

Uso de SIG e do classificador ISODATA na geração do mapa de regiões geoambientais homogêneas na bacia do rio São Francisco, como um produto estratégico de apoio às ações de planejamento de redes fluviométricas

Wougran Soares Galvão¹
Paulo Roberto Meneses²

¹ Agência Nacional de Energia Elétrica - ANEEL
SGAN Quadra 603 / Módulos I e J - 70830-030 - Brasília - DF, Brasil
wougran@aneel.gov.br

² Instituto de Geociências - UnB
Caixa Postal 04457 – 70919-970 - Brasília - DF, Brasil
pmeneses@unb.br

Abstract. The proposed methodology in this work was based on the integrated analysis of a set of geo-referenced geo-environment variables and the unsupervised classification statistical technique ISODATA, with the use of Geographic Information System (GIS), for identification of homogeneous geo-environment regions that provided the basis for planning and management of fluviometric networks in the São Francisco river basin.

Palavras-chave: homogeneous geo-environment regions, geographic information system, river basin, streamgaging network, regiões geoambientais homogêneas, sistema de informação geográfica, bacia hidrográfica, rede fluviométrica.

1. Introdução

Uma rede de dados hidrológicos constitui um conjunto de atividades de coleta e processamento de dados, que segue uma planificação anteriormente definida e visa atender a um objetivo específico ou a um conjunto de objetivos compatíveis entre si. Frequentemente, os objetivos associam-se a uma forma particular de utilização dos recursos hidrológicos, sendo que esta definição antecede à efetivação da coleta dos dados.

Em geral, se não existe um profundo conhecimento hidrológico da região na qual uma rede se estabelecerá, a probabilidade de que a rede resultante proporcione informações efetivamente úteis e representativas é bastante reduzida. Tal conhecimento hidrológico advém do estudo e da experiência, não havendo substituto ao tempo de observação quando se está iniciando uma rede hidrológica em uma zona onde os dados históricos disponíveis são escassos ou não existem. A incerteza hidrológica sempre existirá devido a erros nas medições e a erros causados pela amostragem no espaço e tempo. A informação hidrológica perfeita não pode existir nunca (OMM, 1994). A redução da incerteza sobre a própria incerteza é o aspecto mais importante que permitirá valorizar ao máximo toda a informação contida nos dados que a rede gerará.

Segundo dados da Organização Meteorológica Mundial (OMM, 1992) a maioria das redes hidrométricas dos países foram criadas em função das necessidades nacionais, e na maior parte das vezes em função de problemas muito particulares. A OMM (1992) também relata, a nível mundial, a existência de uma documentação muito abundante sobre o planejamento de redes, todavia com muitos aspectos pouco claros (ausência de exemplos práticos de aplicação real, ausência da definição clara dos objetivos a serem atingidos ou dos critérios que devem ser aplicados, ausência de embasamento técnico, etc.). No Brasil, a documentação sobre o planejamento e histórico da evolução da rede hidrométrica nacional é relativamente escassa.

A rede atual da Agência Nacional de Águas (ANA), herdada da Agência Nacional de Energia Elétrica (ANEEL), com aproximadamente 5.200 estações, é composta em quase sua totalidade de estações que coletam informações de interesse energético, com cerca de 20 % das estações coletando dados para diversos fins, como agricultura, transporte, alerta hidrológico e outros.

Diante da complexidade e da subjetividade do processo de planejamento de redes hidrométricas torna-se cada vez mais necessária uma análise científica do problema. O presente estudo avaliou o potencial da aplicação da técnica de classificação digital não supervisionada ISODATA para a geração do mapa de regiões geoambientais homogêneas entre sub-bacias hidrográficas, como um produto estratégico de apoio ao planejamento e gerenciamento de uma rede de estações fluviométricas.

2. Área de Estudo

O estudo foi realizado na bacia hidrográfica do rio São Francisco (**Figura 1**), na escala base de 1:1.000.000 (IBGE, 2003), em função de seu potencial de geração de energia hidrelétrica, pela extensão do rio, pela variedade fisiográfica apresentada pela bacia, pela quantidade de municípios (503) inseridos na bacia e pela existência e disponibilidade de bases cartográficas e temáticas, digitais e georreferenciadas, geradas pela CODEVASF (2003).

