Classification of Texas Reservoirs in Relation to Limnology and Fish Community Associations

WILLIAM B. DOLMAN
Texas Parks and Wildlife Department, Inland Fisheries Branch
4200 Smith School Road, Austin, Texas 78744, USA

Abstract.—I used cluster analysis to examine associations among 20 fish species to develop a classification scheme for 132 large Texas reservoirs. Five major groups of reservoirs were identified by cluster analysis based on species associations. Of 29 reservoirs surveyed previously, 76% were classified into the same species associations from one survey to the next. When 19 environmental variables were used in canonical correlation analysis of the five reservoir groups, I found a general east-to-west separation of species associations by water quality and a northwest-to-southeast separation by surface elevation and growing season. A discriminant functions model based on a reduced set of nine environmental variables had an unbiased error rate of 18% for predicting the species association in unclassified reservoirs. A stratified sampling scheme based on the classification model decreased the variance of statewide electrofishing catch per effort up to 43% for bluegill *Lepomis macrochirus* and 23% for largemouth bass *Micropterus salmoides* over a simple random sample of reservoirs.

There are about 800,000 hectares of impounded public waters and over 250 public lakes and impoundments larger than 50 hectares in Texas. The geographic size and diversity of the state, combined with limited resources for managing all public waters, make it impossible to monitor fishery resources annually in all bodies of water throughout the state. An empirically derived reservoir classification scheme would allow managers to stratify monitoring programs by reservoir types. This could be accomplished by identifying reservoir groups from patterns of fish community composition and by deriving relationships within these groups among physical and biotic factors. Such a classification scheme might improve the efficiency of a monitoring program for assessing the status of the resource throughout the state.

Fish community analyses that involve various classification methods have been performed by Echelle and Schnell (1976), Johnson et al. (1977), and Tonn et al. (1983). For example, Echelle and Schnell (1976) identified six species associations by factor analysis of 48 fish species in a southern Oklahoma river system. Johnson et al. (1977) classified 2,496 Ontario lakes on the basis of combinations of four major sport fishes present in each lake and found that seven limnological factors defined the lake type in which each species combination was found. Tonn et al. (1983) described three fish assemblages from 18 northern Wisconsin lakes and predicted the fish assemblage type in 11 additional lakes based on five limnological variables.

Previous studies of Texas reservoirs have described associations between species characteristics and physicochemical factors. Farquhar et al. (1982) found significant correlations between species abundance and physicochemical factors in small flood-prevention lakes in north-central Texas. For largemouth bass *Micropterus salmoides* in Texas reservoirs, significant relationships have been found between fish biomass and abundance of submerged vegetation (Durocher et al. 1984), and between fish growth and environmental factors (Miranda and Durocher 1986).

An objective classification scheme for Texas reservoirs would provide a framework for describing relationships between fish associations and environmental variables within groups of reservoirs and for improving the efficiency of reservoir surveys. The objectives of my study were to (1) develop an objective classification scheme for large Texas reservoirs based on fish community associations; (2) relate the groups of reservoirs identified with species associations to combinations of environmental variables; (3) demonstrate the consistency of species associations over time in reservoirs that had been surveyed in earlier years; (4) calculate unbiased error rates for classifying reservoirs based on a reduced set of environmental variables; and (5) illustrate the use of the classification for improving the statistical precision of...
population estimates obtained from statewide resource-monitoring surveys.

Methods

Fishery-monitoring surveys of selected reservoirs have been conducted annually since 1975 by the Inland Fisheries Branch of the Texas Parks and Wildlife Department under the Federal Aid in Fish Restoration Act. I obtained cove-rotenone and limnological survey data for this study from Federal Aid reservoir management reports.

Rotenone data were obtained on 132 reservoirs surveyed during 1975–1984. Twenty-nine reservoirs had previous rotenone surveys during this period. Cove-rotenone samples between May and September usually included one to three coves per reservoir, but one reservoir had five cove samples, and another had six. Cove areas ranged from 0.16 to 6.08 hectares each and were isolated after sunset with block nets. On the next morning, 5% rotenone was applied. On that day and the following day all fish were collected, separated by species, and counted.

