



## METODOLOGIA COMPUTACIONAL PARA CLASSIFICAÇÃO NO SETOR PRIMÁRIO

<sup>1</sup>Emerson M. Furtado, <sup>2</sup>A. Chaves Neto, <sup>3</sup>Z. H. Domingues, <sup>4</sup>R. T. Hosokawa

**RESUMO:** A constituição de 1988 da República Federativa do Brasil estabelece que cada estado da federação pode definir critérios de distribuição dos recursos do ICMS – Imposto sobre Circulação de Mercadorias e Serviços – a que os municípios do Estado têm direito. O Estado do Paraná, pioneiramente, definiu pela Lei Complementar N.º 59/91, conhecida como Lei do ICMS Ecológico, que 2,5% dos recursos sejam destinados aos municípios que possuem Áreas Especialmente Protegidas, caracterizadas pelas Unidades de Conservação. O Instituto Ambiental do Paraná – IAP – realiza anualmente a avaliação do nível de qualidade de conservação das Áreas Especialmente Protegidas. O objetivo deste trabalho foi construir uma metodologia, para a partir de variáveis definidas pelo IAP, estabelecer um ranking classificatório de todas as áreas avaliadas. Além disto, a metodologia foi automatizada por meio de um programa computacional em linguagem Visual Basic a ser utilizado pelo IAP. Na solução para o problema de classificação das Áreas Especialmente Protegidas foram utilizados escores obtidos através da técnica estatística multivariada Análise Fatorial. O trabalho descreve o procedimento estatístico, e apresenta os resultados obtidos com a aplicação do programa.

**Palavras-chave:** classificação, áreas especialmente protegidas, análise fatorial.

## COMPUTATIONAL METHODOLOGY FOR CLASSIFICATION IN PRIMARY SECTOR

**ABSTRACT:** The 1988 constitution of the Brazil Federative Republic establishes that each state of the federation can define distribution approaches of the ICMS resources - Tax on Circulation of Goods and Services - the one that the municipal districts of the State are entitled. The State of Paraná, pioneerly, defined for the Complemental Law N.º 59/91, known as Law of Ecological ICMS, that 2,5% of the resources are distributed to municipal districts that possess especially protected areas, characterized by the Units of Conservation. The Environmental Institute of Paraná - EIP - it accomplishes annually the evaluation of the conservation quality level of the Areas Especially Protected. The evaluation facilitates one it reviews financial to the municipal district. The objective of this work went build a methodology, for starting from variables defined by EIP, to establish a classificatory ranking of all the appraised areas. Besides, the methodology was automated by means of a computerized program in Visual Basic language to be used by EIP. In the solution for the problem of classification of the Areas Especially Protected were used scores obtained through of the multivariate statistical technique Fatorial Analysis. This work describes the statistical procedure, and shows the results obtained with the application of the software.

**Key words:** classification, especially protected areas, factorial analysis

---

<sup>1</sup>Matemático, Professor – PpgMNE/UFPR – UnicenP – R.: Des. Motta, 2219, ap. 704 – CEP: 80420-190 – Curitiba – Paraná – Brasil – e-mail: efurtado@bsi.com.br, <sup>2</sup>Eng. Civil e Matemático, Professor e Pesquisador – PpgMNE/DEST-UFPR, <sup>3</sup>Eng. Florestal – PpgMAD/UFPR, <sup>4</sup>Eng. Florestal, Professor e Pesquisador – PpgEngFlo/UFPR

## 1 INTRODUÇÃO

Segundo a Federação brasileira, cada estado pode definir um perfil de distribuição dos recursos do ICMS – Imposto sobre Circulação de Mercadorias e Serviços – a que os municípios têm direito. O Estado do Paraná, pioneiramente, definiu pela Lei Complementar N.º 59/91, conhecida como Lei do ICMS Ecológico [8], que 5% dos recursos devem ser distribuídos segundo um critério ambiental. Por essa lei, metade dos 5% são distribuídos aos municípios que possuem em seus territórios mananciais, e a outra metade, aos municípios que possuem Áreas Especialmente Protegidas, caracterizadas pelas Unidades de Conservação. O Instituto Ambiental do Paraná – IAP realiza anualmente a avaliação do nível de qualidade de conservação das Áreas Especialmente Protegidas. A avaliação possibilita um repasse financeiro ao município.

