Environmental Biology of Fishes Vol. 17, No. 4, p. 267–279, 1986

Fish ecoregions of Kansas: stream fish assemblage patterns and associated environmental correlates

Clifford L. Hawkes[1], David L. Miller[1,2] & William G. Layher[3]
[1] *USDA Forest Service, Rocky Mountain Forest and Range Experiment Station, 240 West Prospect, Fort Collins, CO 80526, U.S.A.*
[2] *Present address: Northrop Services, Inc., 200 S.W. 35th, Corvallis, OR 97333, U.S.A.*
[3] *Kansas Fish and Game, Box 54A, Rural Route 2, Pratt, KS 67124, U.S.A.*

Keywords: Distribution, Multivariate analysis, Binary data, Principal components analysis, Cluster analysis, Canonical discriminant analysis, Discriminant function analysis

Synopsis

Principal components analysis was performed on fish presence/absence data for 39 common fish species from 410 stream sites in Kansas. The analysis confirmed ten ecologically meaningful fish assemblages, based on species associations. Factor scores based on these assemblages were then clustered into six geographic areas or fish ecoregions. Canonical discriminant analysis identified environmental variables that distinguished the derived fish ecoregions. Mean annual runoff, mean annual growing season, and discharge appear most important. Mean width, mean depth, chloride concentration, water temperature, substrate type, gradient, and percent of pool habitat were less important. Correspondence exists between these fish ecoregions and the patterns of physiographic regions, river basins, geology, soil, and potential natural vegetation in Kansas. The multivariate statistical approach used to classify fish ecoregions should have considerable potential value for fish assessment and management purposes in areas other than the state of Kansas.

Introduction

Assessments, management, and ecological knowledge of the fish fauna of a large area such as a state can be aided by stratifying the area into several regions that are relatively homogeneous with respect to ichthyofaunal assemblage patterns. Typically, studies of individual fish species do not address the community or assemblage within which a species exists. How an individual species integrates into an assemblage of coevolved species, and how this assemblage interacts with the environment are not considered. Nor, with a few exceptions (e.g. Huet 1954, Balon & Stewart 1983), has much attention been given to how differently an assemblage responds because of environmental changes that occur across large geographic areas. However, the limitations of relatively simple, single-species approaches are being recognized (McHugh 1970, Regier & Henderson 1973, Gulland 1977, Larkin 1977, 1978, Regier 1978, Kerr 1982).

For fish assessments or impact analyses, predictive models built for each assemblage-defined homogeneous area can be expected to perform better than single-species models built for large heterogeneous areas. Layher & Maughan (1984, 1985) found that, unless the data were stratified by sample method, it was not possible to build satisfactory single-species, multiple regression models to predict biomass of Micropterus salmoides, Lepomis cyanellus, or Ictalurus punctatus in Kansas. It is also very inefficient to build many, site-specific models to assess a large area. The use of fish assemblages to identify regions of homogeneity provides an effective mechanism to avoid this problem. Furthermore, when correlations can be determined between environmental variables and these regions, the usefulness of the regionalization is enhanced.

Recently, the increased speed and capacity of computers have allowed the use and further development of multivariate statistical techniques that can simplify large, complex data sets. These techniques have been used to classify the distribution of fish in Kansas (Smith & Fisher 1970); to analyze fish distribution patterns in western and central Oklahoma (Stevenson et al. 1974); to examine fish associations in the Kiamichi River, Oklahoma (Echelle & Schnell 1976); to examine fish associations in the Little River, Texas (Rose & Echelle 1981); to determine patterns of stream fish assemblages in Missouri (Pflieger et al. 1981); to identify fish assemblages in northern Wisconsin lakes (Tonn & Magnuson 1982) and to identify ichthyogeographic regions in the Quebec peninsula of Canada (Legendre & Legendre 1984). Classification techniques in these studies vary considerably, depending on the available data and the use of the classification.

The objective of this study is to use the multivariate statistical techniques of ordination and classification to partition the state of Kansas into a relatively small number of homogeneous areas, or fish ecoregions, based on the distribution pattern of stream fish assemblages. We identify environmental variables correlated with the fish ecoregions, and we evaluate these fish ecoregions by comparing them with patterns of climate, landform, geology, soil, potential natural vegetation, land use, and land cover. We also compare our approach to classifying aquatic ecosystems of Kansas with that of others reported in the literature.

Methodology

Data

Data used in this study were collected from 1972 to 1978 by the Kansas Fish and Game Commission.[1] The sampling scheme specified that there be one representative stream collection site for each county in Kansas. Flow conditions during the sampling interval and variability among streams within a county resulted in some counties having no sites while others had several sites. Because of the sampling scheme, several regions of the state were over represented with stream sites. Thus, the sampling was not simple random nor systematic. This condition may affect the multivariate statistical analyses we describe below. A total of 410 stream sites, representing all major watersheds in Kansas, were surveyed for fish presence/absence (86 species collected), numbers, biomass, and 30 physiochemical characteristics. Fish were sampled with eight different methods, depending on conditions at the site. Complete descriptions of all survey methods are contained in Layher (1983) and Layher & Maughan (1984).

In large data sets, quantitative data can have more variability than the presence/absence data. Indeed, the use of eight different fish sampling methods introduces a compromising attribute to this data set: both abundance and biomass might present misleading results. Therefore, the presence/absence matrix was used. Binary data can often be used with good results in multivariate statistical procedures intended for quantitative data (Green 1979). In some cases, results are even better than when proportions or densities are used (Buzas 1972, Petersen 1976, Thorpe 1976). Smart et al. (1974) found that qualitative data emphasize species richness and diversity, and are best used when the interest is in species-environment relationships. Because of the differential susceptibility of species to various sampling methods, the presence/absence data for each method were standardized to zero mean and unit variance for the following analysis. This compensates to some extent for the disparity in sample sizes and number of species taken by the different methods.