Although fishery managers have traditionally used biomass estimates from cove-rotenone surveys, I used density (number/hectare) as the measure of catch per unit effort (C/f) in this study. My intent was to develop a classification scheme based on density that could be extended to other sampling methods, such as electrofishing, from which estimates of biomass are not obtained. In many ecological studies, the unit of measure may have little effect on the general conclusions provided by multivariate analysis (Moore 1974).

I obtained limnological and morphometric data on 155 reservoirs surveyed during 1975–1984. Water chemistry was analyzed quarterly at one to three stations in each reservoir. The first station was located near the dam, and additional sites were located upstream from the dam to the upper areas of the reservoir. Measurements of water quality taken from April to September were used to reduce the within-sample variance relative to the variance among surveys. Such measurements taken at 1 m below surface, mid-depth, and 1 m above bottom at each sample site were used to calculate mean values for each reservoir. Because morphometric parameters may vary seasonally in reservoirs, I used annual means.

I assembled two data sets for species abundance and environmental variables from reservoir surveys. The first consisted of the reservoir \( \times \) species matrix formed from \( C/f \) of \( s \) species (columns) observed in cove-rotenone surveys in each of \( r \) reservoirs (rows) in Texas. The second consisted of a reservoir \( \times \) environmental attribute matrix formed from the values of \( p \) (columns) limnological and morphometric variables.

Species \( C/f \) and most of the environmental variables (except pH) were log\(_{10}\)-transformed to reduce skewness and better approximate univariate normality. I transformed abundance of submerged vegetation (as a proportion of total surface area) with the arcsine of the square root. I identified groups of reservoirs based on species associations with a nonhierarchical cluster analysis that used the \( k \)-means clustering method provided by the FASTCLUS procedure (SAS 1985). The number of reservoir groups retained for interpretation was determined by iteration as the number of clusters for which the within-cluster sum of squares was maximized (cubic clustering criterion; Milligan and Cooper 1985). I characterized associations in each reservoir group based on species whose density was significantly different (\( P < 0.05 \)) from that in all other reservoir groups (Tukey–Kramer comparisons of multiple means; GLM procedure: SAS 1985).

The consistency of reservoir fish associations over time was estimated by repeating the cluster procedure on rotenone surveys from previous years. With the \( r \) original reservoirs in the data file, I (1) replaced data from reservoir survey \( i \) (the most recent) with data from a previous survey \( i^* \) one at a time; (2) repeated the cluster procedure for each new set of \( r^* \) reservoirs to determine if each \( i^* \) reservoir survey was classified into the same group as the most recent; and (3) calculated the proportion of all \( i^* \) reservoir surveys that were classified into the same groups as the \( i \) surveys for the original set of \( r \) reservoirs.

I used canonical correlation (CANDISC procedure: SAS 1985) to analyze the relation between the reservoir species associations and a set of 19 limnological and morphometric variables usually measured or easily obtained by fishery managers. This technique is appealing because it can aid managers by illustrating objectively the extent to which one set of variables (e.g., environmental attributes) is related to another set (e.g., species associations). Canonical correlation produces linear functions (canonical variables) of the environmental variables with the highest possible multiple correlation with the reservoir groups (clusters). Variables with the highest canonical coefficients in each of these linear functions are assumed to define that function. Therefore, the key variables relating the two data sets may be assessed from
each pair of coefficient vectors (canonical equations).

I used a reduced set of the environmental variables to develop a classification model for predicting species associations in unclassified reservoirs with discriminant functions analysis (DISCRIM procedure: SAS 1985). When a discriminant model based on \( n \) observations is used to classify the same \( n \) observations, the classification error rate obtained is likely to underestimate the true error rate (Efron and Gong 1983). Therefore, I used leave-out-one cross-validation (Efron and Gong 1983) with the discriminant function model to calculate unbiased error rates for predicting species associations in unclassified reservoirs. This method avoids bias by fitting a discriminant model to \( n - 1 \) observations and then using this function to classify the remaining observation.

Finally, I compared statewide electrofishing \( C/f \) and variance obtained from a stratified random sample based on the classification suggested with an estimate from a simple random sample of reservoirs. I used the bootstrap method (Efron and Gong 1983) to obtain nonparametric estimates of the mean and variance of each sampling design. From 78 reservoirs sampled with electrofishing in 1987, I drew 200 independent random samples of 30 reservoirs as simple and stratified samples by use of equal, proportional, and optimal allocation (Cochran 1977). With each random sample, I estimated the mean and variance of \( C/f \).