O Vale do São Francisco possui uma área de drenagem da ordem de 640.000 Km², abrangendo os Estados de Minas Gerais, Bahia, Goiás, Pernambuco, Sergipe, Alagoas e Distrito Federal, onde aproximadamente 53,8% de sua área (343.784 Km²) encontra-se incluída no Polígono das Secas.

3. Metodologia

Considera-se, para fins deste estudo, a execução das seguintes atividades realizadas seqüencialmente:

3.1 Seleção do sistema de classificação e codificação de bacias hidrográficas

Os sistemas de classificação e codificação de bacias hidrográficas adotados pelo DNAEE/ANEEL, Pfafstetter/SRH, CNRH e IBGE, foram avaliados para a seleção daquele que seria o mais adequado ao processo de planejamento de redes hidrométricas.

3.2 Identificação e seleção de variáveis geoambientais associadas ao processo de locação de estações fluviométricas.

O termo “variável geoambiental” foi empregado no presente estudo para caracterizar um dado de natureza fisiográfica, hidrológica, meteorológica e sócio-econômica. Há entre os planejadores de redes divergências na identificação e seleção de variáveis geoambientais prioritárias ao processo de configuração e desenho das mesmas, bem como das técnicas empregadas para a concepção, em função da criação da maioria das redes para atendimento de necessidades nacionais.

A partir da revisão bibliográfica efetuada por Galvão (2004) sobre o planejamento de redes hidrometeorológicas, foram selecionadas de forma empírica as variáveis mais diretamente relacionadas com a definição de regiões geoambientais homogêneas.

3.3 Aplicação da técnica de classificação digital não supervisionada ISODATA

Neste tipo de classificação o algoritmo baseia-se na análise de agrupamentos onde são identificadas no espaço de atributos as nuvens (clusters) formadas por “pixels” com características similares. Neste tipo de classificador, é comum especificar o número destas nuvens, isto é, o número de classes possíveis de serem encontradas na área de estudo, mesmo

não sendo este o número ideal destas classes. Define-se então, através de um limiar estabelecido pelo analista, quais as concentrações de “pixels” que devem ser tratadas como grupos distintos (Venturieri e Santos, 1998).