Analysis

The basic approach involved a three-stage process with a combination of the multivariate statistical techniques of ordination and classification, similar

---

[1] Kansas Fish and Game Commission. 1972–1980. Kansas Stream Survey. (A series of river basin reports in two parts: Phase I – preliminary inventory and Phase II – fisheries assessments. There are sixteen Phase II Basin Reports, one for each Kansas river basin.)

to that of Tonn et al. (1983) and Omi et al. (1979). Ordination and classification complement each other; using both strategies, one can often determine the extent to which distinct community types exist, or whether species assemblages occur along a continuum. Multivariate statistical procedures such as principal components analysis (PCA) and canonical discriminant analysis (CDA) require several assumptions, some of which are difficult to meet. However, for descriptive studies, large departures from ideal data structure are tolerable. All statistical procedures were performed using the SAS data analysis system (SAS Institute 1982). Because our objective was to identify fish eco-

Table 1. Species associated with the 10 rotated factors. Species are grouped together under the factor for which they have the highest loading. The species are listed in decreasing order of factor loading on each of the factors. Blanks occur where loadings were less than an absolute value of 0.4. Four species (*Lepomis humilis, Noturus flavus, Phenacobius mirabilis, Morone chrysops*) out of the original 39, loaded lower than 0.4. Scientific names are according to Robbins et al. (1980).

Species	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
Micropterus punctulatus	0.81									
Labidesthes sicculus	0.77									
Fundulus notatus	0.73									
Percina phoxocephala	0.72									
Lepomis megalotis	0.60									
Pimephales vigilax	0.46									
Percina caprodes	0.45				0.40					
Moxostoma erythrurum	0.43									
Pimephales notatus		0.67								
Campostoma anomalum		0.66								
Etheostoma spectabile		0.66								
Notropis umbratilis		0.44								
Ictalurus punctatus			0.76							
Pylodictis olivaris			0.68							
Carpoides carpio			0.65							
Cyprinus carpio			0.47							
Pomoxis annularis				0.67						
Aplodinotus grunniens			0.40	0.59						
Ictiobus bubalus				0.57						
Dorosoma cepedianum				0.55						
Ictalurus natalis					0.60					
Noturus exilis		0.43			0.59					
Etheostoma nigrum					0.49		0.40			
Notropis stramineus					−0.53					
Micropterus salmoides						0.69				
Lepomis macrochirus						0.66				
Pimephales promelas							0.67			
Catostomus commersoni						0.45	0.54			
Semotilus atromaculatus		0.47					0.48			
Moxostoma macrolepidotum								0.83		
Lepisosteus osseus								0.66		
Ictalurus melas									0.71	
Lepomis cyanellus									0.68	
Fundulus zebrinus										0.74
Gambusia affinis	0.41									0.55
Eigenvalue	4.12	2.74	2.38	2.05	1.89	1.78	1.63	1.63	1.59	1.58
% variance explained by each factor and overall variance explained	10.6	7.0	6.1	5.3	4.8	4.6	4.2	4.2	4.1	4.1 = 55

regions that are characteristic of common species, we eliminated from the analysis those species occurring at fewer than 5 percent of the sites. This reduced the number of species in the analysis to 39 (Table 1).

Principal components analysis was performed on the 39 fish species presence/absence matrix to reduce the dimensionality and maximize the variance of the data set. Gower (1966) validated the use of binary data in PCA by showing that distances in multidimensional space are equal to the square root of the complement of the matching coefficient. However, the nature of the sampling scheme increases the possibility of distortion of the first principal components in favor of any over-represented areas.

Ten factors in the analysis had eigenvalues greater than or equal to one. Interpreting a factor with an eigenvalue less than one increases the likelihood of that factor explaining less than a randomly generated variable (Legendre & Legendre 1983). A random variable introduced into a data set would come out of a PCA on standardized variables as the major load on an axis whose eigenvalue would be one. Varimax rotation was performed on these factors to aid in biological interpretation (Cooley & Lohnes 1971). Factor loadings (= correlations) between the original variables (species) and the rotated factors were used to compute factor scores for each stream site. Factor scores were estimated along each factor; thus, each site was represented as a unique point in 10-dimension space as opposed to the original 39-dimension problem. Various assumptions and uses of this multivariate statistical procedure, as applied to biological data, have been reviewed by Poole (1971) and Gauch (1982).

Groups of similar sites were identified from an agglomerative hierarchical cluster analysis of factor scores using Ward's method (SAS Institute 1982). This classification process involves the determination of similar sites based on a systematic comparison of factor scores for all sites. The approach seeks to minimize within-cluster sums of squares, which are indexed by r^2 [= between-cluster sums of squares divided by total (corrected) sums of squares]. Cluster analysis can be performed on raw data; however, use of PCA to derive factor scores clarifies relationships between variables, reduces the effect of redundant variables, and helps infer structure in the original data. Virtually all clustering procedures provide little, if any, information as to the number of clusters in a data set. Milligan & Cooper (1985) performed a simulation study on several stopping rules (= quantitative indices used to evaluate the number of clusters) and found that the stopping rule used in SAS clustering procedures (= cubic clustering criterion) performed competitively. However, in our analyses, this stopping rule indicated fewer clusters than we would have subjectively assigned. If fewer clusters are indicated by a stopping rule when more are present, then the error is more serious than when K clusters are indicated with less than K clusters actually existing (Milligan & Cooper 1985). We therefore decided on an ecological evaluation of the number of clusters. A cutoff point of six clusters was chosen. Examination of more than six clusters, while improving the variance explained, added new clusters whose sites were distinctly more geographically scattered and interspersed with the sites of other clusters. Furthermore, the relationship between the distribution pattern of the sites of these new clusters and environmental patterns was much less distinct. Choosing fewer clusters left out the large river group (Fig. 1). Because of the unique physical and chemical habitat characteristics of large rivers, a large river group is biologically and ecologically meaningful. Pflieger et al. (1981) previously identified a distinct large river group from Missouri fish distribution patterns.