Results
Species Associations

In all, 104 fish species were collected with rotenone from 132 reservoirs. The 20 most abundant species (Table 1) over all samples combined were used in my analysis. These represented 94% of the total abundance of all species in cove-rotenone samples and included many of the species important to fishery managers in Texas. Cluster analysis of species \( C/f \) from the 132 reservoirs yielded six groups of reservoirs representing different associations among the 20 species.

Group 1: Orangespotted sunfish.—Thirty-one reservoirs (Table 2) had high densities of orangespotted sunfish relative to redbear sunfish, threadfin shad, warmouth, bluegill, and largemouth bass (Table 1). Reservoirs included in this group are located at higher elevations over a large area from north-central Texas to the Panhandle and in a small, disjunct area in west Texas (Figure 1).

Group 2: Redbreast sunfish–green sunfish.—The 10 reservoirs in the second group (Table 2) were characterized by associations of redbreast sunfish and green sunfish (Table 1). Blue tilapia, threadfin shad, inland silverside, and logperch were moderately abundant in many of these reservoirs, which are located from south-central to southwest Texas (Figure 1).

Group 3: Longear sunfish.—The third group of 27 reservoirs (Table 2) had fish communities characterized by combinations of longear sunfish, bullhead minnow, freshwater drum, logperch, and orangespotted sunfish (Table 1). These reservoirs are located from central to northeast Texas (Figure 1).

Group 4: Blue tilapia–threadfin shad.—Blue tilapia and threadfin shad dominated the species associations in six reservoirs (Table 2), with white crappie and warmouth usually scarce or absent (Table 1). Five of the impoundments in this group are power plant cooling lakes, and the sixth, Falcon Reservoir, is located in south Texas.

Group 5: Redear sunfish.—Thirty reservoirs (Table 2) typically had high densities of redbear sunfish, largemouth bass, and bluegill, and the lowest densities of gizzard shad, channel catfish, and freshwater drum (Table 1). These reservoirs are located in east Texas (Figure 1).

Group 6: Unclassified.—The remaining 28 reservoirs did not exhibit a common pattern of species association and were not classified by the present analysis. Further analysis of these reservoirs would produce groups with fewer members and greater overlap in species associations and so would contribute little to the classification of Texas impoundments. Unclassified reservoirs should be included in any application of the present classification scheme. However, these reservoirs were not included in further evaluation of species associations in this study.

Consistency of Species Associations

Among the five species associations described by cluster analysis, 29 reservoirs had been surveyed previously by rotenone sampling. The time between consecutive surveys ranged from 1 to 8 years. Of the 29 reservoirs with multiple surveys, 22 (76%) were classified into the same species association from one survey to the next by cluster analysis (Table 3). Species associations were most consistent in reservoirs with blue tilapia–threadfin shad (Group 4) and redbreast sunfish–green sunfish (Group 2) associations. However, it is difficult to make conclusions about the within-group consistency of fish communities because of the small
Table 1.—Mean abundances (number/hectare) of 20 fish species classified into five species assemblages by cluster analysis of cove-rotenone data from 132 Texas reservoirs, 1975–1984. Species are listed in descending order of overall abundance for all rotenone surveys combined. Underlined values indicate species contributing to group separation. Twenty-eight reservoirs were not classified by the current analysis.

