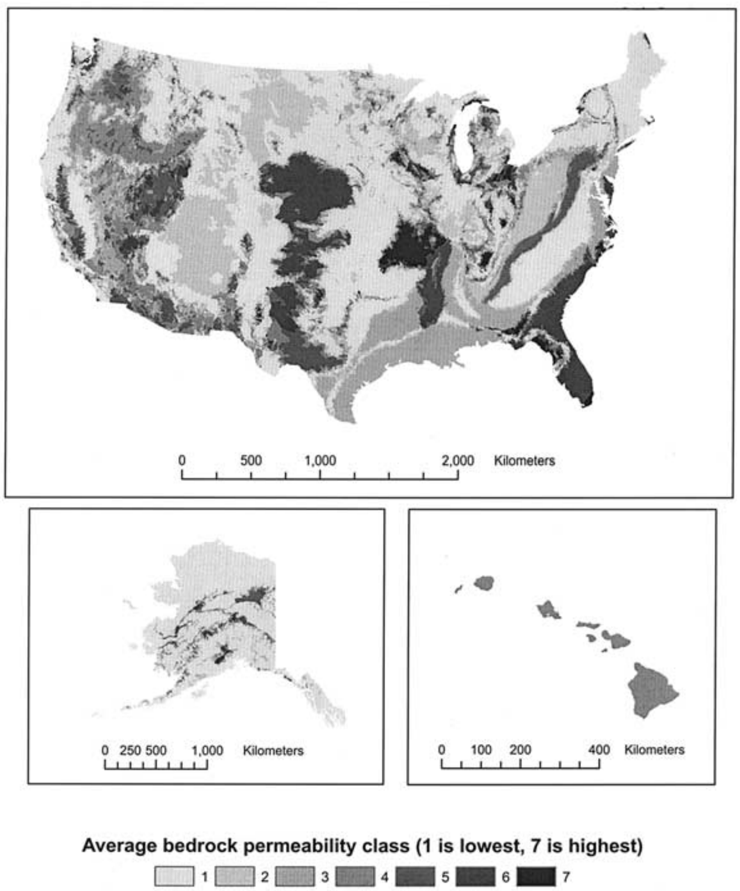

Figure 5. Average bedrock permeability class (1 is lowest, 7 is highest) for small watersheds in the United States.

are similar to those in Hammond (1964) and presented in the USGS National Atlas of the United States "Classes of Land-Surface Form" map (U.S. Geological Survey 2001).

Geologic texture is described by two measures of permeability—one for soil and one for bedrock material. Soil permeability is estimated as the percentage of sand in the soil given in the U.S. Department of Agriculture's STATSGO database (U.S. Department of Agriculture 1994). The percentage of sand in the soil is directly related to permeability; that is, permeability increases as percentage of sand in the soil increases.

Bedrock permeability is quantified by assigning permeability classes to the general lithologic groups (Table 1) in the USGS National Atlas map of principal aquifer groups (U.S. Geological Survey 2001). These bedrock texture divisions are consistent with the primary lithologies used to develop regions for the USGS Ground-Water Resources Program (U.S. Geological Survey 1998). Some of the bedrock texture groups have permeability values ranging over several orders of magnitude, such as basalt and volcanic rocks, but "typical" permeability is the basis for the relative permeability class.

The digital maps representing geologic texture groups were intersected with the delineated watersheds, and average percentage of sand in the soil and average bedrock permeability class values were computed for each of the 43,931 watersheds (Figures 4 and 5).

Climate characteristics are described by mean annual precipitation minus potential evapotranspiration (PET). PET was estimated from mean monthly temperature and latitude using the Hamon equation (Hamon 1961). Mean monthly temperature and precipitation data from first-order meteorological station data for 1961–1990 (Owensby and Ezell 1992) were used to compute the mean annual precipitation and PET estimates. The station data were interpolated to a 1-km-resolution grid using the ARC/INFO inverse-distance weighting method. The mean value of precipitation–PET (Figure 6) was computed for each of the 43,931 watersheds by intersecting the digital precipitation–PET data with the delineated watersheds.

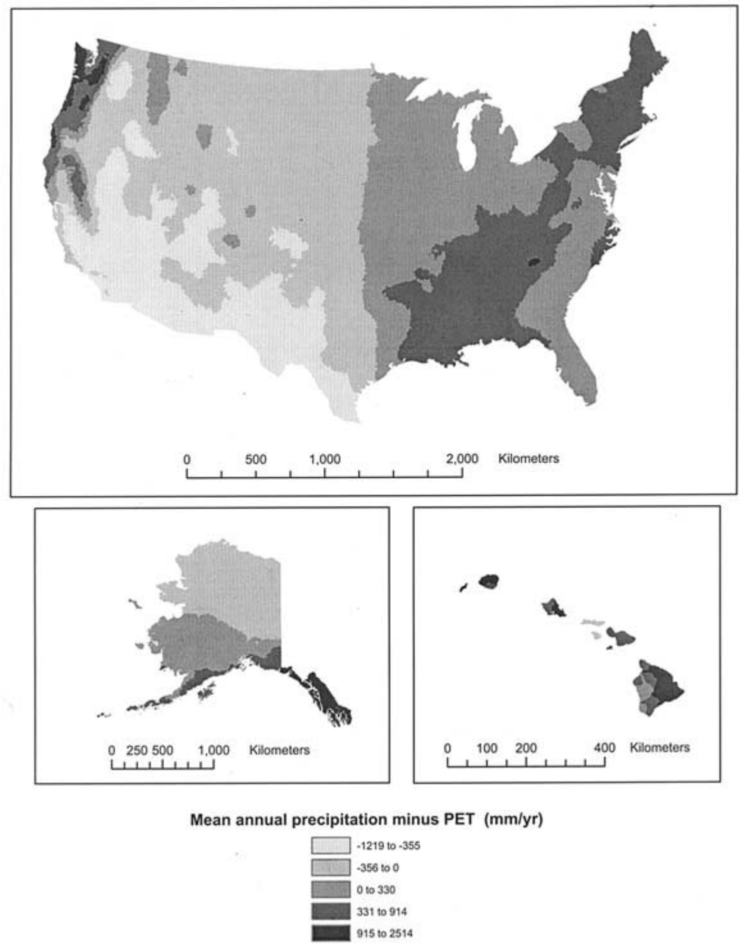

Figure 6. Mean annual precipitation–PET for small watersheds in the United States. PET = potential evapotranspiration.