O objetivo deste trabalho foi o de construir um algoritmo e programá-lo para a partir das variáveis definidas em conjunto com o IAP, estabelecer um *ranking* classificatório de todas as áreas avaliadas. Neste caso, as áreas especialmente protegidas são faxinais do estado do Paraná. A importância histórica, sócio-econômica e cultural dos faxinais pode ser encontrada em CHANG [1].

O programa é absolutamente geral, servindo para a elaboração de qualquer tipo de classificação baseada em critérios quantificáveis. Na solução desse problema de classificação foram utilizados escores fatoriais obtidos da técnica estatística multivariada conhecida como Análise Fatorial. Este trabalho descreve na seção 2, de Material e Métodos, o procedimento estatístico multivariado, o processo de coleta e compilação dos dados e o programa. Na seção 3, de Resultados e Discussão, é apresentada a classificação e são discutidos os resultados obtidos. Finalmente, faz-se uma conclusão na seção 4.

## 2 MATERIAL E MÉTODOS

### 2.1 Introdução

Na abordagem do problema de classificação quanto a qualidade de Áreas Especialmente Protegidas trabalhou-se com 80 variáveis quantitativas e qualitativas. Essas variáveis são de interesse do Instituto Ambiental do Paraná – IAP e a sua definição foi feita de acordo com os técnicos. Os valores foram obtidos através de pesquisa de campo realizada por DOMINGUES [2].

A metodologia aplicada para organizar os dados e obter a classificação foi a Análise Multivariada, especificamente trabalhou-se os escores obtidos da Análise Fatorial. Assim, cada uma das 20 áreas especialmente protegidas (faxinais) é representada por um vetor de dimensão 80 com as variáveis definidas como componentes. Aplicando-se a Análise Fatorial a esses vetores obtém-se 17 fatores que representaram um grau de explicação de 97,8%. Desta maneira, reduz-se a dimensão do problema que antes dependia de 80 variáveis e passou, então, a depender apenas de 17, perdendo pouco mais de 2% da informação da matriz de dados. Prosseguindo-se com a metodologia multivariada, obtém-se um escore univariado para cada uma das áreas em cada um desses 17 fatores que resumem a informação. Ponderando-se, adequadamente, esses 17 fatores, obtém-se um escore único para representar a área. Esses 20 escores, são escalonados adequadamente ao uso do IAP. Nas seções seguintes tem-se um maior detalhamento.

### 2.2 A obtenção da matriz de dados de entrada no algoritmo

A aplicação que motivou este trabalho foi feita em amostra obtida do universo de faxinais de todos os municípios do Estado do Paraná com este tipo de manejo agropastoril. O grau de confiança estimado foi de 90% e a precisão especificada na estimação dos parâmetros foi de 3%. O número de famílias entrevistadas, para obtenção dos valores das variáveis, foi dimensionado partindo-se da estimação do parâmetro de uma variável aleatória Bernoulli. E, definido por:

$$n = \frac{0,25N}{0,25 + \frac{d}{z}(N-1)} \quad (1)$$

onde:  $n$  é o número de amostras coletadas do universo;  $N$  é o número total de famílias;  $d$  é a precisão da estimativa e  $z$  é corresponde  $\Phi^{-1}(z)$  na distribuição normal padrão. O número de famílias a serem entrevistadas em cada Faxinal foi dimensionado, segundo a amostragem estratificada proporcional ou seja de acordo com a seguinte formulação:

$$n_i = \frac{n.N_i}{N} \quad (2)$$

na qual:  $n_i$  é o número de amostras a serem coletadas no Faxinal  $i$ ;  $N_i$  é o número de amostras a serem coletadas do

universo;  $N_i$  é o número de famílias do Faxinal  $i$  e  $N$  é o número total de famílias do universo pesquisado.

Os dados foram obtidos por DOMINGUES [2] e coletados em 4 municípios do Estado do Paraná. As variáveis foram definidas em conjunto com o IAP segundo critérios que envolvem meio ambiente, força de trabalho, desenvolvimento social, economia, educação, tecnologia, saúde, organização comunitária, agricultura, produção, infra-estrutura e compõem dois questionários de pesquisa. As principais variáveis consideradas foram: renda média; áreas de lavoura; áreas de floresta nativa; áreas de floresta plantada; áreas no criadouro comunitário; plantação de araucária; plantação de erva mate; conservação de solos; calagem de solo; sementes selecionadas; crédito agrícola; assistência técnica; participação em atividades comunitárias; luz de rede pública; esgoto com fossa séptica; rede pública para assistência médica; produtores alfabetizados; melhoramento genético da criação; conhecimento do ICMS Ecológico e densidade.