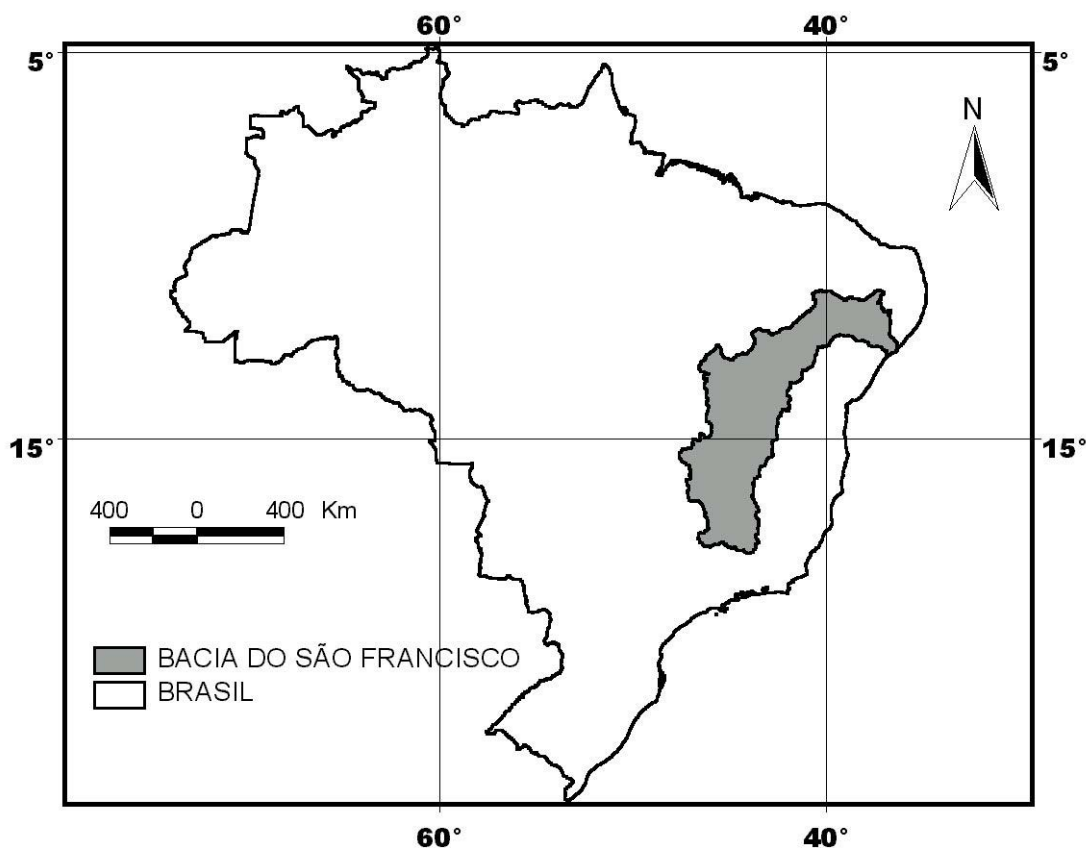


Figura 1 – Mapa de localização da área de estudo.

Hargrove e Luxmoore (1997) utilizaram a técnica ISODATA para estudos de criação de ecorregiões geográficas, baseados no crescimento de vegetação de bosques, através do Sistema de Informações Geográficas Arc/Info. Ressaltaram que a classificação ISODATA raramente tem sido aplicada a dados primários não espectrais.

A partir da padronização das diferentes escalas de valores de representação dos atributos das variáveis fisiográficas e ambientais selecionadas, foi aplicada a técnica ISODATA ao conjunto de dados, através da utilização do Sistema de Tratamento de Imagens Digitais ENVI 4.0, para avaliar a homogeneidade entre as sub-bacias do rio São Francisco, segundo as variáveis analisadas.

3.4 Aquisição de dados hidrométricos

Foi utilizado o Sistema HIDRO, desenvolvido pela ANEEL, para o acesso aos dados fluviométricos da rede nacional do DNAEE/ANEEL/ANA associados à bacia hidrográfica do rio São Francisco. Em 2003 encontravam-se sob responsabilidade da ANA cerca de 445 estações fluviométricas na Bacia do Rio São Francisco, estando 210 estações ativas e 235 estações inativas (ANA, 2003).

Para fins de validação do modelo proposto, as regiões geoambientais homogêneas definidas pelo classificador ISODATA foram avaliadas com respeito à locação espacial das

estações da rede fluviométrica da ANA, bem como através da análise de séries de dados de vazão média mensal das mesmas.

4. Resultados e Discussão

4.1 Seleção do sistema de classificação e codificação de bacias hidrográficas

Para o presente estudo foi adotado o método de classificação e codificação de bacias hidrográficas desenvolvido por Pfafstetter (1989), por tratar-se de um método natural, hierárquico, baseado na topografia da área drenada e na topologia da rede de drenagem (Verdin, 1997), na codificação de bacias com a economia de dígitos, na informação topológica embutida nos dígitos (Silva, 1999), de fácil implementação por técnicas de programação (Furnans e Olivera, 2001), de aplicabilidade global e de fácil integração com os Sistemas de Informação Geográfica - SIG. E, em especial, por permitir para o Brasil a individualização e representação, na escala 1:1.000.000, no quinto nível de classificação, de sub-bacias com um tamanho médio de aproximadamente 6.200 Km², constituindo uma excelente unidade de área para fins de planejamento e gerenciamento de recursos hídricos. No caso da Bacia do Rio São Francisco, as áreas das sub-bacias no quinto nível de classificação de Pfafstetter (1989) variam de 5 Km² a 19.716 Km², com um tamanho médio de 2.137 Km², permitindo a delimitação de 297 (duzentos e noventa e sete) sub-bacias hidrográficas.

4.2 Seleção das variáveis geoambientais mais relevantes

As seguintes variáveis geoambientais foram selecionadas de forma empírica, com base no levantamento bibliográfico realizado, como as mais relevantes para a identificação das regiões geoambientais homogêneas entre as sub-bacias da Bacia do Rio São Francisco: hidrografia - distância (IBGE - 1:1.000.000), densidade de drenagem por Ottobacia (1:1.000.000), geologia (CPRM - 1:2.500.000), fraturas - distância (CPRM - 1:2.500.000), densidade de fraturas por Ottobacia (1:2.500.000), uso do solo (USGS - 1:1.000.000), solos (CODEVASF - 1:1.000.000), índice topográfico composto (1:1.000.000), hidrogeologia (CODEVASF - 1:1.000.000), tectônica (CPRM - 1:2.500.000), erosão atual (CODEVASF - 1:1.000.000), aspecto (1:1.000.000), declividade (1:1.000.000), modelo numérico do terreno (1:1.000.000), direção de fluxo (1:1.000.000) e acumulação de fluxo (1:1.000.000).

