

# UERJ

Dissertação de Mestrado em Engenharia de Computação

**Integração Temática de Dados Espaciais para Avaliação de Processos Erosivos em  
Bacias Hidrográficas Utilizando Lógica Fuzzy e Pesos de Evidência**

Autor: Halisson Belintani

Orientadora: Margareth Simões Penello Meirelles

Programa de Pós-Graduação em Engenharia de Computação  
Área de Concentração: Geomática

Setembro - 2007



**Faculdade de Engenharia**

**INTEGRAÇÃO TEMÁTICA DE DADOS ESPACIAIS PARA AVALIAÇÃO DE  
PROCESSOS EROSIVOS EM BACIAS HIDROGRÁFICAS UTILIZANDO LÓGICA  
FUZZY E PESOS DE EVIDÊNCIA**

**Halisson Belintani**

Dissertação submetida ao corpo docente da Faculdade de Engenharia da Universidade do Estado do Rio de Janeiro – UERJ, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia de Computação – Área de concentração Geomática.

Orientadora: Margareth Simões Penello Meirelles

Programa de Pós-Graduação em Engenharia de Computação  
Área de Concentração: Geomática

Rio de Janeiro  
Setembro – 2007

BELINTANI, HALISSON

Integração Temática de Dados Espaciais para Avaliação de Processos Erosivos em Bacias Hidrográficas Utilizando Lógica Fuzzy e Pesos de Evidência, [Rio de Janeiro] 2007.

XII, 133p. 29,7 cm (FEN/UERJ, M.Sc., Engenharia de Computação – Área de Concentração Geomática, 2007).

Dissertação – Universidade do Estado do Rio de Janeiro – UERJ.

1. Sistema de suporte a decisão. 2. Lógica “*Fuzzy*”. 3. Pesos de Evidência. 4. Método de Bayes 5. Degradação ambiental.

I. FEN/UERJ. II. Título (série).

## FOLHA DE JULGAMENTO

Título: Integração Temática de Dados Espaciais para Avaliação de Processos Erosivos em Bacias Hidrográficas Utilizando Lógica Fuzzy e Pesos de Evidência.

Candidato: Halisson Belintani

Programa de Pós-Graduação em Engenharia de Computação

Área de Concentração: Geomática

Data da defesa: 06 de Setembro de 2007

Aprovada por:

---

Orientadora: Margareth Simões Penello Meirelles, D.Sc., UERJ

---

Heitor Luiz da Costa Coutinho, Ph.D., EMBRAPA Solos

---

Lucieta Guerreiro Martorano, D.Sc., EMBRAPA Solos

---

Ubiratan Porto dos Santos, Ph.D., NCE/UFRJ

## **DEDICATÓRIA**

Aos meus pais Maria José e  
João Batista e à minha irmã Thais.

Resumo da Dissertação apresentada a FEN/UERJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

## **INTEGRAÇÃO TEMÁTICA DE DADOS ESPACIAIS PARA AVALIAÇÃO DE PROCESSOS EROSIVOS EM BACIAS HIDROGRÁFICAS UTILIZANDO LÓGICA FUZZY E PESOS DE EVIDÊNCIA**

Halisson Belintani

Setembro/2007

Orientadora: Margareth Simões Penello Meirelles

Programa de Pós-Graduação em Engenharia de Computação – Área de Concentração em Geomática – Mestrado.

Atualmente a ação do homem tem impactado e contribuído para o acelerado processo de erosão e degradação das terras em diversas bacias hidrográficas, devido ao processo de retirada da cobertura vegetal geralmente associado à expansão das atividades agropecuárias, ocasionando mudanças no uso/cobertura do solo.

Tendo em vista esse problema o presente trabalho aborda duas técnicas de análise espacial com a integração de quatro indicadores (uso do solo, erosividade da chuva, solo e conflito de uso) para avaliar o processo erosivo em bacias hidrográficas. Uma das técnicas é baseada em dados dirigidos pelo conhecimento de especialistas (“*knowledge-driven*”), usando o método de Lógica “*Fuzzy*” e a outra técnica é baseada em dados dirigidos por formulações estatísticas (“*data-driven*”), usando o modelo de “Pesos de Evidência” (*Weights of evidence*).

Com a utilização de interfaces eficientes de visualização e análise dos Sistemas de Informação Geográfica, o trabalho visa através do método de Lógica “*Fuzzy*” gerar um mapa de identificação de áreas de ocorrência de erosões a partir da integração temática e através do método de “Pesos de Evidência” gerar um mapa de potencial de ocorrência de erosão.

A partir daí, pretende-se sugerir uma metodologia de integração temática e avaliar os resultados da aplicação destas duas técnicas. Escolheu-se como estudo de caso a aplicação destas técnicas na Bacia do Alto Taquari onde o processo erosivo acentuado vem trazendo conseqüências ao Pantanal Matogrossense.

Palavras-chave: erosão, Lógica “*Fuzzy*”, Pesos de Evidência, análise espacial, ArcSDM, degradação ambiental.

Abstract of Dissertation presented to FEN/UERJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.).

**THEMATIC INTEGRATION OF SPACE DATA FOR EVALUATION OF EROSION PROCESSES IN RIVER BASINS USING FUZZY LOGIC AND WEIGHTS OF EVIDENCE**

Halisson Belintani

September/2007

Advisor: Margareth Simões Penello Meirelles

Program of Computing Engineering – Geomatic.

Nowadays man's action has impact and contributed to the accelerated erosion process and degradation of the lands in several basins, due to the process of vegetation removal usually associate with the implementation of agricultural activities, inducing land use/land cover change.

The present work approaches two techniques of space analysis with the integration of four indicators (land use, rainfall erosivity, soil and land use conflict) to evaluate the erosion process in river basins. One of the techniques is based on specialists knowledge (“knowledge-driven”), using Fuzzy Logic and the other technique is based on statistical formulations (“data-driven”), using Weights of Evidence.

Efficient interfaces of visualization and spatial analysis were achieved by Geographical Information System. The research seeks through the use of Fuzzy Logic to generate erosion map using thematic integration and through the use Weights of Evidence technique the generation of erosion potential map was achieved.

A methodology was suggested aiming the thematic integration and the evaluation of the results of the application of these two techniques. It was chosen as case study the application of these techniques in the Alto Taquari watershed where an intensive erosive process is causing several consequences to the Pantanal area.

Keywords: Erosion, Fuzzy Logic, Weights of evidence, Spatial analysis, ArcSDM, Environmental degradation.

## SUMÁRIO

<b>CAPÍTULO 1 – INTRODUÇÃO.....</b>	<b>1</b>
1.1 Visão geral.....	1
1.2 Motivação.....	2
1.3 Justificativa.....	3
1.4 Objetivo geral.....	3
1.5 Objetivo específico.....	3
1.6 Estrutura da dissertação.....	4
<b>CAPÍTULO 2 – PROCESSO EROSIVO.....</b>	<b>5</b>
2.1 Erosão.....	5
2.2 Tipos de Erosão.....	6
2.3 Causas do processo erosivo.....	7
2.4 Indicadores do processo erosivo em bacias hidrográficas.....	8
2.5 Impactos ambientais associados ao processo erosivo.....	10
2.6 Considerações finais.....	13
<b>CAPÍTULO 3 – INTEGRAÇÃO TEMÁTICA.....</b>	<b>14</b>
3.1 Ferramentas de integração de dados espaciais.....	14
3.1.1 Sistemas de Suporte a Decisão.....	15
3.1.2 Sistemas de Informação Geográfica.....	16
3.1.3 Análise Multicritério.....	18
3.1.4 Processo de Análise Hierárquica (AHP).....	21
3.2 Técnicas de integração temática.....	25
3.2.1 Lógica “ <i>Fuzzy</i> ”.....	26
3.2.1.1 Conjuntos “ <i>Fuzzy</i> ”.....	27
3.2.1.2 Funções de pertinência .....	28
3.2.1.3 Representações das funções de pertinência .....	28
3.2.1.4 Operações entre conjuntos “ <i>Fuzzy</i> ” .....	31
3.2.1.5 Sistema “Expert <i>Fuzzy</i> ” .....	32
3.2.2 Pesos de Evidência.....	34
3.2.2.1 Descrição do método de Pesos de Evidência .....	35

3.2.2.1 Independência condicional dos dados.....	45
3.2.3 Redes Neurais.....	49
3.2.4 Regressão Logística.....	51
3.3 Técnicas de avaliação espaço-temporal.....	52
3.3.1 Modelagem Dinâmica.....	52
3.3.2 Autômatos celulares.....	54
3.4 Considerações finais.....	56
<b>CAPÍTULO 4 – METODOLOGIA PROPOSTA.....</b>	<b>57</b>
4.1 Fluxograma metodológico.....	57
4.2 Descrição das etapas da metodologia.....	58
<b>CAPÍTULO 5 – ESTUDO DE CASO.....</b>	<b>67</b>
5.1 Escolha da área de estudo.....	67
5.2 Análise do problema.....	68
5.3 Escolha dos indicadores.....	69
5.4 Identificação de materiais necessários.....	70
5.4.1 Mapas digitais e dados cartográficos.....	71
5.4.2 Sistemas computacionais.....	71
5.5 Estruturação da base de dados.....	73
5.5.1 Uso do solo.....	73
5.5.2 Erosividade da chuva.....	75
5.5.3 Solo.....	76
5.5.4 Conflito de uso.....	77
5.5.5 Pontos de localização de voçorocas.....	79
5.6 Aplicação da metodologia de classificação “ <i>Fuzzy</i> ”.....	80
5.6.1 Definição do conjunto de parâmetros de análise.....	80
5.6.2 Geração dos membros “ <i>Fuzzy</i> ”.....	82
5.6.3 Geração dos planos de informação “ <i>Fuzzy</i> ”.....	83
5.6.3.1 Planos de informação “ <i>Fuzzy</i> ” para os indicadores categóricos.....	84
5.6.3.2 Planos de informação “ <i>Fuzzy</i> ” para os indicadores ordenados.....	87
5.6.4 Escolha do operador “ <i>Fuzzy</i> ”.....	90
5.6.5 Definição do grau para o operador gamma.....	90
5.6.6 Integração dos planos de informação “ <i>Fuzzy</i> ”.....	91

5.6.7	Geração do mapa final de pertinência “ <i>Fuzzy</i> ” .....	92
5.7	Aplicação do método de predição por Pesos de Evidência.....	93
5.7.1	Definição do conjunto de parâmetros de análise.....	93
5.7.2	Cálculo dos pesos para cada mapa de predição.....	95
5.7.3	Generalização dos temas de evidência.....	99
5.7.4	Combinação dos temas de evidência.....	103
5.7.5	Avaliação das incertezas e geração do mapa de favorabilidade.....	105
5.7.6	Verificação da independência condicional.....	110
5.8	Considerações finais.....	112
	<b>CAPÍTULO 6 – RESULTADOS E DISCUSSÃO.....</b>	<b>113</b>
6.1	Resultados.....	113
6.2	Análise comparativa com outros resultados.....	115
6.2.1	Mapa de perda de solo.....	115
6.2.2	Mapa de impacto na BAT a partir do SAEB (Sistema de Análise do Risco de Impactos Ambientais Derivados da Erosão em Bacias Hidrográficas).....	117
6.2.3	Utilização de pontos de controle.....	118
	<b>CAPÍTULO 7 – CONCLUSÕES E RECOMENDAÇÕES.....</b>	<b>120</b>
7.1	Conclusões sobre a análise.....	120
7.2	Recomendações.....	121
7.3	Considerações finais.....	122
	<b>REFERÊNCIAS.....</b>	<b>123</b>