Os dados foram coletados por meio dos questionários em entrevistas com  $n = 316$  famílias amostradas em um universo de  $N = 1947$  famílias e referem-se ao ano agrícola de agosto de 1997 a julho de 1998. A partir os dados coletados e compilados, foi criada uma matriz de ordem  $20 \times 80$  que corresponde aos dados de entrada no algoritmo.

## 2.3 A Análise Fatorial

### 2.3.1 Introdução

A Análise Fatorial é uma técnica estatística multivariada, que procura explicar o relacionamento entre  $p$  variáveis aleatórias, que compõem um vetor, por meio de  $m$  fatores ( $m \ll p$ ) comuns a essas variáveis, sendo que esses fatores são combinações lineares das variáveis originais e obtidas a partir da matriz de covariância  $\Sigma$  ou da matriz de correlação  $\rho$  do vetor. Os fatores estão direcionados nos eixos de maior variabilidade. Inicialmente foi desenvolvida por Spearman com trabalhos intitulados “*General intelligence objectively determined and measured*”, de 1904 e “*The Abilities of Man*” de 1926. Mais tarde teve contribuições de psicólogos tais como Thomson, Thurstone e Burt, que a desenvolveram como uma técnica pioneira, tratando de hipóteses sobre a organização da habilidade mental, sugerida pelo exame das matrizes de correlação ou covariância para conjuntos de testes de variáveis cognitivas. A essencial proposta da Análise Fatorial é descrever, se possível, a estrutura

de covariância do relacionamento diferentes várias variáveis em termos de um número menor de variáveis não observáveis, denominadas fatores. Os fatores podem ser entendidos como tipos de indicadores que resumem e sintetizam grupos de variáveis. Isto é, certas variáveis que tem similaridades altas podem ser reunidas em um único grupo denominado fator que passa então a representá-las.

### 2.3.2 Modelo Fatorial Ortogonal e a determinação do número de fatores comuns

Sendo o vetor aleatório observável  $\underline{X}$ , com  $p$  componentes,  $E[\underline{X}] = \underline{\mu}$  e matriz de covariância  $V[\underline{X}] = \Sigma$ , tem-se que o chamado Modelo Fatorial postula que o vetor  $\underline{X}$  é linearmente dependente de variáveis aleatórias não observáveis  $F_1, F_2, \dots, F_m$ , chamadas fatores comuns e  $p$  fontes de variação  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ , chamadas erros ou fatores específicos. O Modelo Fatorial Ortogonal é

$$\begin{matrix} \underline{X} \\ (px1) \end{matrix} = \begin{matrix} \underline{\mu} \\ (px1) \end{matrix} + \begin{matrix} L \\ (pxm) \end{matrix} \cdot \begin{matrix} \underline{F} \\ (mx1) \end{matrix} + \begin{matrix} \underline{\varepsilon} \\ (px1) \end{matrix} \quad (3)$$

com  $m < p$  e, ainda  $\mu_i$  é a média da variável  $i$ ;  $i = 1, 2, 3, \dots, p$ ,  $\varepsilon_i$  é o  $i$ -ésimo fator específico,  $F_j$  é o  $j$ -ésimo fator comum;  $j = 1, 2, 3, \dots, m$ , e  $\lambda_{ij}$  é o carregamento da  $i$ -ésima variável no  $j$ -ésimo fator. Os não observáveis vetores aleatórios  $\underline{F}$  e  $\underline{\varepsilon}$  satisfazem as propriedades:  $\underline{F}$  e  $\underline{\varepsilon}$  são independentes;  $E(\underline{F}) = \underline{0}$ ,  $Cov(\underline{F}) = I$ ;  $E(\underline{\varepsilon}) = \underline{0}$ ,  $Cov(\underline{\varepsilon}) = \Psi$ , onde  $\Psi$  é uma matriz diagonal com a variância específica  $\psi_i$  na posição  $i$  da diagonal principal. Partindo-se das suposições do modelo, tem-se  $\Sigma = Cov(\underline{X}) = LL' + \Psi$  e  $Cov(\underline{X}, \underline{F}) = L$  e a porção da variância da variável aleatória  $X_i$ , advinda como contribuição dos  $m$  fatores comuns, é denominada comunalidade  $i$ , e denotada por  $h_i^2$ . Já a porção da variância desta variável atribuída ao fator específico  $\varepsilon_i$  é chamada de variância específica e é denotada por  $\psi_i$ . Em particular, a variável  $X_i$ , tem  $V(X_i) = (comunalidade\ i) + (variância\ específica\ i)$ . Logo,  $\sigma_i^2 = \sigma_{ii} = h_i^2 + \psi_i$ ,  $i = 1, 2, \dots, p$ . Maiores detalhes em RENCHER [10].