The cluster membership of sites was indicated on a map of Kansas that located all the sites. Lines were then drawn to form the fish ecoregions. This gives geographically contiguous fish ecoregions in place of the interspersed cluster sites. The elimination of this interspersion of cluster sites provides for a more useful classification. Contiguous fish ecoregions are useful in management and planning, whereas clusters with interspersed sites are inconvenient. With interspersed sites, new sites not included in the original cluster analysis cannot be assigned to a cluster without reanalysis. However, with homogeneous, contiguous geographic units

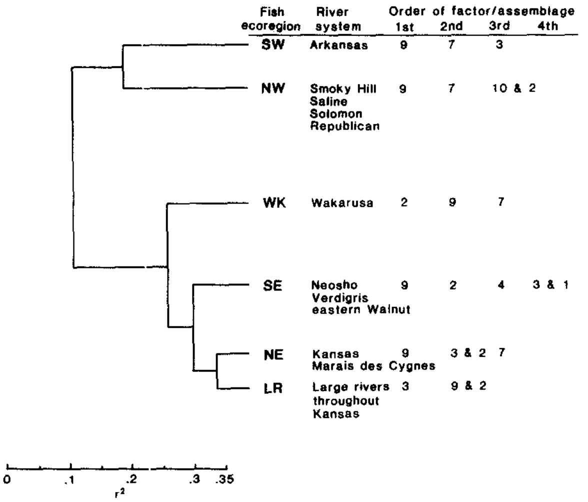

Fig. 1. Cluster dendrogram of stream sites based on rotated factor scores. The factors (Table 1) that distinguish each cluster are listed in order. Factors having more sites with factor scores greater than 1.5 are listed first. Where two factors appear together, they have an equal number of sites with factor scores greater than 1.5. The ordering gives a rough indication of the tendency of each assemblage to contribute toward characterizing the fish composition of each of the clusters and also the derived ecoregions. Each cluster is given a designation that corresponds to the area where most sites in that cluster are located. The horizontal axis represents improvement in r^2. Abbreviations are as follows: NE = northeastern fish ecoregion, SW = southwestern fish ecoregion, NW = northwestern fish ecoregion, SE = southeastern fish ecoregion, WK = Wakarusa River drainage fish ecoregion, LR = large river fish group.

this problem does not exist.

The lines were drawn to enclose the major concentration of sites from a cluster in a non-overlapping way. Between those clusters where interspersion of sites from different clusters occurred most, the lines were drawn by giving equal weighting to each site. For example, a single site from one cluster located in the center of a group of sites from another cluster carried less influence on line location. In this case, the line was drawn to include all of the sites with the cluster of the more numerous sites as well as the single site from the other cluster. Using these rules, most of the lengths of these lines follow watershed boundaries. One interesting exception is in the Walnut Creek drainage where the line was drawn to split the drainage in half along the main stem.

Sites belonging to an original cluster, but falling outside the newly constructed fish ecoregion, were reclassified to the new fish ecoregion associated with their location. The classification table from a discriminant function analysis (DFA) was used to determine the error that resulted from our fish ecoregion demarcation and subsequent site reclassification. Use of within-group covariance matrices as a basis for the measure of generalized square distance in DFA results in a perfect classification of all sites to their original clusters when clusters are determined by the above cluster analysis (SAS Institute 1982). Therefore, the percentage of sites misclassified by DFA represents the error attributable to our new fish ecoregion demarcations.

The extent of environmental differences among the six fish ecoregions was examined with canonical discriminant analysis (CDA). This analysis detects the maximum amount of multivariate variation in across-group means, relative to within-group variation of environmental variables for each site. The derived axes of CDA represent mathematical variables that maximize the differences among fish ecoregions. The individual loadings of environmental variables on the axes give some indication of the multivariate relationships or trends, and indicate which variables are most likely important. However, they do not indicate causation for, nor univariate correlation with, the fish ecoregion differences. Because of missing data, only 24 environmental variables were used (Table 2). This reduced the number of sites in this procedure to 246. However, this reduction did not compromise the sample size of each fish ecoregion beyond an acceptable level. We also feel that these 246 sites provide adequate representation of the fish ecoregions. The data for many of the variables are not distributed normally and have large variances. To alleviate these problems, all environmental data were log transformed except proportions, which were arcsin transformed.

Table 2. Environmental variables loaded on canonical variables 1–3. Environmental variables are grouped together under the canonical variable for which they have the highest loading. The loadings of environmental variables are ordered from highest to lowest. Blanks occur where loadings are less than an absolute value of 0.3. Those environmental variables loading less than 0.3 include percent riffle, percent run, silt substrate, calcium hardness, magnesium hardness, nitrates, phosphates, dissolved oxygen, total alkalinity, and pH.

Environmental variables	Canonical variables		
	1	2	3
Runoff	0.77	− 0.39	
Growing season	0.63		
Mean width	0.55		0.41
Chloride	− 0.52		
Mean depth	0.46		
Bedrock substrate	0.45		
Cobble-rubble substrate	0.44		
Water temperature	0.44		
Percent pool		0.35	
Gravel substrate		0.35	
Discharge			0.60
Gradient			− 0.46
Sand substrate			0.41
Canonical correlation	0.83	0.63	0.53

Results

Fish assemblages

The principal components analysis followed by varimax rotation reduced the 39 common species in the original data set to ten factors, or ten fish assemblages, consisting of fish species that tend to occur together at a site (Table 1). The degree to which they occur together is indicated by the loading of a species on its factor or assemblage. Ten assemblages of fish species provide a more comprehendible set of entities than do 39 separate species.