| Fish species                          | 1         | 2         | 3         | 4         | 5         |
|---------------------------------------|-----------|-----------|-----------|-----------|-----------|
| Gizzard shad \textit{Dorosoma cepedianum} | 4,691.7   | 1,037.1   | 1,015.4   | 6,839.4   | 464.5     |
| Bluegill \textit{Lepomis macrochirus}   | 1,162.3   | 2,871.5   | 2,339.9   | 1,402.3   | 5,962.8   |
| Threadfin shad \textit{Dorosoma petenense} | 41.5      | 4,565.8   | 2,639.1   | 11,223.9  | 1,527.2   |
| Blue tilapia \textit{Tilapia aurea}     | 6.2       | 114.1     | 0.4       | 24,443.5  | 0.0       |
| Orangespotted sunfish \textit{Lepomis humilis} | 3,888.8   | 7.4       | 175.4     | 6.8       | 7.7       |
| Longear sunfish \textit{Lepomis megalotis} | 336.8     | 1,012.3   | 1,569.9   | 433.1     | 813.9     |
| Redear sunfish \textit{Lepomis microlophus} | 34.9      | 332.8     | 158.9     | 100.5     | 2,644.6   |
| Largemouth bass \textit{Micropterus salmoides} | 241.7     | 680.8     | 311.5     | 999.3     | 1,041.1   |
| Warmouth \textit{Lepomis gulosus}       | 35.0      | 676.8     | 357.4     | 19.4      | 1,098.8   |
| Bullhead minnow \textit{Pimephales vigilax} | 144.8     | 322.7     | 1,420.3   | 456.0     | 81.5      |
| Green sunfish \textit{Lepomis cyanellus} | 404.4     | 1,347.5   | 296.5     | 100.0     | 181.9     |
| White crappie \textit{Pomoxis annularis}  | 243.3     | 91.0      | 389.1     | 0.7       | 63.5      |
| Freshwater drum \textit{Aplodinotus grunniens} | 146.5     | 126.2     | 616.6     | 12.3      | 4.8       |
| Channel catfish \textit{Ictalurus punctatus} | 189.7     | 259.5     | 226.4     | 344.1     | 64.5      |
| Logperch \textit{Percina caprodes}       | 89.2      | 288.1     | 381.8     | 28.3      | 100.1     |
| Inland silverside \textit{Menidia beryllina} | 67.2      | 249.2     | 185.2     | 1,519.1   | 12.9      |
| Black bullhead \textit{Ictalurus melas}   | 137.6     | 10.3      | 185.5     | 0.0       | 310.7     |
| Mosquitofish \textit{Gambusia affinis}    | 243.6     | 27.2      | 69.0      | 642.5     | 27.3      |
| Redbreast sunfish \textit{Lepomis auritus}  | 0.0       | 1,270.6   | 2.3       | 0.0       | 120.7     |
| Golden shiner \textit{Notemigonus crysoleucas} | 136.1     | 2.1       | 158.3     | 94.5      | 94.2      |
| Number of reservoirs                     | 31        | 10        | 27        | 6         | 30        |

Figure 1.—Geographic distribution of 104 Texas reservoirs classified into five groups (Table 2) by community analysis of 20 fish species (Table 1). All reservoirs classified into groups designated by large numbers are located within more-or-less discrete geographic regions. Small numbers designate the group numbers and locations of reservoirs outside of their group's typical geographic region. Group 4 species associations were not located within a specific geographic region. Unclassified reservoirs occurred in areas of the state outside the four geographic regions.
CLASSIFICATION OF TEXAS RESERVOIRS

Table 2.—Reservoirs classified into five species assemblages by cluster analysis of the relative abundance of 20 species in cove-rotenone surveys of 132 Texas reservoirs, 1975–1984. Twenty-eight reservoirs were not classified.

| Species association | 1     | 2     | 3     | 4     | 5     |
|---------------------|-------|-------|-------|-------|-------|
| Abilene             | Amistad | Bardwell | Braunig | Amon Carter |
| Alcoa               | Buchanan | Brady Creek | Calaveras | Bob Sandlin |
| Arrowhead           | Canyon  | Bridgeport | Fairfield | Bryan Utilities |
| Baird               | Georgetown | Brownwood | Falcon  | Conroe |
| Balmorhea           | Inks    | Cedar Creek | Lake Creek | Cypress Springs |
| Baylor Creek        | Lyndon B. Johnson | Farmers Creek | Tradinghouse | Fayette County |
| Buffalo Springs     | Medina  | Grapevine | Hawkins | |
| Champion Creek      | Oak Creek | Hords Creek | Holbrook | |
| Clyde               | Travis  | Hubert H. Moss | Houston County | |
| Colorado City       | Whitney | Laron   | Jacksonville | |
| Daniel              |         |         | Lake Fork | |
| Greenbelt           |         |         | Lake o’ the Pines | |
| Hubbard Creek       |         |         | Long | |
| J. B. Thomas        |         |         | Martin Creek | |
| Kemp                |         |         | Monicello | |
| Kickapoo            |         |         | Murvaul | |
| Kirby               |         |         | Nacogdoches | |
| Mackenzie           |         |         | Palestine | |
| McClellan           |         |         | Pinkston | |
| Meredith            |         |         | Quitman | |
| Miller Park         |         |         | Sam Rayburn | |
| Moss Creek          |         |         | Squaw Creek | |
| O. C. Fisher        |         |         | Steinhagen | |
| Pauline             |         |         | Teague | |
| Red Bluff           |         |         | Toledo Bend | |
| Rita Blanca         |         |         | Tyler East | |
| Scarborough         |         |         | Tyler West | |
| Tammen              |         |         | Welsh | |
| Valley Creek        |         |         | Wilkes | |
| Waskashachie        |         |         | Winnsboro | |
| White River         |         |         |         | |