Table 2. Variance explained by the principal components[a]

	Principal component number						
	1	2	3	4	5	6	7
Eigenvalue	2.93	1.16	0.98	0.77	0.74	0.38	0.00
Proportion of variance explained	0.41	0.16	0.14	0.11	0.10	0.05	0.00
Cumulative variance explained	0.41	0.58	0.72	0.83	0.94	1.00	0.00

[a]Eigenvalue, proportion of variance explained, and cumulative variance explained for principal components analysis of land-surface form, geologic texture, and climate characteristics for all watersheds.

Identification of Hydrologic-Landscape Regions by Grouping the Watersheds

The land-surface form, geologic texture, and climate characteristics were used to assign the watersheds to similar groups that define the HLRs for the 50 States. First, a principal components analysis (PCA), based on the correlation matrix, was used to remove co-dependence among the land-surface form, geologic texture, and climate characteristics for the 43,931 watersheds. Five principal components (PCs) were retained for further analysis using the criterion that a component must explain at least 10% of the total variance in the data (Table 2). The loading values (Table 3) show the relations between the watershed characteristics and the principal components. PC 1 groups the land-surface form characteristics; PC 2 groups the geologic texture characteristics; PC 3 mostly reflects precipitation–PET and percent sand; PC 4 is dominated by bedrock per-

Table 3. Principal component loadings[a]

Variable	Principal component number						
	1	2	3	4	5	6	7
Bedrock permeability	0.35	−0.63	0.07	0.69	0.02	0.00	0.00
Percent sand	−0.13	−0.71	−0.56	−0.38	0.09	−0.15	0.00
Relief	−0.82	−0.19	−0.10	0.01	0.02	0.53	0.00
Total percent flatland	0.97	−0.04	−0.04	−0.15	−0.02	0.19	0.00
Percentage flatland in upland	0.72	0.05	0.01	−0.08	0.67	0.14	0.00
Percentage flatland in lowland	0.81	−0.10	−0.05	−0.15	−0.54	0.16	0.00
Precipitation–PET	0.13	0.46	−0.81	0.33	−0.02	0.02	0.00

[a]Principal component loadings from principal components analysis of land-surface form, geologic texture, and climate characteristics for all watersheds.
PET = potential evapotranspiration.

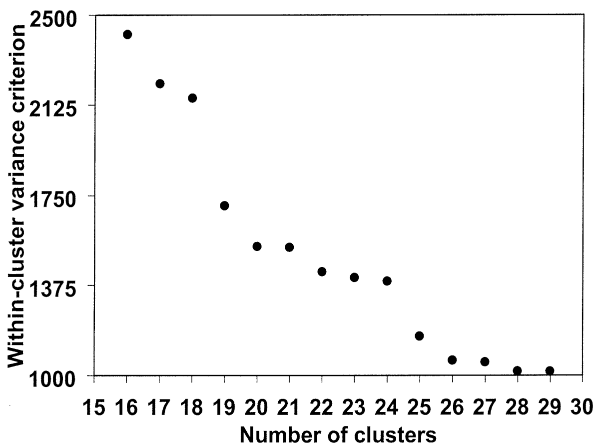

Figure 7. Relation of the within-cluster variance criterion to the number of clusters.

meability; and PC 5 shows whether the flatland is in upland or lowland areas.

A cluster analysis of the scores of the five retained principal components was used to assign each watershed to one of 20 groups (HLRs). The clustering program uses a minimum variance criterion and the nearest neighbor chain algorithm (Murtagh 1985). Several points of relatively abrupt change in the within-cluster variance criterion are apparent in a graph of the within-cluster variance criterion versus the number of clusters (Figure 7); 20 groups were selected as a compromise between low within-cluster variance and the number of regions considered suitable for NAWQA Program design needs.

The GIS and statistical analyses resulted in a map of 20 HLRs (Figure 8). The mean principal-component scores for each of the 20 HLRs are given in Table 4. The original landscape and climate characteristics (before the PCA) and associated HLR numbers are summarized in Tables 5 to 7. Table 5 lists the mean values of the land-surface form, geologic texture, and climate characteristics for all the watersheds in each HLR; for instance, the mean percentage of total flatland for all the watersheds grouped in HLR 1 was 92.5%.

Table 6 shows the relative differences among the HLRs for each of the land-surface form, geologic texture, and climate characteristics. The values in Table 6 were computed by normalizing (scaling) the mean values in each column in Table 5 to the minimum and maximum mean values in the column. In effect, the range of values in a column was rescaled from the original minimum and maximum (Table 5) to vary from 1 to 20 (Table 6). This is similar to ranking but causes very similar values to have the same normalized value. For example, HLR 1 and HLR 6 have nearly identical (and very high) values for the total percentage of flatland (92.5% and 90.6%, respectively) and, therefore, are both assigned a normalized mean value of 20.

Table 7 gives the standard deviations of the land-surface form, geologic texture, and climate characteristics for all the watersheds in each region. Watersheds in HLR 9, for example, have the most variable percentage of total flatland.

The identification numbers and associated colors assigned to the HLRs were chosen according to the overall similarity in hydrologic-landscape characteristics (land-surface form, geologic texture, and climate characteristics) among regions. For example, HLR 1 and HLR 2 are similar in their average hydrologic-landscape characteristics except for bedrock permeability class (Tables 5 and 6). In contrast, HLR 1 and HLR

Figure 8. Hydrologic-landscape regions in the United States.