A brief description of the physiochemical habitat and distribution characteristics of these assemblages gives an understanding, at the assemblage level, of Kansas fish and their habitat. Information extracted from Cross (1967) and Cross & Collins (1975) provides a description of the assemblages obtained (Table 1). Assemblage 1 contains species that occur primarily in the southeastern part of Kansas, bounded to the west by the Arkansas River mainstem and north by the Kansas River mainstem (Missouri River drainage). However, they occur to a lesser extent in Missouri drainage streams. This assemblage is characterized by species that occur in small- to medium-sized, permanent streams. These streams are clear and usually have good pool development and gravel substrates. The species in Assemblage 2 also inhabit small permanent streams, but are found throughout eastern Kansas. *Campostoma anomalum, Etheostoma spectabile,* and *Semotilus atromaculatus* are also found in northwestern Kansas. Assemblage 3 consists of four species that live in large rivers with sandy or rocky bottoms. All of these species are widely distributed except for the *Pylodictis olivaris*, which is found mainly in eastern Kansas. Assemblage 4 species also inhabit large rivers, but tend to be restricted to eastern Kansas. The river populations of these species (especially *Pomoxis annularis*) have undoubtedly been increased by migrants from stocked impoundments. The Osage River system and eastern portions of the Kansas river system have the species of Assemblage 5. These species, especially *Etheostoma nigrum* and *Noturus exilis* tend to occur in small, permanent streams with gravel bottoms. *Ictalurus natalis* occurs in small, turbid streams of intermittent flow. It is able to survive droughts. The ubiquitous species, *Micropterus salmoides* and *Lepomis macrochirus*, are the only species in Assemblage 6. Stream populations of this favorite impoundment predator/prey stocking combination are increased by migrants from impoundments. *Pimephales promelas* of Assemblage 7 is a ubiquitous species but is most likely to occur in intermittent creeks with mud bottoms in western Kansas. *Semotilus atromaculatus* and *Catostomus commersoni* occur primarily in the Kansas river system. *Semotilus atromaculatus* prefer intermittent streams in this system. *Lepisosteus osseus* and *Moxostoma macrolepidotum* that constitute Assemblage 8 are restricted to large streams and rivers in the eastern half of Kansas. *Ictalurus melas* and *Lepomis cyanellus* making up Assemblage 9 are ubiquitous but more common in western Kansas. They occur in small, turbid streams of intermittent flow. Both are tolerant of drought conditions. Finally, *Fundulus zebrinus* of Assemblage 10 is very common in western Kansas. *Gambusia affinis* is limited to southwestern drainages due to a limited tolerance of cold temperatures. Both species prefer the shallow, calm streams found in these areas.

Fifty-five percent of the variation is explained by these ten factors or assemblages. The relatively low percentage of variance explained by the factors, both overall and individually, should not be considered to detract from the value of the information provided by the analysis (Table 1). Gauch (1982) points out that in some cases, particularly with large and 'noisy' data sets, the first few PCA axes may account for as little as 5 percent of total variance and yet be quite informative ecologically. In such cases, percent variance explained has not been found to be a reliable indicator of the quality of results.

Classification of fish ecoregions

The 410 sites were divided into six clusters. Each cluster is distinguished as consisting of a collection of sites with a unique pattern of factor scores derived from the ten factors or fish assemblages (Fig. 1). This pattern of site factor scores is partly indicated by the 'Order of factors/assemblages' in Figure 1. Each of these clusters is assumed to represent a relatively homogeneous collection of sites compared to the original complete set of sites. Inherent in this assumption is that a large heterogeneous area is made up of geographically contiguous units. While the r^2 improvement achieved at the six-cluster level is low, we believe that the clusters are ecologically meaningful at this level. When the sites of each cluster are mapped (Fig. 2), it is apparent that some sites of a cluster occur in slightly scattered geographic locations. Thus, while patterns are evident, it is also apparent that much 'noise' was inherent in the data. However, based on the mapped sites of each cluster and the assumption that there is a definite homogeneity and contiguous pattern to the distribution of sites, a map of fish ecoregions of Kansas was derived (Fig. 2). DFA missclassified only 4.7 percent of the 246 sites. This provides strong support for the location of the lines drawn and the assumption of contiguity and homogeneity of our fish ecoregions.

The southwestern fish ecoregion (Fig. 2) encompasses the western side of the Arkansas River drainage, except for the lower halves of the Chikaskia and Medicine Lodge Rivers, which are included as part of the southeastern fish ecoregion. The northwestern fish ecoregion contains the upper Kansas River drainage (Smoky Hill, Saline, Solomon) and upper and middle Republican drainages. The separation between the northwestern and southwestern fish ecoregions follows along the major drainage divide separating the Missouri and Mississippi drainages. The major differences in fish assemblages between these fish ecoregions involve Assemblage 3 associated with the southwestern ecoregion and Assemblages 2 and 10 associated with the northwestern fish ecoregion (Fig. 1).

The delineation between southwestern and southeastern fish ecoregions is very unusual because it splits the Walnut Creek drainage virtually into equal halves with eastern tributaries to the southeastern fish ecoregion and western tributaries to the southwestern fish ecoregion. The southeastern fish ecoregion also contains the Verdigris, Neosho, Spring, lower Chikaskia, and Medicine Lodge drainages. This is the only ecoregion where Assemblages 1 and 4 are involved in distinguishing an ecoregion (Fig. 1). Cross (1967) stated that the rivers in this area harbor more species than do rivers in other areas of Kansas. The streams in this region provide diverse and persistent habitats.

The northeastern fish ecoregion includes the lower Republican, Big Blue, Missouri, lower Kansas (except Wakarusa River drainage), Marais des Cygnes, Little Osage, and Marmaton drainages. The sites of this ecoregion differ from those of the northwestern ecoregion in that Assemblage 3 is associated with the northeastern ecoregion while Assemblage 10 is associated with the northwestern ecoregion, and in the overall differences in the pattern of site factor scores between the two ecoregions.

It is unusual that a small drainage like the Wakarusa River should define a fish ecoregion,