tween the 19 environmental variables and the four canonical equations were based on 52 observations. Therefore, individual correlations of at least 0.27 were significantly different from zero (P ≤ 0.05; Steel and Torrie 1980). Although each canonical equation is a function of all variables, the level of significance provides an objective minimum criterion for determining the most important variables in each canonical variable. However, a correlation of ±0.27 accounts for only 7% (r² = 0.07) of the within-group variation. Therefore, correlations of at least 0.31 (≥10% of within-group variation) in this study were used to interpret the most important variables contributing to separation of species associations.

Environmental Analysis of Species Associations

The 19 environmental variables chosen for analysis of species associations illustrate the wide range of limnological and morphometric characteristics among Texas reservoirs (Table 4). Of the 104 reservoirs classified into five species associations by cluster analysis, 52 had complete data for all variables. Canonical correlation analysis of these 52 reservoir surveys yielded four canonical variables with correlations significantly greater than zero (P ≤ 0.05; Table 4).

The 76 within-group correlations generated between the 19 environmental variables and the four canonical equations were based on 52 observations. Therefore, individual correlations of at least 0.27 were significantly different from zero (P ≤ 0.05; Steel and Torrie 1980). Although each canonical equation is a function of all variables, the level of significance provides an objective minimum criterion for determining the most important variables in each canonical variable. However, a correlation of ±0.27 accounts for only 7% (r² = 0.07) of the within-group variation. Therefore, correlations of at least 0.31 (≥10% of within-group variation) in this study were used to interpret the most important variables contributing to separation of species associations.

Only nine variables in the canonical correlation were important for describing species associations in the five reservoir groups. The first canonical equation illustrated the distribution of species associations among reservoirs in relation to water chemistry and surface elevation (Table 4). For example, Group 5 species associations would more likely occur in lower-elevation reservoirs with low sample sizes. Over the period covered by these surveys, the changes in species associations were not related to the length of time between surveys (Table 3). In reservoirs where species associations shifted between surveys, changes were due to a collective increase in the relative abundance of key species, with a corresponding decrease in the most abundant species from the earlier survey (Figure 2).
FIGURE 2.—Relative abundances (% of total number) of key fish species in four reservoirs with multiple surveys that were not classified into the same species association as their most recent survey used in this study. The bottom species in each group was the most abundant species from the earlier survey, but did not contribute to group separation in the present analysis.

hardness, pH, total alkalinity, and conductivity (Figure 3). Other species associations were uncommon in reservoirs with similar environmental characteristics.

The second canonical equation illustrated the separation of Group 1 species associations from Groups 2 and 3 in relation to growing season and pH (Table 4; Figure 3). The third described the separation of Group 4 species associations along growing season and turbidity gradients (Table 4; Figure 3). The separation of Groups 2 and 3 along reservoir depth (deeper in Group 2) and temperature (higher in Group 3) was described by the fourth canonical equation (Table 4).

Discriminant functions analysis was used to measure the ability of the nine important environmental variables to predict the species association in 82 reservoirs. The full-model discriminant analysis classified 90% of the reservoirs into their original species associations obtained from the earlier cluster analysis. However, this probability is upwardly biased for predicting species associations in unclassified reservoirs because it is derived from a model that used all of the data. Leave-out-one cross-validation correctly classified 82% of the reservoirs into their original species association (Table 5). Predictive success of the model ranged from 67% in Group 4 to 88% in Group 3 (Table 5).

Application of the Classification Scheme

Of the 132 reservoirs used to obtain the present classification scheme, electrofishing surveys were conducted on 57 in 1987. Electrofishing surveys from an additional 21 reservoirs not in the classification model were included with Group 6 reservoirs (unclassified) for a total of 78 reservoirs with electrofishing data. Within-group variances of electrofishing data were not homogeneous (P < 0.01). This indicated that a stratified sampling design that used groups as strata would yield a lower variance of \( \bar{C} \) estimates compared with that from a simple random sample of reservoirs.