4.3 Aplicação do classificador ISODATA

Com o objetivo de identificar e delimitar as regiões geoambientais homogêneas existentes entre as sub-bacias (Ottobacias - Nível 05) hidrográficas do rio São Francisco, como produto de apoio ao planejamento e gerenciamento de redes hidrométricas, foi aplicado ao conjunto das 16 variáveis geoambientais selecionadas, a partir da padronização das diferentes escalas de valores de representação dos atributos das variáveis selecionadas, o algoritmo de classificação digital não supervisionada ISODATA.

Os resultados obtidos com a aplicação do classificador ISODATA, a partir de diversos testes efetuados ao conjunto de 16 variáveis, produziram imagens com muitos ruídos, resultantes da classificação ao longo de todas as Ottobacias, e zonas homogêneas ocupando larga extensão de área. Este resultado reflete o fato de que a integração do conteúdo de informação existente no conjunto das 16 variáveis não contribui positivamente para a identificação das regiões geoambientais homogêneas, conforme previsto por Thomaz (1999). Fato provavelmente explicado pela existência de índices de correlação relativamente altos entre alguns dos pares de variáveis utilizados na classificação.

Considerando que Luz (2003) destaca a importância da utilização da técnica estatística de Análise por Componentes Principais para a eliminação do espaço de variáveis explicativas de

uma modelagem hidrológica daquelas dimensões nas quais ocorrem um grande número de variáveis correlacionadas, foi aplicada a Técnica de Transformação por Componentes Principais ao conjunto das 16 variáveis geoambientais selecionadas, na tentativa de avaliar os resultados do classificador ISODATA a partir de um conjunto de variáveis sem redundância de informações, conforme proposto por Hargrove e Luxmoore (1997).

As imagens componentes principais (CP) resultantes apresentaram os seguintes autovalores acumulados: CP1(50,6%), CP2(59,4%), CP3(65,8%), CP4(71,4%), CP5(76,4%), CP6(80,9%), CP7(84,6%), CP8(87,6%), CP9(90,3%), CP10(92,7%), CP11(94,8%), CP12(96,2%), CP13(97,6%), CP14(98,6%), CP15(99,3%) e CP16(100%).

Os resultados obtidos com a aplicação do classificador ISODATA a partir de diversos testes efetuados ao conjunto de 16 imagens componentes principais evidenciaram que 81% da informação (CP1 a CP6) foi encontrado como sendo a quantidade de informação não correlacionada ideal para a obtenção de uma imagem classificada contendo uma boa definição das regiões geoambientais homogêneas (**Figura 2**), sendo constituída por 15 regiões homogêneas. As imagens classificadas obtidas com a quantidade de informação superior a 81% mostraram forte tendência ao incremento de ruídos de classificação ao longo de todas as Ottobacias, resultando em maior fragmentação da classificação em função direta ao aumento da quantidade de informação.

4.4 Validação do modelo de regiões geoambientais homogêneas

Com o propósito de avaliar as regiões geoambientais homogêneas definidas pelo classificador ISODATA foi feita uma análise da locação das estações fluviométricas por Ottobacia e sua relação com as unidades geoambientais homogêneas, tomando como base dados hidrológicos

Das 297 Ottobacias constituintes da bacia do rio São Francisco, somente 132 possuem estações fluviométricas ativas administradas pela ANA e outras entidades. Seis Ottobacias situadas na cabeceira do rio São Francisco chegam a possuir de 17 a 44 estações. Cerca de 100 Ottobacias possuem entre 1 e 3 estações, e 26 Ottobacias possuem entre 4 e 10 estações.

A análise visual do resultado apresentado pela classificação realizada também permite destacar a forte relação existente entre a localização de estações fluviométricas e as áreas limítrofes entre diferentes zonas homogêneas, resultado corroborado pelos dados apresentados na **Tabela 1**.