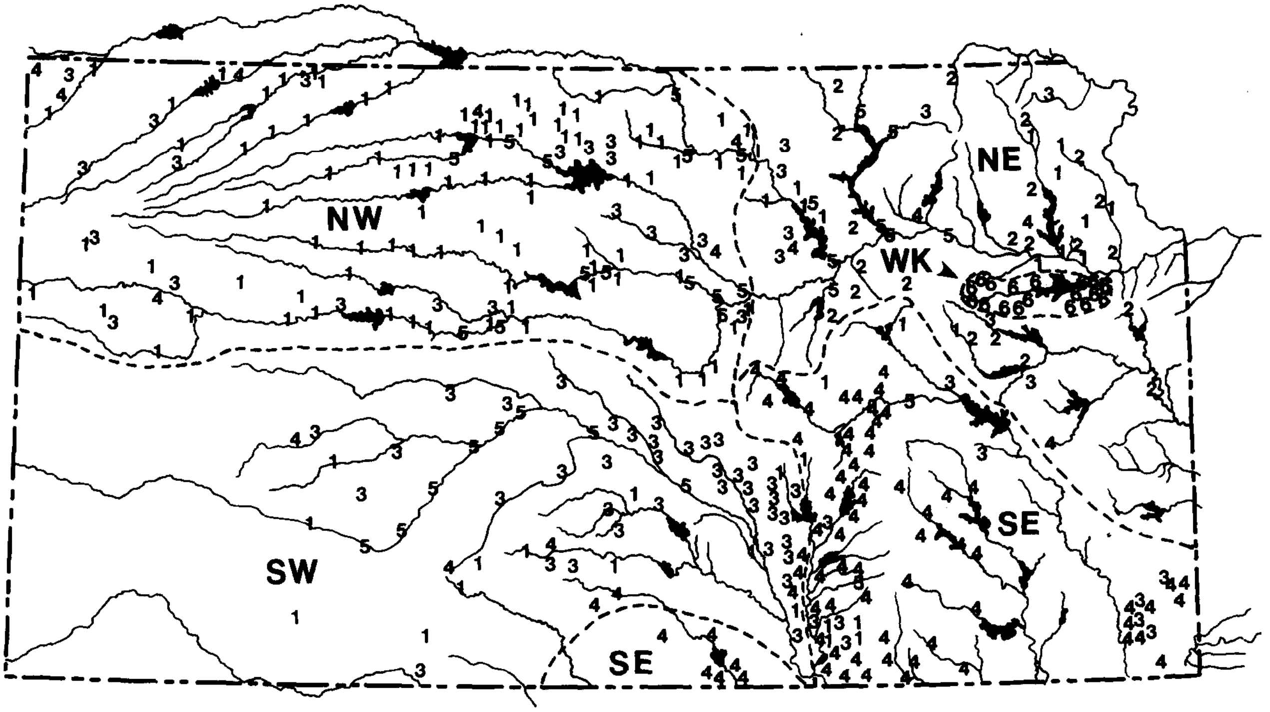

Fig. 2. Location of stream sites based on cluster analysis. Numbers refer to clusters as follows: northwestern cluster (1), northeastern cluster (2), southwestern cluster (3), southeastern cluster (4), large river cluster (5), and Wakarusa drainage cluster (6). Intuitively-derived delineation of Kansas fish ecoregions indicated by dashed lines.

especially to the point of being the first fish ecoregion distinguished among the drainages of eastern Kansas (Fig. 1). The Wakarusa is the only fish ecoregion that has a site factor score pattern with a predominance of sites with high factor scores for Assemblage 2 (Fig. 1). This is probably a primary factor in accounting for the uniqueness of the Wakarusa ecoregion. Also, the Wakarusa contains a rather unique assortment of species and it crosses terrain formed by rocks of three different escarpments (Deacon & Metcalf 1961, Metcalf 1966). The lower part of the drainage includes 13 species, the middle has 27 species, and the upland plain area has 15 species (Metcalf 1966). An examination of the environmental data (Appendix) in conjunction with our analysis of the fish data strongly supports the distinctiveness of this ecoregion.

The large river fish group transcends the fish ecoregions and has no boundaries. Pflieger et al. (1981) also demonstrated a large river group in Missouri. This fish group has sites in most major rivers and major tributaries, as well as in locations close enough to reservoirs to be influenced by species migrating from them. This group, as expected, is the only group that has a factor score pattern with a predominance of sites with high factor scores for Assemblage 3, consisting of *Ictalurus punctatus, Pylodictis olivaris, Carpoides carpio,* and *Cyprinus carpio* (Fig. 1, Table 1).

Environmental correlates of fish ecoregions

Loadings of the environmental variables on the canonical variables, derived from the canonical discriminant analysis, are given in Table 2. These canonical variables maximize differences among the fish ecoregions. The loading level of an environmental variable on a canonical variable is indicative of the degree to which it assisted in maximizing the differences among the fish ecoregions along the axis of that particular canonical variable. However, it is important to recognize that the environmental variables operate in concert with each other as a multivariate system and not as isolated univariate variables. A loading is dependent on quantitative interactions with the other environmental variables that load highly on a canonical variable. Also, the overall character of a particular data set used in canonical discriminant analysis, can affect the relative loading levels of an environmental variable on a canonical variable.

Runoff and growing season loaded highest on the first canonical variable (CV). Mean width and chloride were moderately loaded, while mean depth, predominance of bedrock and cobble-rubble substrates, and water temperature were weakly loaded. Chloride concentration shows an inverse relationship with the other variables. Percent pool, and percent of sites with predominance of gravel substrate loaded positively, though weakly, on CV 2. Runoff is inversely related to these environmental variables. Discharge loaded highest on CV 3. Mean width, percent of sites with a predominance of sand substrate, and gradient loaded weakly on CV 3. Gradient is inversely related to the other three variables, as would be expected.

The mean canonical variable value (= group centroid) for each fish region is plotted for the first three CV's (Fig. 3). The environmental variables that loaded highest are shown with their general trends along the appropriate CV axis. The pattern of environmental trends along a CV axis is relative only to that axis and again, must be considered in an interactive multivariate context. CV 1 strongly separates the northwestern fish ecoregion from the other fish ecoregions and the large river group. The general trend of environmental variables along this particular axis is for the northwestern ecoregion sites to have lower runoff, shorter growing season, lower mean widths, and higher chloride concentrations, while the reverse is true for the other ecoregions.

CV 2 distinguishes the southwestern and southeastern fish regions from the other groups. By its extreme position, the Wakarusa fish ecoregion is also distinguished by CV 2. The general environmental trends along this axis are for the Wakarusa, at one extreme, to have sites with lower percent pools, fewer sites with gravel as the predominant substrate, and higher runoff. The reverse is true for sites of the southwestern and southeastern ecoregions at the other extreme along this axis.

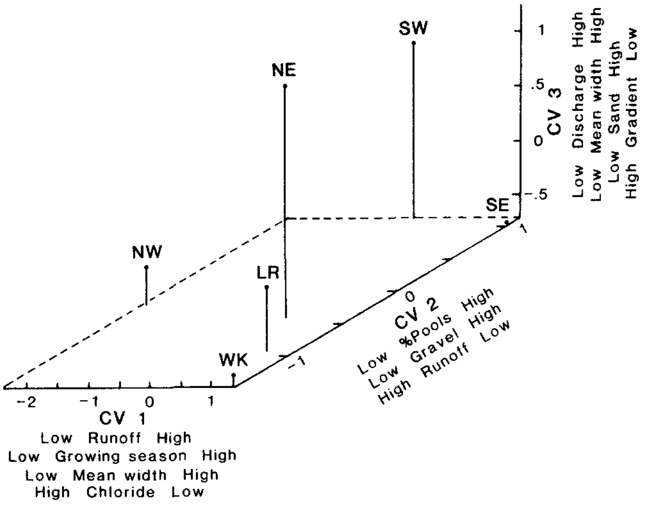

Fig. 3. Mean canonical variable (= group centroid) values derived from CDA of environmental data for northwestern (NW), southeastern (SE), southwestern (SW), northeastern (NE), Wakarusa River drainage (WK) fish ecoregions, and the large river (LR) fish group. The primary environmental variables that distinguished among the fish ecoregions and the large river fish group are listed in order of their loading on each respective canonical variable axis. The trends of these environmental variables along the axes are also given.