O número de fatores comuns fixado na construção do algoritmo desse estudo foi baseado no critério da igualdade entre o número de fatores  $m$  e o número de autovalores maiores que a média. Uma vez que a análise foi efetuada a partir da matriz de correlação amostral  $R$ , a média é 1. Este critério foi proposto por Guttman em 1954 e a sua eficácia foi atestada por Kayser, em trabalhos datados de 1960.

### 2.3.3 Estimação dos carregamentos dos fatores, rotação e escores fatoriais

O método adotado na estimação dos carregamentos dos fatores, foi o das Componentes Principais. A Análise de Componentes Principais conduzida a partir da matriz de correlação amostral  $R$  é feita em função dos pares de autovalor-autovetor de  $R$ , que são  $(\hat{\lambda}_1, \hat{e}_1), (\hat{\lambda}_2, \hat{e}_2), \dots, (\hat{\lambda}_p, \hat{e}_p)$ , onde  $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_p$ . Sendo  $m$  o número de fatores comuns, a matriz dos carregamentos fatoriais estimados  $\{\hat{\ell}_{ij}\}$  é dada por

$$\hat{L} = \begin{bmatrix} \sqrt{\hat{\lambda}_1} \cdot \hat{e}_{11} & \dots & \sqrt{\hat{\lambda}_2} \cdot \hat{e}_{21} & \dots & \sqrt{\hat{\lambda}_m} \cdot \hat{e}_{m1} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \sqrt{\hat{\lambda}_1} \cdot \hat{e}_{1n} & \dots & \sqrt{\hat{\lambda}_2} \cdot \hat{e}_{2n} & \dots & \sqrt{\hat{\lambda}_m} \cdot \hat{e}_{mn} \end{bmatrix} \quad (4)$$

As variâncias específicas são fornecidas pelos elementos diagonais de  $R - \hat{L}\hat{L}'$  (5). Desta forma,  $\hat{\psi}_i = 1 - \sum_{j=1}^m \hat{\ell}_{ij}^2$  (6) e as comunalidades são  $\hat{h}_i^2 = \hat{\ell}_{i1}^2 + \hat{\ell}_{i2}^2 + \dots + \hat{\ell}_{im}^2$  (7).

Os carregamentos fatoriais estão sujeitos a uma rotação de fatores, sem alteração das comunalidades e variâncias específicas, afim de se alcançar uma estrutura mais simples. O objetivo é o de facilitar a interpretação e, conseqüentemente, ter-se uma visão mais clara dos agrupamentos das variáveis nos fatores. Basicamente os carregamentos são multiplicados por uma matriz ortogonal  $T$  sem prejudicar a reprodução da matriz de correlação  $R = \hat{L}\hat{L}' + \hat{\Psi}$ . As comunalidades  $\hat{h}_i^2$ ,  $i = 1, 2, \dots, p$ , ficam inalteradas com a transformação  $L^* = LT$ . Neste trabalho, optou-se pelo método de rotação de Kayser denominado Método Varimax Normal [6].

Escores fatoriais são estimativas dos valores para os vetores fatoriais aleatórios não observáveis  $F_j$ ,  $j = 1, 2, \dots, n$ . A estimação é complicada pelo fato de que as quantidades  $f_j$  e  $\epsilon_j$  não são observadas como  $x_j$ . Como solução na estimação usou-se o Método dos Mínimos Quadrados Ponderados. JOHNSON & WICHERN [5] sugere outros métodos que poderiam ser utilizados. Tomando-se  $\hat{L}, \hat{\Psi}$  e  $\hat{\mu} = \bar{x}$ , o  $j$ -ésimo caso fica

$$\hat{f}_j = \begin{pmatrix} \hat{L}' \hat{\Psi} & \hat{L}' \\ \hat{L}' \hat{\Psi} & \hat{L}' \end{pmatrix}^{-1} \cdot \begin{pmatrix} \hat{L}' \hat{\Psi} \\ \hat{L}' \end{pmatrix} \cdot (x_j - \bar{x}) \quad (8)$$

Maiores detalhes da técnica Análise Fatorial podem ser encontrados em FURTADO [3], RENCHER [9] e LAWLEY [7].