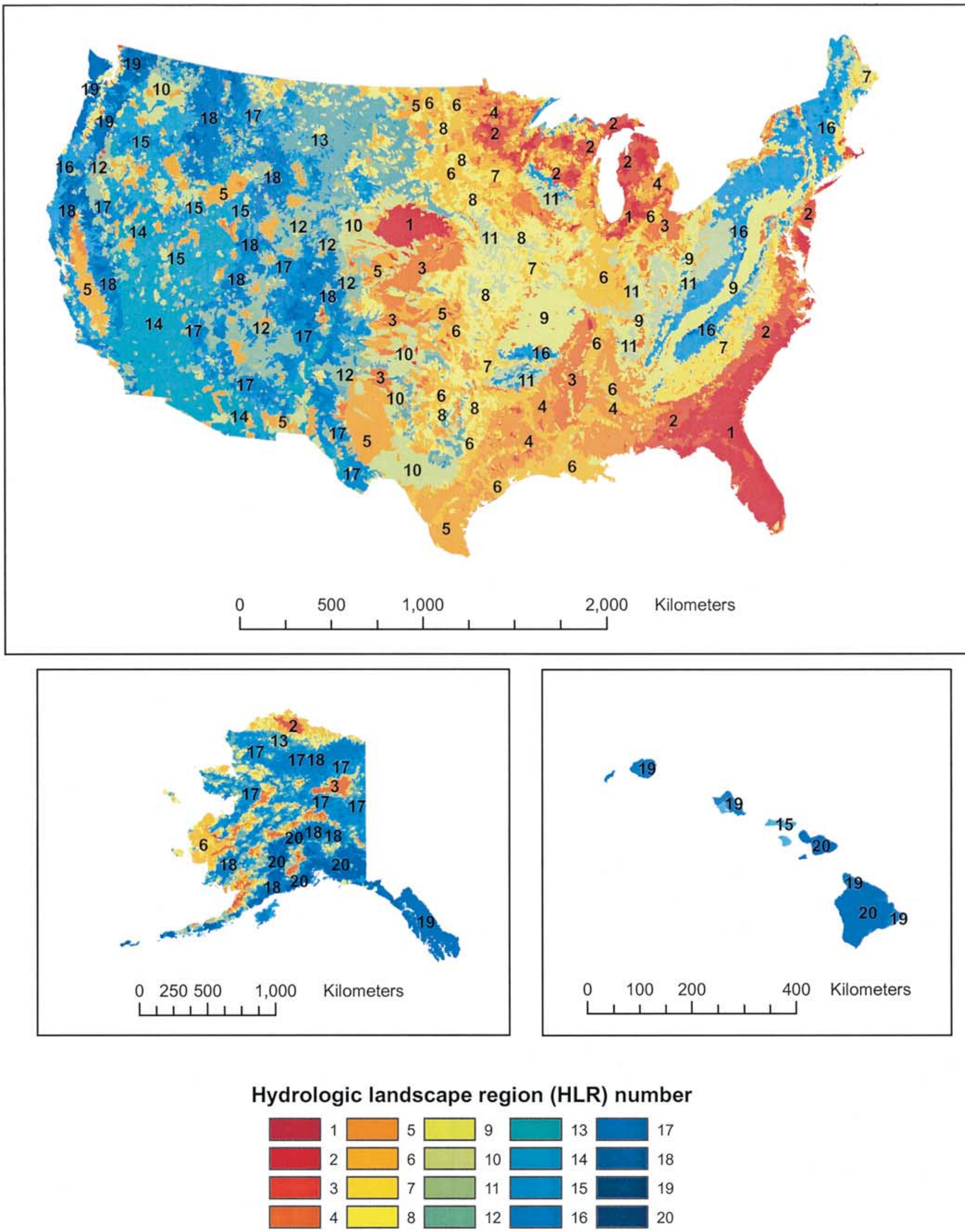
0 500 1,000 2,000 Kilometers
0 250 500 1,000 Kilometers
0 100 200 400 Kilometers
Hydrologic landscape region (HLR) number
1 5 9 13 17
2 6 10 14 18
3 7 11 15 19
4 8 12 16 20

Table 4. Mean principal component scores for each of the 20 hydrologic-landscape regions (HLRs)

HLR number	Principal component number				
	1	2	3	4	5
1	2.3	-2.4	-1.5	0.1	0.4
2	1.6	-1.0	–1.4	–0.9	0.0
3	2.2	–0.6	0.4	1.4	–0.3
4	1.7	0.2	–0.7	–0.2	0.0
5	1.4	–1.3	0.7	–0.3	–0.4
6	1.9	0.8	0.3	–0.4	–0.8
7	0.9	0.8	–0.6	–0.6	0.3
8	1.1	0.9	0.8	–0.6	0.2
9	0.5	0.1	–0.2	1.7	0.1
10	0.1	–1.0	1.1	0.8	0.3
11	0.2	1.5	0.1	0.3	0.2
12	–1.2	–0.8	0.1	–0.6	0.0
13	–0.4	0.5	0.9	–0.3	0.1
14	–1.8	–2.1	0.9	–0.1	0.0
15	–1.4	–0.7	1.0	1.0	0.0
16	–1.2	0.9	–1.0	0.3	0.0
17	–1.7	0.2	0.6	–0.3	0.0
18	–2.4	–0.1	–0.2	–0.4	0.2
19	–1.9	0.9	–2.7	1.2	0.0
20	–3.1	–0.1	–1.1	0.0	0.2

20 are very different in all characteristics except precipitation–PET.

An important attribute of HLRs is that the same region can occur in different parts of the United States. A notable example of this feature is HLR 1, which is located in the Southeast, the western Great Lakes area, and the Sand Hills of Nebraska (Figure 8). This indicates that the hydrologic-landscape characteristics of watersheds in these geographically disparate sections of the United States are similar. A visual inspection of the individual hydrologic-landscape characteristics (Figures 3 to 6) confirms that these areas of the United States have a flat land-surface form and highly permeable soils and bedrock. The Sand Hills of Nebraska has lower precipitation–PET than the Southeast and western Great Lakes area. However, the overall similarity in the three regions is greater than their differences.

Hydrologic System Hypotheses

The particular combination of land-surface form, geologic texture, and climate characteristics can be used to develop hypotheses on how the hydrologic system might function for a given hydrologic landscape. Table 8 provides a short qualitative description of each HLR and some hydrologic hypotheses that are based on the HLR characteristics. The qualitative descriptions summarize the mean hydrologic-landscape characteristics (Tables 5 and 6). Note that to simplify the HLR descriptions, low permeability soils are called impermeable soils, low permeability bedrock is called impermeable bedrock, and shallow subsurface flow paths are called shallow ground-water flow paths.

Several examples of how differences in the hydrologic-landscape characteristics are expected to result in different hydrologic systems are given in the following paragraphs.

