Revista Brasileira de Climatologia
ISSN: 1980-055x (Impressa) 2237-8642 (Eletrônica)

APLICAÇÃO DA LÓGICA NEBULOSA (FUZZY CLUSTER) NA DEFINIÇÃO DE UNIDADES CLIMÁTICAS: ESTUDO DE CASO NA BACIA DO RIO PARAIBUNA-MG/RJ

FERREIRA, Cássia de Castro Martins – cassia.castro@ufjf.edu.br
UFJF – Universidade Federal de Juiz de Fora

RESUMO. Este trabalho tem como objetivo à classificação de grupos climáticos segundo variáveis climáticas utilizando o método de classificação multivariada baseado na lógica nebulosa, visando identificar classes similares, além de avaliar a variabilidade espacial e temporal dos agrupamentos. A área de análise e classificação das categorias climáticas consiste na Bacia do Rio Paraibuna – MG/RJ, na qual foram delimitados três anos de análise (janeiro/2008 a dezembro/2010), tendo como base de dados para o trabalho, 20 postos de coleta de dados de superfície, que consistiram nos dados de temperatura média diária, temperatura mínima diária, temperatura máxima diária, amplitude térmica diária, precipitação total diária, dias de chuva, deficiência e excedente hídricos e evapotranspiração. A metodologia utilizada para classificar as categorias climáticas foi a Lógica *Fuzzy,* os dados foram agrupados através do algoritmo FCM ("*Fuzzy-C-Means*") que estabeleceu o maior grau de similaridade entre as variáveis. Como resultado, obteve-se a classificação da área da bacia em cinco categorias climáticas, mostrando que a lógica *fuzzy* é uma metodologia que pode ser utilizada na classificação de categorias climáticas.
Palavras-chave: lógica *fuzzy*; classificação climática; análise de agrupamento.

APPLICATION OF LOGIC NEBULA (FUZZY CLUSTER) TO THE CLIMATE UNITS DEFINITION: A CASE STUDY IN RIVER BASIN PARAIBUNA-MG/RJ
ABSTRACT. This work main the classification of group climate according climate variables using the method of multivariate classification based on fuzzy logic, main identify similar classes also available the spatial variability and temporal of groups. The area of analysis and classification of climate categories is the Paraibuna – MG/RJ river basin, in that was delimited three years of analysis (2008 January to 2010 December), having 20 collection points of data of surface, how database to the work, that consisted in data of medium diary temperature, minimum diary temperature, maximum diary temperature, temperature range diary, total precipitation diary, rain days, water deficiency and excess and evapotranspiration. The methodology that was used to classify the climate categories was the logical fuzzy, the data was grouped using the FCM (Fuzzy-C-Means) algorithm that establish the major degree of similarity between the variables. How result, we have the classification of the basin area in five climate categories, showing that the fuzzy logical is a methodology that can be used in the classification of climate categories.
Key-words: fuzzy logical; climate classification; analysis grouping.

INTRODUÇÃO

O conhecimento das condições climáticas de uma determinada localidade, como um dos elementos a ser considerado nas tomadas de decisões, seja a nível econômico, de planejamento territorial ou para fins de manejo dos recursos terrestres é de extrema importância, pois propicia um melhor ordenamento do território, podendo desta forma, minimizar, por exemplo, eventos de escorregamentos de massa e enchentes, melhorar o aproveitamento das precipitações, ventos, energia solar como recurso energético, para a agricultura e uma melhor ambiência urbana, dentre outros.

Sendo enfatizado por Aspiazú *et al.* (1990) que o estudo dos fatores ambientais é fundamental para o planejamento do uso da terra e para entender, explicar e prever o crescimento e o desenvolvimento dos recursos naturais. E dentre esses fatores, o clima é considerado o regulador central, exercendo influências diretas ou indiretas sobre os fatores bióticos e abióticos.

Como, independentemente da ação humana, o clima difere-se de uma região para outra, ou até mesmo dentro de alguns poucos quilômetros. As variações em quantidade intensidade e distribuição estacional de temperatura, precipitação, umidade, pressão atmosférica e ventos, associados às variáveis do local, vão produzir inúmeras situações e tipos climáticos diferentes. Tornando, importante obter mecanismos que nos permitam conhecer e identificar estas diferentes unidades climáticas em escala regional.

Assim sendo, haveria uma necessidade crescente de busca, evolução e adaptação de métodos de classificação, capazes de sintetizar e agrupar características similares, através de critérios próprios. Com isto, torna-se mais exequível o mapeamento e a identificação de diferentes unidades climáticas de uma região.

Para isso a Teoria de Conjuntos *Fuzzy criada* em 1965, por L. A. Zadeh foi estabelecida a partir de conceitos já aceitos da lógica clássica objetivando criar uma metodologia matemática para o tratamento de informações de caráter impreciso ou vago. Os métodos clássicos, tradicionalmente utilizados em classificações climáticas, criam classes homogêneas e discretas

com limites bem delineados ou mesmo abruptos. Estas classes são comumente representadas em mapas climáticos, isto é, uma localidade no qual a temperatura se enquadra no intervalo de 19^0 a 21^0C nunca poderá estar no mesmo grupo, de outro local, que apresenta temperaturas de $21,1^0$C, mesmo que $21,1^0$C esteja mais próximo das localidades que apresentam a temperatura de 21^0C do que de outro grupo no qual o intervalo será de $21,1^0$C até 24^0C.

Desta forma, nas teorias clássicas um conceito pode ser verdadeiro ou falso, como por exemplo, um elemento pode pertencer ou não a um conjunto. Não há nada errado neste conceito, entretanto pode haver alguma limitação uma vez que dados exatos podem ser vistos como relativos. A limitação se refere à dificuldade em incorporar satisfatoriamente à teoria matemática palavras que carregam sentido "nebuloso, vago" (*fuzzy*).

Zadeh propôs uma caracterização mais ampla, generalizando a função característica típica da lógica clássica, de modo que ela pudesse assumir um número infinito de valores no intervalo [0,1]. Ela se baseia em que uma proposição lógica não é necessariamente falsa ou verdadeira, mas possui de fato, graus de verdade. Quando dizemos que o tempo está "ensolarado" não queremos dizer que não haja nuvens. Da mesma forma, em um tempo parcialmente nublado podemos ter sol. Geralmente aceitamos como ensolarado um céu com até 20 ou 30% de nuvens (KLIR, 1995).