A locação de 128 estações fluviométricas distantes até 1 Km dos limites entre regiões geoambientais homogêneas, e aumentando para 172 estações se considerado o afastamento até 2 Km, e para 192 estações até 3Km, comprovaram ser o mapa de regiões geoambientais homogêneas um poderoso instrumento para apoiar as atividades de planejamento e gerenciamento de redes fluviométricas.

Tabela 1 – Quantificação das estações fluviométricas ativas da ANA com relação à proximidade dos limites entre regiões geoambientais homogêneas.

DISTÂNCIA (Km)	Nº ESTAÇÕES
0 - 1	128
1 - 2	44
2 - 3	20
3 - 4	11
4 - 5	4
> 5	3

A relação da locação geográfica das estações fluviométricas em zonas limítrofes de regiões geoambientais homogêneas ou em suas proximidades também ficou constatado no

caso da rede fluviométrica operada por outras entidades na bacia do rio São Francisco, como pode ser visualizado na **Tabela 2**.

Tabela 2 – Quantificação das estações fluviométricas ativas operadas por outras entidades na bacia do rio São Francisco, com relação a proximidade dos limites entre regiões geoambientais homogêneas.

DISTANCIA (Km)	Nº ESTAÇÕES
0 - 1	146
1 - 2	58
2 - 3	22
3 - 4	15
4 - 5	7
> 5	7

Outra importante constatação foi feita a partir da análise de 76 pares de séries históricas de dados consistidos de vazão média mensal. Onde, estações fluviométricas localizadas numa mesma região geoambiental homogênea apresentaram séries de dados com coeficientes de correlação relativamente altos, enquanto os dados de estações fluviométricas localizadas em classes distintas de regiões geoambientais homogêneas apresentaram coeficientes de correlação substancialmente mais baixos. Estes resultados evidenciaram que da cabeceira à foz do rio São Francisco a vazão média mensal tende a apresentar um padrão hidrológico também diferenciado em função das características fisiográficas e ambientais das regiões geoambientais homogêneas.

O mapa de regiões geoambientais homogêneas mostrou-se um produto cartográfico de grande potencial para apoiar as ações de planejamento e gerenciamento de redes hidrometeorológicas, pela importância das zonas limítrofes entre regiões geoambientais homogêneas para a locação de estações fluviométricas, bem como pelo estabelecimento da relação de que dados consistidos de vazão média mensal para estações fluviométricas localizadas numa mesma região geoambiental homogênea apresentam para as suas séries de dados coeficientes de correlação relativamente altos, e que estações fluviométricas localizadas em regiões geoambientais homogêneas distintas apresentam para as suas séries de dados coeficientes de correlação relativamente mais baixos. Estes dados corroboram as informações de Solomon (1976) que relatou que o modelo de regiões homogêneas, definido nos estudos de planejamento de redes do Canadá como regiões que possuem a mesma relação entre os parâmetros de um modelo hidrológico e as características do terreno, podem ser usadas no planejamento de redes para apoiar a identificação do número mínimo de estações requeridas em cada região homogênea, bem como para sinalizar a locação aproximada de novas estações a serem incorporadas à rede para aumentar a eficiência da mesma. Além de que, destacou que o planejamento da rede hidrométrica do Canadá obteve uma melhor eficiência através do uso de dados do terreno, incluindo dados topográficos, dados de uso do solo e da cobertura vegetal, e que imagens de satélite deveriam ser utilizadas para a atualização periódica dos mapas de uso e ocupação do solo.

5. Conclusões

O método de classificação e codificação de bacias hidrográficas desenvolvido por Pfafstetter (1989) mostrou-se muito adequado para apoiar o planejamento e gerenciamento de redes hidrométricas, em seu nível 5 de classificação/codificação.

O algoritmo de classificação digital não supervisionada ISODATA mostrou-se eficiente para a geração do mapa de regiões geoambientais homogêneas, como produto de apoio ao planejamento e gerenciamento de redes hidrométricas, quando aplicado ao subconjunto das seis primeiras imagens componentes principais derivadas do conjunto das 16 variáveis geoambientais.