- HLR 1 has very flat terrain indicative of a "plains" land-surface form, very permeable soils and bedrock, and a surplus of precipitation over PET typical of a subhumid climate. Given these geologic characteristics, precipitation is expected to infiltrate through the soil and recharge the ground-water system. Both shallow and deep ground-water flow are expected to be important hydrologic flow paths because of the permeable soils and bedrock.
- HLR 6 has a "plains" land-surface form impermeable soils and bedrock, and a subhumid climate. Infiltration of precipitation through the soil is expected to be minimal because of impermeable soils. Overland flow is the primary hydrologic flow path also because of the impermeable soils, and ground-water recharge likely will be limited.
- HLR 20 has a "mountains" land-surface form, permeable soils, impermeable bedrock, and a humid climate. Infiltration of precipitation through the soil will be high, but the impermeable bedrock will limit deep ground-water flow. The steep terrain and shallow recharge will cause shallow ground water to be the primary hydrologic flow path.

Inferences can be made about water-quality issues from the hydrologic system hypotheses (Table 8). For example, HLRs with substantial recharge to ground water are areas where transport of contaminants from

Table 5. Mean values of watershed characteristics[a]

HLR number	Total flatland (%)	Flatland in upland (%)	Flatland in lowland (%)	Relief (m)	Bedrock permeability class	Sand (%)	Precipitation PET (mm/yr)
	Mean values						
		Land-surface form			Geologic texture		Climate
1	92.5	36.5	56.0	50	5.8	69.7	213
2	82.8	28.2	54.6	73	2.7	61.9	264
3	86.0	28.3	57.8	98	5.8	20.1	119
4	82.7	29.8	52.9	73	2.6	36.0	323
5	78.0	20.2	57.8	233	4.0	39.0	–338
6	90.6	22.7	67.9	68	1.8	18.8	119
7	64.0	27.5	36.5	110	1.2	33.5	320
8	71.2	30.6	40.6	132	1.3	17.9	–58
9	41.7	17.4	24.3	213	4.9	22.4	394
10	37.5	16.5	21.1	290	4.9	30.1	–305
11	41.9	19.4	22.5	130	1.6	15.8	333
12	22.3	4.1	18.2	641	2.1	46.3	–191
13	30.3	11.4	18.9	257	1.5	24.0	–180
14	16.0	0.4	15.6	1225	4.1	47.9	–582
15	8.2	1.3	7.0	769	4.2	26.4	–249
16	10.8	3.3	7.5	452	1.5	33.6	505
17	7.6	1.2	6.4	665	1.5	29.6	–173
18	2.0	0.2	1.8	1174	1.2	40.7	–8
19	4.8	0.3	4.4	1129	2.2	39.5	1156
20	1.7	0.1	1.7	1966	1.4	41.7	323

[a]Mean values of land-surface form, geologic texture, and climate characteristics for all watersheds in each hydrologic-landscape region (HLR). m = meters; PET = potential evapotranspiration.

the land surface to ground water would be expected. In contrast, HLRs expected to generate overland flow could readily transport contaminants from the land surface to streams.

Evaluation of the Statistical Power of Hydrologic-Landscape Regions and Comparison with Other Regional Frameworks

The statistical power, as measured by the R-squared value in an analysis of variance (ANOVA), of HLRs to separate the conterminous United States into distinct areas of land-surface form, geologic texture, climate, land cover, and water-quality characteristics was quantified. The statistical power of HLRs was compared to that of two other regional frameworks (Figure 9): ecoregions, a regional framework widely used in water-resources management; and square geometric regions, a regional framework with no physical meaning. The square geometric regions serve as a baseline with which to compare the HLRs and ecoregions. (Square geometric regions are described below in more detail.)

In this analysis, data from 493 NAWQA water-quality sampling sites (Figure 10) and the associated drainage basins of those sites were used. The GIS steps in the evaluation used 1-km resolution grids in the Albers Equal Area projection (central meridian, –96°; standard parallels, 29.5° and 45.5°; latitude of origin, 23°; datum, North American Datum of 1983).

The land-surface form characteristics used to define HLRs are highly correlated with land-surface slope. Therefore, the NAWQA sampling-site drainage basins were characterized by land-surface slope instead of percentage of flatland and relief. The geologic texture variables used to test the regional frameworks were bedrock permeability class and percentage of sand, and the climate characteristic was precipitation–PET.

Land-cover characteristics used to test the regional frameworks were taken from USGS GIRAS (Geographic Information Retrieval and Analysis System) data (U.S. Geological Survey 1990). The land-cover categories were forest, rangeland, urban land, and agricultural land. Kuchler's potential natural vegetation (Kuchler 1964) map also was used as a metric of land-cover characteristics.

Table 6. Normalized mean values of watershed characteristics[a]

	Normalized mean values						
		Land-surface form			Geologic texture		Climate
HLR numbe	Total flatland	Flatland in upland	Flatland in lowland	Relief	Bedrock permeability class	Sand	Precipitation–PET
1	20	20	17	1	20	20	10
2	18	16	16	1	7	17	10
3	19	16	17	1	20	3	9
4	18	17	16	1	7	8	11
5	17	11	17	3	13	9	4
6	20	13	20	1	4	2	9
7	14	15	11	2	1	7	11
8	16	17	12	2	1	2	7
9	9	10	7	3	16	3	12
10	8	10	7	3	16	6	4
11	9	11	7	2	3	1	11
12	5	3	6	7	5	12	5
13	7	7	6	3	2	4	5
14	4	1	5	13	13	12	1
15	2	2	3	8	13	5	5
16	3	3	3	5	2	7	13
17	2	2	2	7	2	6	5
18	1	1	1	12	1	10	7
19	2	1	2	12	5	9	20
20	1	1	1	20	2	10	11

[a]Normalized mean values of land-surface form, geologic texture, and climate characteristics for each hydrologic-landscape region (HLR) ranging from 1 to 20 (lowest mean to highest mean among the 20 HLRs).