Como podemos então definir os limites de nossas proposições? Pode-se dizer que, quando o ser humano utiliza tais termos, ele infere um grau de verdade. Na zona intertropical, por exemplo, pode-se aceitar nuvens em maior quantidade, visto que uma pouca quantidade de radiação direta poderia ser suficiente para o processamento das atividades que dela dependem. Um profissional que dependa do cálculo da radiação direta para executar o seu trabalho, já estabeleceria limites mais rígidos e precisos. Os valores que definem o grau de verdade de uma proposição podem não ser conscientes, mas de alguma forma definidos pelo contexto social, linguístico, ou uma base de referência experimental. A definição e a modelagem destes termos e valores constituem a base da teoria e do estudo dos sistemas nebulosos.

Uma relevante propriedade da modelagem *fuzzy*, salientado por (RUHOFF *et al.*, 2005) é a capacidade de codificar conhecimentos inexatos, de tal forma que se aproxime dos processos decisórios. Os sistemas de inferência *Fuzzy* proporcionam a apreensão do conhecimento próximo ao modelo "cognitivo" muito aplicado na análise de problemas de previsão e classificação. Logo, o processo de obtenção do conhecimento é simplificado, mais preciso e com menor probabilidade de erros.

Desta forma, dentre os problemas relacionados à lógica *boolena* são os limiares nítidos e rígidos, que podem não representar os fenômenos naturais corretamente. Segundo Burrough e MacDonnell (1998), na natureza os fenômenos não são representados por limites rígidos ou estáticos. Em função destas colocações, é preciso buscar metodologias de análise espacial, que possam produzir novas informações que representassem de forma mais complexa os fenômenos naturais. Assim, o uso da inferência *fuzzy,* que se caracteriza na indefinição de fronteiras ou limiares entre classes, seria um caminho a ser experimentado e explorado.

Devido ao próprio caráter de imprecisão intrínseco à classificação de unidades climáticas, pode-se então, utilizar a lógica *fuzzy* como uma metodologia alternativa para este processo.

Um conjunto nebuloso é definido por determinado grupo de elementos do universo de discurso X sendo que cada elemento pertence a este conjunto com certo grau de pertinência (KLIR *et al.*, 1995).

A função característica que associa cada elemento a um grau de relevância ou pertinência é chamada de função de pertinência ("*membership*"). Esta função é normalizada assumindo valores reais no intervalo [0,1] e descrita por:

$\mu A: X \rightarrow [0,1]$, onde A é o conjunto nebuloso e X é o universo de discurso (BRAGA, 1998).

Os conjuntos *Fuzzy* para um determinado universo recebem nomes, isto é dados qualitativos, representados por variáveis linguísticas. Estas variáveis permitem uma maior sistematização para se definir uma característica aproximada de fenômenos não precisos, destacando aqui o clima, ou a sucessão de tipos de tempo. O uso da descrição linguística e não de variáveis quantitativas, permite o tratamento de situações que são muito complexas para serem analisadas por meio de termos matemáticos convencionais.

Agrupamentos nebulosos

Agrupamentos Nebulosos ("*Fuzzy Clustering*") são técnicas de agrupamentos onde cada elemento pertence a um grupo com um determinado grau de pertinência.

As formas de agrupar dados são básicas para muitas classificações. O objetivo dos agrupamentos é identificar grupos de dados similares de um grande conjunto de dados e produzir uma concisa representação do comportamento do sistema.

Os algoritmos de agrupamentos clássicos ("*crisp" ou "hard" "clusters"*) geram grupos nos quais cada dado é destinado para um grupo somente. Mas existem situações nas quais um dado não pertence somente a um grupo, mas a mais de um grupo, quando, por exemplo, o elemento está equidistante entre dois grupos. Nestes casos, os métodos de agrupamentos nebulosos (*"fuzzy clustering"*) são uma ferramenta mais adequada para representar estruturas de elementos reais (ZIMMERMANN, 1985).

Portanto, os agrupamentos nebulosos (*"fuzzy clustering"*) podem oferecer melhores descrições dos dados quando a sobreposição de classes é possível, ou seja, quando os elementos pertencem a todos os grupos, mas em diferentes graus, e isto pode ser quantificado por uma função de pertinência ("*membership*") correspondente (ROUSSEEUW, 1995).

O algoritmo de agrupamentos nebulosos mais simples e que foi utilizado no presente trabalho é o "*fuzzy-c-means*" (FCM). Esta técnica foi originalmente introduzida por Jim Bezdek em 1981 e trouxe grandes melhorias aos métodos de agrupamentos, pois cada dado pertence a um grupo com um grau de pertinência específico.

A vantagem do FCM em relação ao HCM (*hard-c-means*: método de agrupamento clássico onde cada elemento pertence a somente um grupo) é que o FCM consegue melhor detalhamento e descrição dos dados.

As formas de agrupar dados são básicas para muitas classificações. O objetivo dos agrupamentos é identificar grupos de dados similares de um grande conjunto de dados e produzir uma concisa representação do comportamento do sistema (FERREIRA, 2000).

Tendo em vista a incerteza intrínseca à classificação climática, a possibilidade de se descrever tanto o componente térmico quanto hídrico por meio de termos linguísticos, permite o uso da lógica *fuzzy* na classificação do clima, visando facilitar a compreensão de sua variação espacial.

Dessa forma, este trabalho aborda o uso de uma metodologia, baseada na lógica nebulosa, utilizando a análise de agrupamento nebuloso, para análise dos tipos de tempo na Bacia do Rio Paraibuna –MG, estabelecendo um quadro de síntese das unidades climáticas da bacia.

MATERIAL E MÉTODOS

As variáveis climáticas

Todos os dados utilizados podem ser subdivididos em:

1- Dados "reais": compreendem os dados climáticos observados, sem nenhum tratamento estatístico, que no presente trabalho, consistiram na fonte geradora de todos os demais dados e resultados.

2- Dados que sofreram um primeiro tratamento estatístico e matemático: são os dados de temperatura média diária, temperatura mínima diária, temperatura máxima diária, deficiência hídrica, excedente hídrico, amplitude térmica e evapotranspiração real, que

ISSN: 1980-055x (Impressa) 2237-8642 (Eletrônica)

foram estimados por meio de procedimentos estatísticos ou modelos matemáticos, baseados em dados de superfície, registrados por meio de aparelhos meteorológicos.

3- Dados derivados de Análise Multivariada, baseada na análise de Agrupamentos Nebulosos, que resultaram de outra etapa do trabalho, que consistiu no uso da estatística, para a geração de grupos mais similares, ocasionando, portanto, na classificação da área estudada em regiões climáticas diferentes, para todo o período analisado.

