*The World's Largest Open Access Agricultural & Applied Economics Digital Library*

# This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

# Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

*No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.*

*Journal of Agricultural and Applied Economics*, 29,2(December 1997):373–383
© 1997 Southern Agricultural Economics Association

# Multivariate Procedures for Identifying Rural Land Submarkets

**Gary A. Kennedy, Steven A. Henning, Lonnie R. Vandeveer, and Ming Dai**

ABSTRACT

Research has developed empirical models and procedures for analyzing factors which influence rural land markets; however, there have been limited efforts in developing procedures for identifying rural land submarkets. Principal component analysis is used here to detect the presence of multiple rural land submarkets. Cluster analysis is then used as a basis for identifying eight contiguous rural land submarkets in Louisiana. As opposed to single-attribute procedures that have been based largely on subjective judgment, multivariate procedures illustrated in this analysis provide a means for capturing the combined effects of physical and socioeconomic influences in delineating rural land submarkets.

**Key Words:** cluster analysis, land market analysis, principal component analysis, rural land submarkets.

Changing economic conditions within the agricultural production sector, along with an increasing demand for nonagricultural uses of land, have increased the interest in land market research. Important questions concern land market trends and the identification and measurement of factors operative in rural land markets. This research requires that rural land markets or submarkets be defined or identified. Classical spatial theory suggests a need to delineate relatively homogeneous areas for conducting market research (Bressler and King). However, no universally accepted method or technique to segment the aggregate rural land market into relatively homogeneous land submarkets has emerged. Similarly, empirical research suggests a need to identify rural land submarkets because values vary substantially in importance and direction by regional location (Elad, Clifton, and Epperson).

Several studies have classified rural land into submarkets based on a single characteristic such as soil or commodity type. Ramsey and Corty analyzed Louisiana rural land values by dividing the state into eight agricultural production areas, while Xu, Mittelhammer, and Barkley examined rural land values in Washington State using six substate regions based primarily on agricultural production patterns. Other studies have used factors such as population density, topography, or climate to define some degree of homogeneity in rural land markets (Spurlock and Adrian; Corty; Herr; Vollink).

Multivariate methods have been employed to identify rural land submarkets. One such method is automatic interaction detection

Gary A. Kennedy is an assistant professor in the Department of Agriculture, Northeast Louisiana University, and an adjunct assistant professor in the Department of Agricultural Economics and Agribusiness, Louisiana State University. Steven A. Henning, Lonnie R. Vandeveer, and Ming Dai are associate professor, professor, and former instructor, respectively, all in the Department of Agricultural Economics and Agribusiness, Louisiana State University.

(AID) (Clifton and Spurlock; Foster). A major limitation of AID is that it is considered appropriate only if there are a large number of explanatory variables and 200 or more observations (Jackson; Sonquist, Baker, and Morgan). Cluster analysis is another multivariate procedure that has been used in several economic studies (Cox, Siskin, and Miller; Solomon and Pyrdol; Fesenmaier, Goodchild, and Morrison) and it is expected to be useful in identifying rural land markets.

Cluster analysis, along with principal component analysis in this study, is extended to group 59 Louisiana parishes into a series of homogeneous rural land submarkets. Rural land submarkets are estimated using 13 physical and socioeconomic characteristics. Principal component analysis is used to test for groupings within the data, while cluster analysis is used to determine the number of groupings (submarkets) and the boundaries of rural land submarkets. These multivariate procedures are described in the next section, followed by a discussion of the variable selection, analysis results, and conclusions. Disaggregation of the rural land market into smaller and more homogeneous submarkets using multivariate procedures is expected to provide an improved basis for analyzing the variation in rural land prices across areas that are similar with respect to geographical, demographical, economic, and agricultural characteristics.

## Multivariate Procedures

Principal component analysis is a multivariate technique for examining relationships among several quantitative variables. It is useful in summarizing data and detecting linear relationships. Plots of principal components are especially valuable in exploratory data analysis (Manly). Principal component analysis is used in this analysis to explore relationships among physical and socioeconomic variables which are thought to be influential in identifying rural land submarkets.

Principal component analysis finds a set of orthogonal axes in the direction of greatest variance among individuals. This procedure extracts from a $p \times n$ data matrix (parishes $\times$ variables) a new set of $r \times n$, where $r$ is a set of newly derived components and $n$ is the original number of variables. These axes (components) are linear combinations of the original variables $X_1, X_2, \ldots, X_n$:

$$(1) \qquad C_i = a_{i1}X_1 + \ldots + a_{in}X_n, \qquad i = 1, 2, \ldots, n,$$

subject to the condition that

$$(2) \qquad a_{i1}^2 + a_{i2}^2 + \ldots + a_{in}^2 = 1.$$

The coefficients, $a$, which are called eigenvectors, are chosen so that the first principal component, $C_1$, has as large a variance as possible, and makes the maximum contribution to the total variance. If the constraint equation (2) is not introduced, the variance of $C_1$ could be increased by simply increasing any one of the $a_{in}$ values. The second component, uncorrelated with the first, makes the maximum contribution to the residual variance, and so on, until all of the variance is taken into account. The components are ranked in order according to the proportion of the total variance for which they account. If the original variables are highly correlated, a single principal component may express a large part of the total variation in the data. When the first two principal components account for a high percentage of variation in the data, a plot of individuals (parishes) against these two components is an effective, unbiased way to visually identify groupings within the data (Manly).