Table 7. Standard deviations of watershed characteristics[a]

	Standard deviations						
		Land-surface form			Geologic texture		Climate
HLR number	Total flatland (%)	Flatland in upland (%)	Flatland in lowland (%)	Relief (m)	Bedrock permeability class	Sand (%)	Precipitation–PET (mm/yr)
1	12.3	21.0	22.5	49	0.7	13.7	168
2	19.0	19.1	21.7	57	1.0	11.1	155
3	16.0	21.9	21.0	99	0.7	11.3	251
4	14.9	18.9	17.1	48	0.7	8.0	206
5	18.8	20.5	16.9	277	1.6	12.2	191
6	10.7	15.6	14.6	54	0.9	7.6	259
7	19.9	23.0	13.4	77	0.4	9.7	163
8	16.9	22.6	12.9	111	0.5	8.9	180
9	25.5	21.3	15.9	187	1.4	10.4	178
10	18.6	19.0	11.8	189	0.9	13.1	147
11	17.3	17.6	11.7	70	0.6	6.2	157
12	19.3	9.0	16.1	340	0.9	8.8	259
13	14.1	13.3	10.2	157	0.5	6.5	142
14	16.2	2.5	15.8	529	1.1	11.3	345
15	7.9	3.2	7.7	271	0.9	7.7	142
16	9.8	6.3	7.5	269	0.7	10.3	178
17	7.5	3.0	7.0	250	0.6	7.1	152
18	3.4	0.6	3.4	337	0.4	11.0	213
19	10.0	2.0	9.5	498	1.4	13.2	368
20	3.6	0.4	3.6	683	0.8	7.7	333

[a]Standard deviations of land-surface form, geologic texture, and climate characteristics for all watersheds in each hydrologic-landscape region (HLR).
m = meters; PET = potential evapotranspiration.

Table 8. Hydrologic landscape region (HLR) descriptions and hydrologic hypotheses (x, indicates that the hydrologic process occurs in the HLR)

HLR number	Description	Primary hydrologic flow paths: Overland flow	Shallow ground water	Deep ground water
1	Subhumid plains with permeable soils and bedrock		x	x
2	Humid plains with permeable soils and bedrock		x	x
3	Subhumid plains with impermeable soils and permeable bedrock	x		x
4	Humid plains with permeable soils and bedrock		x	x
5	Arid plains with permeable soils and bedrock		x	x
6	Subhumid plains with impermeable soils and bedrock	x		
7	Humid plains with permeable soils and impermeable bedrock		x	
8	Semiarid plains with impermeable soils and bedrock	x		
9	Humid plateaus with impermeable soils and permeable bedrock	x		x
10	Arid plateaus with impermeable soils and permeable bedrock	x		x
11	Humid plateaus with impermeable soils and bedrock	x		
12	Semiarid plateaus with permeable soils and impermeable bedrock		x	
13	Semiarid plateaus with impermeable soils and bedrock	x		
14	Arid playas with permeable soils and bedrock		x	x
15	Semiarid mountains with impermeable soils and permeable bedrock	x		x
16	Humid mountains with permeable soils and impermeable bedrock		x	
17	Semiarid mountains with impermeable soils and bedrock	x		
18	Semiarid mountains with permeable soils and impermeable bedrock		x	
19	Very humid mountains with permeable soils and impermeable bedrock		x	
20	Humid mountains with permeable soils and impermeable bedrock		x	

The drainage basin for each water-quality sampling site was used with GIS data to calculate mean or majority values for the land-surface form, geologic texture, climate, and land-cover characteristics. The drainage basin boundaries also were used to determine the HLR, ecoregion, and geometric region to which each basin belonged.

In addition to the drainage-basin characteristics, two metrics of water quality were estimated at subsets of the NAWQA sites where adequate data were available. These water-quality measures were fish species richness (266 sites), collected using methods described by Meador and others (1993), and nitrogen transport efficiency (350 sites), which is the estimated percentage of total nitrogen inputs to the basin that is exported from the basin in streamflow.

The amount of variance in landscape, climate, and water-quality characteristics among the NAWQA sites that could be attributed to each regional framework was computed. The percentage of the total variance among the sites that was explained by each regional framework was expressed as an R-squared value. This R-squared value measures the power of the regional framework to separate the conterminous United States into distinct areas in terms of each basin characteristic. A regional framework is effective when the R-squared value is high.

The regionalization power of square geometric regions (Figure 9B) was computed as a "baseline" with which to compare HLRs and ecoregions. The size and number (19) of geometric regions were selected to match the approximate size and number of the other regional frameworks. Unlike the other regional frameworks, however, the boundaries of the geometric regions are based only on simple geometry; that is, the boundaries have no physical meaning derived from underlying maps or conceptual models.

The ANOVA results show that HLRs explain 73% to 83% of the variance among the NAWQA watersheds in the land-surface form, geologic texture, and climate variables (Table 9). These HLR R-squared values are higher than the geometric region values (22% to 63%) and the ecoregion values (31% to 79%) for the land-surface form, geologic texture, and climate characteristics.

The R-squared values for HLRs ranged from 15% to 68% for the land-cover variables (Table 9). The land-cover R-squared values for geometric regions (9% to 79%) and ecoregions (19% to 89%) were in a comparable range to the land-cover R-squared values for HLRs.

For the water-quality metrics, the R-squared values were 31% (fish species richness) and 34% (nitrogen transport efficiency) for HLRs (Table 9). The geomet-