## LISTA DE FIGURAS

Figura 3.1	Estrutura hierárquica básica do AHP.....	22
Figura 3.2	Exemplo de aplicação do AHP.....	22
Figura 3.3	Representação gráfica da função de pertinência.....	30
Figura 3.4	Critério do contraste máximo.....	41
Figura 3.5	Autômato celular.....	55
Figura 4.1	Fluxograma metodológico.....	57
Figura 5.1	Localização da BAT.....	68
Figura 5.2	Seqüência para representação do mapa temático de uso do solo por microbacia.....	74
Figura 5.3	Seqüência para representação do mapa temático de erosividade da chuva por microbacia.....	75
Figura 5.4	Seqüência para representação do mapa temático de solo por microbacia.....	77
Figura 5.5	Seqüência para representação do mapa temático de conflito de uso por microbacia.....	78
Figura 5.6	Espacialização de áreas de voçorocamento na BAT.....	80
Figura 5.7	Iniciando “ <i>Set Analysis Parameters</i> ”.....	81
Figura 5.8	Caixa de diálogo “ <i>Set Analysis Parameters</i> ”.....	82
Figura 5.9	Iniciando “ <i>Define Fuzzy membership</i> ”.....	82
Figura 5.10	Caixa de diálogo “ <i>Define Fuzzy membership</i> ”.....	83
Figura 5.11	Entrada de parâmetros para dados categóricos.....	84
Figura 5.12	Mapa de pertinência do indicador uso do solo.....	85
Figura 5.13	Mapa de pertinência do indicador solo.....	86
Figura 5.14	Mapa de pertinência do indicador erosividade da chuva.....	88
Figura 5.15	Mapa de pertinência do indicador conflito de uso.....	89
Figura 5.16	Inicando “ <i>Fuzzy Logic</i> ”.....	91
Figura 5.17	Caixa de diálogo “ <i>Gamma Operator</i> ”.....	91
Figura 5.18	Mapa de potencial final “ <i>Fuzzy</i> ”.....	92
Figura 5.19	Iniciando “ <i>Set Analysis Parameters</i> ”.....	94
Figura 5.20	Caixa de diálogo “ <i>Set Analysis Parameters</i> ”.....	94
Figura 5.21	Iniciando “ <i>Calculate Theme Weights</i> ”.....	95
Figura 5.22	Caixa de diálogo “ <i>Calculate Theme Weights</i> ”.....	96
Figura 5.23	Iniciando “ <i>Generalize evidential themes</i> ”.....	99
Figura 5.24	Caixa de diálogo “ <i>Generalize evidential themes</i> ”.....	100
Figura 5.25	Caixa de diálogo do método “ <i>Map Class</i> ” de generalização.....	101
Figura 5.26	Caixa de diálogo “ <i>Group Classes</i> ”.....	102
Figura 5.27	Iniciando “ <i>Calculate Response Theme</i> ”.....	104
Figura 5.28	Caixa de diálogo “ <i>Inputs to Weights of Evidence – Themes</i> ”.....	104
Figura 5.29	Caixa de diálogo “ <i>Inputs to Weights of Evidence - Classes</i> ”.....	105
Figura 5.30	Mapa de favorabilidade da modelagem Pesos de Evidência ( <i>Woef</i> ).....	109
Figura 5.31	Caixa de diálogo “ <i>Check Condicional Independence</i> ”.....	110
Figura 5.32	Caixa de diálogo “ <i>Assessment of Condicional Independence</i> ”.....	112
Figura 6.1	Mapas gerados no estudo de caso da BAT.....	113
Figura 6.2	Mapa de distribuição de perdas de solos na BAT em 1994.....	116
Figura 6.3	Mapa de impacto na BAT a partir do SAEB.....	118

## LISTA DE TABELAS

Tabela 3.1	Escala de valores AHP.....	23
Tabela 3.2	Matriz de comparação par a par.....	24
Tabela 3.3	Representação tabular da função de pertinência.....	29
Tabela 3.4	Exemplo de pesos para dado com 10 classes ordenadas.....	40
Tabela 3.5	Tabela de contingência.....	46
Tabela 3.6	Tabela de contingência entre dois temas.....	48
Tabela 3.7	Exemplo de pesos para três mapas.....	49
Tabela 4.1	Síntese do modelamento “ <i>Fuzzy</i> ”.....	62
Tabela 4.2	Síntese do modelamento por Pesos de Evidência.....	65
Tabela 5.1	Classificação do indicador uso do solo.....	74
Tabela 5.2	Classificação do indicador erosividade da chuva.....	76
Tabela 5.3	Tipos de solos da BAT.....	76
Tabela 5.4	Classificação do indicador solo.....	77
Tabela 5.5	Tipos de uso da terra da BAT.....	78
Tabela 5.6	Classificação do indicador conflito de uso.....	78
Tabela 5.7	Coordenadas dos pontos de localização de voçorocas na BAT.....	79
Tabela 5.8	Valores dos pesos para geração dos membros “ <i>Fuzzy</i> ” de classe de uso..	84
Tabela 5.9	Valores dos pesos para geração dos membros “ <i>Fuzzy</i> ” de solo.....	86
Tabela 5.10	Valores dos pesos para geração dos membros “ <i>Fuzzy</i> ” de erosividade da chuva.....	88
Tabela 5.11	Valores dos pesos para geração dos membros “ <i>Fuzzy</i> ” de conflito de uso.....	89
Tabela 5.12	Cálculo dos ponderadores de evidência para o indicador classe de uso....	98
Tabela 5.13	Cálculo dos ponderadores de evidência para o indicador solo.....	98
Tabela 5.14	Cálculo dos ponderadores de evidência para o indicador erosividade da chuva.....	98
Tabela 5.15	Cálculo dos ponderadores de evidência para o indicador conflito de uso.	98
Tabela 5.16	Tabela de pesos (WOE) .....	105
Tabela 5.17	Tabela de variância (WOEVAR) .....	107
Tabela 5.18	Tabela de pesos de evidência (WOFE) .....	107
Tabela 5.19	Tabela de condição única para BAT.....	108
Tabela 5.20	Tabelas SDMUC_TBL e SDMUC_CI.....	111
Tabela 6.1	Tabela resumo dos pesos positivos e negativos para cada indicador.....	114

# **CAPÍTULO 1 - INTRODUÇÃO**

## **1.1 Visão Geral**

A importância da conservação e do uso sustentável das bacias hidrográficas vem sendo pauta de pesquisas e tratados intergovernamentais. O estabelecimento de diretrizes para avaliação ambiental sido prioridade em estudos desenvolvidos por organizações nacionais e internacionais voltadas ao meio ambiente [1].

Algumas bacias hidrográficas vêm sofrendo alterações ambientais devido a processos naturais, como alterações climáticas, regimes hidrológicos, entre outros, e por ações antrópicas como a agropecuária (com lançamentos de agroquímicos, desmatamentos), construções civis (com a construção de barragens, diques, estradas) e ações de indústrias (poluição química).

Os problemas gerados por esses fatores vêm se agravando ao longo dos anos contribuindo para um grave problema nessas regiões, a erosão. Segundo Correa [29], a erosão pode ser definida como um desgaste mecânico das terras, onde atualmente a ação do homem pode ser considerada o principal elemento causador desse processo.

A Bacia do Alto Taquari (BAT), localizada na região Centro-Oeste do Brasil, é uma área onde o processo erosivo se encontra em estágio avançado. Dentre as ações do homem que contribuem para o processo de erosão nesta região, destaca-se o processo de desmatamento, queimada e retirada da cobertura vegetal para fins agropecuários, sem a preocupação com o replantio ou recuperação das áreas degradadas, o que provoca um desequilíbrio do meio ambiente.

Um agravante na BAT para o processo de erosão é o fato da maioria dos desmatamentos terem sido feitos em áreas vulneráveis à erosão, onde na maior parte das vezes os solos são arenosos [29]. Há ainda um constante acúmulo de material que provoca um crescente assoreamento, observado ao longo do rio Taquari, fazendo com que o rio passe a inundar permanentemente grandes áreas de pastagens nativas na planície do Pantanal Sul Matogrossense, gerando impactos não só na pecuária com a degradação de pastagens, mas também neste bioma com a degradação do solo, da água, diminuição da fauna, entre outros.

Os problemas sociais e econômicos nesta região também vêm se agravando com o aumento de terras desvalorizadas e improdutivas, em função da aceleração dos processos de erosão no planalto, do assoreamento e de constantes inundações [1]. Além disto, como a base da economia nessa região é a pecuária extensiva de corte, desenvolvida em campos de pastagens naturais, o aumento da inundação em áreas do baixo curso do rio Taquari tem

diminuído a rentabilidade da pecuária, à medida que extensas áreas de campo passaram a ser inundadas vários meses durante o ano, desde 1974 [39].

A crescente influência das atividades humanas no funcionamento e equilíbrio de ecossistemas como este tem merecido cada vez mais atenção por parte da comunidade científica, pois várias vezes já se estabeleceram relações entre as ações do homem e degradações intensivas do solo em diversas regiões, além de perda de biodiversidade, mudanças em regimes hidrológicos e mesmo alterações climáticas [51].

Tendo em vista esses problemas, o presente trabalho irá abordar como estudo de caso a região da BAT para a aplicação e avaliação de duas técnicas de análise espacial baseadas em “dados dirigidos pelo conhecimento de especialistas” e “dados dirigidos por formulações estatísticas”, com a utilização de interfaces eficientes de visualização dos Sistemas de Informação Geográfica (SIGs), visando a geração de mapas de localização de erosões a partir de integração temática de dados e a geração de mapa de favorabilidade, isto é, mapa de potencial de novos pontos de erosão para auxiliar sistemas de suporte a decisão.

O trabalho visa auxiliar no processo de tomada de decisão através destas técnicas, monitorando e prevendo a expansão do processo erosivo na região, possibilitando ações mitigadoras para uma melhor gestão integrada desta importante bacia hidrográfica e a consequente redução dos impactos no pantanal.

Com a identificação de áreas mais favoráveis à ocorrência de erosão a partir dos mapas criados pode-se diminuir o fenômeno das voçorocas (fendas provocadas por processos erosivos) e apontar áreas prioritárias para prevenção ao surgimento de erosão, intensificando as práticas de conservação e manejo do solo e auxiliando as atividades agropecuárias sem prejuízos para o meio ambiente.

## **1.2 Motivação**

A principal motivação para o desenvolvimento do trabalho foi subsidiar a avaliação dos processos erosivos através da utilização de geotecnologias, obtendo maior conhecimento em algumas técnicas de análise espacial, a fim de agregar conhecimento que pudesse auxiliar na compreensão de problemas ambientais, na tomada de decisão e na gestão e planejamento territorial.

### **1.3 Justificativa**

A importância de se desenvolver um trabalho voltado para o estudo do processo erosivo em uma bacia hidrográfica está relacionada diretamente à necessidade de identificação e gestão destas áreas no Brasil. Por ser um país de grande extensão territorial e por falta de um planejamento efetivo do uso da terra, existem áreas no Brasil com grandes perdas de solo e acentuado processo erosivo, sendo necessária à busca por ferramentas e metodologias que auxiliem na localização destas áreas e na identificação de áreas com potencial a ocorrência de erosão.

Além do estudo em uma bacia hidrográfica, este trabalho propõe avaliar algumas ferramentas de integração temática, levando-se em consideração as modernas técnicas de análise espacial disponíveis nos Sistemas de Informação Geográfica.

### **1.4 Objetivo Geral**

O objetivo deste trabalho é propor técnicas de integração temática e suporte à tomada de decisão através da ferramenta de análise espacial “*Spatial Data Modeller*” (ArcSDM), que possibilite a indicação de áreas que são potencialmente favoráveis a ocorrência de erosão em uma bacia hidrográfica.

A integração temática de dados espaciais poderá orientar esforços na geração de mapas para auxiliar equipes multidisciplinares a identificar ocorrências de erosão além de favorecer a prevenção de novos casos, auxiliando, desta forma, o planejamento territorial e a tomada de decisão.

### **1.5 Objetivos específicos**

Análise, aplicação e avaliação de duas técnicas da ferramenta ArcSDM (Lógica “*Fuzzy*” e Peso de Evidências) aplicado à modelagem de incertezas e modelagem com pesos calculados estatisticamente para a geração de mapas que permitam identificar áreas com maior favorabilidade à ocorrência de erosão em um estudo de caso na bacia hidrográfica do Alto Taquari, MS. Pretende-se ainda avaliar se a técnica de Pesos de Evidência será satisfatória para aplicações na área ambiental, pois esta técnica foi desenvolvida originalmente para outras aplicações (aplicações não-espaciais em diagnósticos médicos), sendo adaptada para o mapeamento de favorabilidade mineral em ambientes geológicos, onde

é atualmente muito utilizada. Como objetivo específico pretende-se ainda elaborar um tutorial que permita a outros usuários utilizarem a ferramenta e a metodologia aplicada neste trabalho em pesquisas semelhantes em outras bacias hidrográficas.

## **1.6 Estrutura da dissertação**

A dissertação foi estruturada em sete capítulos, com seus conteúdos resumidos na forma abaixo, acrescidos das referências bibliográficas.

O primeiro capítulo introduz o tema do trabalho procurando apresentar também a motivação, justificativa e objetivos do mesmo.

O segundo capítulo descreve os fundamentos teóricos sobre o processo erosivo destacando suas causas e conseqüências para o meio ambiente, bem como os indicadores relacionados a esse processo.

O terceiro capítulo aborda as ferramentas e técnicas de integração de dados espaciais procurando destacar os conceitos relativos aos métodos utilizados no estudo de caso deste trabalho. Este capítulo detalha o método de inferência espacial baseada nas técnicas de Lógica Nebulosa (Lógica “*Fuzzy*”) e o método de Pesos de Evidência (*Weight of Evidence - Woef*), abordando as formulações matemáticas utilizadas para explicar a geração dos mapas de favorabilidade a erosão. Ainda neste capítulo, procura-se destacar as tendências de novos métodos utilizados na integração de dados espaciais.

O quarto capítulo apresenta a metodologia utilizada para a elaboração do trabalho, apresentando o fluxograma metodológico para a avaliação de processos erosivos em uma área de estudo.

O quinto capítulo descreve a aplicação das técnicas escolhidas para a análise a ser efetuada, utilizando a ferramenta “*Spatial Data Modeller*” tomando como estudo de caso o processo erosivo na Bacia do Alto Taquari, MS. Este capítulo busca detalhar cada passo para a realização da análise espacial procurando descrever um pequeno tutorial para a utilização destas técnicas.

O sexto capítulo apresenta as discussões dos resultados do trabalho, tecendo comentários a respeito da análise, bem como estabelecendo uma comparação entre as técnicas apresentadas e uma comparação com outros trabalhos.

O sétimo capítulo procura destacar as conclusões e os pontos mais relevantes na análise, apresentando ainda as recomendações para trabalhos futuros nessa linha de pesquisa.

## **CAPÍTULO 2 – PROCESSO EROSIVO**

Neste capítulo é apresentada a fundamentação teórica sobre o processo erosivo. Assim sendo, serão vistos conceitos relacionados à degradação das terras, fatores que contribuem para o processo erosivo e as conseqüências ambientais causadas pela erosão.