Ficou muito bem definida, para grande parte das estações fluviométricas, em vários trechos da bacia de montante a jusante, a relação da locação geográfica das mesmas em zonas limítrofes de regiões geoambientais homogêneas ou em suas proximidades. Comprovou-se através da análise de séries de dados consistidos de vazão média mensal, que estações fluviométricas localizadas numa mesma região geoambiental homogênea apresentam para as suas séries de dados coeficientes de correlação relativamente altos, e que estações fluviométricas localizadas em regiões geoambientais homogêneas distintas apresentam para as suas séries de dados coeficientes de correlação relativamente mais baixos.

Os resultados obtidos comprovaram ser o mapa de regiões geoambientais homogêneas um poderoso instrumento para apoiar as atividades de planejamento e gerenciamento da rede fluviométrica instalada na bacia do rio São Francisco.

O uso de Sistema de Informação Geográfica (SIG), para a espacialização, integração e análise de dados fisiográficos, ambientais, sócio-econômicos e hidrológicos das Ottobacias do rio São Francisco, mostrou-se eficiente, constituindo uma valiosa ferramenta para apoiar as ações de planejamento e gerenciamento de redes fluviométricas, podendo ser empregado para otimizar a aplicação de recursos financeiros, contribuindo para a identificação de novos locais a serem monitorados e para a eliminação de estações desnecessárias.

Referências

- ANA. Agência Nacional de Águas. **Relatório Estatístico da Rede Básica Hidrometeorológica Nacional**. Agência Nacional de Águas, Superintendência de Informações Hidrológicas – SIH, Brasília, ANA, Março, 2003, 20 p.
- CODEVASF. COMPANHIA DE DESENVOLVIMENTO DO VALE DO SÃO FRANCISCO. **Base de Dados Digitais Georreferenciados da Bacia do São Francisco**. Brasília, CODEVASF, Área de Planejamento, Coordenadoria de Projetos Especiais, Supervisão de Geoprocessamento. Documento publicado em CD-ROM, 2003.
- Furnans, J.; Oliveira, F. Watershed Topology – The Pfafstetter System. In: Annual ESRI User Conference, 21. **Proceedings...** Califórnia, ESRI, 2001. Acesso em: 10/01/2002. Disponível em: <http://www.esri.com/library/userconf/proc01/professional/abstracts/a1008.html>.
- Galvão, W. S. **Uso de SIG na geração de modelos de potencial à locação de estações fluviométricas e de regiões geoambientais homogêneas na bacia do rio São Francisco**. 2004. 302. p. Tese (Doutorado em Geologia) – Universidade de Brasília, Brasília, 2004.
- Hargrove, W. W.; Luxmoore, R. J. **A spatial clustering technique for the identification of customizable ecorregions**. 1997. Acesso em: 10/07/2002. Disponível em: <http://research.esd.ornl.gov/~hnw/esri>.
- IBGE. Instituto Brasileiro de Geografia e Estatística. **Base Cartográfica Integrada Digital do Brasil ao Milionésimo**. Versão 1.0 para ArcGis Desktop/ArcView. 2003. Departamento de Cartografia (DECAR). Documento publicado em CD-ROM, 2003.
- Luz, J. A. G. O uso de componentes principais nos modelos de regressão linear: uma abordagem teórica. In: XV Simpósio Brasileiro de Recursos Hídricos. **Anais**. Curitiba: ABRH, 2003. 16 p.
- OMM. ORGANIZACIÓN METEOROLÓGICA MUNDIAL. **Proceedings of the International Workshop on Network Design Practices**. 1992. 11-15 de noviembre de 1991, Coblensa, Alemanha.
- OMM. ORGANIZACIÓN METEOROLÓGICA MUNDIAL. **Guía de Prácticas Hidrológicas. Adquisición y Proceso de Datos, Análisis, Predicción y otras Aplicaciones**. OMM, N° 168, Quinta Edición, 1994, 781 p.

Pfafstetter, O. **Classificação de Bacias Hidrográficas – Metodologia de Codificação**. Rio de Janeiro, RJ: Departamento Nacional de Obras de Saneamento (DNOS), 1989, p. 19. Manuscrito não publicado.

Silva, P. A. Classificação e codificação de bacias hidrográficas brasileiras segundo o método Pfafstetter, com uso de geoprocessamento. Encontro de Las Aguas, 2, 1999, Montevideo. **Proceedings...** Montevideo, Uruguay: IICA, 1999. Acesso em: 15/02/2002. Disponível em: <http://www.iica.org.uy>.