Although principal component analysis is useful in providing information on variable relationships, it may not be used to determine the number of rural land submarkets or to determine the boundaries of such submarkets. In this analysis, cluster analysis is employed to determine the number of rural land submarkets and to delineate submarket boundaries.

Cluster analysis is an analytical technique that can be used to identify subgroups of individuals or objects. The objective is to classify a sample of individuals (parishes) into a small number of mutually exclusive groups, based on the properties (physical and socioeconomic characteristics) of the individuals. Groups are not predefined. Clustering techniques measure some form of similarity or association of individuals to determine how the individuals should be grouped. This generally is achieved by relating the Euclidean distance between individuals or groups of individuals.

Because the agreement of different clustering methods supports the validity of clustering classifications, four clustering methods were used in identifying rural land submarkets. These methods include: (*a*) average linkage (Sokal and Michener), (*b*) complete linkage (Sorensen), (*c*) flexible-beta (Lance and Williams), and (*d*) Ward's minimum-variance (Ward). The following notation is used to describe these methods, with lowercase symbols generally pertaining to observations, and with uppercase symbols denoting clusters:

$x_i$ = the $i$th observation;

$\mathbf{x}_i$ = the $i$th row vector;

$C_K$ = the $K$th cluster, subset of $\{1, 2, \ldots, n\}$;

$N_K$ = the number of observations in $C_K$;

$\bar{\mathbf{x}}$ = sample mean vector;

$\bar{\mathbf{x}}_K$ = mean vector for cluster $C_K$;

$\|\mathbf{x}\|$ = Euclidean length of the vector $\mathbf{x}$, i.e., the square root of the sum of the squares of the elements of $\mathbf{x}$;

$W_K = \sum_{i \in C_K} \|\mathbf{x}_i - \bar{\mathbf{x}}_K\|^2$;

$B_{KL} = W_M - W_K - W_L \quad \text{if } C_M = C_K \cup C_L$;

$d(\mathbf{x}, \mathbf{y})$ = any distance or dissimilarity measure between observations or vectors $\mathbf{x}$ and $\mathbf{y}$;

and

$D_{KL}$ = any distance or dissimilarity measure between clusters $C_K$ and $C_L$.

With average linkage, the distance between two clusters is the average distance between pairs of observations, one in each cluster. This distance is defined by

$$D_{KL} = \sum_{i \in C_K} \sum_{j \in C_L} d(x_i, x_j)/(N_K N_L). \tag{3}$$

If $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|^2$, then

$$D_{KL} = \|\bar{\mathbf{x}}_K - \bar{\mathbf{x}}_L\|^2 + W_K/N_K + W_L/N_L. \tag{4}$$

The combinatorial function is

$$D_{JM} = (N_K D_{JK} + N_L D_{JL})/N_M. \tag{5}$$

Average linkage tends to join clusters with small variances and is slightly biased toward producing clusters with the same variance.

Complete linkage defines the distance between the two clusters as the maximum distance between an observation in one cluster and an observation in the other cluster. Distance here is defined by

$$D_{KL} = \max_{i \in C_K} \max_{j \in C_L} d(x_i, x_j). \tag{6}$$

The combinatorial formula is

$$D_{JM} = \max(D_{JK}, D_{JL}). \tag{7}$$

Complete linkage is strongly biased toward producing clusters with roughly equal diameters and can be severely distorted by moderate outliers.

With the flexible-beta method, the distance between two clusters is adjusted by a beta parameter. The combinatorial formula is

$$D_{JM} = (D_{JK} + D_{JL})(1 - \beta)/2 + D_{KL}\beta, \tag{8}$$

where $\beta$ is the value of the beta option. This method allows adjustment of the beta parameter for data with many outliers.

With Ward's minimum-variance method, the distance between two clusters is the ANOVA sum of squares between the two clusters added up over all of the variables. This distance is defined by

$$D_{KL} = B_{KL} = \|\bar{\mathbf{x}}_K - \bar{\mathbf{x}}_L\|^2/(1/N_K + 1/N_L). \tag{9}$$

If $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|^2/2$, then the combinatorial formula is

$$D_{JM} = [(N_J + N_K)D_{JK} + (N_J + N_L)D_{JL} - N_J D_{KL}]/(N_J + N_M). \tag{10}$$

At each generation, the within-cluster sum of squares is minimized over all partitions obtainable by merging two clusters from the previous generation. The sums of squares are easier to interpret when they are divided by the total sum of squares to give proportions of variance. Clusters are joined to maximize the likelihood at each level of the hierarchy under assumptions of a multivariate mixture, equal spherical covariance matrices, and equal sampling probabilities. Ward's method tends to join clusters with a small number of observations and is biased toward producing clusters with roughly the same number of observations. It is also sensitive to outliers.