A partir da manipulação e organização dos dados e informações geradas, foram seguidas as seguintes etapas:

- Realização de matrizes operacionais com uma listagem dos dados dos agrupamentos nebulosos;
- Matrizes de pertinência;
- Matrizes diárias de dados finais do agrupamento nebuloso, para cada uma das variáveis.
- Agrupamentos nebulosos finais diários, para todo o grupo de variáveis;
- Agrupamentos nebulosos de síntese: mensais, sazonais e anuais, para todo o grupo de variáveis.
- Dados climáticos diários associados às informações da circulação atmosférica.

Análise de agrupamentos nebulosos

A área de análise e classificação climática consiste na Bacia do Rio Paraibuna – MG, no qual foi delimitado três anos de análise (janeiro/2008 a dezembro/2010), este período foi escolhido, não se baseando em anos-padrão, anos mais quentes ou frios, ou mesmo outros critérios científicos. Sendo o único requisito utilizado para escolha dos três anos a integridade e a disponibilidade de dados da série em todas as localidades. Verificando se neste período os dados estavam completos, sem interrupções nas observações, além de verificar o período em que um maior número de estações apresentavam estas características.

O período de três anos foi estabelecido de forma tal, que nos permitisse ter uma abrangência maior dentro da série, mesmo, sabendo que ainda seria considerado um período curto de análise quando se trata de uma classificação climática. Tendo como base de dados para o trabalho, 20 postos de coleta de dados de superfície pertencentes ao Quinto e Sexto Distrito de Meteorologia, que consistiram nos dados de temperatura média diária, temperatura mínima diária, temperatura máxima diária, amplitude térmica diária, precipitação total diária, dias de chuva, evapotranspiração deficiência e excedente hídricos. Foram escolhidos os postos do Quinto e Sexto Distrito por terem uma homogeneidade no tipo de instrumento, na coleta da informação e por conterem todos os dados utilizados na pesquisa.

Estes elementos foram escolhidos por permitirem uma distinção entre o componente térmico (temperatura mínima diária, temperatura máxima diária e amplitude térmica diária) e hídrico (precipitação total diária, dias de chuva, evapotranspiração deficiência e excedente hídricos) para a bacia em análise, integrando os indicadores climáticos de cada localidade.

Os domínios e termos linguísticos atribuídos às variáveis de entrada e saída estão relacionadas na tabela 1. Os intervalos das curvas de pertinência foram caracterizados após exploração dos dados com auxílio do software livre R (*R Development Core Team*, 2005).

ISSN: 1980-055x (Impressa) 2237-8642 (Eletrônica)

Tabela 1

Domínios, termos linguísticos e características das variáveis de entrada e saída

Domínio Temporal	Janeiro de 2008 a Dezembro de 2010
Domínio Espacial	Bacia do rio Paraibuna
População	20 Estações Climatológicas
Antecedentes	temperatura média diária, temperatura mínima diária, temperatura máxima diária, amplitude térmica diária, precipitação total diária, dias de chuva, deficiência, excedente hídricos e evapotranspiração
Antecedente	Temperatura Média Diária
Termos Linguísticos	Mais frio (MF), Frio (F), Intermediário (I), Quente (Q), Mais Quente (MQ)
Universo (T^0C)	5,9 a 31,3^0 C
Antecedente	Temperatura Mínima Diária
Termos Linguísticos	Mais frio (MF), Frio (F), Intermediário (I), Quente (Q), Mais Quente (MQ)
Universo ($Tmin^0C$)	1,7 a 31,7 ^{0}C
Antecedente	Temperatura Máxima Diária
Termos Linguísticos	Mais frio (MF), Frio (F), Intermediário (I), Quente (Q), Mais Quente (MQ)
Universo ($Tmax^0C$)	37,5 a 11,2 ^{0}C
Antecedente	Amplitude Térmica Diária
Termos Linguísticos	Mais baixa (MB), Baixa (B), Intermediária (I), Alta (A), Mais Alta (MA),
Universo ($Tatd^0C$)	19,1 a 0,7 ^{0}C
Antecedente	Precipitação Total Diária
Termos Linguísticos	Mais seco (MS), Seco (S), Intermediário (I), Úmido (U), Mais Úmido (UM)
Universo (Pmm)	0 a 128,5 mm
Antecedente	Dias de Chuva
Termos Linguísticos	Mais baixa (MB), Baixa (B), Intermediário (I), Alta (A), Mais Alta (MA)
Universo (DCmm)	0 a 31
Antecedente	Deficiência Hídrica
Termos Linguísticos	Mais seco (MS), Seco (S), Intermediário (I)
Universo (DHmm)	0 a 17,8 mm
Antecedente	Excedente Hídrico
Termos Linguísticos	Úmido (U), Mais Úmido (UM), Intermediário (I)
Universo (EH mm)	0 a 306,7 mm
Antecedente	Evapotranspiração
Termos Linguísticos	Mais baixa (MB), Baixa (B), Intermediário (I), Alta (A), Mais Alta (MA)
Universo (ETPmm)	O,4 a 23,1 mm
Consequente	Tipos de tempo
Termos Linguísticos	Mais Seco e Mais Frio (MSF), Mais Úmido e Mais Frio (MUF), Seco e Mais Frio (SMF), Úmido e Mais Frio (UMF), Intermediário e Mais Frio (IMF), Intermediário e Frio (IF), Intermediário e Mais Quente (IMQ), Intermediário e Quente (IQ), Intermediário e Intermediário (II), Úmido e Intermediário (UI), Seco e Intermediário (SI), Mais Seco e Mais Quente (MSQ), Mais Úmido e Mais Quente (MUQ), Úmido e Mais Quente (UMQ), Seco e Quente (SQ), Seco e Frio (SF), Úmido e Quente (UQ), Úmido e Frio (UF)...
Universo (TT)	[0,...,1]%

O sistema de inferência *Fuzzy* foi desenvolvido com o uso do Matlab 6.5 e o *Fuzzy Logic Toolbox*, Matsumoto (2004).

O modelo é composto por quatro módulos fundamentais: Módulo de Entrada (*fuzzificação*), onde ocorre a representação das entradas numéricas de dados através de conjuntos difusos; Base de conhecimento, que representa o conjunto de regras estabelecidas com base na intuição e/ou na experiência, de um ou mais peritos no sistema a modelar; Motor de

inferência**,** que determina o valor da variável do conjunto difuso de saída, com base nos parâmetros estabelecidos nos módulos de *fuzzificação*; e o Módulo de *desfuzzificação*, onde se tem a obtenção de um valor real (determinístico) para a saída.