### **2.1 Erosão**

A erosão pode ser entendida como um processo de degradação do solo e seu transporte, feito geralmente pela água da chuva e pelo vento. A erosão causada pela água, também chamada de erosão hídrica, é o tipo de erosão mais importante e preocupante no Brasil, pois desagrega e transporta o material erodido com grande facilidade, principalmente em regiões de clima úmido onde seus resultados são mais drásticos [5].

Na erosão hídrica as gotas de chuva ao impactarem um solo desprovido de vegetação desagregam partículas que, conforme seu tamanho são facilmente carregadas pela enxurrada. Esse tipo de erosão resulta da interação de forças ativas, como as características da chuva, a declividade do terreno e a capacidade do solo em absorver água, e de forças passivas, como a resistência que o solo exerce à ação erosiva da água, os métodos de cultivo e a densidade da cobertura vegetal. A resistência do solo determina a sua erodibilidade, que é a tendência inerente do solo de erodir-se em diferentes proporções, devido unicamente às diferenças peculiares de cada classe de solo [68].

A erosão em geral compreende os processos de desagregação, transporte e deposição do solo. A fase de desagregação é o processo de quebra e individualização das partículas da massa de solo podendo ser determinada em função da força com que as gotas de chuva caem no solo, da intensidade das chuvas e do estado de agregação do solo. O transporte das partículas de solo ocorre em função da topografia, do tamanho das partículas do solo e do volume de água das chuvas. A fase de deposição é determinada pelo volume das enxurradas, pela densidade das partículas do solo e pela topografia [46].

Segundo Álvares e Pimenta [5] o processo de erosão resulta de uma combinação de fatores que são dependentes e estão interligados entre si, e apresentam grande variabilidade espacial e temporal, tornando este fenômeno difícil de equacionar e quantificar. Os fatores que influenciam os processos erosivos são a erosividade da precipitação, medida pela sua intensidade e energia cinética, a erodibilidade dos solos, definida pelas suas características

físicas e químicas, a cobertura vegetal, pela sua maior ou menor proteção do solo, os declives e comprimentos das encostas e as práticas de conservação existentes.

Segundo Antônio Guerra [45], para se entender como se dá a dinâmica do processo erosivo deve-se, ainda, levar em conta fatores climáticos, geomorfológicos, pedológicos, hidrológicos e também os relativos ao uso da terra e cobertura vegetal, que são de grande importância na detonação desses processos.

A modelagem quantitativa da erosão hídrica tem sido habitualmente avaliada a partir de métodos empíricos. Um dos modelos utilizados é a Equação Universal de Perda de solo - USLE (Wischmeier e Smith [94], citado por [36]) que agrupa as variáveis que influenciam a erosão em seis fatores genéricos, que se relacionam de acordo com a seguinte fórmula:

$$A=R \times K \times L \times S \times C \times P$$

onde A indica a perda de solo, R a erosividade das chuvas, K a erodibilidade do solo, LS o fator topográfico, C o fator de uso/manejo do solo e P o fator de práticas conservacionistas de solo.

## **2.2 Tipos de erosão**

Existem três principais tipos de erosão: laminar, sulcos e voçorocas. Na erosão laminar, o solo desagregado pelas gotas de água é removido do local, de forma mais ou menos uniforme. A erosão laminar se caracteriza por retirar finas camadas ou lâminas de solo, daí o nome laminar. O aparecimento de raízes e pedras na superfície do solo cultivados com culturas permanentes, como são os casos do cacau, do café, da pastagem, e de certas fruteiras podem ser sinais de ocorrência dessa erosão [11].

Quando a água se concentra em pequenos canais, tem-se o processo erosivo em sulcos. A erosão em sulco é o tipo de erosão que mais chama atenção do homem. Porém, provoca menos desgaste do que a erosão laminar. Locais com grande declividade e com solo descoberto é favorável à ocorrência desse tipo de erosão [11]. Vários fatores influem para o surgimento da erosão em sulcos, como por exemplo, a aração quando acompanha o declive ou o surgimento de trilhas feitas pelo gado em movimento, resultando em desgaste, empobrecimento do solo e posterior dificuldade para manejo com sulcos já formados.

A erosão em voçoroca provoca um grande impacto visual. A sua origem é a erosão em sulco não controlada. É muito cara e difícil de ser controlada. Em áreas de ocorrência de voçoroca fica impossibilitado o uso de máquinas e equipamentos agrícolas [11].

Antônio Guerra [45] descreve as voçorocas como características erosivas que possuem paredes laterais íngrimes, em geral fundo chato, ocorrendo fluxo de água no seu interior durante os eventos chuvosos. As voçorocas são processos erosivos de maior intensidade provocados pelo agravamento do processo de erosão onde esse desgaste do terreno indica a perda total ou parcial da sua fertilidade.

Segundo Pereira [72] a voçoroca é formada pela combinação do escoamento superficial e o escoamento subsuperficial, em ambientes desequilibrados pelo desmatamento, fogo e endurecimento do solo.

Quase sempre a causa do surgimento de voçorocas está associada à modificação do ambiente pela intervenção humana, levando a uma acentuação do processo erosivo devido à retirada da cobertura vegetal, num solo susceptível à erosão, sendo mal utilizado com atividade agrícola desordenada, queimadas, pisoteio do gado em fortes declividades, excesso de animais na pastagem, que aliado à ação das chuvas culmina em situação mais extrema da degradação natural [72].

### **2.3 Causas do processo erosivo**

Apesar do processo erosivo estar associado à interação de forças ativas e passivas como dito anteriormente, muitas ações do homem alteram o ambiente e aceleram esse processo. Ações como a retirada da cobertura vegetal com a destruição de florestas, o uso inadequado do solo, utilização de técnicas agrícolas inadequadas, desmatamento, queimadas ou terraplanagem, deixam o solo desprotegido fazendo com que a erosão seja mais severa [74]. A seguir é feito um breve comentário, segundo Portugal [74], sobre algumas das ações humanas e algumas características ambientais que podem causar o processo erosivo:

\* Retirada da cobertura vegetal - a cobertura vegetal impede as águas das chuvas agirem diretamente no solo, destruindo a camada superficial e carregando os nutrientes. A cobertura vegetal é fator muito importante no controle do desgaste do solo. A água da chuva inicia a sua ação erosiva pela atuação direta da gota de chuva sobre o terreno, soltando as partículas de terra. Quando se retira a cobertura vegetal com que a natureza protege o solo,

para em seu lugar colocar uma cultura ou pastagem, o terreno fica exposto à ação direta da água da chuva.

\* Uso do solo – muitas vezes a erosão pode estar associada ao mau uso do solo, como o excesso de pastoreio, o uso de maquinários pesados e até mesmo com o plantio feito de maneira incorreta, que torna o solo empobrecido. O tipo de exploração que recobre o terreno influencia o processo erosivo. Modificações de grandes áreas de cobertura natural, como florestas e cerrados, que cedem lugar à pastagens e agricultura provocam a acentuação da erosão, levando à aceleração de perda de terras produtivas [33].

\* Características do solo – características como textura, estrutura, profundidade do solo, permeabilidade, entre outros, também influenciam o processo de erosão.

\* Pecuária - a intensiva criação de gado pode produzir erosão. O gado, ao mesmo tempo que se alimenta das gramíneas, pisa sobre ele e pode diminuir a cobertura vegetal do solo iniciando um processo de erosão superficial.

\*Relevo – o relevo também constitui mais um fator que pode favorecer a erosão quando a topografia é acentuada ou sem cobertura vegetal. Outros fatores ligados ao relevo que podem influenciar no processo de erosão são comprimento da encosta e o grau de declive (declividade).

\* Outros - chuvas intensas, queimadas, infiltração de água no solo, infertilidade do solo, a ação dos ventos que provocam o desgaste do solo, etc.

## **2.4 Indicadores do processo erosivo em bacias hidrográficas**

No Brasil e demais países tropicais, a intensidade, o volume e a concentração das chuvas acarretam um intenso processo de erosão dos solos, que se torna mais forte quando se trata de áreas com topografia inclinada ou sem cobertura vegetal. O processo erosivo também tem sido presenciado ao longo de bacias hidrográficas onde tem ocorrido muita interferência humana [27].

Segundo Jenkins et. al. [49], em termos ambientais, a bacia hidrográfica é uma unidade ecossistêmica e morfológica que melhor reflete os impactos das interferências

antrópicas. Constitui um sistema aberto que recebe energia fornecida pela atuação do clima e dos agentes geológico-geomorfológicos locais e elimina fluxos energéticos pela saída de água, sedimentos e materiais solúveis.

Em função das mudanças de entrada e saída de energia ocorrem ajustes internos nos elementos das formas e nos processos associados. Mudanças significativas no comportamento das condições naturais de uma bacia, causadas por processos naturais ou atividades antrópicas, podem gerar alterações e impactos nos seus fluxos energéticos, desencadeando desequilíbrios ambientais. Dentre os processos causadores dessa degradação destaca-se a erosão [49].

Para avaliar o efeito das perturbações numa bacia hidrográfica é necessário antes conhecer as características do ecossistema natural e as características de intervenções antrópicas. A análise desses aspectos do ecossistema envolve características ambientais como solo, relevo, clima, e intervenções antrópicas como desmatamentos e pecuária, que podem ser utilizadas para analisar os processos erosivos em uma bacia.

Algumas dessas características podem ser definidas como indicadores, ou seja, como parâmetros de uma análise para avaliar o processo erosivo. Os indicadores são variáveis que são escolhidas a fim de se definir um ambiente ou condições mais favoráveis para a ocorrência de um determinado evento como, por exemplo, o processo de erosão. Os indicadores representam um determinado evento auxiliando na tarefa de descrever um ambiente [58].

Os indicadores são componentes essenciais em processos de tomada de decisão envolvendo múltiplos critérios, como é o caso da análise do processo erosivo em bacias hidrográficas, pois permitem agregar informações consideradas relevantes pelos decisores e pela comunidade [24].

Moldan e Bilharz [65] afirmam que os indicadores desejáveis são aqueles que resumem ou simplificam informações, quantificando, medindo ou comunicando dados, de forma a tornar um determinado fenômeno compreensível por todos. Conforme Tunstall [91], os indicadores podem desempenhar várias funções, sendo as principais:

- estudar condições e tendências;
- comparar lugares, situações ou alternativas;
- proporcionar antecedência ao advertir sobre algum efeito ou impacto de uma ação;
- prever futuras condições e tendências.

## **2.5 Impactos ambientais associados ao processo erosivo**

A erosão dos solos causa uma gama de impactos ambientais, desde sua própria degradação até problemas ambientais de forma geral. Os problemas iniciam no local onde ocorrem, estendendo-se para grandes áreas [64].

Os impactos ambientais causados pela erosão se agravam em ecossistemas mais sensíveis à atividade antrópica. Os impactos ambientais decorrentes desta ocupação desordenada das terras pelo homem têm causado a degradação dos recursos naturais como o solo, a água, a vegetação e a diminuição da fauna, devido à destruição da vegetação nativa e o uso indiscriminado de agroquímicos e pesticida [10].

Em algumas regiões, como o pantanal, observa-se uma aceleração do processo erosivo, causada por profundas modificações do ambiente devido à ocupação humana, com a substituição da vegetação original por pastagens e por cultivos agrícolas. Esse fenômeno foi acentuado com o processo de expansão agrícola causando uma série de impactos ambientais [10].

Alterações do ambiente como estas podem contribuir para o agravamento do processo erosivo causando impactos ambientais ainda piores que podem comprometer a qualidade e quantidade de água, alterar regimes hidrológicos, gerar processos de desertificação, entre outros.

A seguir é feito um breve comentário sobre os principais impactos causados pelo processo erosivo destacando-se o comprometimento da qualidade da água de rios e lagoas, o assoreamento, o empobrecimento do solo e a desertificação.

### **Qualidade da água**

O processo erosivo é responsável por importantes desajustes ambientais. Um deles é o transporte de material para dentro dos rios que, muitas vezes, causa a poluição dos mananciais, comprometendo a qualidade da água para consumo humano e animal, podendo provocar até a sua morte [11].

Pode-se considerar dois aspectos negativos no transporte de material para dentro do rio: o primeiro é o preenchimento do volume original dos rios e lagos e como conseqüência, vindas as grandes chuvas, os corpos d'água extravasam, causando inundações; o segundo é que a instabilidade causada nas partes mais elevadas podem levar a deslocamentos repentinos de grandes massas de solo e rochas que desabam talude abaixo, causando, no geral, grandes tragédias [74].

Considerando os efeitos poluidores, pode-se citar que os arrastamentos podem encobrir porções de terrenos férteis e sepultá-los com materiais áridos; podem causar a morte da fauna e flora do fundo dos rios e lagos por soterramento; podem causar turbidez nas águas, dificultando a ação da luz solar na realização da fotossíntese, importante para a purificação e oxigenação das águas; podem arrastar biocidas e adubos até os corpos d'água e causarem, com isso, desequilíbrio na fauna e flora nesses corpos d'água [74].

### **Assoreamento**

As águas da chuva quando arrastam o solo, quer ele seja rico de nutrientes e materiais orgânicos, quer ele seja árido, provocam o enchimento dos leitos dos rios e lagos com esses materiais. Esse fenômeno de enchimento se chama assoreamento. Da mesma forma que a erosão, o assoreamento também é um fenômeno geológico natural e contínuo, estando associado à atuação dos processos erosivos.