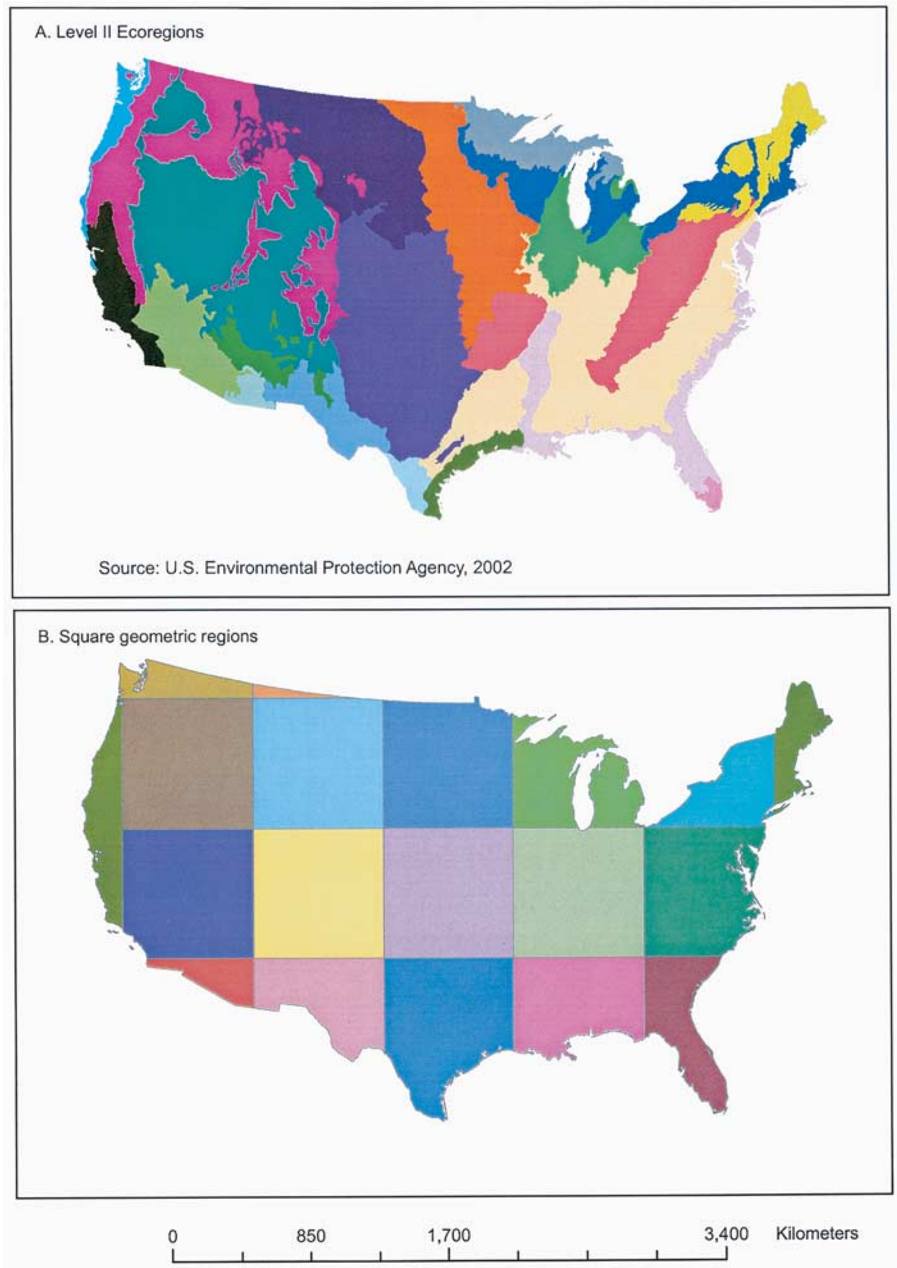

Figure 9. (A) Ecoregions and (B) square geometric regions in the conterminous United States.

ric region and ecoregion values were higher for fish species richness (42% and 45%, respectively), and lower for nitrogen transport efficiency (23% and 24%, respectively).

Some of the R-squared values for geometric regions are high even though the geometric region boundaries have no physical meaning. For example, geometric regions explain 63% of the variance in precipitation–PET among the watersheds and 79% of the variance in potential natural vegetation. These high R-squared values for geometric regions can be understood in terms of the spatial autocorrelation lengths of the landscape, climate, and water-quality variables. Precipitation–PET, for example, varies gradually over long distances (Figure 6). The NAWQA basin-average values reflect this long spatial autocorrelation pattern (Figure 11). The variability in precipitation–PET of NAWQA basins within any one of the geometric regions, therefore, is small compared to the variability in precipitation–PET of watersheds among the geometric regions. Put another way, precipitation–PET has a long spatial autocorrelation length relative to the size of the geometric

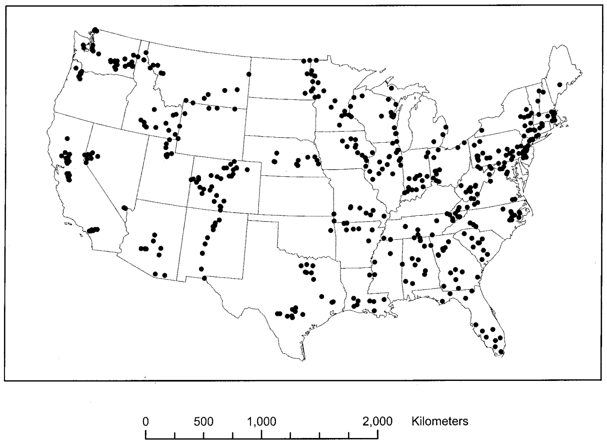

Figure 10. Location of National Water-Quality Assessment water-quality sampling sites in the conterminous United States.

regions. In contrast to precipitation–PET is bedrock permeability class (Figure 5), which has a relatively short autocorrelation length relative to the size of geometric regions (Figure 11) and a low R-squared value. These results suggest that if the autocorrelation length of a landscape, climate, or water-quality variable is long relative to the typical size of a region, then the regional framework will have some power to differentiate relatively distinct areas.

It is not unexpected that the regionalization power (expressed in terms of R-squared values) of the HLRs was greater than the regionalization power of geometric regions or ecoregions for the land-surface form, geologic texture, and climate characteristics. The statistical analyses used to delineate the HLRs ensured that the variance in land-surface form, geologic texture, and climate characteristics within regions was minimized and that the variance among regions was maximized. In fact, it would be surprising if the other regional frameworks were superior to HLRs in differentiating spatial variability in physical and climate factors. Unlike the HLRs, the boundaries of ecoregions are not statistically optimized to provide the greatest differentiation in physical and climate variables. Ecoregion boundaries are drawn using expert judgment where there are distinct breaks in the patterns of abiotic and biotic characteristics such as physiography, land use, and vegetation.

Hypothesis tests for comparing HLRs or ecoregions to geometric regions, in terms of the percentage of explained variance in each of the watershed characteristics, were conducted using nested F-tests (Helsel and Hirsch 2002). The nested F-test was structured such that a nested combination of two regional frameworks was compared to one of the regional frameworks alone. For example, the nested combination of ecoregions and geometric regions was compared to geometric regions alone. The hypothesis being tested in this example is that ecoregions explain a component of the variability in a watershed characteristic that is in addition to, or different from, the explanatory power of geometric regions alone. The F-statistic is computed from the R-squared value of the combined ecoregions and geometric regions model (R_{e+g}^2), the R-squared value of the geometric regions model (R_g^2), the number of observations (n), the number of ecoregions (N_e), and the number of geometric regions (N_g):

$$F = \frac{(R_{e+g}^2 - R_g^2)/N_e}{(1 - R_{e+g}^2)/(n - N_e - N_g - 1)} \qquad (1)$$

The F-statistic is distributed with N_e degrees of freedom in the numerator and $n - N_e - N_g - 1$ degrees of freedom in the denominator.