As variáveis difusas de entrada se referem às características dos elementos climáticos (térmicos e hídricos), estas entradas tiveram valores precisos derivados de medições (temperaturas média, mínima e máxima e amplitude térmica diária, precipitação e dias de chuva) ou estimadas (evapotranspiração, deficiência e excedente hídricos). Após a organização da base de dados, foi efetivado o mapeamento destes dados junto aos conjuntos *fuzzy*, que é realizado no estágio de *fuzzificação.* Devido ao tipo de dado e variável utilizada, a função aplicada para *fuzzificar* os universos foi a Gaussiana, que é baseada na média e no desvio-padrão.

A inferência é o procedimento no qual foi realizada a combinação dos conjuntos *Fuzzy* (tabela 2), e utilizado o operador de implicação (se... então...) e o conectivo lógico (e) para caracterizar qual a consequência para cada combinação de antecedentes (Tabela 3).

Após estabelecer as regras, que consiste na estratégia de classificação dos tipos de tempo na bacia, para cada dia do ano e realizar a inferência, seguiu-se para outro passo, que consiste na *defuzzyficação,* que pode ser definida como a etapa em que ocorre a transformação dos resultados *fuzzy* em um valor numérico, através do método do centro de gravidade.

O agrupamento *"fuzzy"* começa de um conjunto de dados (vetores reais) que serão classificados e agrupados através do algorítmo FCM ("*Fuzzy-C-Means")*. Obtém, então, os agrupamentos nebulosos ("*Fuzzy Clustering*"), depois, é realizada a "*Fuzzyficação"* e as funções de pertinência são extraídas a partir destes agrupamentos.

O algorítmo FCM ("*Fuzzy-C-Means*") aplica a distância Euclidiana, procurada por grupos esféricos de aproximadamente mesmo tamanho e o número de grupos pode ser determinado previamente ou não. O coeficiente de distância está baseado no conceito de que a medida quantificada do grau de similaridade entre duas ou mais variáveis, pode ser demonstrada pela distância que separa uma variável da outra, em um sistema de coordenadas retangulares, portanto quanto maior for o grau de similaridade entre duas ou mais variáveis, menor será a distância entre elas.

Mapeamento e interpolação dos dados

Para verificarmos a espacialização das unidades climáticas e seu comportamento espacial, detectamos a necessidade de mapear diariamente, semanalmente, mensalmente, sazonalmente e no final estabelecer um mapa síntese, no qual as unidades climáticas foram delimitadas em função de sua maior ocorrência em todo o período analisado. Isto é, durante os 3 anos analisados e mapeados, a estação de Juiz de Fora, por exemplo, em 71% dos dias estava inserida na classe 4, 20% dos dias inserida na classe 3, 2% dos dias inserida na classe 2, 6% inserida na classe 5 e 1% dos dias inserida na classe 1. Caracterizando desta forma, como sendo uma localidade pertencente à classe 4. Sendo os mapas diários ou de sínteses criados a partir da maior ocorrência de determinada estação meteorológica, durante todo o período analisado, em uma dada classe climática. A partir deste dado foram realizados os mapas semanais, mensais, sazonais e de síntese.

Como trabalhamos com pontos amostrais e ao colocarmos estas informações no mapa ela passa a inferir uma conotação não mais pontual e sim espacial, e para isso foi necessário realizar uma interpolação dos dados.

A interpolação é um procedimento para a prognose de valores desconhecidos, utilizando alguns valores conhecidos para as regiões vizinhas. Os pontos das regiões vizinhas podem ser regularmente ou irregularmente espaçados (BURROUGH, 1986).

Assim, a qualidade de uma interpolação de dados resulta na distribuição dos pontos conhecidos usados no cálculo e na correção dos modelos matemáticos com o fenômeno. A interpolação assume o fato de que o fenômeno a ser interpolado seria rigorosamente

aproximado pela função matemática usada. Os valores desconhecidos são calculados de acordo com esta função. Desta forma, a escolha de um modelo apropriado é essencial para obter resultados razoáveis (ARANOFF, 1989). Os melhores resultados são obtidos quando a função matemática comporta-se de maneira similar ao fenômeno.

Existem vários métodos de interpolação dos dados, no presente trabalho optou-se pela técnica do kriging, que foi a que melhor respondeu a espacialização dos dados, tendo sido outros métodos experimentados. A definição de utilizar este método foi em função da distribuição dos pontos amostrais, que se encontravam distribuídos de forma irregular por toda a área de estudo. Na krigeagem o procedimento é semelhante ao de interpolação por média móvel ponderada, exceto que neste método os pesos são determinados a partir de uma análise espacial, baseada no semivariograma experimental. Em média, a krigeagem fornece estimativas nas quais a diferença entre valores estimados e verdadeiros para o mesmo ponto deve ser nula e estes estimadores possuem a menor variância dentre todos os estimadores não tendenciosos SILVA et al. (2003).

A krigeagem engloba um vasto conjunto de métodos de estimação e nesse projeto optou-se pela utilização do método da Krigeagem Ordinária. Este é um método interpolador exato no sentindo de que, quando as equações de combinação linear e de estimador não tendencioso forem utilizadas, os valores interpolados irão coincidir com os valores dos pontos amostrais.

RESULTADOS E DISCUSSÃO

As variáveis climáticas

A análise discriminante evidenciou, como mais importantes na separação das classes, as variáveis temperatura média diária, precipitação, deficiência e excedente hídrico.

Logo, a variável temperatura mínima diária, a temperatura máxima diária e a amplitude térmica, pouco contribuíram na discriminação dos grupos, pois mantinha o mesmo comportamento da temperatura média diária. Sendo que, a temperatura média diária teve um alto grau de discriminação dos grupos, e foi a variável mais discriminante quando as variáveis: precipitação total diária e deficiência e excedente hídricos eram pouco representativas entre os grupos principalmente no período de maio a setembro, quando os valores de precipitação são muito baixos ou ausentes e consequentemente os de excedente hídrico), isto é, possuíam o valor zero em quase todas as localidades, não conseguindo, por meio desta variável identificar os grupos, pois os valores eram iguais, em quase todas as localidades. As variáveis: precipitação, deficiência e excedente hídricos foram boas para discriminar os grupos, quando representativas (no período de outubro a abril). A variável evapotranspiração têm uma grande correlação com os valores de temperatura média e precipitação total, geralmente ela discriminou os grupos conjuntamente a estas variáveis.

Como foi verificado uma correlação muito grande entre a precipitação e os valores de excedente e deficiência hídrica, optou-se por utilizar os mesmos termos linguísticos, definindo se uma região é mais ou menos úmida do que a outra.