Os processos de assoreamento estão intimamente associados aos processos erosivos que atuam nas encostas e nos vales, constituindo um dos principais impactos decorrentes da erosão acentuada em uma bacia hidrográfica. Correspondem a fenômenos de deposição acentuada de sedimentos, ocorrendo quando a energia do agente transportador (geralmente a água) é suplantada pela energia do agente gravitacional, ou quando há condições favoráveis à deposição de partículas sólidas [47].

Os efeitos do assoreamento em uma bacia de drenagem podem gerar uma série de problemas, tais como: perda de volume total de água dos reservatórios; diminuição das zonas de recarga dos aquíferos; problemas operacionais nas usinas; aumento das taxas de erosão e dos locais atingidos por enchentes nas áreas marginais; comprometimento da navegabilidade dos canais; perda de qualidade das águas por turbidez e por veiculação de poluentes pelos sedimentos em suspensão; comprometimento da piscicultura e dos habitats aquáticos, etc [47].

O assoreamento de lagos, lagoas e dos principais canais fluviais prejudica seriamente a navegabilidade de alguns rios e, inclusive, ameaçando atividades agropastoris com o alagamento de áreas produtivas [33]. Além das nascentes dos rios sofrerem constante processo de assoreamento, em volta delas geralmente ocorre desmatamento para implementação de monoculturas, principalmente de soja, fazendo com que a cada ano, os rios vão ficando cada vez mais rasos.

## **Empobrecimento do solo**

A erosão como um processo de desgaste acelerado do solo provoca o empobrecimento do solo e a perda da capacidade produtiva com o passar dos tempos. A superfície do solo, não castigado, é naturalmente coberta por uma camada de terra rica em nutrientes inorgânicos e materiais orgânicos que permitem o crescimento das vegetações; se essa camada é retirada, esses materiais desaparecem e o solo perde a propriedade de fazer crescer vegetações e pode-se dizer que, no caso, o terreno ficou árido ou que houve uma desertificação.

Segundo Haridasan [46] quando a cobertura vegetal e parte do solo são retirados, as gotas da chuva entram em contato direto com o solo, provocando compactação da superfície, desagregação e espalhamento das partículas, tornando-as, assim, disponíveis para o escoamento superficial e posterior deposição. A compactação, além de diminuir o oxigênio e a água disponível para as raízes das vegetações, impede sua penetração no solo, constituindo, portanto, o principal fator de impedimento da revegetação de uma área degradada.

Além disso, a compactação do solo, ao impedir a infiltração da água, aumenta o escoamento superficial que arrasta as partículas de solo, depositando grande quantidade de solo nas baixadas. Quanto maior for a intensidade de chuva, maior será o impacto sobre o solo desnudo, ou seja, maior será a compactação e a perda de partículas das camadas superficiais. [46].

## **Desertificação**

A Convenção de Combate à Desertificação – CCD ([www.unccd.int](http://www.unccd.int)) definiu a desertificação como a degradação do solo em áreas áridas, semi-áridas e sub-úmidas secas resultantes de variações climáticas e atividades humanas”. Este processo consiste na perda da produtividade biológica e econômica das terras agrícolas, das pastagens e das áreas de matas nativas devido às variabilidades climáticas e às atividades humanas [50].

O processo de desertificação pode ser entendido como a combinação de vários fenômenos: diminuição das chuvas, elevação das temperaturas, empobrecimento dos solos entre outros, causando acentuada diminuição da biodiversidade.

A verificação do processo de desertificação gera impactos negativos no setor agrícola com o comprometimento da produção de alimentos. Além do enorme prejuízo causado pela quebra de safras e diminuição da produção, existe o custo quase incalculável de recuperação da capacidade produtiva de extensas áreas agrícolas e da extinção de espécies nativas, algumas com alto valor econômico e outras que podem vir a ser aproveitadas na agropecuária, inclusive no melhoramento genético, ou nas indústrias farmacêutica, química e outras. [50].

## **2.6 Considerações finais**

Neste capítulo foram apresentados conceitos sobre o processo erosivo, suas causas e sua influência sobre o meio ambiente a fim de se entender como esse processo pode interferir em ecossistemas como bacias hidrográficas.

No próximo capítulo serão apresentados os conceitos sobre integração temática de dados e algumas técnicas de análise espacial *dirigida pelos dados* e *dirigida pelo conhecimento* visando conhecer como podem ser correlacionados alguns indicadores ambientais vistos neste capítulo visando a análise do processo erosivo.

## **CAPÍTULO 3 – INTEGRAÇÃO TEMÁTICA DE DADOS**

Serão apresentados neste capítulo conceitos sobre as técnicas de integração temática de dados, que constituem uma importante ferramenta de auxílio para visualização, tratamento e análise dos dados.

No processo de integração temática os dados são organizados em forma de planos de informação (*layers*), ou seja, como uma série de camadas, cada uma das quais contendo feições gráficas espacialmente relacionadas. Cada camada, que representa um tema ou uma classe de informação, é um conjunto de feições homogêneas que estão posicionalmente relacionadas às outras camadas [73].

A organização dos planos de informação é definida segundo os temas de interesse no estudo. Para o estudo de impactos ambientais é realizada a análise de dados do meio físico-biótico, onde os planos de informação serão representados por dados como drenagens, clima, vegetação, solo, conflito de uso, etc.

A viabilidade de sobrepor informações integrando os planos de informação, bem como a visualização das mesmas em um ambiente de fácil manipulação, é extremamente útil para gerar hipóteses, avaliação de possíveis associações entre os eventos estudados e possibilidades de análises.

Os métodos de análise de distribuições espaciais são especialmente úteis na situação em que um evento em estudo é gerado por fatores ambientais de difícil detecção e onde os fatores relacionados têm distribuição espacialmente condicionada como, por exemplo, na modelagem estatística [73]. A área ambiental é muito influenciada pelos padrões de ocupação do espaço. Assim, os planos de informação devem descrever as características do local ou evento de maneira bem detalhada, além de ser necessário localizar o mais precisamente possível onde estão acontecendo os eventos estudados, locais de potencial risco ambiental e as áreas onde se concentram situações vulneráveis.

### **3.1 Ferramentas de integração de dados espaciais**

Ferramentas de integração de informações espaciais, como SIGs - Sistemas de Informação Geográfica e SDSSs – Sistemas de Suporte à Decisão Espaciais, permitem que a informação seja integrada e disposta em camadas, através da combinação de dados espaciais, tipicamente encontrados em mapas, com dados não espaciais, como informações demográficas e de saúde [17].

Essas ferramentas ajudam na caracterização das áreas, organizando e integrando os dados e apresentando-os de uma forma visual e mais compreensível. Ferramentas de integração de informações espaciais também evidenciam os relacionamentos entre diferentes tipos de dados e podem ser utilizadas para responder a uma série de questões sobre impactos e cenários alternativos. Dependendo da quantidade e do tipo de informação necessária, essas ferramentas podem ser utilizadas em conjunto. Por exemplo, um SIG pode ser acoplado a um modelo de dispersão do ar para criar um modelo espacial e/ou modelos de decisão para criar um SDSS, que apresentará com maior clareza a situação e as alternativas possíveis [17].

### **3.1.1 Sistema de Suporte à Decisão**

O processo de tomada de decisão normalmente envolve dados, múltiplos critérios, seleção de alternativas e vários agentes de decisão. Esse quadro piora quando a questão é de natureza multidisciplinar, pois certamente haverá envolvimento de valores sócio-econômicos, ambientais e políticos; envolvimento de um maior número de agentes tomadores de decisão e um número maior de grupos de interesses com objetivos conflitantes [13].

Os Sistemas de Suporte à Decisão combinam a tecnologia de Sistema de Informação Geográfica com modelos matemáticos, de modo que os resultados de várias ações possam ser simulados e comparados com situações existentes, permitindo avaliar uma série de questões hipotéticas no gerenciamento de decisões [64].

Pode-se definir que o Sistema de Suporte à Decisão é um esforço para resolver o dilema dos objetivos conflitantes, cuja presença impede a existência da “solução ótima” e conduz para a procura da “solução de melhor acordo” [82], onde quanto maior for o entendimento do sistema pelo tomador de decisão melhor será a solução.

Os Sistemas de Suporte à Decisão são sistemas interativos, baseados em computadores, que tem como objetivo principal ajudar os decisores a utilizar dados e modelos para identificar e resolver problemas, assim como tomar decisões [20]. Esses sistemas ajudam os decisores a utilizar e manipular dados, bem como construir e utilizar modelos matemáticos.

Segundo Turban [92], os sistemas de suporte a decisão têm quatro características principais: incorporam dados e modelos, são sistemas desenhados para ajudar os gestores nos seus processos de decisão, auxiliam avaliações de gestão e têm como objetivo melhorar a eficácia das decisões.

A eficácia máxima desses sistemas é atingida quando se conseguem incorporar conhecimentos do decisor e programas baseados em computador, no processo de tomada de

decisão. Para ser eficaz o sistema deve ser de fácil utilização, ajudando a explorar o problema de decisão de um modo interativo e simples. Segundo Cabral [19], esses sistemas nunca devem substituir o decisor (ou decisores), apenas devem fornecer elementos para apoiar e fundamentar a decisão.

Segundo Malczewski [56], os problemas de decisão espacial caracterizam-se normalmente por permitirem a inclusão de um grande número de alternativas de decisão. Cada alternativa é avaliada com base em múltiplos critérios que podem ser qualitativos ou quantitativos. Na resolução desses problemas existe frequentemente mais do que um decisor envolvido no processo de tomada de decisão.

Nos últimos anos tem-se assistido a um incremento na investigação e no desenvolvimento de sistemas de suporte a decisão. Os Sistemas de Suporte à Decisão passaram a analisar também a componente espacial, onde os conceitos de dados espaciais, informação e tomada de decisão encontram-se inter-relacionados.

Segundo Pina [73], os dados espaciais possuem uma posição no espaço, isto é, uma localização geográfica conhecida. Quando os dados espaciais são organizados, apresentados, analisados, interpretados e considerados úteis para um problema de decisão específico, transformam-se em informação. Essa informação é que poderá ser processada de modo a se tornar significativa e auxiliar no processo de tomada de decisão.

Há sistemas que são capazes de utilizar dados espaciais por incluírem a capacidade adicional de executar operações espaciais. Esses sistemas de informação tornam possível a realização de análises espaciais e são chamados de Sistemas de Informação Geográfica.

### **3.1.2 Sistemas de Informação Geográfica**

Um Sistema de Informação Geográfica (SIG) é um sistema baseado na integração de dados de entrada, gerenciamento, manipulação, análise e saída de dados e atributos para ajudar nas atividades de suporte à decisão [73]. Os Sistemas de Informação Geográfica são marcados pela crescente capacidade de análise, tratamento de dados e pela facilidade do acesso à informação, permitindo observar a distribuição espacial de situações de risco e problemas que envolvem as mais diversas áreas incluindo a área ambiental. A abordagem espacial permite a integração de dados promovendo a interação de informações diversas, fornecendo elementos para construir uma importante ferramenta de apoio que auxilie o entendimento, aumentando o poder de orientar o processo de tomada de decisão fazendo uso de mapas a fim de visualizar a distribuição geográfica de dados.

A atual geração de um Sistema de Informação Geográfica configura uma tecnologia estabelecida para armazenar, organizar, recuperar e modificar informações sobre a distribuição espacial de recursos naturais, dados geo-demográficos, redes de utilidade pública e muitos outros tipos de dados localizados na superfície da terra [19].

Segundo Cabral [19], um dos objetivos de um SIG é fornecer informação para auxiliar a tomada de decisão que envolve dados espaciais. A tecnologia de um SIG integra operações convencionais de bases de dados, como captura, armazenamento, manipulação, análise e apresentação de dados, com possibilidades de seleção e busca de informações e análise estatística, juntamente com a possibilidade de visualização e análise geográfica oferecida pelos mapas. Essa capacidade distingue os SIGs dos demais sistemas de informação e torna-os úteis para organização no processo de entendimento da ocorrência de eventos, predição e simulação de situações e planejamento de estratégias.

Os Sistemas de Informação Geográfica têm sido amplamente utilizados para a estruturação e organização de variáveis espaciais na geração de alternativas para vários tipos de problemas. Seu uso intenso justifica-se pelo fato de constituir uma poderosa ferramenta que integra o conjunto de rotinas de programação desenvolvidas para representar e manipular grandes quantidades de dados armazenados em bancos de dados, os quais contêm informações representativas do mundo real, através de coordenadas geográficas, possibilitando análises espaciais [89].

Os SIGs permitem a realização de análises espaciais complexas através da rápida formação e alteração de cenários que propiciam aos planejadores e administradores em geral, subsídios para a tomada de decisões. A opção por esta tecnologia busca melhorar a eficiência operacional e permitir uma boa administração das informações estratégicas, tanto para minimizar os custos operacionais quanto para agilizar o processo decisório.

Nos Sistemas de Informação Geográfica a figura do usuário é considerada muito importante nas análises, pois é o responsável por interpretar os resultados gerados, colocando sua experiência para analisar e gerar informações e decisões que auxiliam no processo de suporte à decisão.