In addition to the agreement of different clustering methods, other measures of validity include the agreement of other multivariate methods (e.g., principal component analysis), the demonstration of stability and robustness, and the agreement with existing classifications (Romesburg). A cluster classification is considered stable if it is not disturbed by the addition of further information (adding additional variables); it is considered robust if removal of one or two variables from the original data matrix does not result in major changes in the classification.

## Variable Selection

Although there does not appear to be widespread agreement in the literature as to exactly which variables define land market or "neighborhood" relationships (Dubin), variables selected to describe submarket areas generally are based on physical and socioeconomic features. The choice of variables to be considered has a large influence on the ultimate results of a multivariate analysis (Anderberg; Manly). Variables that are similar for all observations have limited discriminating power, whereas those with consistent differences from one subgroup to another have the ability to induce strong distinctions.

Due to significant differences in climate, soil types, and topography within Louisiana, agricultural production is largely specialized in particular regions of the state. Although productivity of the soil is expected to have a large impact on rural land markets, no continuous soil productivity variable was available for Louisiana. However, the production areas of certain commodities tend to relate implicitly to the productivity of soil. While commodities such as wheat, corn, and soybeans are grown throughout the state, others, such as cotton, rice, sugar cane, and timber, tend to be area specific. Socioeconomic variables such as changes in population, size of the rural labor force, unemployment, and income also influence the structure of Louisiana rural land markets. Socioeconomic variables were selected on their ability to allow grouping of parishes with similar characteristics, while not segmenting parishes on the basis of rural versus urban.

For this study, variables selected to delineate Louisiana rural land submarkets are presented in table 1. Secondary parish-level data for 59 of 64 parishes were collected on each variable from the *1992 Census of Agriculture* (U.S. Department of Commerce), the *1992 Statistical Abstract of Louisiana* (University of New Orleans), the Louisiana Department of Agriculture and Forestry, and the Louisiana Population Data Center (Louisiana State University and LSU Agricultural Center). Five parishes from the New Orleans metropolitan area were not included in this analysis because of limited agricultural activity as reported by the *1992 Census of Agriculture.*

Parish-level data are not expected to precisely define rural land market boundaries because changes in topography or soil type rarely occur along parish boundaries (Reiling and Wiegmann). However, since socioeconomic variables tend to be more homogeneous with respect to parish boundaries, parish-level data are expected to provide a generally useful indication of rural land submarket areas within the state.

## Analysis

Before performing a cluster analysis, it is necessary to consider scaling or transforming the variables, since variables with large variances tend to have a greater impact on the resulting clusters than those with small variances. Vari-

**Table 1.** Summary of Variables Selected, Louisiana Rural Land Submarket Analysis, 1994

| Variable | Definition |
|---|---|
| **Crop Variables:**[a] | |
| *PERCANE* | Acres of sugar cane harvested/total parish cropland acres |
| *PERCOTT* | Acres of cotton harvested/total parish cropland acres |
| *PERRICE* | Acres of rice harvested/total parish cropland acres |
| *PERCROP* | Total acres of cropland/parish acres of land in farms |
| *PERPAST* | Total acres of pastureland/parish acres of land in farms |
| **Timber Variables:**[b] | |
| *SOFTSAW* | Landowner income from sales of pine sawtimber ($) |
| *HARDSAW* | Landowner income from sales of hard sawtimber ($) |
| *SOFTPULP* | Landowner income from sales of pine pulpwood ($) |
| *HARDPULP* | Landowner income from sales of hard pulpwood ($) |
| **Socioeconomic Variables:**[c] | |
| *PERABOVE* | Population above poverty/total population for parish |
| *PERLABOR* | Agricultural labor/total labor force for parish |
| *PERPOP* | Percent population change, 1980–90 |
| *UNEMP* | Unemployment rate |

[a] U.S. Department of Commerce.
[b] Louisiana Department of Agriculture and Forestry.
[c] Louisiana State University and LSU Agricultural Center; University of New Orleans.

ables used in this study were standardized to remove any possible arbitrary effects of differing units of measurement. Variables standardized to a zero mean and unit variance are transformed as follows:

$$(11) \qquad z_i = (x_i - \bar{x})/SD_i, \qquad i = 1, 2, \ldots, n,$$

where $z_i$ is the standardized value of variable $i$, $x_i$ is the unstandardized value of variable $i$, $\bar{x}$ is the mean value of variable $i$, and $SD_i$ is the standard deviation of variable $i$. Applying equation (11) to the 13 original variables allows all characteristics to contribute more equally to the similarities among individuals (parishes); however, equal weighting may not be the most appropriate approach (Anderberg). Some of the original variables can be expected to be of greater importance than others in defining rural land submarket areas. Variables expected to have a greater influence or importance can be weighted more heavily by standardizing a higher standard deviation. While determining each variable's influence on the rural land market was largely subjective in nature, economic theory suggests that soil productivity (implicitly represented in the commodity variables selected) has a large influence in rural land markets. Therefore, selected commodity variables were weighted more heavily. Means and standard deviations for the original and standardized variables are presented in table 2.