Os agrupamentos nebulosos ("*Fuzzy Cluster*")

A aplicação do método de agrupamentos nebulosos (FCM) para as sete variáveis gerou cinco grupos: 1, 2, 3, 4 e 5. A opção de trabalhar com 5 classes, foi estabelecido por apresentar uma melhor definição pelo programa utilizado, assim como, optamos pela seleção de classes a partir de uma mesma quantidade, pressupondo diferentes níveis de perda de informação e qualidade, e não pelos valores indicimétricos das ligações, que pressupõem quantidades variadas de classes mantendo porém um bom nível qualitativo. De forma empírica, poderíamos também determinar níveis de corte que definissem apenas as mesmas cinco classes obtidas por meio estatístico.

A declaração condicional *se T é Q; P é S; DH é S; EH é S então TP é MFS*, tem uma função de pertinência (T, P, DH, EH) que mede o grau de verdade da relação de implicação entre T, P,

ISSN: 1980-055x (Impressa) 2237-8642 (Eletrônica)

DH, EH, onde T é a Temperatura média, P é a precipitação total, DH é a deficiência hídrica e EX é o excedente hídrico. Desta forma a regra pode ser apresentada como: *se* baseado na precipitação total a região é muito seca; *se* segundo os dados de temperatura a região é muito fria; *se* segundo os dados de deficiência e excedente hídrico a região é muito úmida, então os tipos de tempo daquela localidade pode ser considerado como Mais Fria e Mais Úmida ou seja, R_{11}: *se* (T é MF); (P é MU); (DF é NULA) *e* (DE é MU), *então* (TP é MFU).

Tabela 2
Regras utilizadas na inferência *fuzzy,* para as variáveis temperatura, precipitação, deficiência e excedente hídricos

Precipitação/deficiência e excedente hídricos	Temperatura (T)/Temperatura Mínima (TMI)/Temperatura Máxima (TMA)/Amplitude Térmica (AT)				
(P, DH e EX)	MF	F	I	Q	MQ
MS	MFMS	FMS	IMS	QMS	MQMS
S	MFS	FS	IS	QS	MQS
I	MFI	IF	II	QI	MQI
U	MFU	FU	IU	QU	MQU
UM	MFMU	FMU	IMU	QMU	MQMU

Tabela 3
Terminologia linguística utilizada para a classificação dos tipos de tempo

TIPOS DE TEMPO	CONJUNTOS
FRIO E ÚMIDO	
FRIO E SECO	
INTERMEDIÁRIO	
QUENTE E ÚMIDO	
QUENTE E SECO	

Foram considerados elementos integrantes de um mesmo conjunto ou classe, um intervalo de valores, que variou em função do dia e de uma maior ou menor proximidade entre os valores extremos (mínimo e máximo), não foi estipulado, portanto, um valor constante de corte entre as classes. Desta forma, para cada matriz, variaram as distâncias consideradas, pois estabelecer distâncias padrões, não corresponderiam à realidade dos dados presentes, mesmo que isto, a princípio, venha a ser considerado um empecilho para uma futura comparação dos dados.

As classes referentes aos agrupamentos foram geradas pelos índices de pertinência sequencial. Foram estabelecidas cinco classes para as matrizes, que foram mantidas por todo o período analisado e para todas as variáveis utilizadas. Proporcionando a comparação das informações da fase final de processamento, ao mesmo nível em que foram agrupadas as informações nas matrizes geradoras.

Assim, foi possível termos um quadro de síntese do tempo atmosférico diário, que obedeceu àquele padrão já exposto anteriormente, que consiste em subdividir a área em que o grupo com menores temperaturas médias diárias, por exemplo, é o grupo 1 < grupo 2 < grupo 3 < grupo 4 < grupo 5 que representa o grupo com as temperaturas mais altas da área da bacia, para um determinado dia do mês.

Os agrupamentos nebulosos levam em consideração exatamente esta flexibilidade para classificarmos seus grupos, se estabelecermos classes fixas estaria trabalhando com o processo booleano e não nebuloso.

ISSN: 1980-055x (Impressa) 2237-8642 (Eletrônica)

A bacia do rio Paraibuna-MG

O rio Paraibuna nasce na Serra da Mantiqueira no município de Antônio Carlos, a cerca de 1.200m de altitude e percorre aproximadamente 166km, dos quais 44km como divisa natural dos Estados de Minas Gerais e Rio de Janeiro. Seus principais afluentes são os rios Peixe, Preto e Cágado, que drenam os 6.777km² da área da bacia, localizada na Zona da Mata de Minas Gerais, onde vivem 771.202 hab. segundo a estimativa da população feita pelo IBGE para 2005 (Figura 1).

Desta população 90,9% estão concentradas em áreas urbanas e somente 9,1% em áreas rurais. A população total do município de Juiz de Fora representa quase 80% da população total da bacia, e é aí também onde concentram-se mais de 90% das indústrias da bacia, distribuídas principalmente entre os setores siderúrgico, químico, têxtil, papel e papelão e alimentícios. A agropecuária destaca-se como outra atividade econômica de importância e está presente em todos os municípios da bacia.