Both sampling designs produced unbiased es-

| Reservoir group | 1 | 2 | 3 | 4 | 5 | Total or mean |
|-----------------|---|---|---|---|---|---------------|
| Number of reservoirs with multiple surveys | 3 | 7 | 4 | 4 | 11 | 29 |
| Previous surveys matching current reservoir classification | | | | | | |
| Number | 2 | 6 | 2 | 4 | 8 | 22 |
| Percent | 67 | 86 | 50 | 100 | 73 | 76 |
| Mean number of years between reservoir surveys when | | | | | | |
| Classification changed | 1.0 | 7.0 | 4.0 | ND | 3.3 | 3.7 |
| ND | ND | (3.0) | ND | (1.5) | (1.1) | |
| Classification was retained | 5.5 | 4.8 | 6.5 | 3.5 | 4.0 | 5.0 |
| (0.5) | (0.7) | (1.5) | (0.6) | (1.0) | (0.5) | |
CLASSIFICATION OF TEXAS RESERVOIRS

Table 4. Summary statistics for environmental variables of Texas reservoirs, 1975–1984, and results of canonical correlation analysis of species associations obtained by cluster analysis. Within-group correlations between the canonical variables and the environmental variables are based on 52 reservoirs with complete data. Important variables describing each canonical variable are underlined.

| Variable                        | Number of reservoirs | Mean | Range                  | Canonical variable |
|---------------------------------|----------------------|------|------------------------|--------------------|
| Reservoir age (years)           | 155                  | 31   | 5–84                   | 0.156              |
| Drainage area (km²)             | 141                  | 13,528 | 0–421,074              | -0.044             |
| Surface area (hectares)         | 155                  | 3,698 | 40–73,522              | 0.109              |
| Surface elevation (m)²          | 138                  | 271   | 2–1,177                | 0.055              |
| Storage (hectare-meters)        | 144                  | 30,634 | 81–552,462             | 0.401              |
| Shoreline length (km)           | 119                  | 146   | 3–1,920                | 0.123              |
| Growing season                 | 141                  | 242   | 185–327                | 0.095              |
| Submerged vegetation (%)        | 132                  | 11    | 0–100                  | -0.018             |
| Mean depth (m)                  | 147                  | 5.7   | 0.9–18.9               | -0.039             |
| Thermocline depth (m)          | 123                  | 5.6   | 0–30.5                 | -0.196             |
| Fluctuation (m)                 | 120                  | 1.3   | 0–6.1                  | -0.059             |
| Outlet depth (m)²              | 153                  | 10.5  | 0–57.9                 | -0.156             |
| Temperature (°C)                | 155                  | 25.2  | 19.0–32.1              | -0.004             |
| Dissolved oxygen (mg/L)         | 155                  | 6.0   | 3.3–9.0                | 0.139              |
| pH                              | 155                  | 7.2   | 6.0–9.1                | 0.333              |
| Total alkalinity (mg/L as CaCO₃) | 155             | 121.7 | 5.8–460.5              | 0.429              |
| Total hardness (mg/L)           | 155                  | 188   | 18–1,617               | 0.376              |
| Conductivity (µS/cm)            | 155                  | 749   | 62–7,313               | 0.520              |
| Turbidity (JTUs)                | 155                  | 32.1  | 0.3–171.6              | 0.037              |

Canonical correlation

Proportion of variance
Cumulative proportion

Discussion

Canonical correlation analysis has been used to illustrate the separation of species or species associations across environmental dimensions (e.g., Cassie and Michael 1968; McIntire 1978; Poore and Mobley 1980). Among Texas reservoirs, canonical correlation illustrated a general east–west separation of species associations based on water chemistry, and a further separation from northeast to southeast by growing season and elevation. Inferences from this classification scheme are limited to the species and environmental variables I used in the model. Important species not included in the classification model were Morone spp., blue catfish Ictalurus furcatus, smallmouth bass Micropterus dolomieu, and walleye Stizostedion vitreum. These are primarily pelagic species whose relative densities are not accurately represented by cove-rotenone surveys. Smallmouth bass and walleye are introduced species restricted in the classification model were Micropterus dolomieu, and walleye Stizostedion vitreum.