Principal component analysis was used to initially explore relationships among the 13 physical and socioeconomic variables. Comparison of the eigenvalues of the correlation matrix and the proportion of the total variation accounted for by the 13 principal components in the analysis are presented in table 3.

The first two principal components account for more than 60% of the total variability. A plot of the 59 parishes against these two components provides a scatter diagram (figure 1) that allows identification of groupings within the data. The positions of the parishes in multidimensional character space have been projected onto the plane of the first and second principal components, so that figure 1 is the most informative display possible of the distribution of the parishes in two-dimensional character space. Each quadrant in figure 1 contains parishes whose first two principal components vary in magnitude and sign. For ex-

**Table 2.** Means and Standard Deviations for Original and Standardized Variables, Louisiana Rural Land Submarket Analysis, 1994

| | Original Variables | | Standardized Variables | |
|---|---|---|---|---|
| Variable | Mean | Standard Deviation | Mean | Standard Deviation |
| **Crop Variables:** | | | | |
| *PERCANE* | 10.1 | 20.6 | 0 | 5 |
| *PERCOTT* | 9.2 | 15.6 | 0 | 5 |
| *PERRICE* | 5.6 | 10.9 | 0 | 5 |
| *PERCROP* | 65.7 | 19.7 | 0 | 2 |
| *PERPAST* | 32.8 | 20.9 | 0 | 2 |
| **Timber Variables:** | | | | |
| *SOFTSAW* | 6,459,581 | 7,923,001 | 0 | 3 |
| *HARDSAW* | 661,792 | 618,524 | 0 | 1 |
| *SOFTPULP* | 1,563,971 | 2,003,332 | 0 | 1 |
| *HARDPULP* | 351,166 | 350,962 | 0 | 1 |
| **Socioeconomic Variables:** | | | | |
| *PERABOVE* | 72.5 | 7.9 | 0 | 2 |
| *PERLABOR* | 5.3 | 4.0 | 0 | 1 |
| *PERPOP* | −0.7 | 9.7 | 0 | 2 |
| *UNEMP* | 8.8 | 2.2 | 0 | 2 |

ample, the northeast quadrant contains those parishes whose first two principal components are positive, the southwest quadrant contains those parishes with both components negative, the northwest quadrant contains those with a positive first component and a negative second component, and the southeast quadrant contains those with a negative first component and a positive second component. Examination of figure 1 suggests that multiple groupings are present within the data. Because principal component analysis only provides an unbiased initial assessment of groupings within the data, cluster analysis is applied as a means to assess the possible number of submarkets and the delineation of submarket boundaries.

**Table 3.** Eigenvalues of the Correlation Matrix and the Proportion of Total Variance

| Principal Component | Eigen-value | Percent of Total Variance | Cumulative Percentage |
|---|---|---|---|
| 1 | 5.06 | 38.95 | 38.95 |
| 2 | 2.93 | 22.52 | 61.47 |
| 3 | 1.13 | 8.73 | 70.20 |
| 4 | 0.85 | 6.52 | 76.73 |
| 5 | 0.73 | 5.64 | 82.37 |
| 6 | 0.68 | 5.26 | 87.62 |
| 7 | 0.47 | 3.59 | 91.22 |
| 8 | 0.36 | 2.74 | 93.96 |
| 9 | 0.29 | 2.26 | 96.20 |
| 10 | 0.21 | 1.60 | 97.80 |
| 11 | 0.14 | 1.07 | 98.87 |
| 12 | 0.08 | 0.59 | 99.46 |
| 13 | 0.07 | 0.54 | 100.00 |

Clustering of Louisiana parishes into homogeneous rural land submarkets results from the combined influence of physical and socioeconomic characteristics. Although identification of the number of clusters present in a data set is somewhat subjective, the use of stopping rules introduces a degree of objectivity to the process. The cubic clustering criteria suggest the presence of eight clusters in the data. The reliability of the cubic clustering criteria was explored by Milligan and Cooper in a simulation study comparing the ability of a variety of stopping rules to identify the number of distinct clusters present in a data set. The cubic clustering criteria ranked seventh among the 30 stopping rules examined, erring more often on the side of too many rather than too few