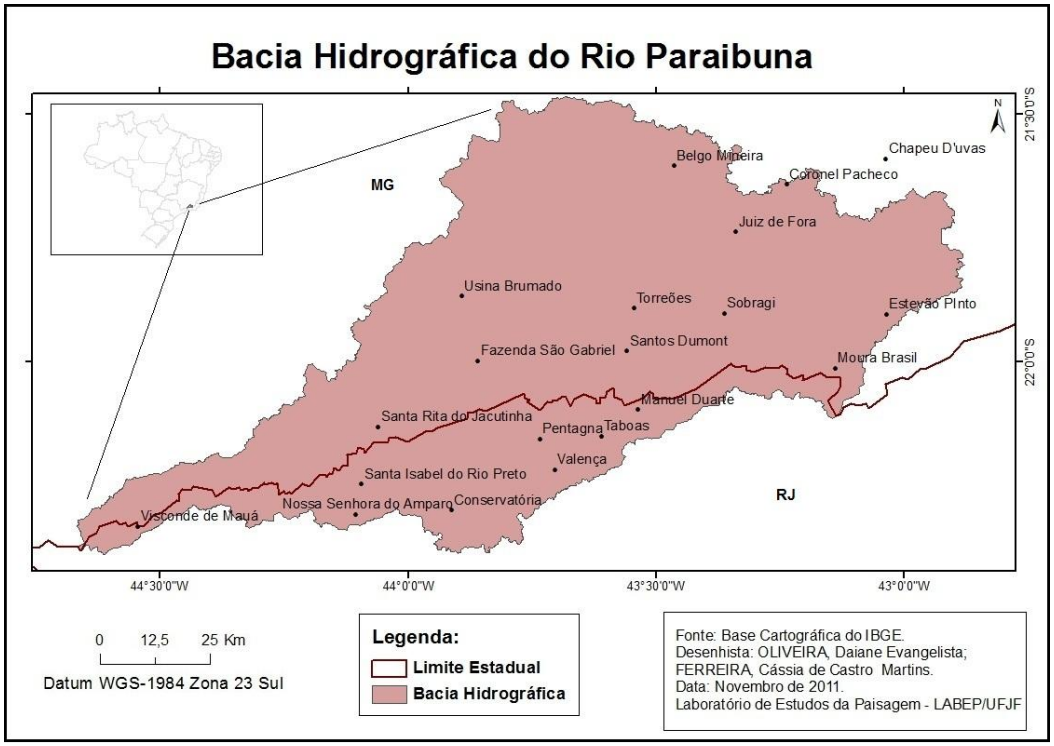

Figura 1. Localização da Bacia Hidrográfica do rio Paraibuna –MG/RJ e das 20 estações meteorológicas utilizadas na análise.

Frente o que foi analisado pode-se considerar que a bacia do rio Paraibuna, quando subdividida nas sub-bacias dos rios Peixe, Preto e Cágado apresentam condições ambientais bastante distintas. Os aspectos mais relevantes na diferença ambiental destas bacias estão associados principalmente na interação das variáveis geomorfológicas com variações nas condições climáticas, vegetacionais e de uso antrópico.

As diferenças geofísicas entre estas bacias, marcadas pelos aspectos geológicos/geomorfológicos, estão ligadas ao processo evolutivo da Serra da Mantiqueira.

As bacias dos rios Paraibuna, Peixe e Preto representam a escarpa interior (vertente norte) da Serra da Mantiqueira voltada para o vale do rio Paraíba do Sul, enquanto que a do rio Cágado representa um planalto com suave inclinação para sul, até atingir o rio Paraibuna.

A bacia do Paraibuna, associada à ocupação do Vale do rio Paraíba do Sul, apresenta um histórico de uso do solo que se inicia, desde o final do século XVIII, iniciada pelo Ciclo Cafeeiro, e posteriormente o seu declínio, no final do século XIX, pela substituição por atividades de pecuária extensiva e leiteira. Além do tempo de uso do solo, a intensidade e forma de suas intervenções garantem à bacia do rio Paraibuna um cenário de degradação. Este

ambiente, hoje, pode ser caracterizado pela ausência de cobertura florestal densa, solos erodidos e pouco férteis e redução dos mananciais de água. Consequentemente as atividades econômicas, associadas ao uso do solo nesta área, tornaram-se pouco produtivas. Este cenário de degradação é também representativo em grandes porções do vale do rio Paraíba do Sul, é fundamental a implementação de metodologias adequadas à reabilitação ambiental que considera a restauração da funcionalidade original dos geossistemas, a fim de minimizar processos erosivos, melhorar a qualidade da água, incrementar a biodiversidade e proporcionar meios de sustentabilidade para comunidade local.

Desta forma, a bacia do rio Paraibuna que originalmente, nos remete a uma visão e percepção de homogeneidade, hoje mostra uma forma de uso e ocupação, que vem levando a contrastes na paisagem e consequentemente a uma diferenciação em várias unidades climáticas, derivada da intensa ação humana nos últimos dois séculos.

Caracterização das classes e categorias climáticas, realizadas por meio da análise de agrupamento nebuloso na bacia do Rio Paraibuna-MG/RJ

Poderíamos destacar que: os grupos foram definidos pelas variáveis: temperatura média diária, evapotranspiração, precipitação total diária, deficiência e excedente hídrico, porém estas variáveis tiveram um maior ou menor poder de discriminação dos grupos em função do período do ano. Este sintoma nos levou a discriminar dois períodos, um que corresponderia ao período de verão, que compreendem os meses de janeiro, fevereiro, março, outubro, novembro e dezembro, e o outro período corresponderia ao período de inverno, que compreendem os meses de abril, maio, junho, julho, agosto e setembro.

No período de verão (outubro a março), temos como variáveis principais de discriminação dos grupos, a precipitação total diária e correlacionada a ela o índice de excedente hídrico, que nos meses de março e outubro, foi mais incidente, porém, apresentando alguns períodos com um baixo índice de deficiência hídrica em algumas localidades. E, no período de inverno (abril a setembro), temos como variáveis principais de discriminação dos grupos a temperatura média diária e correlacionada a ela a evapotranspiração e os índices de deficiência hídrica, a precipitação e o excedente hídrico quando incidentes, possuía um menor grau de discriminação, devido o número muito baixo de dias e estações meteorológicas com ocorrência de chuvas neste período do ano, e sem a presença de excedente hídrico. Assim, estas variáveis (precipitação e excedente hídrico), foram muito pouco utilizadas para a definição dos grupos neste período.

A partir, das análises diárias, mensais, sazonal, foi possível estabelecer um quadro da síntese das situações de tempo da Bacia, no qual foram mantidas as cinco classes definidas desde o início da classificação da área, consistindo nos cinco seguimentos principais e distintos de unidades climáticas:

Unidade Climática 1 – Mantiqueira 1 (envolve a classe 1)

Esta é uma unidade bem discriminada climaticamente das demais, devido ao relevo. É uma área que predomina a participação da onda de leste, com interferências das Frentes Polares, mas que, devido ao relevo têm um comportamento climático diferenciado.

A região mais ao Sul da bacia se localiza nos contrafortes ocidentais da Serra da Mantiqueira, aonde possui volumes de chuva maiores do que das demais regiões, chegando a 2000 mm/a em Visconde de Mauá, as precipitações são intensificadas devido ao fator orográfico, que é representativo na região. Devido às altitudes mais elevadas, temos a ocorrência de temperaturas mais baixas. Possui no inverno uma redução na quantidade de chuvas, mas que ainda é superior as outras áreas.

Podemos destacar que a Massa Tropical Atlântica predominante no inverno, sofre um resfriamento, levando a um aumento significativo da sua estabilidade. O mau tempo, seguido

ISSN: 1980-055x (Impressa) 2237-8642 (Eletrônica)

de precipitações restringe-se normalmente as passagens frontais, que chegam já um pouco enfraquecidas, principalmente no inverno.
O que mais diferencia a classe 1 e 2, das demais, é a quantidade de chuvas e a temperatura, que chega a ser de quase 1000 mm/a de diferença. Devido principalmente às diferenças altimétricas. No qual, Visconde de Mauá está localizada a 1030 m de altitude e as demais estações se encontram em torno de 550 m de altitude.