Solomon, S. I. **Regional mapping and climatic influence in data transfer methods**. World Meteorological Organization, Operational Hydrology, Report N° 8, WMO – N° 433, Geneva, Switzerland, 1976, pp. 57-62.

Thomaz, C. E. **Estudo de classificadores para o reconhecimento automático de faces**. Dissertação de Mestrado em Engenharia Elétrica. Rio de Janeiro, Pontifícia Universidade Católica, Departamento de Engenharia Elétrica. 1999, 104 p.

Venturieri, A.; Santos, J. R. Técnicas de classificação de imagens para análise de cobertura vegetal. In: Assad, E. D.; Sano, E. E. **Sistema de Informações Geográficas, Aplicações na Agricultura**. Brasília, EMBRAPA-SPI/EMBRAPA-CPAC, Parte III, Capítulo 18, 1998, 434 p.

Verdin, K. L. A System for Topological Coding Global Drainage Basin and Stream Networks. In: Annual ESRI User Conference, 17, San Diego, California. **Proceedings...** California, ESRI, 1997, 6 p.

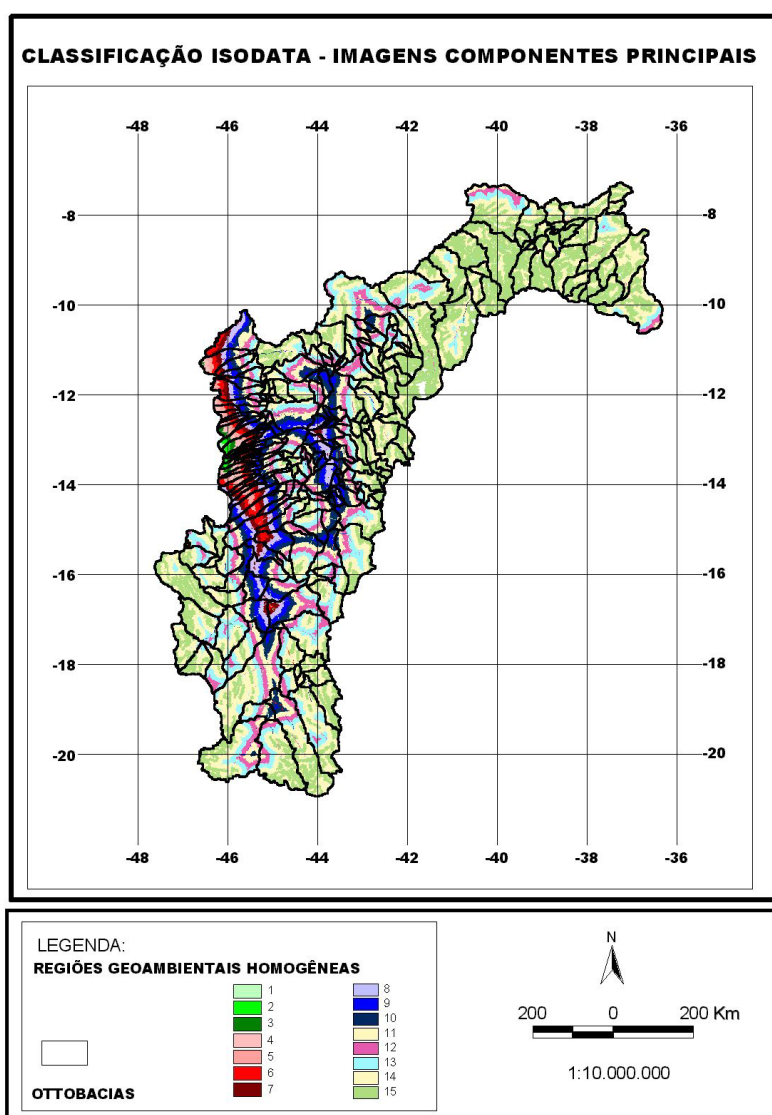


Figura 2 – Imagem resultante da classificação ISODATA a partir das imagens componentes principais CP1 a CP6, geradas do conjunto de 16 variáveis geoambientais selecionadas, constituída por 15 regiões geoambientais homogêneas.