### 2.3.4 Aplicação da metodologia desenvolvida

A compreensão de todos os passos, desde os dados brutos, até os escores classificatórios finais pode ser visualizada no Fluxograma a seguir (Figura 1).

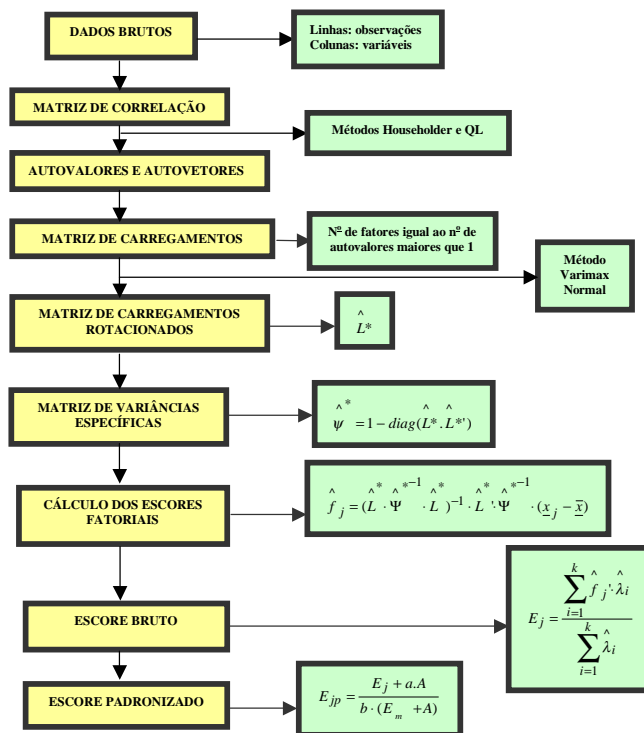


Figura 1 – Fluxograma do cálculo do escore padronizado

A partir dos dados coletados, foi criada uma matriz de dados de ordem 20x80. Dessa matriz, calculou-se as correlações existentes entre cada par de variáveis, criando-se a matriz de correlação. Em seguida, determinou-se os pares de autovalores e autovetores dessa matriz através dos Métodos de Householder e QL [11]. Foram considerados 17 fatores que representaram um grau de explicação de 97,8%. A matriz de carregamentos fatoriais,  $L$ , foi formada e os fatores foram, então, sem alteração do grau de explicação, rotacionados pelo Método Varimax Normal, sendo obtida, assim, a matriz de carregamentos fatoriais rotacionados. Em seguida, determinou-se as comunalidades e as variâncias específicas de cada variável. Observou-se que as comunalidades eram todas com valores próximos de 1 e, por outro lado, as variâncias específicas com valores próximos de 0, indicando portanto que as variáveis consi-

deradas eram todas adequadas ao modelo fatorial. Os escores fatoriais,  $f_j$ ,  $j=1,2,\dots,n=20$ , foram estimados pelo Método de Mínimos Quadrados Ponderados. A seguir calculou-se os escores classificatórios brutos  $E_j$  para os 20 faxinais ponderando-se os escores fatoriais originais pela importância de cada fator, representada pelo seu respectivo autovalor,

$$E_j = \frac{\sum_{i=1}^k \hat{f}_j \cdot \hat{\lambda}_i}{\sum_{i=1}^k \hat{\lambda}_i}, j = 1, 2, 3, \dots, 20 \quad (9)$$