Unidade Climática 2 – Mantiqueira 2 (classe 2)
Esta unidade climática em muito se assemelha a Unidade climática 1, no qual o relevo tem uma função primordial seja nos totais de precipitação ou nas temperaturas mais amenas. Possui uma participação significativa da onda de leste, com interferências durante todo o ano das Frentes Polares, normalmente ocasionando chuvas e quedas representativas da temperatura.
Esta região se localiza nos contrafortes ocidentais da bacia, já evidenciando precipitação menor devido ao efeito orográfico a barlavento da Serra da Mantiqueira, no qual recebe o ar úmido vindo do oceano. Com altitudes ainda elevadas as temperaturas são mais baixas do que nas demais unidades climáticas, excetuando a Unidade climática 1. Possui uma sazonalidade das precipitações melhor definida, com um inverno um pouco mais seco.
A Massa Tropical Atlântica tem uma atuação predominante no inverno, no qual ocorre maior estabilidade atmosférica, sendo alternada pelas incursões das frentes frias que podem provocar precipitações e queda da temperatura

Unidade Climática 3 – Intermediária (Classe 3)
Esta região é marcada pela sazonalidade das precipitações, no qual a definição de um período seco que se estende de maio a setembro. Este período está associado à atuação da Massa Tropical Atlântica, que propicia uma estabilidade atmosférica, geralmente associado a um campo de Alta Pressão, com a predominância de um ar seco, podendo inclusive apresentar altas temperaturas durante o dia, devido a baixa nebulosidade. Este sistema é impactado pela atuação das Frentes Frias, que normalmente não provocam precipitações e quando ocorrem não representam volumes superiores a 70 mm para todo o período seco, mas que implicam em um decréscimo das temperaturas.
É a região que possui um padrão médio entre as regiões anteriores 1 e 2; 4 e 5, em que tanto em termos térmicos e hídricos estão incluídos em uma classe intermediária, que pode também estar associado a altitude, que compreende uma classe intermediária, entre a Região da Serra da Mantiqueira, e a região da foz do Rio Paraibuna, estando localizada na média a 800m de altitude.
O que mais diferencia das classes 1 e 2 é a quantidade e distribuição das precipitações, além de apresentarem temperaturas mais elevadas. Com relação a Unidade 4 e 5 é a diferença de temperatura entre as duas, no qual a Unidade climática 5 as temperaturas são mais elevadas não variando muito a quantidade e frequência das precipitações. Esta Unidade climática tem uma alta similaridade a Unidade 4.

Unidade Climática 4 - Juiz de Fora (envolve a classe 4)
Esta região possui um período seco muito bem definido, que deve a participação da Massa Tropical Atlântica, e que mesmo com a passagem da massa polar, que aí não estaria associada a altas altitudes, não conseguem provocar chuvas com volumes de precipitação significativos, raramente ultrapassando os 50 mm, durante todo o período de inverno, apenas uma pequena queda de temperatura.
Além de destacar como uma área aonde a atividade humana é mais ativa, com o estabelecimento de um complexo urbano industrial mais representativo, quando analisado sob o conjunto da região, levando a uma maior individualização deste espaço a nível climático.

ISSN: 1980-055x (Impressa) 2237-8642 (Eletrônica)

É a região que possui um padrão médio entre as regiões anteriores, em que tanto em termos térmicos e hídricos estão incluídos em uma classe intermediária, que pode também estar associado a altitude, que também está numa classe intermediária, entre a Região da Serra da Mantiqueira, e a região da foz do rio Paraibuna.
O que mais diferencia as classes três e quatro é a diferença de temperatura entre as duas, não variando muito a quantidade e frequência das precipitações.

Unidade Climática 5 - Foz do Rio Paraibuna
As características geomorfológicas da região mais próxima à foz do rio Paraibuna, com o rio Paraíba do Sul possui características distintas.
É uma região que se destaca como um faixa de menor umidade, tendo inclusive um maior número de incidência de deficiência hídrica, e suas temperaturas são mais elevadas, associado às altitudes mais baixas. Está localizada entre as Serras do Mar e da Mantiqueira, em que os valores de precipitação em torno de 1000 mm/a podem ser justificadas, por ser uma área aonde predomina a atuação da onda de leste e a descida, a sotavento, do ar vindo do oceano, que já perdeu grande parte da sua umidade, devido a precipitação forçada, quando o ar teve que ultrapassar a Serra do Mar.
Apresentando, como em toda a região, duas estações bem definidas uma seca e outra úmida.
Esta subdivisão da região estudada em cinco unidades climáticas pode ser visualizada no mapa final (Figura 2) de definição das unidades climáticas para a Bacia do rio Paraibuna, baseada na sucessão dos tipos de tempo diários mediatizados pela análise de agrupamento, utilizando a lógica *fuzzy*.