Segundo Pina [73] de uma maneira geral pode-se identificar como objetivos no SIG a visualização das informações através da integração de mapas, gráficos e imagens, a organização dos dados permitindo combinar vários tipos diferentes destas informações, mas principalmente a integração de dados vindos de diversas fontes, formatos e escalas, permitindo sempre associar novas informações que contribuam no processo de análise dos

dados permitindo transformar os dados em informações úteis no processo de tomada de decisões.

A avaliação de processos de erosão a partir da integração temática de dados pode pressupor a utilização dos Sistemas de Informação Geográfica já que são sistemas computacionais usados para o entendimento dos fatos e fenômenos que ocorrem no espaço geográfico. A sua capacidade de reunir uma grande quantidade de dados convencionais de expressão espacial, estruturando-os e integrando-os adequadamente, torna os SIGs importantes ferramentas neste tipo de análise [73].

Depois de vários anos de desenvolvimento, o SIG tem sido aplicado em alguns importantes campos de conhecimento, como o monitoramento ambiental, gerenciamento de recursos naturais, aplicações em comércio e negócios e diferentes utilidades. Um dos propósitos do SIG é fazer evoluções ou predições com diferentes dados específicos de integração de modelos para combinar dados e atributos espaciais de várias origens para descrever e analisar interações e prover auxílio no suporte à decisão.

### **3.1.3 Análise Multicritério**

A Análise Multicritério pode ser vista como uma classe de Sistema de Suporte a Decisão com a particularidade de se basearem em vários critérios para a tomada de decisão espacial [20]. A Análise Multicritério caracteriza-se principalmente pela capacidade de analisar problemas incorporando critérios, tanto quantitativos como qualitativos, e que, muitas vezes são conflitantes entre si. Esse tipo de análise baseiam-se no princípio que para a tomada de decisão a experiência e o conhecimento são tão valiosos quanto os dados utilizados [81].

A decisão espacial baseada em múltiplos critérios requer uma articulação dos objetivos do decisor e a identificação de atributos úteis para indicar o grau com que estes objetivos são efetivamente atingidos. O objetivo e os atributos formam uma estrutura hierárquica de critérios de avaliação para um problema de decisão específico [57].

Os critérios deverão representar adequadamente a natureza multi-critério de um problema de decisão. Os critérios que compõem estes tipos de problemas devem ser completos, operacionais, divisíveis em componentes individuais, não redundantes e minimalistas. Uma vez estabelecida a estrutura hierárquica dos objetos e atributos, cada critério deve ser representado num nível de informação da base de dados geográfica [57].

Em alguns casos na Análise Multicritério podem ser considerados, ao mesmo tempo, aspectos econômicos, sociais, ambientais e políticos bem como quaisquer outros que se

mostrarem relevantes à comparação de um determinado conjunto de alternativas [24]. Além disso, há frequentemente a intervenção de múltiplos decisores, com conflitos de interesses, pontos de vista conflitantes e diferentes juízos de valores.

Os métodos multicritérios de análise de decisão aparecem como uma opção para que estas diversidades sejam integradas. O objetivo da Análise Multicritério é ajudar o decisor a analisar os dados que são intensamente complexos e buscar a melhor estratégia de gestão do seu problema [13].

Para Castro et. al. [24], apesar da Análise Multicritério ter a vantagem de poder integrar vários critérios podendo ser bem mais percebidas e comparadas às particularidades de cada alternativa do problema, há a desvantagem desse tipo de análise necessitar de um grande número de informações para a avaliação de cada alternativa, além do fato dos decisores e a escolha dos critérios avaliados terem bastante importância no processo, tornando a análise, às vezes, muito subjetiva.

A análise de multicritérios de apoio à decisão permite fornecer uma classificação dos problemas analisados, do melhor para o pior, através de uma análise dos indicadores, que são as características de cada problema. A Análise Multicritério é desenvolvida em etapas que, de modo geral, podem ser representadas da seguinte maneira, segundo Soares [85]:

a) Formulação do problema.

De um modo bastante simplista, corresponde a saber sobre o que se quer decidir.

b) Determinação de um conjunto de ações potenciais.

Os atores envolvidos na tomada de decisão devem constituir um conjunto de ações (alternativas) que atendam ao problema colocado.

c) Elaboração de uma família coerente de critérios.

Definição de um conjunto de critérios que permita avaliar os efeitos causados pelo problema a ser analisado. Segundo Soares [85], esta é uma tarefa longa, com sucessivas aproximações entre os objetivos desejados e a possibilidade de atendimento com os recursos, tempo e conhecimentos disponíveis. Para a construção de critérios, normalmente, são utilizados elementos denominados parâmetros e indicadores. Hierarquicamente, os parâmetros, que são dados mais diretos e simples, estariam na base da estrutura de construção. Em nível intermediário, se encontrariam os indicadores, representando conjuntos de dados de

natureza diferente agregados em uma característica mais sintética, seguidos em um nível superior pelos critérios.

d) Avaliação dos critérios.

Esta etapa é, geralmente, formalizada através de uma matriz de avaliações ou tabela de performances, na qual as linhas correspondem às ações a avaliar e as colunas representam os respectivos critérios de avaliação previamente estabelecidos. Os critérios podem ser diferentemente percebidos por diferentes analistas, devendo ser definidos por especialista(s) no problema em questão.

e) Determinação de pesos dos critérios e limites de discriminação.

Nesta etapa, estabelecem-se os pesos associados a cada critério. Os pesos traduzem numericamente a importância relativa de cada critério. A ponderação de critérios pode ser realizada através de várias técnicas como: hierarquização de critérios com a comparação de critérios par a par, notação, distribuição de pesos, taxa de substituição, regressão múltipla, jogos de cartas, etc.

f) Agregação dos critérios.

Consiste em associar, após o preenchimento da matriz de avaliação e segundo um modelo matemático definido, as avaliações dos diferentes critérios para cada ação. As ações serão em seguida comparadas entre si por um julgamento relativo do valor de cada ação.

Os métodos de Análise Multicritério são muitos e a escolha do melhor deles para a resolução de um determinado problema não é trivial. A escolha de um deles irá depender do problema a ser analisado, da familiaridade do analista com determinado método e da existência dos recursos necessários a sua aplicação [24].

Os principais métodos de Análise Multicritério para o auxílio a tomada de decisão são o MACBETH - Measuring Attractiveness by a Categorical Based Evaluation Technique, o PROMETEE - Preference Ranking Organization Method for Enrichment Evaluations, o ELECTRE - Elimination et Choix Traduisant la Réalité e o AHP - Analytic Hierarchy Process [13], que será descrito com mais detalhes a seguir.

### 3.1.4 Processo de Análise Hierárquica (AHP)

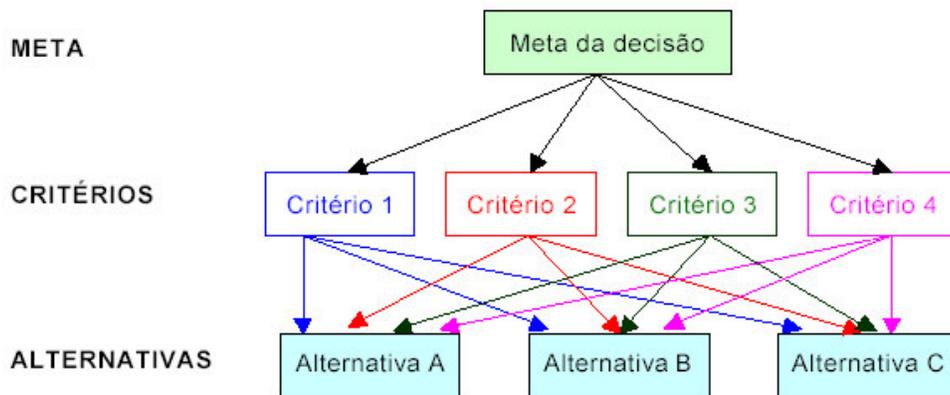
Saaty [78] desenvolveu o Processo de Análise Hierárquica (AHP - Analytic Hierarchy process). Sua origem data de 1971, quando o Dr. Thomas Saaty trabalhava no Departamento de Defesa dos Estados Unidos; foi desenvolvida em 1972, num estudo sobre o racionamento de energia para indústrias; e chegou em sua maturidade aplicativa com o Estudo dos Transportes do Sudão, em 1973, havendo um grande enriquecimento teórico entre 1974 e 1978.

O AHP é uma técnica de tomada de decisão matemática que sempre considera os aspectos qualitativos e quantitativos de uma decisão. Isto reduz decisões complexas a uma série de comparações par a par, sintetizando os resultados. Comparado com outras técnicas dirigidas pelo conhecimento, o AHP usa a habilidade humana para comparar alternativas, contribuindo para escolher a melhor alternativa para a tomada de decisão [82].

Schmidt [81] define que o AHP caracteriza-se pela capacidade de analisar o problema de tomada de decisão a sintetizar as informações através da construção de níveis hierárquicos e a comparação par-a-par dos níveis da hierarquia. Em outras palavras o AHP reflete o que parece ser um método natural de funcionamento da mente humana. Ao defrontar-se com um grande número de elementos, controláveis ou não, que abrangem uma situação complexa, ela os agrega em grupos, segundo propriedades comuns (estruturação de um problema em uma hierarquia).

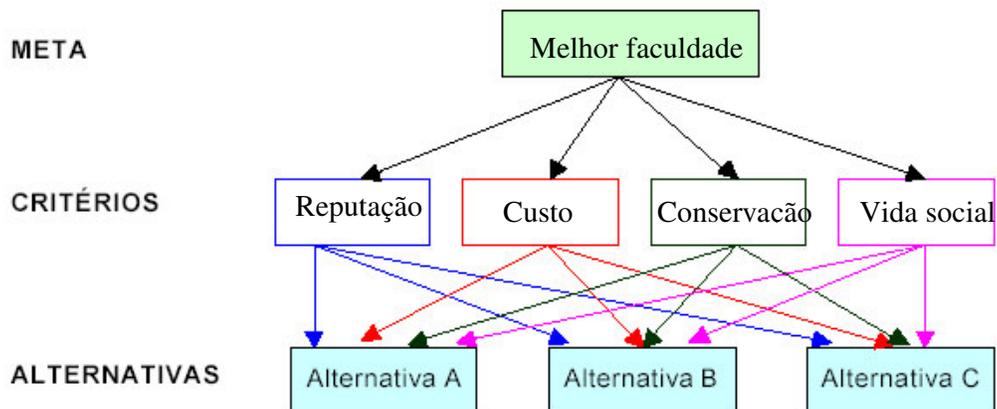
A hierarquia pode ser entendida como um sistema de níveis estratificados, cada um constituído de elementos, ou fatores. A questão central, em termos dessa hierarquia, é a seguinte: com que peso os fatores individuais do nível mais baixo da hierarquia influenciam em seu fator máximo, o objetivo geral? Desde que esta influência não seja uniforme em relação aos fatores, chega-se à identificação de sua intensidade, ou, às suas prioridades. Esta determinação das prioridades dos fatores mais baixos com relação ao objetivo pode reduzir-se a uma seqüência de problemas de prioridade, um para cada nível, e cada um desses problemas de prioridade a uma seqüência de comparações por pares. Essas comparações continuam a ser o ingrediente central desta teoria, mesmo que o problema original se complique por relações de *feedback* entre vários níveis ou fatores [81].

Assim, o primeiro passo para a solução de um problema com o auxílio do AHP é a construção da hierarquia, baseada na identificação de níveis e elementos importantes para a avaliação. Gartner [41] propõe o seguinte fluxograma para exemplificar a estrutura hierárquica básica do AHP, arranjando os elementos de um problema dentro de uma estrutura hierárquica que inclui metas, critérios de avaliação e alternativas para auxiliar o processo de tomada de decisão.



**Figura 3.1** Estrutura hierárquica básica do AHP  
Fonte: adaptado de [41]

Schmoldt e Peterson [82] propuseram como exemplo de aplicação do AHP considerar que a hierarquia na figura 3.2 é projetada para selecionar a "melhor faculdade" para estudar. A meta, "melhor faculdade", aparece ao topo da hierarquia. Os critérios aparecem no próximo nível: reputação acadêmica, custo, conservação do campus e vida social. As faculdades a serem consideradas estão rotuladas pelas letras A, B, e C e estão no nível mais baixo. Os critérios são comparados par a par com respeito à importância deles em produzir uma boa faculdade.



**Figura 3.2** Exemplo de aplicação do AHP  
Fonte: adaptado de [82]

Em um segundo momento ocorre a fase de avaliação, sendo realizado pares de comparação que são executados para determinar a prioridade dos critérios com respeito ao objetivo ou meta, ou seja, expressar a importância relativa dos critérios analisados. Para medir o grau de importância de um critério de um determinado nível, sobre aqueles de um nível inferior é utilizada uma escala de valores. A escala de valores utilizada para a comparação varia de 1-9 [79], como mostrada na tabela abaixo.