The nested F-tests indicate that HLRs and ecoregions both explain a component of variability in the watershed landscape, climate, and water-quality characteristics (except for fish-species richness) that is not captured in the geometric regions (Table 9). This implies that the regional frameworks have some power to identify distinct areas that is in addition to, or different from, the effect of the size of the regions relative to the autocorrelation length of the basin characteristics. Unlike geometric regions, the boundaries of the HLRs and ecoregions are irregular and, in the case of HLRs, not contiguous. These complex boundaries provide statistical power to the HLR and ecoregion frameworks in terms of the ability to delineate distinct regions, indicating that the boundaries conform to abrupt changes in landscape and climate characteristics. Landscape and climate characteristics change across the boundary from one HLR to another because the cluster analysis maximizes the ratio of among-region variance to within-region variance of landscape variables. Similarly, landscape and climate characteristics change from one ecoregion to another because the boundaries are drawn where maps indicate distinct breaks in physiography and land cover.

Caveats

The approach used to define HLRs is objective in the sense that it is based on statistical methods applied to digital geospatial data. Subjective expert judgement, however, is required in making several choices required for the analysis: (1) the particular set of variables used, (2) details in the statistical analyses, and (3) details in the GIS analyses. Using a different set of watershed characteristics

Table 9. Percentages of total variance explained[a]

;	Percentage of total variance explained (R-squared value) by the individual regional frameworks			Percentage of total variance explained (R-squared value) by the combined regional frameworks	
Landscape, climate, and water-quality characteristics	HLRs	Geometric regions	Ecoregions	HLRs nested with geometric regions	Ecoregions nested with geometric regions
Land-surface form					
Slope	79	53	67	82	73
Geologic texture					
Bedrock permeability class	80	22	31	82	46
Percent sand	73	48	37	79	62
Climate					
Precipitation–PET	83	63	79	89	87
Land cover					
Forest	49	18	52	59	57
Rangeland	60	53	61	66	71
Urban	15	9	19	26	24
Agriculture	50	39	48	60	56
Potential natural vegetation	68	79	89	85	91
Water quality					
Fish species richness	31	42	45	46	48
Nitrogen transport efficiency	34	23	24	43	38

[a]Percentages of total variance explained (R-squared values) for the regional frameworks individually and in combination. R-squared values in bold italics indicate a statistically significant (at a significance level of 0.01) increase in explained variance for the combined regional frameworks compared to geometric regions alone.
PET = potential evapotranspiration; HLR = hydrologic-landscape region.

would have affected the derived HLR map. In addition, averaging the variables over smaller areas would have produced a map with finer spatial detail. Changing any of these factors likely would lead to other regional maps that could be equally valid and useful.

The methods used to define HLRs are expected to be sensitive to the spatial scale of the analysis. In the study described herein, the spatial extent of the analysis covered all 50 States; this was the appropriate spatial scale for the purpose of identifying HLRs to help design a national water-quality assessment. Satisfying a different objective may require a different spatial scale of analysis and might result in a different set of regions.

The statistical approach used to evaluate the regional frameworks also is expected to be sensitive to the spatial extent of the analysis. If the spatial domain of the analysis were limited to only a small area of the United States and regional frameworks with finer spatial resolution were used, such as level III or level IV ecoregions, then the ANOVA results for the regional frameworks probably would be different. Most likely, the results would reflect, in part, the interaction of the spatial autocorrelation length of the landscape and climate variable of interest with the typical region size of the regional framework.

Implications for Environmental Management

Regional frameworks, such as HLRs and ecoregions, are useful tools for environmental management. They can be used for water-quality sampling network design, synthesis of information collected from diverse locations, and extrapolation of this information to unmonitored locations within and among regions.

The methodology used to define HLRs is a relatively objective and flexible approach for defining a regional framework. The set of variables and spatial scale used in the analysis can be customized to suit specific environmental management needs. For example, Preston (2000) grouped land cover, soil type, slope, and geology variables to delineate "hydrochemical response units" in Maryland. The State then used the regions to help develop a statewide water-quality management plan. On a broader scale, Hargrove and others (2003) used a cluster analysis of climatic and physiographic factors for the conterminous United States to analyze the representativeness of a meteorological network.

The finding that geometric regions are as good (in terms of R-squared values) as HLRs or ecoregions at delineating distinct areas for some landscape, climate, and water-quality characteristics does not mean that geometric regions is an equally useful regional frame-