Discussion

Canonical correlation analysis has been used to illustrate the separation of species or species associations across environmental dimensions (e.g., Cassie and Michael 1968; McIntire 1978; Poore and Mobley 1980). Among Texas reservoirs, canonical correlation illustrated a general east–west separation of species associations based on water chemistry, and a further separation from northeast to southeast by growing season and elevation. Inferences from this classification scheme are limited to the species and environmental variables I used in the model. Important species not included in the classification model were Morone spp., blue catfish Ictalurus furcatus, smallmouth bass Micropterus dolomieu, and walleye Stizostedion vitreum. These are primarily pelagic species whose relative densities are not accurately represented by cove-rotenone surveys. Smallmouth bass and walleye are introduced species restricted in distribution to a few reservoirs as a result of stocking programs by the Texas Parks and Wildlife Department.

Cove-rotenone surveys give only an instantaneous picture of community composition; therefore, seasonal species-abundance data were not available for describing changes in community as-
Cluster analysis of data from earlier surveys classified most reservoirs into the same species associations as occurred for the most recent survey. This suggests that summer species associations are relatively constant among years in most Texas reservoirs. Changes in species compositions between multiple surveys were attributed to a shift in relative abundance of key species that contributed to the original group classification. Therefore, fishery managers should be able to detect real changes in species composition that may require closer investigation and possible changes in management practices.

Any classification model should include an unbiased estimate of the error associated with classifying new observations. The classification model developed for Wisconsin lakes by Tonn et al. (1983) had an implied classification error rate of 0% but an unbiased rate of 14% estimated by jackknife techniques (W. M. Tonn, University of Al-

---

**Table 5.** Classification summary from leave-out-one cross-validation of the discriminant functions model used to predict species associations in Texas reservoirs, based on nine environmental variables.

| Original classification group | Cross-validation classification group | Total | Number correct (%) |
|-------------------------------|--------------------------------------|-------|--------------------|
|                              | 1  2  3  4  5                        |       |                    |
| 1                             | 19 1 3 0 1                          | 24    | 19 (79)            |
| 2                             | 0  6 2 0 8                          | 8     | 6 (75)             |
| 3                             | 2  0 22 0 25                         | 22    | 22 (88)            |
| 4                             | 0  0 1 2 3                          | 3     | 2 (67)             |
| 5                             | 0  1 3 0 18                         | 22    | 18 (82)            |
| Total                         | 21 8 31 2 20                        | 82    | 67 (82)            |
Table 6.—Mean catch per effort (C/f) and variance for bluegill and largemouth bass collected during electrofishing surveys in Texas reservoirs. Estimates are based on 200 bootstrap samples of 30 reservoirs chosen at random from 78 reservoirs surveyed by electrofishing in 1987. Strata were the reservoir groups observed from cluster analysis of species assemblages (see Table 2).

| Species       | Sample design or allocation | C/f  | Variance | Relative variation |
|---------------|-----------------------------|------|----------|--------------------|
| Bluegill      | Simple random               | 147.3| 679.2    | 100                |
|               | Stratified random           |      |          |                    |
|               | Equal                       | 143.9| 389.7    | 57                 |
|               | Optimal                     | 141.8| 408.9    | 60                 |
|               | Proportional                | 146.6| 455.4    | 67                 |
| Largemouth bass| Simple random               | 68.3 | 141.4    | 100                |
|               | Stratified random           |      |          |                    |
|               | Equal                       | 68.2 | 109.3    | 77                 |
|               | Optimal                     | 67.4 | 111.7    | 79                 |
|               | Proportional                | 68.9 | 126.1    | 89                 |

*a C/f = number of fish/hectare.

*b Under equal allocation, 29 reservoirs were sampled because Group 4 had only 4 reservoirs surveyed by electrofishing in 1987.

In the present study, the reservoir classification model had a biased error rate of 10% versus 18% obtained by cross-validation. These validation techniques illustrate the importance of obtaining unbiased error rates because the error rate obtained with a classification model that uses all of the data implies that the model is more precise than it actually is. In the absence of new data for testing the classification model, cross-validation or jackknife methods are satisfactory for estimating unbiased error rates for classifying reservoirs. Newly surveyed reservoirs can be classified into species associations with known reliability by these models, but the models' error rates may prompt fishery managers to search for more precise models.