Valor	Definição	Explicação
1	<b>Igual importância</b>	os dois critérios contribuem de forma idêntica para o objetivo
3	<b>Pouco mais importante</b>	a análise e a experiência mostram que um critério é um pouco mais importante que o outro
5	<b>Muito mais importante</b>	a análise e a experiência mostram que um critério é claramente mais importante que o outro
7	<b>Bastante Importante</b>	a análise e a experiência mostram que um dos critérios é predominante para o objetivo
9	<b>Extremamente mais importante</b>	sem qualquer dúvida um dos critérios é absolutamente predominante para o objetivo
2,4,6,8 Valores recíprocos dos anteriores	<b>Valores intermediários</b>	também podem ser utilizados

1/9	1/7	1/5	1/3	1	3	5	7	9
extrema-mente	bastante	muito	pouco	igual	pouco	muito	bastante	extrema-mente
MENOS IMPORTANTE				MAIS IMPORTANTE				

**Tabela 3.1** Escala de valores AHP  
Fonte: adaptado de [79]

Assim é produzida uma matriz de comparação par a par obtida para os critérios analisados. Segundo Zambon [98] a matriz de comparação é uma matriz quadrada  $n \times n$ , onde as linhas e colunas correspondem aos  $n$  critérios analisados para o problema em questão. Assim, o valor  $a_{ij}$  representa a importância relativa do critério da linha  $i$  face ao critério da coluna  $j$ . Como esta matriz é recíproca, apenas a metade triangular inferior necessita ser avaliada, já que a outra metade deriva desta e a diagonal principal assume valores iguais a 1. Segundo Câmara [22] a matriz de comparação par a par é a elaboração de uma relação de importância relativa entre as evidências.

A matriz de comparação par a par obtida para os critérios do exemplo mencionado anteriormente, que analisa a melhor faculdade, pode ser apresentada na tabela abaixo:

	Reputação	Custo	Conservação	Vida social
Reputação	1	3	5	7
Custo	1/3	1	5	9
Conservação	1/5	1/5	1	3
Vida social	1/3	1/9	1/3	1

**Tabela 3.2** Matriz de comparação par a par  
Fonte: adaptado de [79]

Alguns autores consideram que as avaliações utilizando o AHP trazem resultados satisfatórios para algumas situações envolvendo a gestão de recursos naturais como auxílio a tomada de decisão. Nesse sentido Schmoltdt e Peterson [82] apresentaram trabalho destinado a utilização do AHP para tomada de decisão no gerenciamento de ecossistemas, mostrando que o AHP pode ser pertinente a qualquer aplicação de gerenciamento de ecossistema que requer múltiplas variáveis, múltiplos assuntos ou um complexo processo de tomada de decisão, já que administração e gerenciamento de um ecossistema pode incluir assuntos relacionados à qualidade e quantidade das águas, gerenciamento de floresta, administração de vida selvagem, entre outros, sendo importante estabelecer uma ordem de prioridades e fazer decisões com base na distribuição dessas variáveis e recursos. Além deste, outros trabalhos já foram desenvolvidos na área ambiental, como o trabalho desenvolvido por Manfro [57], que utilizou o AHP como instrumento de apoio à decisão no gerenciamento costeiro.

O AHP é um método útil para analistas e tomadores de decisão, na resolução de problemas complexos. Ele é útil também, quando diversos interesses, ambientais, sociais, políticos, entre outros, estão envolvidos e o número de pessoas que participam do processo é grande, já que é um método simples e fácil de ser explicado aos decisores, permitindo a interação entre o analista e o decisor [81].

### 3.2 Técnicas de integração temática

Nesta seção são apresentadas algumas técnicas de integração temática de dados. De acordo com Bonham-Carter [14], os modelos de integração de dados nos Sistemas de Informação Geográfica são divididos em duas categorias, os dirigidos pelos dados e os dirigidos pelo conhecimento.

Os modelos de integração de dados dirigidos pelo conhecimento são baseados em diferentes modelos para a estimação de pesos para os diferentes planos de informação (mapas de evidências) utilizadas em uma análise. A abordagem baseada em conhecimento geralmente emprega: auxílio de especialistas, definição de um modelo espacial, lógica ou simultaneidade *Booleana* e lógica ou possibilidade “*Fuzzy*”. Nesse tipo de modelo os pesos são baseados em opiniões de especialistas (subjetivo).

No modelo dirigido pelos dados, são calculados pesos usando métodos estatísticos e evidências encontradas em uma área de estudo para estimar a relação espacial entre planos de informação (mapas de evidências) e o mapa de resposta final gerada na análise. Para isso a abordagem dirigida pelos dados geralmente emprega as associações de medidas entre planos de informação e evidências já conhecidas (ocorrências).

Os modelos de análise espacial dirigido pelos dados geralmente visam combinar dados espaciais para descrever e analisar interações, de modo a fazer previsões por meio de modelos prospectivos empíricos, que forneçam apoio para a definição de áreas de maior potencial para a ocorrência de um determinado evento ou fenômeno. A combinação de dados a partir desses modelos permite reduzir a ambigüidade de interpretação que normalmente ocorre na análise individual dos mesmos [87].

Segundo Filho e Crosta [87], no modelo dirigido pelos dados os mapas são combinados por diferentes técnicas, tais como, regressão logística, ponderação por evidências (teorema bayesiano), razões de probabilidade e redes neurais (inteligência artificial) e os modelos dirigidos pelo conhecimento incluem o uso da lógica booleana (ou lógica crisp), média ponderada, lógica difusa (nebulosa) ou possibilidade “*Fuzzy*”, e teoria da crença de Dempster-Shafer.

A seguir é feito um breve comentário a respeito de algumas dessas técnicas dando destaque às técnicas incluídas no sistema computacional utilizada no estudo de caso: Lógica “*Fuzzy*”, Pesos de Evidência, regressão logística e Redes Neurais.

### 3.2.1 Lógica “Fuzzy”

A Lógica “Fuzzy” faz parte de um novo paradigma conhecido por sistemas inteligentes. Estes sistemas procuram fornecer respostas que solucionam problemas, de forma apropriada às situações específicas destes problemas, mesmo sendo novas ou inesperadas [83].

A Lógica “Fuzzy”, também chamada lógica nebulosa, é uma técnica que abrange a generalização da teoria dos conjuntos clássica e visa uma aproximação do pensamento humano. Essa teoria foi introduzida, em 1965, pelo matemático Lotfi A. Zadeh, com a principal intenção de dar um tratamento matemático a certos termos lingüísticos subjetivos, como “aproximadamente”, “em torno de”, dentre outros. Esse seria um primeiro passo para se representar e armazenar informações incertas, tornando possível o cálculo com essas informações, a exemplo do que faz o ser humano [48].

A Lógica “Fuzzy” representa uma forma inovadora de traduzir informações vagas, imprecisas, em valores numéricos. Possibilita a inclusão da experiência humana em controle computadorizado, tornando possível decisões em problemas complexos. Na maioria das vezes, o pensamento humano não pode ser classificado e reduzido a dois estados binários, verdade ou falso, branco ou preto, alto ou baixo, ou seja, reduzir a somente dois valores (característica da teoria clássica). Contrapondo a isto, a Lógica “Fuzzy” é multivalorada, isto é, não atribui a uma afirmação o valor verdadeiro ou o falso, e sim a um grau de pertinência dentro de um intervalo numérico [83].

Na abordagem “Fuzzy” as imprecisões caracterizam as classes que podem ter ou não fronteiras bem definidas. A teoria “Fuzzy” pode ser considerada como uma metodologia para generalizar qualquer teoria específica da forma *crisp* (discreto) para uma forma contínua (“Fuzzy”). Segundo Burrough e McDonnell [18], é recomendada a utilização dessas técnicas para tratar de fenômenos ambíguos, vagos ou ambivalentes, em modelos matemáticos ou conceituais.

A Lógica “Fuzzy” traz a possibilidade de modelar e manipular matematicamente informações vagas e imprecisas, naturais da linguagem humana e, portanto, as fornecidas por especialistas ao caracterizar alguns estudos. Esta manipulação é feita a partir da composição de variáveis escolhidas para modelar matematicamente o processo em questão, quando a implicação das variáveis é estabelecida por um conjunto de regras lingüísticas baseadas no conhecimento de especialistas [7].

### 3.2.1.1 Conjuntos “Fuzzy”

Na teoria clássica dos conjuntos (baseada na lógica *Booleana*), os conjuntos são ditos "*crisp*", de tal forma que um dado elemento do universo em discurso (domínio) pertence ou não pertence a um referido conjunto. Assim o elemento pertence ou não ao conjunto sendo atribuído à condição de verdadeiro ou falso, assumindo os valores 1 ou 0 respectivamente (determinístico) [96].

Diferentemente da teoria clássica de conjuntos, na teoria dos conjuntos nebulosos existe um grau de pertinência de cada elemento a um determinado conjunto. O conjunto “Fuzzy” estabelece uma caracterização de classes, que por várias razões não tem ou não pode definir limites rígidos (bordas) entre classes. Essas classes, definidas de maneira inexata, são chamadas de conjunto “Fuzzy” [22]. Assim o grau de pertinência do conjunto “Fuzzy”, geralmente representado por  $\mu(x)$ , é expresso em termos de uma escala que varia continuamente entre 0 e 1 que reflete o grau de verdade de uma proposta ou hipótese (probabilístico).

Resumidamente os conjuntos “Fuzzy” são uma extensão dos conjuntos convencionais, que permitem somente que elementos sejam verdadeiros ou falsos (lógica booleana, bivalente). Os conjuntos “Fuzzy” permitem que seus elementos possuam um certo “grau de pertinência” associado, sendo esta propriedade conhecida como “multivalência”. Isto permite a aproximação com o mundo real que não é bivalente, é na realidade multivalente, com um vasto número de opções ao invés de somente duas [12].

Segundo Meirelles [61], o conjunto de valores “Fuzzy”  $A$  em  $X$  é definido matematicamente como o par ordenado  $A = \{ (x, \mu_A(x)) \}$ ,  $x \in X$ , onde  $X = (x)$  é o universo de atributos,  $\mu_A(x)$  é conhecido como grau de pertinência de  $x$  em  $A$ . O valor  $\mu_A(x)$  é um número que pertence ao intervalo  $[0,1]$ , onde 1 representa a associação total do conjunto, e 0 a não associação. Os graus de associação de  $x$  em  $A$  refletem uma espécie de ordem que não é baseada na probabilidade, mas na possibilidade. Estes valores são obtidos por funções de pertinência.

Em outras palavras a pertinência na teoria “Fuzzy” é definida por uma tabela de pares ordenados, onde todo o valor de  $x$ , é associado com um valor do grau de pertinência  $\mu(x)$ . O par ordenado  $(x, \mu(x))$  é conhecido como conjunto “Fuzzy”.

### 3.2.1.2 Funções de pertinência

Para caracterizar formalmente um conjunto “*Fuzzy*”, pode-se recorrer ao conceito de função de pertinência. A função de pertinência é uma função que, dado o valor de um atributo  $z$ , ela determina se o elemento avaliado pertence ou não a um determinado conjunto em análise [22].

A função de pertinência associa cada elemento  $x$  de um conjunto em análise, (pertencente a um universo  $U$ ), a um número real  $\mu(x)$  no intervalo de 0 a 1, que representa o grau de possibilidade que cada elemento venha a pertencer ao conjunto isto é, o quanto é possível para cada elemento pertencer ao conjunto [18]. Em outras palavras a função de pertinência atribui valores entre 0 e 1 para cada elemento de um determinado universo, discriminando os elementos que pertencem dos que não pertencem a um determinado conjunto em análise. Quanto mais o valor de  $x$  for próximo ao conceito central mais os valores da função de pertinência serão próximos de 1, e os que estão distantes receberão valores menores, próximos de 0.

De acordo com Burrough [18], a função de pertinência para conjuntos “*Fuzzy*” deve assegurar que o grau de pertinência seja igual a 1 no centro do conjunto e que caia de forma adequada das regiões de fronteira até as regiões externas do conjunto onde o valor seria 0. O ponto onde o grau de pertinência é 0,5 é chamado ponto de *crossover*, ou ponto de cruzamento.

A função de pertinência pode assumir qualquer forma analítica ou arbitrária, apropriada a cada problema específico. Existem várias funções que podem ser utilizadas para determinar o valor de pertinência do conjunto “*Fuzzy*”. As funções mais comuns utilizadas para determinar valores de membros “*Fuzzy*” e mais comumente usadas em SIG são as funções lineares e a quadrática [22].

### 3.2.1.3 Representações das funções de pertinência

As representações das funções que definem os elementos de um conjunto “*Fuzzy*” facilitam a visualização deste conjunto, e podem ser feitas na forma tabular, graficamente e na forma analítica [48].

Para conjuntos finitos, as funções podem ser representadas por tabelas. A tabela representando um conjunto “*Fuzzy*” lista todos os elementos do conjunto com seus respectivos graus de pertinência.

Jafelice [48] propõe o seguinte exemplo para ilustrar a representação da função de pertinência por tabelas, considerando que A seja o conjunto dos alunos “estudiosos” de uma sala de aula de uma faculdade, e que sejam os alunos desta sala: Fernando, Carlos, Márcia e André. Este conjunto A é um subconjunto do conjunto universo X com todos os alunos da faculdade. Nem todos os alunos do conjunto A estudam diligentemente, logo alguns têm um grau de mais estudioso, outros menos estudiosos variando entre os valores 0 e 1. Para os alunos citados tem-se a representação na tabela abaixo:

<i>Estudante</i>	<i>Grau de estudo</i>
Carlos	0,3
Márcia	0,7
Fernando	0,8
André	0,9

**Tabela 3.3** Representação tabular da função de pertinência  
Adaptado de [48]

Uma outra forma de se representar um conjunto “*Fuzzy*” é feita graficamente. A representação gráfica é a mais usada na literatura “*Fuzzy*” por ter uma interpretação mais intuitiva. No caso de se fazer representação em duas dimensões, o eixo vertical representa o grau de pertinência no intervalo [0,1], e o eixo horizontal contém a informação a ser modelada.