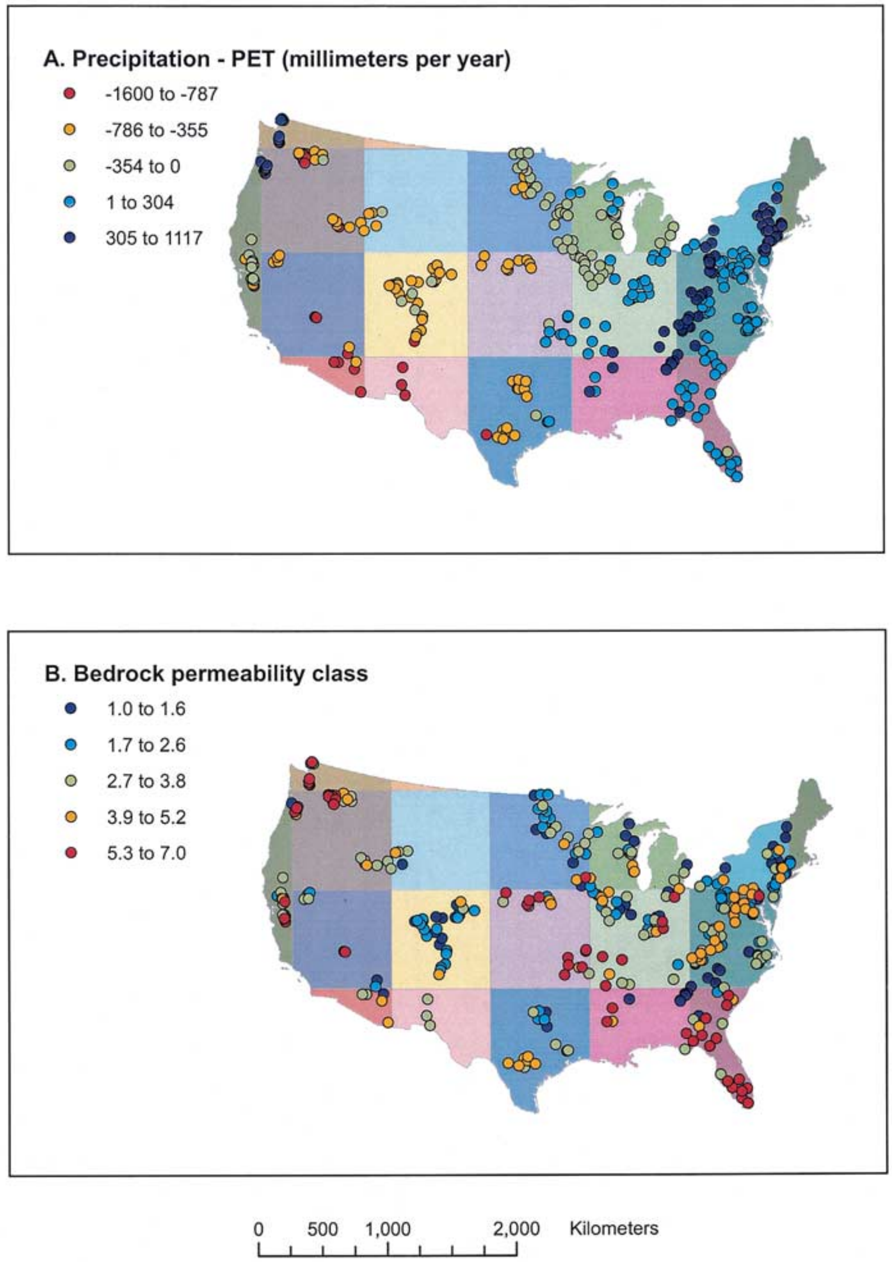

Figure 11. (A) Average precipitation–PET and (B) average bedrock permeability classes for NAWQA drainage basins. The drainage-basin average values, indicated by filled circles at the location of basin outlets, are overlain on square geometric regions. PET = potential evapotranspiration; NAWQA = National Water-Quality Assessment.

work for environmental management. Geometric regions are simple to define but have no conceptual basis. A conceptual basis may be a crucial requirement for a regional framework, depending on the application. If, for example, the goal of using a regional framework is to help understand differences among regions in terms of land use and water quality, then a conceptually based framework provides some advantages. Hypotheses (like those in Table 8) about important differences among regions can be more readily posed and tested if the regional framework is based on some conceptual model.

References

Gilliom, R. G., W. M. Alley, and M. E. Gurtz. 1995. Design of the National Water-Quality Assessment Program—Occurrence and distribution of water-quality conditions. U.S. Geological Survey Circular 1112, Reston, Virginia, 33 pp.

Hammond, E. H. 1964. Classes of land-surface form in the 48 states, USA. Annals of the Association of American Geographers 54, no. 1964, map supplement no. 4, 1:5,000,000 scale.

Hamon, W. R. 1961. Estimating potential evapotranspiration. Journal of the Hydraulics Division. *Proceedings of the American Society of Civil Engineers* 87:107–120.

Hargrove, W. W., Hoffman, F. M., Law, B. E. 2003. New analysis reveals representativeness of the Ameriflux network. Eos 84:529, 535.

Hargrove, W. W., and F. M. Hoffman. 1999. Using multivariate clustering to characterize ecoregion borders. *Computing in Science and Engineering* 1:18–25.

Helsel, D. R., and R. M. Hirsch. 2002. Statistical methods in water resources. U.S. Geological Survey Techniques of Water-Resources Investigations Book 4, Chapter A3 (http://water.usgs.gov/pubs/twri/twri4a3/)

Kuchler, A. W. 1964. Manual to accompany the map, potential natural vegetation of the conterminous United States. American Geographical Society, Special Publication No. 36, New York 143.

Meador, M. R., T. F. Cuffney, and M. E. Gurtz. 1993. Methods for sampling fish communities as part of the National Water-Quality Assessment Program. U.S. Geological Survey Open-File Report 93–104, 40 pp.

Murtagh, F. 1985. Multidimensional clustering algorithms. Physica-Verlag, Vienna.

Omernik, J. M. 1987. Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* 77:118–125.

Omernik, J. M., and G. E. Griffith. 1991. Ecological regions versus hydrologic units: frameworks for managing water quality. *Journal of Soil and Water Conservation* 46:334–340.

Owensby, J. R., and D. S. Ezell. 1992. Climatography of the United States No. 81—monthly station normals of temperature, precipitation, and heating and cooling degree days, 1961–90. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Climatic Data Center. Ashville, North Carolina.

Preston, S. D. 2000. Statistical identification of hydrochemical response units for hydrologic monitoring and modeling in Maryland. U.S. Geological Survey Water-Resources Investigations Report 00–4232. Reston, Virginia, 8 pp.

U.S. Department of Agriculture. 1994. State Soil Geographic (STATSGO) data base: data use information. Miscellaneous Publication Number 1492, 35 pp.

U.S. Environmental Protection Agency. 2002. Level III ecoregions of the continental United States (revision of Omernik 1987). U.S. Environmental Protection Agency, National Health and Environmental Effects Laboratory, Western Ecology Division, Corvallis, Oregon. Map M-1, various scales.

U.S. Geological Survey. 1990. Land use and land cover digital data from 1:250,000- and 1:100,000-scale maps: Data users guide 4 (US GeoData—National Mapping Program Technical Instructions). U.S. Department of Interior, U.S. Geological Survey. Reston, Virginia, 33 pp.

U.S. Geological Survey. 1998. Strategic directions for the U.S. Geological Survey Ground-Water Resources Program: A Report to Congress, 30 November 1998. (http://water.usgs.gov/ogw/gwrp/stratdir/stratdir.html).

U.S. Geological Survey. 2001. National atlas of the United States maps. U.S. Geological Survey Fact Sheet 086-01 (http://mac.usgs.gov/mac/isb/pubs/factsheets/fs08601.html).

Verdin, K. L., and S. K. Greenlee. 1996. Development of continental scale digital elevation models and extraction of hydrographic features. In Proceedings, Third International Conference/Workshop on Integrating GIS and Environmental Modeling, Santa Fe, New Mexico, 21–26 January 1996. Santa Barbara, California. National Center for Geographic Information and Analysis.

Winter, T. C. 2001. The concept of hydrologic landscapes. *Journal of the American Water Resources Association* 37:335–349.