My study demonstrates that an objective classification scheme can be derived from annual resource-monitoring data taken over large environmental and geographic ranges. As stated by Tonn et al. (1983), the prediction success of the model may diminish when it is used to classify other reservoirs that fall outside the range of the limnological variables in the original set of reservoirs. However, the goal in my study was to develop a comprehensive classification scheme for all major reservoirs in Texas. Therefore, because reservoirs are surveyed each year, the classification scheme could be updated with new survey data.

My model provides a strategy for monitoring statewide fishery trends based on surveys of representative reservoirs with each species association. For example, I found that a stratified sampling scheme based on my classification of Texas reservoirs improved the precision of statewide C/f estimates for bluegills and largemouth bass sampled by electrofishing.

Managers may use fishery variables other than species abundance to develop models with specific objectives. Classification models may also be used to allocate available management resources based on the number of reservoirs with each species association or to develop regional (within-group) objectives for fishery population characteristics such as growth, condition, and recruitment rates.

Acknowledgments

Funding was provided through Federal Aid in Fish Restoration Project F-30-R by the U.S. Fish and Wildlife Service and the Texas Parks and Wildlife Department. Manuscript reviews were provided by R. W. Luebke, W. C. Provine, P. P. Durocher, N. E. Carter, G. C. Matlock, L. E. Miranda, W. M. Tonn, and an anonymous referee. Special thanks to S. J. Gutreuter for his reviews and helpful suggestions at various stages of this project.

References

Cassie, R. M., and A. D. Michael. 1968. Fauna and sediments of an intertidal mud flat: a multivariate analysis. Journal of Experimental Marine Biology and Ecology 2:1-23.

Cochran, W. G. 1977. Sampling techniques, 3rd edition. Wiley, New York.

Durocher, P. P., W. C. Provine, and J. E. Kraai. 1984. Relationship between abundance of largemouth bass and submerged vegetation in Texas reservoirs. North American Journal of Fisheries Management 4:84-88.

Echelle, A. A., and G. D. Schnell. 1976. Factor analysis
of species associations among fishes of the Kiamichi River, Oklahoma. Transactions of the American Fisheries Society 105:17-31.

Efron, B., and G. Gong. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. American Statistician 37:36-48.

Farquhar, B. W., R. L. Noble, and C. E. Cichra. 1982. Factors affecting fish populations in flood prevention lakes of north-central Texas. Proceedings of the Annual Conference Southeastern Association of Fish and Wildlife Agencies 34:292-306.

Johnson, M. G., J. H. Leach, C. K. Minns, and C. H. Olver. 1977. Limnological characteristics of Ontario lakes in relation to associations of walleye (Stizostedion vitreum), northern pike (Esox lucius), lake trout (Salvelinus namaycush), and smallmouth bass (Micropterus dolomieui). Journal of the Fisheries Research Board of Canada 34:1592-1601.

McIntire, C. D. 1978. The distribution of estuarine diatoms along environmental gradients: a canonical correlation. Estuarine and Coastal Marine Science 6:447-457.

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

Miranda, L. E., and P. P. Durocher. 1986. Effects of environmental factors on growth of largemouth bass in Texas reservoirs. Pages 115-121 in G. E. Hall and M. J. Van Den Avyle, editors. Reservoir fisheries management: strategies for the 80's. American Fisheries Society, Southern Division, Reservoir Committee, Bethesda, Maryland.

Moore, P. G. 1974. The kelp fauna of northeast Britain. III. Qualitative and quantitative ordinations, and the utility of a multivariate approach. Journal of Experimental Marine Biology and Ecology 16:257-300.

Poore, G. C. B., and M. C. Mobley. 1980. Canonical correlation analysis of marine macrobenthos survey data. Journal of Experimental Marine Biology and Ecology 45:37-50.

SAS. 1985. SAS user's guide: statistics, version 5 edition. SAS Institute, Cary, North Carolina.

Steel, R. G. D., and J. H. Torrie. 1980. Principles and procedures of statistics: a biometrical approach, 2nd edition. McGraw-Hill, New York.

Tonn, W. M., J. J. Magnuson, and A. M. Forbes. 1983. Community analysis in fishery management: an application with northern Wisconsin lakes. Transactions of the American Fisheries Society 112:368-377.

Received August 8, 1988
Accepted December 29, 1989