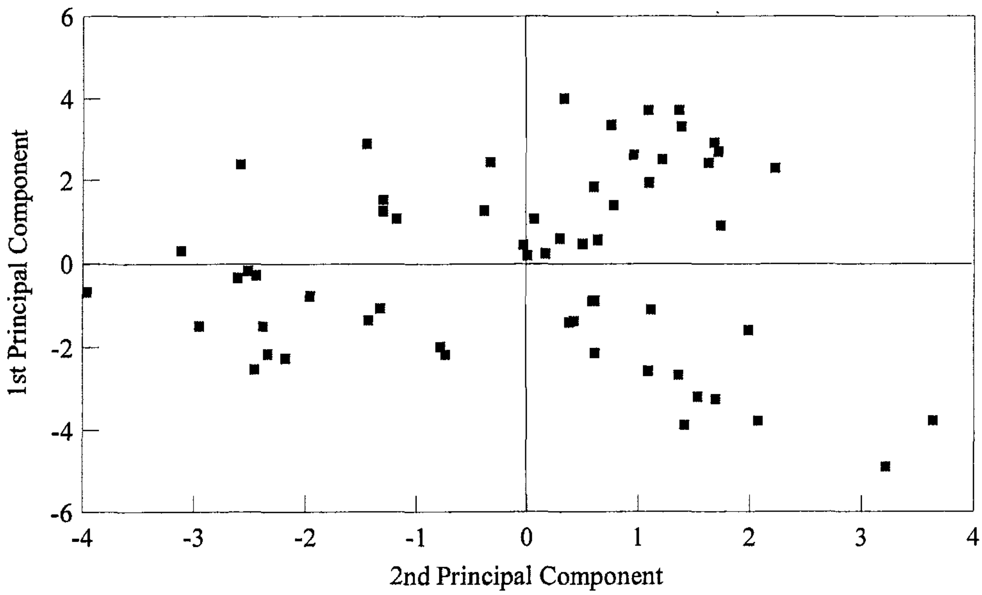


**Figure 1.** Location of parishes in character space defined by the first two principal components for the data, Louisiana rural land submarket analysis, 1994

clusters. For this study, the presence of eight clusters is geographically plausible, given the number of relatively homogeneous soil type areas used by Ramsey and Corty in a previous study.

The average linkage, complete linkage, flexible-beta, and Ward's minimum-variance clustering methods gave similar results. Examination of the results generated by each method suggested the presence of several distinct groups. The flexible-beta and Ward's minimum-variance methods provided identical results, illustrated in figure 2.

The application of the four clustering methods largely resulted in parishes being grouped contiguously. However, given the presence of eight clusters, comparison of the spatial locations of parishes with the cluster assignment illustrated in figure 2 reveals three anomalies. First, cluster 4 includes parishes in the north central as well as the western sections of the state. Parishes in these areas are on a gently sloping to hilly coastal plain, primarily produce pine and hardwood timber, and are similar with respect to socioeconomic characteristics. However, parishes in cluster 4 are separated by parishes of cluster 7, which consist of alluvial soils of the Red River and produce crops such as cotton and soybeans. Second, cluster 8 consists of Calcasieu and Cameron parishes in the southwest corner of the state, in addition to Lafayette which is noncontiguous to the former two. All three of these parishes are very similar to cluster 3 with respect to soils and commodities produced. However, because the parishes of cluster 8 have large urban areas and higher income levels, they differ from cluster 3 with respect to socioeconomic characteristics. The third anomaly is the fact that Madison Parish did not group with other parishes in the upper Mississippi Delta region (cluster 6).

Points identified by cluster groups in figure 2 allow a comparison to be made between the results of the cluster analysis and the plot of parishes against the values of the first two principal components that are illustrated in figure 3. The locations of parishes demonstrated strong clustering in the space defined by the first two principal components for the data. The relationships among parishes ranged from three-parish clusters to 12-parish clusters. The location of Madison Parish, identified in figure 3, illustrates Madison's close relationship to both clusters 5 and 6 in terms of physical and socioeconomic characteristics.