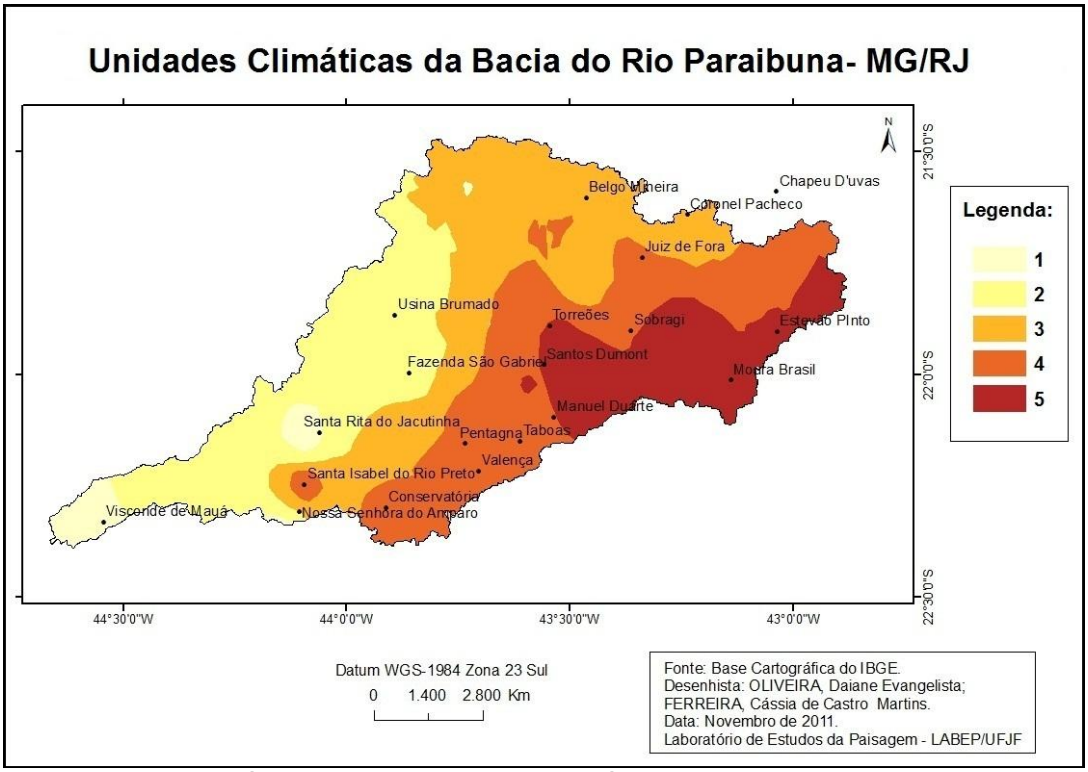

Figura 2. Esquema síntese final das unidades climáticas para a Bacia do Rio Paraibuna – MG.

CONSIDERAÇÕES FINAIS
A estatística, a estatística espacial (SIG), e o próprio modelo utilizado, possibilitaria uma visão de que a análise geográfica do clima é ao mesmo tempo qualitativa, mas também quantitativa e possível de ser analisada, interpretada, classificada, utilizando modelos que permitam não só a descrição quântica do fenômeno, mas que também trabalha de forma a nos fornecer uma melhor e maior capacidade na interligação, conexão, similaridade, pertinência entre um número muito grande de dados numéricos, que sozinhos não dizem muita coisa, mas quando

juntos e comparados podem nos dar grandes informações para serem analisadas, questionadas e interpretadas.

A inferência climática nebulosa mostra uma grade de valores relativos, que supera tecnicamente o processo de intersecção de conjuntos espaciais, como as operações baseadas na lógica booleana, que apresentam a mesma ordem de grandeza. Possibilita a análise do espaço geográfico de forma contínua, e não por meio de limites rígidos. Permitindo descrever e modelar estas variáveis, que são nebulosas por si só, em termos do quanto cada uma contribui para a formação do grupo, estabelecendo grupos com uma determinada caracterização climática e mais representativa da realidade. Viabilizando a criação de uma superfície de decisão, que representa uma variação da grandeza avaliada, oferecendo uma flexibilidade maior na definição das categorias climáticas.

AGRADECIMENTO

Agradecimento a FAPEMIG pelo incentivo e financiamento da pesquisa.

REFERÊNCIAS

.ARANOFF, S. **Geographic Information Systems: a management perspective**. Ottawa: WDL Publications, 1989. 294p.

.ASPIAZÚ, C.; RIBEIRO, G.A.; VIANELLO, R.L.; RIBEIRO, J.C.; VALENTE, O.F.; PAULA NETO, F. de. **Análise dos componentes principais, aplicada na classificação climática do Estado de Minas**. Teste Metodológico. *Revista Árvore*, v.14, n.1, p.1-15, 1990.

.BRAGA, A.L. **Ferramentas de Manipulação Nebulosa de Dados com Aplicação em Sistemas de Informações Geográficas.** Rio de Janeiro, UFRJ, 1987 (Dissertação de Mestrado – COPPE-UFRJ), 1998. 169p.

. BURROUGH, P.A. **Principles of Geographical Information Systems for Land Resources Assessment.** New York: Oxford University Press, 1986. 193p.

.BURROUGH, P.A.; McDONNELL, R.A. **Principles of geographic information systems**. Oxford: Oxford University, 1998. 333p.

.DPI. Desenvolvido pela DPI (Divisão de Processamento de Imagens) – INPE. São José dos Campos, 2004. **Apresenta ajuda ao usuário na maneira como utilizar as ferramentas de Geoestatística do *Software* SPRING.** Disponível em: http://www.dpi.inpe.br/spring/usuario/geoestati.htm. Acesso em: 11 nov 2005.

.FERREIRA, C.C.M. ***Tipos de tempo e categorias climáticas na Bacia do Rio Paraibuna-MG***. São Paulo, USP, 2002 (Tese de Doutorado – FFLCH-USP), 2002. 246p.

.FERREIRA, L.C.M. ***Aplicação de lógica nebulosa na análise de grupos populacionais***. Rio de Janeiro, UFRJ, 2000 (Dissertação de Mestrado – Engenharia Biomédica-UFRJ), 2000. 141p.

.KLIR, G.G. e YUAN, B. ***Fuzzy Sets and Fuzzy Logic: Theory and Applications.*** New Jersey: Prentice Hall, 1995. 574p.

.KLIR, G.G. **Fuzzy Databases and Information Retrieval Systems**. In: *Fuzzy Sets and Fuzzy Logic: Theory and Applications.* New Jersey: Prentice Hall, 1995, p. 49-78.

.MATSUMOTO, E.Y. *MATLAB 6.5:* ***Fundamentos de Programação*****.** 2.ed. São Paulo: Érica, 2004. 342p.

.R DEVELOPMENT CORE TEAM (2005). R: **A language and environment for statistical computing.** R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org.

.ROUSSEEUW, P.J. **Discussions: Fuzzy Clustering at the Intersection**. *Technometrics*, v.37, n.3, p, 283-286, 1995.

.RUHOFF, A.L. et. al. Lógica **Fuzzy e zoneamento ambiental da Bacia do Arroio Grande**. In: *Anais* XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, p. 2355-2362, 2005.

.SANT'ANNA NETO, J.L. **Por uma geografia do clima: antecedentes históricos, paradigmas contemporâneos e uma nova razão para um novo conhecimento.** *Terra Livre: Publicação da Associação dos Geógrafos Brasileiros*, São Paulo, n. 17, Segundo semestre, p, 49-62, 2001.

.SILVA, S. A. B., SIMÕES, P. M., MARTINS, F. R., PEREIRA, E. B**. Aplicação da geoestatística no desenvolvimento de uma base de dados climatológicos para uso no modelo de transferência radiativa BRASILSR**, Belo Horizonte, Abril 2003.

.ZADEH, L.A. Fuzzy Sets. **Information and Control**, v. 08, n.08, p, 338-353, 1965.

.ZIMMERMANN, H.J. **Fuzzy Set Theory and its Applications***.* Boston – Dordrecht – Lancaster: Kluwer Nijhoff Publishing, 1985. 363 p.