A classificação final é estabelecida com base no escore padronizado correspondente ao escore bruto  $E$ , mas em uma escala especificada pelo usuário. Esta escala é obtida em função do escore bruto  $E$ , da amplitude  $A$  e dos parâmetros  $Z_1$  e  $Z_2$  especificados pelo usuário. Então, tem-se:

$$E_p = \frac{E + a.A}{b.(E_{m\acute{a}x} + A)} \quad (10)$$

onde  $A = e_{m\acute{a}x} - e_{m\acute{i}n}$ , e os parâmetros  $a$  e  $b$  são obtidos a partir de:

$$a = \frac{1,2Z_1 E_{m\acute{a}x} - 0,9Z_2 E_{m\acute{i}n}}{A.(0,9Z_2 - 1,2Z_1)} \quad (11)$$

$$b = \frac{A}{(0,9Z_2 - 1,2Z_1).(E_{m\acute{a}x} + A)} \quad (12)$$

O escore padronizado final é dado por:

$$E_{jp} = \frac{E_j + a.A}{b.(E_{m\acute{a}x} + A)}, j = 1, 2, 3, \dots, 20 \quad (13)$$

## 2.4 Descrição do Programa

O programa foi construído em linguagem *Visual Basic* e em ambiente *Windows*. Mais detalhes sobre a linguagem de programação podem ser encontrados em HALVORSON [4].

A entrada de dados no algoritmo desenvolvido e programado neste estudo é completamente geral e utiliza uma matriz inicial de entrada em que nas linhas, tem-se as observações e, nas colunas, as variáveis. Pode ser feita via teclado, por meio de banco de dados em formato

*Microsoft Access* ou até mesmo através da planilha eletrônica *Microsoft Excel*. Além da classificação final, outras importantes informações podem ser obtidas através do programa tais como a correlação entre as variáveis, a importância de cada variável na classificação, o grau de explicação do número de fatores escolhido, a importância de cada variável em cada fator e os fatores mais significativos.

O programa não é do tipo autoexecutável, ou seja, para a instalação há a necessidade de selecionar o *drive* que é colocado o cd-rom ou disquete e, em seguida, executar o programa *setup.exe*. Além disso, pode ser usado em qualquer computador que possua o sistema operacional *Windows*.

O resultado da análise pode ser obtido através de um relatório impresso que conterá a classificação de todas as observações consideradas, registrando data, hora, local e responsável técnico da análise.

## 3 RESULTADOS E DISCUSSÃO

Utilizando-se o algoritmo que resume a metodologia desenvolvida na seção anterior e seguindo-se as etapas propostas pelo fluxograma da fig. 1, obteve-se como resultado na aplicação a seguinte classificação:

Faxinal	Escore Bruto	Escore Padronizado
1° São Pedro	406,0385	1,8000
2° Ivaí-Anta Gorda	255,5732	1,4586
3° Ponte Nova	163,4198	1,2495
4° Linha Brasília	150,4757	1,2201
5° Patos Velhos	116,3835	1,1427
6° Papanduva de Baixo	92,0459	1,0875
7° Queimadas	71,1442	1,0401
8° Cachoeira do Palmital	62,966	1,0215
9° Rio dos Couros	2,5779	0,8845
10° Rio do Meio	-9,2271	0,8577
11° Tijuco Preto	-11,4835	0,8526
12° Paraná-Anta Gorda	-19,0331	0,8355
13° Guanabara	-71,2905	0,7169
14° Salto	-101,915	0,6474
15° Taboãozinho	-105,585	0,6391
16° Dos Mellos	-113,344	0,6215
17° Marmeleiro de Baixo	-123,331	0,5988
18° Dos Krieger	-202,28	0,4197
19° Marmeleiro de Cima	-264,043	0,2795
20° Rio Bonito	-299,092	0,2000

Tabela 1 – Classificação final

Essa classificação correspondeu aproximadamente à expectativa subjetiva dos especialistas do IAP. Esperava-se de início pelas observações realizadas e pesquisas de anos anteriores, que o faxinal S. Pedro colocar-se-ia dentre os mais bem classificados. Por outro lado, os

faxinais Rio Bonito e Marmeleiro Grande notoriamente apresentavam maiores dificuldades, com baixa qualidade de vida dos seus moradores e má conservação das áreas.

O programa desenvolvido possibilitará, mediante atualização dos dados pelo IAP, classificações periódicas confiáveis baseadas em critério científico e não político. Este é o resultado principal do trabalho, ou melhor, a classificação dos faxinais feita por um algoritmo programado e aberto a posteriores atualizações de dados corresponde ao objetivo principal do trabalho.