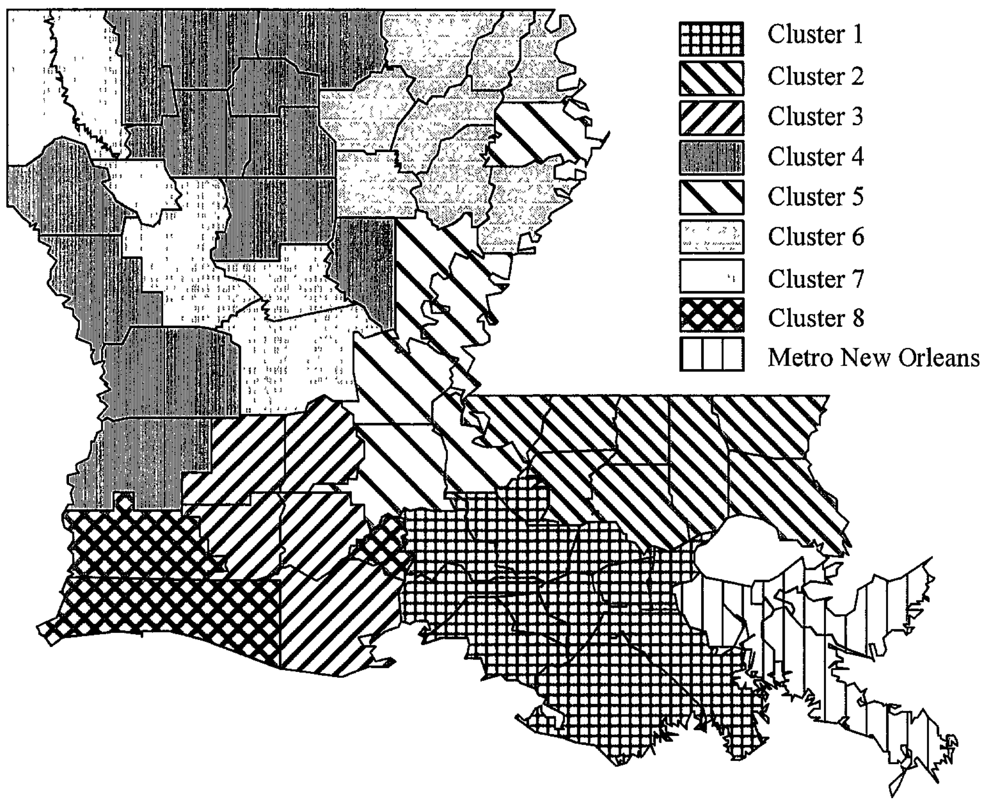


**Figure 2.** Clustering of parishes, Louisiana rural land submarket analysis, 1994

Cluster analysis is a useful tool in attempting to reveal the structure and relations within a data set. However, because the selection and weighting of variables used in the analysis and the selection of clustering method are largely subjective, results of a cluster analysis are more of an aid in exploring hypotheses about the data (Anderberg). Given eight discrete clusters for this study, the issue of boundary delineation among rural land submarkets cannot be resolved on the basis of cluster analysis alone.

The results of the principal component analysis and cluster analysis, along with additional information, were used to identify the Louisiana rural land submarkets illustrated in figure 4. Submarkets B, G, and H (figure 4) are consistent with clusters 7, 2, and 1, respectively (figure 2). Submarket H is comprised primarily of the sugar cane production area of the state and is homogeneous with respect to socioeconomic characteristics. Parishes in submarket G have limited agricultural production and are largely influenced by the Baton Rouge and New Orleans metropolitan areas. Submarket B contains parishes that are located along the Red River, with a mix of crop and timber production.

Submarkets A and C, identified in figure 4, are defined by the geographic separation of cluster 4 by cluster 7 (figure 2). Most of the parishes in the western portion of cluster 4 border the Toledo Bend Reservoir, a large lake primarily used for recreation. Although both areas of cluster 4 are large producers of soft and hardwood timber, landowners in the western area tend to receive slightly higher prices for soft timber and slightly lower prices for hardwood timber, as compared with parishes of the north central area of cluster 4 (Louisiana Department of Agriculture and Forestry).

Submarket E (figure 4) is formed by the combination of clusters 3 and 8 (figure 3). Although the parishes of cluster 8 were grouped separately into two noncontiguous areas due to socioeconomic factors, both areas are pri-

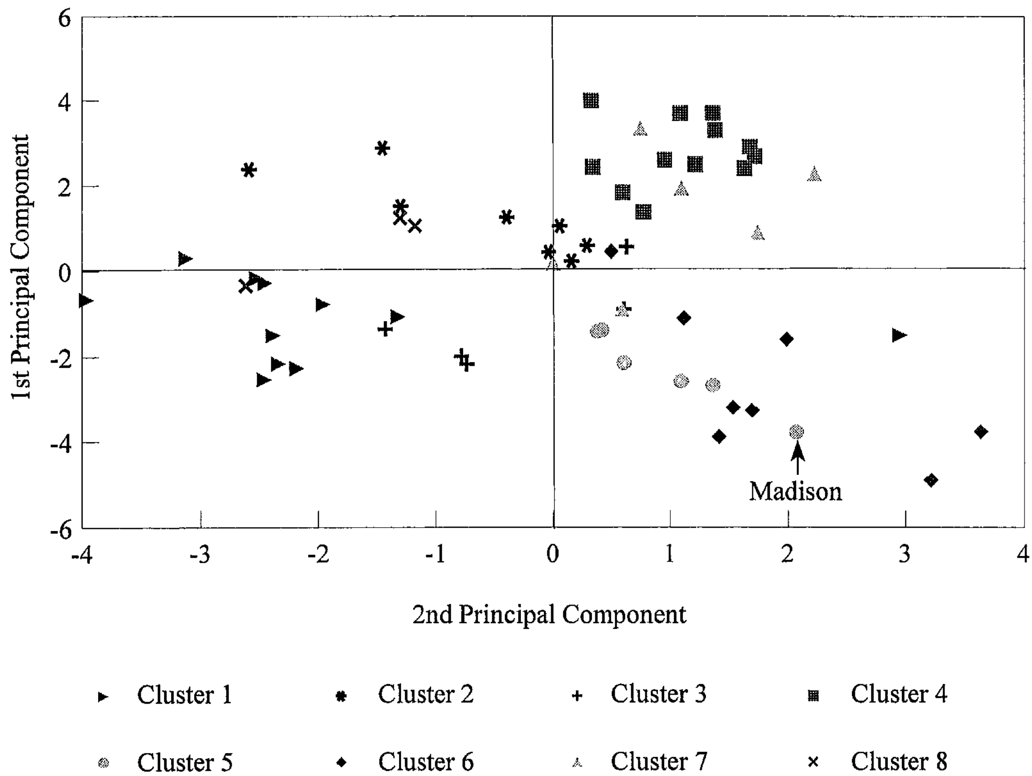


**Figure 3.** Parishes grouped by clusters in character space defined by the first two principal components for the data, Louisiana rural land submarket analysis, 1994

marily rice production regions, containing soils of the coastal prairies. While Madison Parish was grouped with cluster 5 (figure 2), its close relationship and contiguity with parishes in cluster 6 suggested that it should be included in submarket D in figure 4.