CV 3 separates the southwestern and northeastern fish ecoregions from the rest, and to a lesser extent the northwestern ecoregion and large river group from the Wakarusa and southeastern regions. Relative to this axis, the southwestern and northeastern ecoregion sites, as opposed to the other ecoregion sites, are characterized by higher discharges, and greater mean widths, lower gradients, and a predominance of sand substrate.

Discussion

A comparison of our classification of Kansas with the multivariate classification of fish distributions done by Smith & Fisher (1970) shows that our Assemblage 1 corresponds fairly closely with their factor IV group. All nine of the species in our Assemblage 1 are included in their factor IV group, which included 21 species. Only three other species that were in assemblages important in distinguishing our southeastern fish ecoregion were also in their factor IV group. Those three species are *Notropis umbratilis, Ictiobus bubalus,* and *Dorosoma cepedianum.* None of the other groups identified by Smith & Fisher matched any of our fish assemblages. However, there was some additional correspondence among geographic patterns. The geographic location of their factor IV group corresponds well with our southeastern fish ecoregion. Their factor I and our northeastern fish ecoregion seemed to be geographically similar. Their factor II was centered across the northern part of the state and would correspond roughly with a combination of our northwestern and northeastern fish ecoregions. Their factor VIII was centered in the west and corresponds roughly with a combination of our northwestern and southwestern fish ecoregions.

The differences in findings between our study and that of Smith & Fisher (1970) probably reflect, primarily, differences in the fish species included, which may have partly reflected the time span of data collection, and secondarily, differences in multivariate analysis procedures used. Their analysis included many relatively rare species. The data set used in their analysis contained 105 fish species. Most of the collections they used were taken from 1915 to 1966 (Cross 1967). In contrast, our analyses included only common fish species, and the data were collected recently.

We found our fish ecoregions to be logical when compared with published patterns of other environmental factors (Self 1978, USGS 1976), including river basin boundaries, major physiographic features, geologic formations, potential natural vegetation, soils, and certain climatic factors. The separation of northern from southern fish ecoregions is along the divide separating the Mississippi and Missouri River drainages. The separation between eastern and western regions follows the line separating major physiographic regions – the Great Plains to the west and the Central Lowlands to the east. It also follows the western border of the Flint Hills and the geologic formations of Permian age. There is a relatively close coincidence of the border separating the northwestern and northeastern ecoregions with changes in soils from the western typic ustolls to the eastern shallow udic ustolls. There is also agreement with the separation between western mixed prairie and eastern tall grass prairie potential natural vegetation types. The southwestern and southeastern fish ecoregion borders match the soil transition from the western sandy udic ustolls and deep udic ustolls, to the eastern shallow udic ustolls. For the most part, the separation of eastern and western fish ecoregions is along watershed divides (USGS 1976). Two exceptions are in the Republican River basin and in the Walnut Creek watershed. Other environmental trends that matched the separation between our western and eastern fish ecoregions are the west to east trends of increasing precipitation and precipitation runoff (Self 1978). No obvious, simple association between the fish ecoregions and land use and land cover (USGS 1970) was apparent.

Our findings confirm the statements of Cross (1967) that aquatic environments vary most from east to west in Kansas, that these differences in habitats for stream fishes are associated with the physiographic and climatic differences that affect substrate, amount of runoff, stream chemistry, groundwater development, and the permanence of streams.

The indicated trends of some of the environmental variables may be misleading. For example, runoff is most meaningful for small streams, which are most influenced by the runoff at a site. However, a large stream or river is much less influenced by the runoff at a site on the river than by the runoff throughout the entire river basin or watershed. Also, stream width and instantaneous discharge are greatly influenced by year, season, and the sampling scheme. Collections for the data set we used were taken over several years and over a wide part of each year. With this in mind, there appears to be a general, multivariate trend for the northwestern ecoregion to have lower runoff, shorter growing season, lower mean stream widths, and higher chloride concentrations than the other ecoregions. Similarly, the southwestern and southeastern ecoregions have more sites with a higher percentage of pools and more sites with gravel as the predominant substrate type, and the sites generally have lower runoff. Northeastern and southwestern ecoregions generally have sites with higher discharge, larger mean stream widths, lower gradients, and more sites having sand as the predominant substrate, in contrast to the other ecoregions.

An alternative approach to classifying aquatic ecosystems based on a landscape or aquatic ecoregion approach has been developed by Hughes & Omernik (1981). They investigated spatial patterns in aquatic ecosystems based on patterns of terrestrial characteristics. Their basic premise is that terrestrial watershed factors control the development and functioning of aquatic ecosystems. However, our fish ecoregions are not congruent with the aquatic ecoregions that Hughes & Omernik have delineated for Kansas (Robert Hughes, personal communication).

By identifying contiguous areas of homogeneity, a fish ecoregion classification should make it possible to extrapolate knowledge gained from a few sites to other sites within an ecoregion. Various management actions, including remedial measures, that might be appropriate for one ecoregion and not others might be more easily recognized, and appropriate planning and management action taken. That is, these ecoregions provide ecologically meaningful management units as opposed to site- or politically-defined management units. Once this classification is in place, it should facilitate subsequent development, by fish ecoregion, of improved biomass prediction models, and development of measures of environmental condition such as the index of biotic integrity, which is based on an analysis of stream fish communities (Karr 1981, Fausch et al. 1984).

Finally, our classification is based on fish association patterns, in contrast to other classifications that are based only on environmental factors assumed to be important in determining the patterns of aquatic ecosystems and their organisms. Because our multivariate statistical approach is not restricted to location nor size of area, it should have wide applicability for establishing fish ecoregions in areas other than the state of Kansas.