A seguir um exemplo adaptado de Franca-Rocha et. al. [34] é utilizado para ilustrar a representação gráfica da função de pertinência, supondo que se esteja avaliando se uma determinada área está sofrendo processo de erosão. Poderia ser utilizada a teoria clássica dos conjuntos para fazer afirmações do tipo:

“A área sofre processo de erosão” = 1 (verdadeiro)

“A área tem pouca intensidade de chuva” = 0 (falso)

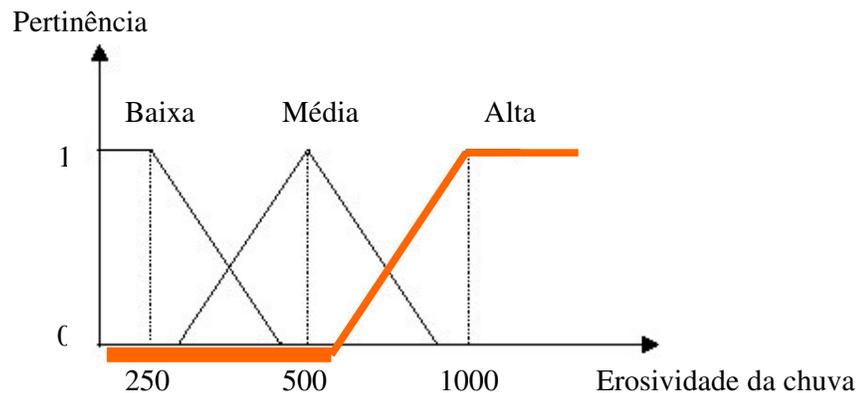
A definição do que pode ser considerado “uma área sofrendo processo de erosão” e “uma área com pouca intensidade de chuva” passa por intervalos de incertezas ou desconhecimentos, em outras palavras, não há limites precisos para caracterizar a partir de que medidas a área pode ser considerada sofrendo ou não processo de erosão. Portanto, o grau de verdade das afirmativas anteriores depende da sua proximidade dos extremos, onde foi bem estabelecida a certeza de que o elemento pertence ou não ao conjunto.

Assim sendo, pode-se re-elaborar as afirmativas precedentes, usando a teoria dos conjuntos nebulosos para avaliar o grau de pertinência, tendo como referência uma escala de valores que variam gradativamente entre os números 0 e 1.

1. “A área sofre processo de erosão” = 0,6 (alguma possibilidade)
2. “A área tem pouca intensidade de chuva” = 0,9 (grande possibilidade)

Assim atribui-se o valor 1 para aquelas áreas sofrendo o processo de erosão, e o valor 0, para aquelas áreas consideradas sem erosão. Para as áreas com medidas intermediárias pode-se atribuir valores compreendidos em uma escala contínua entre 0 e 1 para designar o grau de possibilidade de a área estar ou não sofrendo processo de erosão.

Considerando-se que a erosividade da chuva seja um dos fatores utilizados para a avaliação se uma área sofre processo erosivo tem-se na figura 3.3 a representação gráfica da função de pertinência. Pode-se atribuir o valor 1 para aquelas áreas que tiverem uma erosividade de chuva muito elevada (acima de 1000 ppm) e que são consideradas, portanto, áreas sofrendo processo erosivo. Da mesma forma pode-se atribuir o valor 0 para aquelas áreas com baixa erosividade de chuva (abaixo de 750 ppm) e que são consideradas áreas que não sofrem processo erosivo. Para as áreas com medidas intermediárias de erosividade de chuva (entre 750 e 1000 ppm) pode-se atribuir valores entre 0 e 1, para designar o grau de possibilidade da área estar sofrendo processo erosivo.



**Figura 3.3** Representação gráfica da função de pertinência  
Fonte: adaptado de [34]

A representação analítica é também bastante utilizada em teoria dos conjuntos “Fuzzy”. O conjunto “Fuzzy” do exemplo dado anteriormente pode ser representado analiticamente da seguinte forma:

$$\mu(x) = \begin{cases} 0 & x < 750 \\ \frac{x-750}{250} & 750 < x < 1000 \\ 1 & x > 1000 \end{cases}$$

Segundo Meirelles [61] os membros “*Fuzzy*” definidos pelos conjuntos “*Fuzzy*” podem ser combinados segundo análises multi-critérios, definidas através de uma seqüência lógica realizada pelos operadores “*Fuzzy*” vistos a seguir.

#### 3.2.1.4 Operações entre conjuntos “*Fuzzy*”

Assim como nos conjuntos *Booleanos*, a modelagem através do uso da Lógica “*Fuzzy*” pode ser realizada utilizando operadores lógicos para combinar dados de diversos conjuntos.

Vários operadores podem ser utilizados de forma a se obter um mapa resultante da integração de vários planos de informação “*Fuzzy*” através dos operadores: interseção lógica (*Fuzzy* AND), união lógica (*Fuzzy* OR), soma algébrica (*Fuzzy* SUM), produto algébrico (*Fuzzy* PRODUCT) e o operador gama (*Fuzzy* GAMMA).

##### ***Fuzzy* “AND” (interseção)**

Esta operação equivale ao “AND” *Booleano*. A interseção é obtida através do operador MIN e é controlada pelo menor valor entre os elementos dos conjuntos “*Fuzzy*” combinados [22].

$$\textit{Fuzzy interseção} = \mu_{\text{comb}} = \text{MIN} (\mu_a, \mu_b, \mu_c, \dots)$$

Os valores de  $\mu_a, \mu_b, \mu_c$  são valores de pertinência.

##### ***Fuzzy* “OR” (união)**

*Fuzzy* “OR” é como o “Or” *Booleano*, onde o valor de pertinência de saída é controlado pelos valores máximos dos mapas de entrada [22] e é dado por:

$$\textit{Fuzzy união} = \mu_{\text{comb}} = \text{MAX} (\mu_a, \mu_b, \mu_c, \dots)$$

##### **Soma algébrica *Fuzzy***

Este operador é complementar ao produto algébrico *Fuzzy* e é representado por:

$$\textit{Fuzzy soma} = \mu_{\text{comb}} = 1 - [(1 - \mu_a) * (1 - \mu_b) * (1 - \mu_c)]$$

O resultado desta operação é sempre um valor maior ou igual ao maior valor de pertinência “*Fuzzy*” de entrada. O efeito desta operação é, portanto, de aumentar o valor de pertinência.

### **Produto algébrico *Fuzzy***

Este operador corresponde ao produto algébrico entre os valores dos elementos “*Fuzzy*” e seu resultado tende a ser pequeno devido ao fato de multiplicar diversos números menores do que 1. A função de pertinência combinada, neste caso, é dada por:

$$Fuzzy \text{ Produto} = \mu_{comb} = (\mu_a * \mu_b * \mu_c)$$

### **Operador Gamma**

O operador gamma inclui o produto entre a soma algébrica e o produto *Fuzzy*, e é representado por:

$$\text{Operador } Fuzzy \text{ Gamma } (\gamma) = \mu_{comb} = (Fuzzy \text{ soma})^\gamma * (Fuzzy \text{ produto algébrico})^{1-\gamma}$$

O parâmetro  $\gamma$  é um valor escolhido entre 0 e 1. Quando  $\gamma$  é igual a 1, a combinação é a mesma que a soma algébrica *Fuzzy*, quando  $\gamma$  é igual a 0, a combinação equivale ao produto algébrico *Fuzzy* [61]. Quando o valor de  $\gamma$  está entre 0 e 0,35 há uma redução do valor “*Fuzzy*” resultando que o valor de saída seja menor do que o menor valor de entrada. Quando o valor de  $\gamma$  está entre 0,8 e 1 há um aumento do valor “*Fuzzy*” resultando que o valor de saída seja maior do que o maior valor de entrada. Por fim quando o valor de  $\gamma$  estiver entre 0,35 e 0,8 os valores de saída cairão sempre entre o menor e o maior valor de entrada [22].

### **3.2.1.5 Sistema “Expert *Fuzzy*”**

Um sistema “*Fuzzy*” pode transformar informações vagas e imprecisas em um formato numérico, preciso e possível de ser compreendido pelo computador. Se tomarmos, por exemplo, a informação de que um copo tem água até a metade, pode-se transformar essa informação em um dado mais preciso, onde o estado “copo com água até a metade” pode possuir um grau de pertinência 0,5. Estes sistemas têm a finalidade de combinar a informação numérica e lingüística para obter sistemas eficientes e aplicáveis a diversas áreas. Uma função de pertinência deve ser definida para mapear elementos reais a um intervalo [0,1].

Embora seja possível modelar fenômenos através da Lógica “*Fuzzy*”, é necessário um processamento computacional sobre esse modelo. Segundo Jafelice [48], um Sistema “*Expert Fuzzy*” pode ser descrito pelo conjunto de operações e regras que possuem quatro componentes: um processador de entrada (*fuzzificação*), uma coleção de regras nebulosas (base de regras), uma máquina de inferência e um processador de saída (*defuzzificação*).

### **Fuzzyficação**

Na etapa de *fuzzyficação* as informações reais (dados de entrada) são convertidas em números difusos (valores “*Fuzzy*”) através de variáveis lingüísticas. Uma variável lingüística é uma variável cujo valor é expresso qualitativamente por um termo lingüístico (que fornece um conceito à variável) e quantitativamente por uma função de pertinência. Por exemplo: a variável lingüística “temperatura” pode ser expressa pelos termos lingüísticos “baixa”, “média” ou “alta” [48].

Esta é a etapa onde se modela matematicamente a informação das variáveis de entrada por meio de conjuntos “*Fuzzy*”. Nesta etapa é mostrada a grande importância do especialista do processo a ser analisado, pois a cada variável de entrada devem ser atribuídos termos lingüísticos que representam os estados desta variável e, a cada termo lingüístico, deve ser associado um conjunto “*Fuzzy*” por uma função de pertinência [7].

### **Base de regras**

As regras “*Fuzzy*” é o que constitui o núcleo do sistema “*Fuzzy*”, pois se referem às variáveis e às classificações lingüísticas utilizadas no sistema “*Fuzzy*”. As regras em um Sistema “*Expert Fuzzy*” são usualmente de forma similar ao seguinte: se  $x$  é baixo e  $y$  é alto então  $z$  é médio, onde  $x$  e  $y$  são variáveis de entrada e  $z$  é uma variável de saída. Baixo, alto e médio são funções de pertinência. O conjunto de regras em um Sistema “*Expert Fuzzy*” é conhecido como uma base de regra.

A base de regras é composta por uma coleção de proposições “*Fuzzy*” na forma Se...então.... Cada uma destas proposições pode, por exemplo, ser descrita lingüisticamente de acordo com o conhecimento de um especialista. A base de regras descreve relações entre as variáveis lingüísticas, para serem utilizadas na etapa de inferência [48].

### **Inferência**

A etapa de inferência consiste em aplicar as regras “*Fuzzy*” que foram inseridas na base de conhecimento do sistema “*Fuzzy*”. A etapa de inferência é onde se definem quais são os conectivos lógicos usados para estabelecer a relação “*Fuzzy*” que modela a base de regras. É deste módulo que depende o sucesso do sistema “*Fuzzy*” já que ele fornecerá a saída (controle) “*Fuzzy*” a ser adotado pelo controlador a partir de cada entrada “*Fuzzy*” [7].

É neste componente que cada proposição “*Fuzzy*” é traduzida matematicamente por meio das técnicas de raciocínio aproximado. Os operadores matemáticos serão selecionados para definir a relação “*Fuzzy*” que modela a base de regras. Desta forma, a máquina de

inferência “*Fuzzy*” é de fundamental importância para o sucesso do sistema, já que fornece a saída a partir de cada entrada e da relação definida pela base de regras [48].

Existem diferentes métodos de inferência “*Fuzzy*”, com diferentes propriedades. Dois métodos particulares de inferência “*Fuzzy*” utilizados em sistemas computacionais como Matlab ([www.mathworks.com](http://www.mathworks.com)) são: o método de Mamdani e o Método de Takagi-Sugeno.

### **Defuzzyficação**

A etapa da *defuzzificação* traduz o estado da variável de saída “*Fuzzy*” para um valor numérico, ou seja, é um processo de se representar um conjunto “*Fuzzy*” por um número real. Em sistemas “*Fuzzy*” deve-se escolher um método para *defuzzificar* a saída e obter um número real que a represente [48].

A *defuzzificação* faz com que a partir dos graus de participação de cada variável de uma regra, retorne o grau de participação da saída e conseqüentemente o valor real da saída. É usualmente utilizada para converter o conjunto “*Fuzzy*” de saída para um número *crisp*.

O método mais utilizado no processo de *defuzzificação* é o do centro de gravidade ou centróide. Neste método a saída do sistema “*Fuzzy*” é dada pelo centro de gravidade dos valores difusos obtidos a partir da inferência [7].

### **3.2.2 Pesos de Evidência**

O modelo de Pesos de Evidência é um método quantitativo para combinar evidências em suporte de uma hipótese, pertencente a um grupo de métodos apropriados para a Tomada de Decisão Multi-Critério [40]. Esse modelo é usado para predizer uma hipótese sobre ocorrências de um evento baseado na combinação de evidências conhecidas em uma área de estudo.