## Conclusions

Previous research has relied largely on subjective, single-attribute procedures to identify homogeneous rural land submarkets. In this analysis, objective procedures, which include the use of principal component and cluster analyses, are used to capture the combined effects of physical and socioeconomic characteristics in delineating rural land submarkets. Results of this study suggest rural land submarkets in Louisiana are well-formed, non-overlapping entities made up of parishes that are primarily contiguous.

Although some judgment is required in cluster analysis, several criteria may be used to evaluate the validity of the results. The general agreement and consistency of different clustering methods in grouping parishes into eight rural submarket areas indicates the attainment of well-structured clusters. The agreement of the plot of parishes against the first two principal components with the clustering method (figure 3) adds validity to the analysis. Stability in rural submarkets was also indicated because addition of relevant socioeconomic or physical variables to the analysis had little or no effect on the clustering outcome. Similarly, general removal of one or two variables from the original data matrix did not produce major changes in clustering results, which suggested that the classification was robust.

Results from procedures used in this analysis are consistent with rural land submarket areas used in previous research. One-half of the submarkets identified through our multivariate procedures correspond exactly to previous analyses, while one other area differs

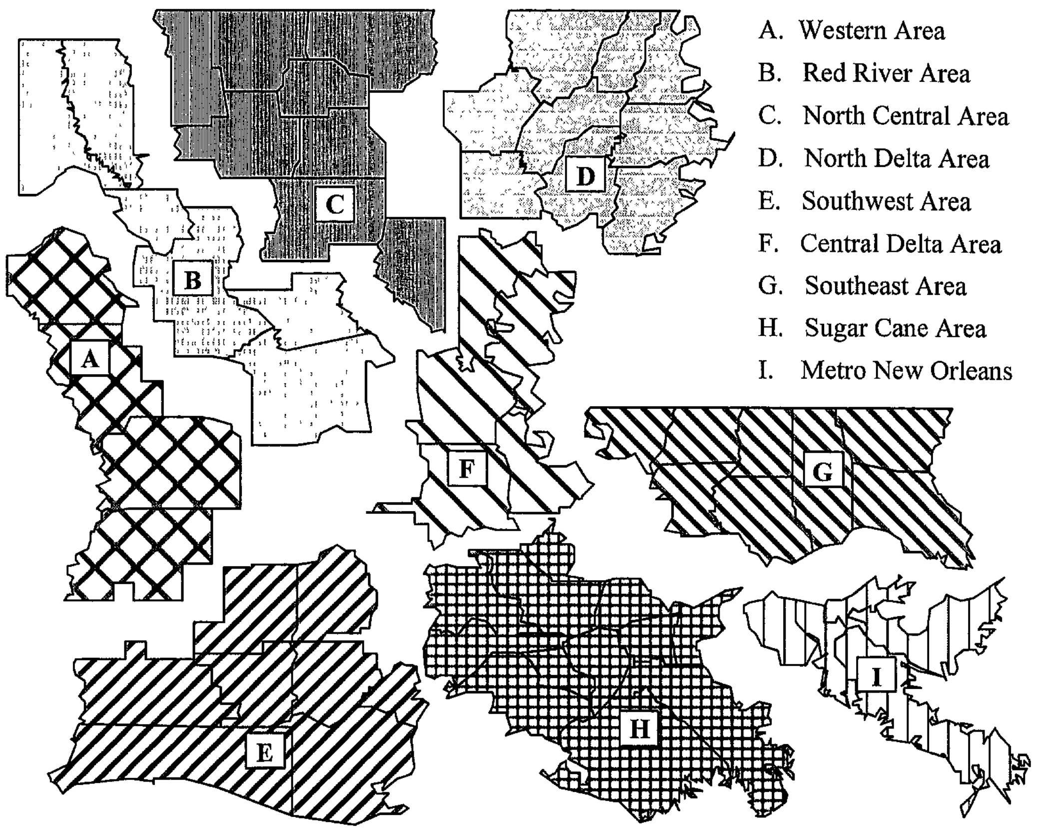


**Figure 4.** Identification of rural land submarkets, Louisiana rural land submarket analysis, 1994

from previous analysis by only one parish. Multivariate procedures led to a regrouping of parishes in the three remaining areas. Newly defined rural submarket areas reflect more homogeneous areas with respect to physical and socioeconomic characteristics, and generally reflect an improvement in the classification of rural submarkets for empirical analysis.

In general, results from this study indicate that the aggregate Louisiana rural land market can be viewed as a conglomerate of smaller rural land submarkets. Principal component analysis and cluster analysis provide a multivariate strategy that allows one to account for the combined effects of physical and socioeconomic characteristics in delineating rural land submarkets. These procedures are expected to be useful in conducting hedonic analyses and other empirical work requiring the definition of rural land submarkets. Multivariate procedures also are expected to be useful in identifying rural land submarkets in other states or areas of the country where rural submarket analyses are required. Procedures may be used to update rural land submarkets as new data become available and as economic conditions change.