Acknowledgements

We wish to thank Robert Hartman of the Kansas Fish and Game who graciously provided the computerized tape of data used in this study. We also thank Robert Behnke, Kurt Fausch, Robert Hughes, Gerald Smith, and two anonymous reviewers for their very helpful suggestions.

References cited

Balon, E.K. & D.J. Stewart. 1983. Fish assemblages in a river with unusual gradient (Luongo, Africa – Zaire System), reflections on river zonation, and description of another new species. Env. Biol. Fish. 9: 225–252.

Buzas, M.A. 1972. Patterns of species diversity and their explanation. Taxon 21: 275–286.

Cooley, W.W. & P.R. Lohnes. 1971. Multivariate data analysis. John Wiley and Sons, New York. 364 pp.

Cross, F.B. 1967. Handbook of fishes of Kansas. Museum of Natural History, University of Kansas, Lawrence. 357 pp.

Cross, F.B. & J.T. Collins. 1975. Fishes in Kansas. University of Kansas, Museum of Natural History, Public Education Series No. 3, Lawrence. 189 pp.

Deacon, J.E. & A.L. Metcalf. 1961. Fishes of the Wakarusa River in Kansas. University of Kansas Publication, Museum of Natural History 13: 309–322.

Echelle, A.A. & G.D. Schnell. 1976. Factor analysis of species associations among fishes of the Kiamichi River, Oklahoma. Trans. Amer. Fish. Soc. 105: 17–31.

Fausch, K.D., J.R. Karr & P.R. Yant. 1984. Regional application of an index of biotic integrity based on stream fish communities. Trans. Amer. Fish. Soc. 113: 39–55.

Gauch, H.G., Jr. 1982. Multivariate analysis in community ecology. Cambridge University Press, Cambridge. 298 pp.

Gower, J.C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53: 325–338.

Green, H.R. 1979. Sampling design and statistical methods for environmental biologists. John Wiley and Sons, New York. 257 pp.

Gulland, J.A. 1977. Goals and objectives of fishery management. FAO Fisheries Technical Paper FIRS/T166, Rome. 14 pp.

Huet, M. 1954. Biologie, profils en long et en travers des eaux courantes. Bull. franc. Piscic. 175: 41–53.

Hughes, R.M. & J.M. Omernik. 1981. A proposed approach to determine regional patterns in aquatic ecosystems. pp. 92–102. *In*: N.B. Armentrout (ed.) Acquisition and Utilization of Aquatic Habitat Inventory Information, Proceedings of a Symposium, American Fisheries Society, Bethesda.

Karr, J.R. 1981. Assessment of biotic integrity using fish communities. Fisheries 6: 21–27.

Kerr, S.R. 1982. The role of external analysis in fisheries science. Trans. Amer. Fish. Soc. 111: 165–170.

Larkin, P.A. 1977. An epitaph for the concept of maximum sustained yield. Trans. Amer. Fish. Soc. 106: 1–11.

Larkin, P.A. 1978. Fisheries management – an essay for ecologists. Annual Review of Ecology and Systematics 9: 57–73.

Layher, W.G. 1983. Habitat suitability for selected adult fishes in prairie streams. Ph. D. Thesis, Oklahoma State University, Stillwater. 333 pp.

Layher, W.G. & O.E. Maughan. 1984. Analysis and refinement of habitat evaluation procedures for eight warm-water fish species. pp. 182–250. *In*: J.W. Terrell (ed.) Proceedings of a Workshop on Fish Habitat Suitability Index Models, U.S. Fish. Wildl. Serv. Biol. Rep. 85(6).

Layher, W.G. & O.E. Maughan. 1985. Relations between habitat variables and channel catfish populations in prairie streams. Trans. Amer. Fish. Soc. 114: 771–781.

Legendre, L. & P. Legendre. 1983. Numerical ecology. Developments in Environmental modelling, No. 3. Elsevier Scientific Publishing Company, Amsterdam. 419 pp.

Legendre, P. & V. Legendre. 1984. Postglacial dispersal of freshwater fishes in the Quebec peninsula. Can. J. Fish. Aquat. Sci. 41: 1781–1802.

McHugh, J.L. 1970. Trends in fishery research. American Fisheries Society Special Publication 7: 25–56.

Metcalf, A.L. 1966. Fishes of the Kansas River system in relation to zoogeography of the Great Plains. University of Kansas Publication, Museum of Natural History 17: 23–189.

Milligan, G.W. & M.C. Cooper. 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika 50: 159–179.

Omi, P.N., L.C. Wensel & J.L. Murphy. 1979. An application of multivariate statistics to land-use planning: classifying land units into homogeneous zones. Forest Science 25: 399–414.

Peterson, C.H. 1976. Relative abundance of living and dead molluscs in two California lagoons. Lethaia 9: 137–148.

Pflieger, W.L., M.A. Schene, Jr. & P.L. Haverland. 1981. Techniques for the classification of stream habitats, with examples of their application in defining the stream habitats of Missouri. pp. 362–368. *In*: N.B. Armentrout (ed.) Acquisition and Utilization of Aquatic Habitat Inventory Information, Proceedings of a Symposium, American Fisheries Society, Bethesda.

Poole, R.W. 1971. The use of factor analysis in modelling natural communities of plants and animals. Illinois Natural History Survey, Biological Notes No. 72, Urbana. 14 pp.

Regier, H.A. 1978. A balanced science of renewable resources. University of Washington Press, Seattle. 108 pp.

Regier, H.A. & H.G. Henderson. 1973. Toward a broad ecological model of fish communities and fisheries. Trans. Amer. Fish. Soc. 102: 56–72.

Rose, D.R. & A.A. Echelle. 1981. Factor analysis of associations of fishes in Little River, central Texas, with an interdrainage comparison. Amer. Mid. Natur. 106: 379–391.

Robins, C.R., R.M. Bailey, C.E. Bond, J.R. Brooker, E.A. Lachner, R.N. Lea & W.B. Scott. 1980. List of common and scientific names of fishes from the United States and Canada. American Fisheries Society Special Publication No. 12, Bethesda. 174 pp.