#### 4 CONCLUSÕES

Pode-se concluir com base nos resultados obtidos que a ferramenta computacional desenvolvida pode ser aplicada com sucesso na construção de classificações e ainda,

- a) a amostra de dados brutos, obtida a partir dos questionários, foi suficiente para se efetuar o ranqueamento, considerando a precisão da estimativa de 3% e o grau de confiança da estimativa de 90%;
- b) os resultados dos escores finais obtidos com o programa, retratam com fidelidade a qualidade de cada faxinal;
- c) as variáveis apresentaram-se moderadamente correlacionadas, fato que justificou a utilização da Análise Fatorial;
- d) a Rotação Varimax, utilizada para uma melhor interpretação dos dados, mostrou-se necessária em função do grande número de variáveis correlacionadas;
- e) a Análise Fatorial utilizada no ranqueamento cumpriu com os objetivos propostos;
- f) a linguagem de programação mostrou-se adequada às pretensões do trabalho;
- g) o critério adotado em relação ao número de fatores mostrou-se adequado para a análise, uma vez que os 17 fatores considerados explicaram 97,8% do total da variância;
- h) os valores altos das comunalidades mostraram que a maior parte da variabilidade apresentada pelas variáveis é devida aos fatores comuns;

- i) os Métodos Householder e QL mostraram-se confiáveis e eficientes na obtenção dos autovalores e autovetores da matriz de correlação;
- j) quando se objetiva uma redução de um número grande de variáveis, em um número bem menor com pouca perda de informação, estando as variáveis relativamente correlacionadas, a Análise Fatorial mostra-se muito confiável;
- k) o programa atendeu as exigências da análise.

É importante ressaltar ainda que, embora o programa tenha sido usado na classificação de áreas de acordo com critérios ambientais pré-definidos, trata-se de um programa completamente genérico. Neste sentido, pode classificar não apenas áreas, mas também pessoas, animais, cidades, organizações, empresas, etc. Assim sendo, para a aplicação da técnica estatística por meio do programa, basta apenas que as variáveis de entrada sejam adequadamente selecionadas e possam ser quantificadas numericamente.

Na agricultura e pecuária de uma forma geral, pode-se vislumbrar inúmeras aplicações desta técnica de classificação, pois é muito comum a comparação entre produtos e também entre animais com base em critérios subjetivos. Este programa eliminaria muito da incerteza da subjetividade. Dificuldades encontradas em classificações com alto grau de subjetividade podem muito bem ser transpostas através desta metodologia.

#### 5 REFERÊNCIAS

- CHANG, M. Y. **Faxinais no Paraná**. Londrina: IAPAR, n.80, 1988. 20 p. (Informe de Pesquisa).
- DOMINGUES, Zilna H. **Hierarquização de faxinais inscritos no cadastro estadual de unidades de conservação e uso especial, visando o ICMS ecológico**. 1999. 143f. Dissertação de Mestrado – Setor de Ciências Agrárias, Curso de Engenharia Florestal, Universidade Federal do Paraná, Curitiba.
- FURTADO, Emerson M. **Automação do Ranqueamento Qualitativo de Áreas Especialmente Protegidas do Estado do Paraná através da Análise Fatorial**. 1999. 220f. Dissertação de Mestrado – Setor de Ciências Exatas – Curso de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná, Curitiba.
- HALVORSON, Michael. **Microsoft visual basic 5 – step by step**. São Paulo: Makron Books, 1997. 379 p.

- JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 4. ed. New Jersey: Prentice Hall, Englewood cliffs, 1998. 816 p.
- KAYSER, Henry F. **The varimax criterion for analytic rotation in factor analysis**. *Psychometrika*, v. 23, n. 3, University of Illinois, 1958.
- LAWLEY, D. N.; MAXWELL, A. E. **Factor analysis as a statistical method**. New York: American Elsevier Publishing Company, 1971.153p.
- LOUREIRO, Wilson. **ICMS Ecológico em perguntas e respostas**. Governo do estado do Paraná, 1998. 34 p.
- RENCHER, Alvin C. **Methods of multivariate analysis**. New York: Wiley InterScience, 1995. 627 p.
- RENCHER, Alvin C. **Multivariate statistical inference and applications**, New York: Wiley InterScience, 1998. 557 p.
- VETTERLING, William T. et al. **Numerical recipes example book (FORTRAN)**. 2. ed. London: Cambridge University Press, 1992. 245 p.