## References

Anderberg, M.R. *Cluster Analysis for Applications.* New York: Academic Press, 1973.

Bressler, R.G., and R.A. King. *Markets, Prices, and Interregional Trade.* New York: John Wiley and Sons, 1970.

Clifton, I.D., and S. Spurlock. "Analysis of Variations in Farm Real Estate Prices over Homogeneous Market Areas in the Southeast." *S. J. Agr. Econ.* 15(July 1983):89–96.

Corty, F.L. "The Relationship of Farmland Values and Regional Population Densities." *Louisiana Rural Economist* 32(August 1970):12–16. Dept. Agr. Econ., Louisiana State University.

Cox, P.T., B. Siskin, and A. Miller. "A More Objective Procedure for Determining Economic Subregions: Cluster Analysis." *S. J. Agr. Econ.* 1(July 1969):37–44.

Dubin, R.A. "Estimation of Regression Coefficients in the Presence of Spatially Autocorrelated Error Terms." *Rev. Econ. and Statis.* 70(August 1988):466–74.

Elad, R.L., I.D. Clifton, and J.E. Epperson. "Hedonic Estimation Applied to the Farmland Market in Georgia." *J. Agr. and Appl. Econ.* 26(December 1994):351–66.

Fesenmaier, D.R., M.F. Goodchild, and S. Morrison. "The Spatial Structure of the Rural-Urban Fringe: A Multivariate Approach." *Canadian Geographer* 23(1979):255–65.

Foster, A.K. "An Economic Application of Hedonic Pricing to Georgia Farmland Submarkets." Unpub. M.S. thesis, University of Georgia, 1986.

Herr, W.M. "The Influence of Farm Mortgage Loan Terms on Farm Real Estate Values." *S. J. Agr. Econ.* 7(July 1975):153–58.

Jackson, B.B. *Multivariate Data Analysis: An Introduction.* Homewood IL: Irwin Publishing Co., 1983.

Lance, G.N., and W.T. Williams. "A General Theory of Classificatory Sorting Strategies." *Computer J.* 9(1967):373–80.

Louisiana Department of Agriculture and Forestry. "1994 Louisiana Timber and Pulpwood Production." Baton Rouge, May 1995.

Louisiana State University and LSU Agricultural Center. "1994 Louisiana Population Data Sheet." Louisiana Population Data Center, Baton Rouge, July 1994.

Manly, B.F. *Multivariate Statistical Methods: A Primer.* New York: Chapman and Hall, 1986.

Milligan, G.W., and M.C. Cooper. "An Examination of Procedures for Determining the Number of Clusters in a Data Set." *Psychometrika* 50(1985):159–79.

Ramsey, A.F., and F.L. Corty. "The Rural Land Market and Rural Real Estate Values in Louisiana, 1974." DAE Res. Rep. No. 504, Louisiana State University Agricultural Center, Louisiana State University, Baton Rouge, 1976.

Reiling, S.D., and F.H. Wiegmann. "Louisiana Agriculture: Economic Trends and Current Status, 1940–1977." Bull. No. 718, Agr. Exp. Sta., Louisiana State University, Baton Rouge, 1979.

Romesburg, C.H. *Cluster Analysis for Researchers.* Belmont CA: Lifetime Learning Publications, 1984.

Sokal, R.R., and C.D. Michener. "A Statistical Method for Evaluating Systematic Relationships." In *University of Kansas Science Bulletin,* No. 38, 1958, pp. 1409–38.

Solomon, B.D., and J.J. Pyrdol. "Delineating Coal Market Regions." *Economic Geography* 62(April 1986):111–24.

Sonquist, J.A., E.L. Baker, and J.M. Morgan. *Searching for Structure.* Ann Arbor MI: Institute for Social Research, the University of Michigan, 1973.

Sorensen, T. "A Method of Establishing Groups of Equal Amplitude in Plant Sociology Based on Similarity of Species Content and Its Application to Analyses of the Vegetation on Danish Commons." *Biologiske Skrifter* 5(1948):1–34.

Spurlock, S.R., and F.L. Adrian. "Variations in Rural Land Values in the Wiregrass Region of Alabama." Bull. No. 504, Agr. Exp. Sta., Auburn University, 1978.

U.S. Department of Commerce. *1992 Census of Agriculture,* Louisiana State and Parish Data (AC-A-18). USDC/Economic and Statistics Administration, Washington DC, September 1994.

University of New Orleans. *1992 Statistical Abstract of Louisiana,* 9th ed. Div. Business and Econ. Res., College of Business Administration, New Orleans, 1994.

Vollink, W.J. "Analysis of Factors Related to Per Acre Prices of Bare Land in North Carolina, 1975–1976." *S. J. Agr. Econ.* 10(July 1978): 15–16.

Ward, J.H. "Hierarchical Grouping to Optimize an Objective Function." *J. Amer. Statis. Assoc.* 58(1963):236–44.

Xu, F., R.C. Mittelhammer, and P. Barkley. "Measuring the Contributions of Site Characteristics to the Value of Agricultural Land." *Land Econ.* 69(1993):356–69.