SAS Institute Inc. 1982. SAS User's Guide: Statistics, 1982 Edition. SAS Institute Inc., Cary. 584 pp.

Self, H. 1978. Environment and man in Kansas. A geographical analysis. The Regents Press of Kansas, Lawrence. 288 pp.

Smart, P.F.M., S.E. Meacock & J.M. Lambert. 1974. Investigations into the properties of quantitative vegetational data. I. Pilot study. J. Ecol. 62: 735–760.

Smith, G.R. & D.R. Fisher. 1970. Factor analysis of distribution patterns of Kansas fishes. pp. 259–277. *In*: W. Dort, Jr. & J.K. Jones, Jr. (ed.) Pleistocene and Recent Environments of the Central Great Plains, University of Kansas, Department of Geology, Special Publication 3, University of Kansas Press, Lawrence.

Stevenson, M.M., G.D. Schnell & R. Black. 1974. Factor analysis of fish distribution patterns in western and central Oklahoma. Systematic Zoology 23: 202–218.

Thorpe, R.S. 1976. Biometric analysis of geographic variation and racial affinities. Biol. Rev. 51: 407–452.

Tonn, W.M. & J.J. Magnuson. 1982. Patterns in the species composition and richness of fish assemblages in northern Wisconsin lakes. Ecology 63: 1149–1166.

Tonn, W.M., J.J. Magnuson & A.M. Forbes. 1983. Community analysis in fishery management: an application with northern Wisconsin lakes. Trans. Amer. Fish. Soc. 112: 368–377.

U.S. Geological Survey. 1970. The national atlas of the United States. U.S. Government Printing Office. Washington, D.C. 427 pp.

U.S. Geological Survey. 1976. Hydrologic unit map. 1974. State of Kansas. Prepared in cooperation with the U.S. Water Resources Council. USDI Geological Survey, Reston.

Received 31.8.1985 *Accepted 12.1.1986*

Appendix

Mean ($\bar{x}$), standard deviation (sd), range, and sample size (n), by fish ecoregion, for the physiochemical characteristics used in CDA.

Fish ecoregion	$\bar{x}$	sd	Range	n
		Runoff (mm yr^{-1})		
NW	15	13	3–51	128
SW	25	18	3–51	184
NE	5	36	1–203	56
SE	86	66	25–254	94
WK	104	13	102–152	15
LR	64	48	8–203	43

Fish ecoregion	x̄	sd	Range	n
Discharge (m^3 sec^{-1})				
NW	0.5	1.3	0–9.6	126
SW	0.5	0.9	0–4.5	69
NE	0.3	0.5	0–1.9	49
SE	0.7	2.9	0–25.2	82
WK	0.1	0.1	0–0.6	14
LR	4.5	8.1	0–28	37
Growing season (days)				
NW	166	12.1	82–188	128
SW	185	4.3	173–191	84
NE	181	6.0	170–194	56
SE	187	5.0	179–193	94
WK	188	5.7	183–194	15
LR	179	7.8	168–194	43
Mean width (m)				
NW	5.2	4.3	0.9–34.4	128
SW	8.4	5.3	1.5–30.5	84
NE	9.5	9.6	2.4–61.0	56
SE	12.5	8.8	2.4–54.9	94
WK	4.9	3.6	2.1–15.6	15
LR	19.6	23.5	3.7–110.0	43
Mean depth (m)				
NW	0.3	0.2	0.1–1.1	128
SW	0.5	0.3	0.1–1.5	84
NE	0.5	0.4	0.1–2.1	56
SE	0.6	0.4	0.1–3.1	94
WK	0.6	0.3	0.2–1.1	15
LR	0.7	0.7	0.1–4.6	43
Gradient (m km^{-1})				
NW	1.4	0.9	0.3–5.0	89
SW	1.6	1.2	0.1–5.0	79
NE	1.4	0.9	0.3–4.0	55
SE	1.7	1.3	0.1–8.0	89
WK	2.2	1.4	0.4–5.0	15
LR	0.8	0.8	0.2–4.2	42
Silt (% of sites with substrate type)				
NW	73	44	–	128
SW	75	44	–	84
NE	80	40	–	56
SE	79	41	–	94
WK	73	46	–	15
LR	72	45	–	43
Sand (% of sites with substrate type)				
NW	65	48	–	128
SW	55	50	–	84
NE	34	48	–	56
SE	18	39	–	94
WK	13	35	–	15
LR	60	50	–	43
Gravel (% of sites with substrate type)				
NW	26	44	–	128
SW	36	48	–	84
NE	34	48	–	56
SE	73	45	–	94
WK	27	46	–	15
LR	32	47	–	43
Cobble-rubble (% of sites with substrate type)				
NW	10	30	–	128
SW	27	45	–	84
NE	48	50	–	56
SE	53	50	–	94
WK	47	52	–	15
LE	26	44	–	43
Bedrock (% of sites with substrate type)				
NW	1	9	–	128
SW	5	21	–	84
NE	17	39	–	56
SE	29	45	–	94
WK	53	52	–	15
LR	6	21	–	43
Pool (%)				
NW	31	40.2	0–100	128
SW	55	41.8	0–100	78
NE	43	35.2	0–100	56
SE	52	40.6	0–100	92
WK	57	26.1	22–100	15
LR	30	34.0	0–100	40
Riffle (%)				
NW	4	10.9	0–64	128
SW	10	19.1	0–81	76
NE	13	16.7	0–88	56
SE	11	13.5	0–55	93
WK	16	17.9	0–100	15
LR	14	23.7	0–70	40
Run (%)				
NW	65	41.1	0–100	128
SW	38	41.7	0–100	76
NE	44	38.2	0–100	56
SE	37	42.9	0–100	92
WK	27	24.1	0–56	15
LR	57	40.1	0–100	41
Water temperature (° C)				
NW	16	8.3	1–21	128
SW	23	4.6	10–32	77
NE	21	7.0	7–32	55
SE	25	7.0	14–30	90
WK	25	2.0	20–29	15
LR	23	6.9	7–29	40
Chlorides (ppm)				
NW	107	150.2	1.3–830	126
SW	179	318.5	2.5–1870	73
NE	29	33.6	0.1–175	56
SE	58	104.5	0.3–700	84
WK	11	3.7	5–20	15
LR	147	171.0	3–560	42