O modelo de Pesos de Evidência é um modelo de predição baseado em dados que difere de modelos de predição baseados em conhecimento. A distinção entre modelo baseado em dados e modelo baseado em conhecimento é aparente no modelo de Pesos de Evidência, que confia na avaliação dos dados de entrada para “estimar a importância relativa das evidências por métodos estatísticos”.

Ao invés de empregar conhecimento dirigido procurando outros modelos de mapa semelhantes, o Peso de Evidência usa ocorrências conhecidas para fazer ajustes do seu próprio modelo. O objetivo deste modelo é formar mapas de predição de probabilidade de

ocorrência de um fenômeno particular em uma área de estudo, baseada em um ou mais dados de evidência.

O modelo de Pesos das Evidências utiliza uma formulação log-linear do Teorema de Probabilidade de Bayes [14] para combinar padrões em planos de informação, com o objetivo de prognosticar a distribuição de objetos espacialmente referenciados para dar suporte a uma determinada hipótese [7].

O método foi originalmente desenvolvido para uma aplicação não espacial em diagnósticos médicos, com evidências obtidas a partir de um conjunto de sintomas e hipóteses do tipo “este paciente tem a doença X”. Para cada sintoma, um par de pesos era calculado, uma para presença do sintoma e uma para a ausência do sintoma. A magnitude dos pesos dependia da associação medida entre o sintoma e o padrão da doença em um grande grupo de pacientes. Os pesos podiam ser usados para estimar a probabilidade que um novo paciente poderia adquirir a doença, baseado na presença ou ausência do sintoma [34].

O modelo de Pesos de Evidência foi adaptado mais tarde em 1980 para o mapeamento de exploração mineral com o SIG [16]. Nesta situação, a evidência consistia em um conjunto de dados de exploração (mapas geológicos, geoquímicos e geofísicos), e as hipóteses eram baseadas em determinar se um local era favorável à ocorrência do depósito do tipo X [40].

Os pesos eram estimados das medidas associadas entre as ocorrências minerais conhecidas e os valores dos mapas usados como preditores. Baseado na combinação dos planos de informação (mapas de evidência) selecionados, o resultado final é extraído como um mapa de potencial representando a probabilidade de ocorrência do evento analisado [34].

Esse método tem sido muito utilizado para mapeamento de potenciais minerais, mas atualmente é usado em outros tipos de predição espacial visando a predição probabilística de ocorrência de objetos pontuais, como por exemplo, a análise baseada na predição de pontos de erosão.

### **3.2.2.1 Descrição do método de Pesos de Evidência**

O método de Pesos de Evidência assume que, para uma região em particular, uma série de mapas binários são conhecidos, e são usados como preditores de um determinado evento. Os mapas preditores binários podem ser usados como mapas de entrada na análise. Assume-se ainda que as localizações de algumas ocorrências são conhecidas (essas ocorrências são chamadas de pontos de treinamento). Os produtos desejados são mapas que mostram a probabilidade de ocorrência e a associação da incerteza da probabilidade estimada.

O método de Pesos de Evidência para a geração de um mapa de potencial envolve a estimação da probabilidade a priori, o cálculo dos pesos positivo e negativo para cada evidência relativa às ocorrências, cálculo da probabilidade posteriori, a conversão de *logits* para probabilidade posteriori, a combinação de conjunto de dados e a independência condicional dos dados. A seguir será feita uma descrição do método de Pesos de Evidência com base em Bonham-Carter [14] e Câmara [21], para que se possa entender o significado dos resultados obtidos por este método de análise.

A probabilidade a priori consiste em determinar a probabilidade de ocorrências de um determinado evento por unidade de área. Considerando, por exemplo, que exista uma área de estudos "T", onde um mapa "B" é usado como tema de evidência (ex. mapa de vegetação) e onde "D" são pontos de ocorrências conhecidas (ex. pontos de erosão). Temos desta forma que a probabilidade a priori é dada por:

$$P(D) = \frac{N(D)}{N(T)} \quad \text{equação 3.1}$$

O termo chance é definido por Bonham-Carter [14] como a razão da probabilidade de que um evento ocorra pela probabilidade de que ele não ocorra. O método de Pesos de Evidência usa o logaritmo natural da chance, conhecido como *log odds* ou *logits* (bem parecido à formulação "*Odds*" que foi desenvolvida por um sistema "*expert*" de prospecção) para facilitar a interpretação dos dados que são diretamente relacionados à probabilidade. Sendo assim temos:

$$\text{Odds} = \frac{\text{probabilidade a priori}}{(1 - \text{probabilidade a priori})} \quad \text{equação 3.2}$$

$$O(D) = \frac{P(D)}{1 - P(D)} = \frac{P(D)}{P(D')} \quad \text{equação 3.3}$$

$$\text{logit}(D) = \ln [O(D)] \quad \text{equação 3.4}$$

O modelamento de Pesos de Evidência usa a localização das ocorrências conhecidas para determinar coeficientes para cada mapa de predição. Contudo, existem dois coeficientes, ou pesos,  $W_+$  e  $W_-$ , para cada mapa de predição. Os mapas de predição são usualmente binários, e  $W_+$  e  $W_-$  referem-se a essas áreas onde o padrão binário está presente ou não está presente, respectivamente. Os pesos são calculados como logaritmos de razões de probabilidade condicional, aplicando-se a formulação do Teorema de Bayes, usando medidas

entre a área do padrão binário e a área de estudo total, o número de ocorrências de um evento com o padrão binário e o número de ocorrências na área de estudo [40].

Os pesos produzem uma medida da associação espacial entre os pontos e o tema evidência. Um peso é calculado para cada classe do tema evidência. Um valor positivo de peso indica que há mais pontos sobre a classe do que ocorreria devido à chance; de forma contrária, um valor negativo indica que muito poucos pontos ocorrem do que seria o esperado. Um valor de zero ou muito próximo de zero, indica que os pontos de treinamento estão aleatoriamente distribuídos, com respeito àquela classe. Para mapas binários com somente duas classes,  $W+$  é usado para os pesos onde o tema evidência está presente, e  $W-$  é usado para os pesos onde o tema evidência encontra-se ausente [40].

Algumas vezes  $W+$  pode ser próximo de zero, ainda que  $W-$  seja fortemente negativo. Esta situação sugere que a presença do tema não é particularmente preditiva dos pontos de treinamento, mas a ausência do tema provoca forte evidência de que os pontos são pouco prováveis de ocorrer. Ao contrário, pode haver um desequilíbrio entre os valores absolutos de  $W+$  e  $W-$ , em outra direção, ou os dois pesos podem ter valores absolutos próximos do mesmo intervalo.

O peso de evidência  $W +$  é o logaritmo natural do termo LS que é conhecido como **razão de suficiência**.

$$W + = \log_e LS \quad \text{equação 3.5}$$

A razão de suficiência é dada pela razão entre a probabilidade do padrão binário B estar presente dada a presença da ocorrência ( $P(B|D)$ ), sobre a probabilidade do padrão binário B estar presente dada a ausência da ocorrência ( $P(B|D')$ ). Assim:

$$LS = \frac{P(B|D)}{P(B|D')} = \frac{N(B \cap D) / N(D)}{N(B) / N(T)} \quad \text{equação 3.6}$$

O peso de evidência  $W -$  é o logaritmo natural do termo LN que é conhecido como **razão de necessidade**.

$$W - = \log_e LN \quad \text{equação 3.7}$$

A razão de necessidade é dada pela razão entre a probabilidade do padrão binário B estar ausente dada à presença da ocorrência ( $P(B'|D)$ ), sobre a probabilidade do padrão binário B estar ausente dada a ausência da ocorrência, ( $P(B'|D')$ ). Assim:

$$LN = \frac{P(B'|D)}{P(B'|D')} = \frac{N(B' \cap D) / N(D)}{N(B') / N(T)} \quad \text{equação 3.8}$$

LS e LN também são chamados **razões de probabilidade**.

De acordo com Bonham-Carter [14] a partir dos valores calculados para  $W+$  e  $W-$  é possível fazer algumas associações espaciais entre o tema de evidência B e os pontos de treinamento D, como as associações mostradas abaixo.

- $W+ > 0$  indica associação positiva dada a presença de B
- $W+ = 0$  indica nenhuma associação
- $W+ < 0$  indica associação negativa dada a presença de B
- $W- > 0$  indica associação positiva dada a ausência de B
- $W- = 0$  indica nenhuma associação
- $W- < 0$  indica associação negativa dada a ausência de B

Segundo Araújo [9], os valores dos pesos podem ser interpretados da seguinte maneira: um peso positivo para um valor particular do mapa de evidência indica que ocorrências são mais frequentes do que o esperado pelo cálculo da chance, enquanto valores negativos indicam o contrário, ou seja, ocorrências são menos frequentes do que o valor esperado pelo cálculo da chance. Um peso de zero indica que não há correlação espacial entre o tema analisado e a ocorrência.

A diferença entre os pesos é conhecida como contraste (C). O contraste é uma medida total da associação espacial entre os pontos de treinamento e o tema evidência, combinando os efeitos dos dois pesos. O contraste mostra quais dos temas de evidência utilizados na análise são mais fortemente associados com os pontos de ocorrência conhecidos [40].

$$C = (W+) - (W-) \quad \text{equação 3.9}$$

Em termos práticos pode-se utilizar um exemplo adaptado de Bonham-Carter [14], onde pretende-se fazer uma avaliação do processo de erosão em uma determinada região de

estudo T, que possui uma área equivalente a 1000 unidades de área. Dentro desta área há a presença de um padrão B que possui área de 500 unidades de área. Nesta área existem 30 pontos identificados como erosão (voçorocas, por exemplo), que serão os pontos de treinamento, sendo que 20 desses pontos encontram-se contido dentro da área B. Assim a seqüência de cálculos para a determinação dos pesos  $W_+$  e  $W_-$  e o contraste C será:

A probabilidade do padrão binário B estar presente dada a presença de um ponto de erosão  
 $P(B|D) = P(B \cap D) / P(D) = 20 / 30 = 0,66$  equação 3.10

A probabilidade do padrão binário B estar presente dada a ausência de um ponto de erosão  
 $P(B|D') = P(B \cap D') / P(D') = (500-20) / (1000-30) = 480/970 = 0,49$  equação 3.11

A probabilidade do padrão binário B estar ausente dada a presença de um ponto de erosão  
 $P(B'|D) = P(B' \cap D) / P(D) = (30-20)/30 = 10/30 = 0,33$  equação 3.12

A probabilidade do padrão binário B estar ausente dada a ausência de um ponto de erosão  
 $P(B'|D') = P(B' \cap D') / P(D') = (1000-500-30+20)/(1000-30) = 490/970 = 0,5$   
 equação 3.13

A taxa de suficiência para o padrão B  
 $LS = P(B|D) / P(B|D') = 0,66 / 0,49 = 1,3472$  equação 3.14

O pesos de evidência positivo para o padrão B  
 $W_+ = \log_e LS = 0,2980$  equação 3.15

A taxa de necessidade para o padrão B  
 $LN = P(B'|D) / P(B'|D') = 0,33 / 0,50 = 0,65$  equação 3.16

O pesos de evidência negativo para o padrão B  
 $W_- = \log_e LN = - 0,4157$  equação 3.17

Contraste  
 $C = (W_+) - (W_-) = 0,7137$  equação 3.18

O exemplo dado anteriormente representa o cálculo para determinação dos pesos para somente um tema de evidência com uma classe. Na análise espacial é comum ter temas de evidência com múltiplas classes, como é o caso de vegetação, ou solo, entre outros. Os temas com múltiplas classes podem ter classes ordenadas ou não ordenadas. Nas classes ordenadas (como por exemplo, mapas de contornos) usualmente o cálculo dos pesos é baseado nos sucessivos níveis. Os temas com classes não ordenadas também são chamados categóricos. No caso de temas com múltiplas classes as fórmulas mencionadas acima serão utilizadas para calcular os pesos para cada classe do tema de evidência.

O modelo de Pesos de Evidência pode ser aplicado tanto a mapas com múltiplas classes de evidência, como mapas com classes binárias (ausência ou presença da evidência). Porém as modelagens através de mapas com classes binárias têm sido utilizadas com mais frequência, pois são mais fáceis de interpretar e pelo fato de conferir maior confiança estatística à modelagem de Pesos de Evidência, principalmente quando o número de pontos de treinamento é pequeno.

Para a utilização de mapas binários faz-se necessária à reclassificação dos mapas com múltiplas classes. O processo de conversão de mapas com múltiplas classes em mapas binários pode ser feito subjetivamente ou através de métodos estatísticos. Nos mapas com classes ordenadas a reclassificação pode ocorrer com base no contraste e observando-se ainda a medida de sua normalização (para verificar o grau de variância), podendo simplificar assim os mapas, reduzindo o número de classes e separando as áreas mais favoráveis daquelas menos favoráveis [34].

Um dos processos utilizados para a reclassificação dos mapas com múltiplas classes é o critério do contraste máximo que separa as classes anômalas (padrão) daquelas não anômalas (fora do padrão). O critério de contraste máximo determina o limiar de corte que maximizaria a associação espacial entre os temas de evidência resultantes da reclassificação binária com as ocorrências conhecidas.

Imagine, por exemplo, que um tema com 10 classes ordenadas foi utilizado para avaliar o processo de erosão, sendo calculados os pesos de evidência para cada uma das classes, conforme a tabela abaixo.

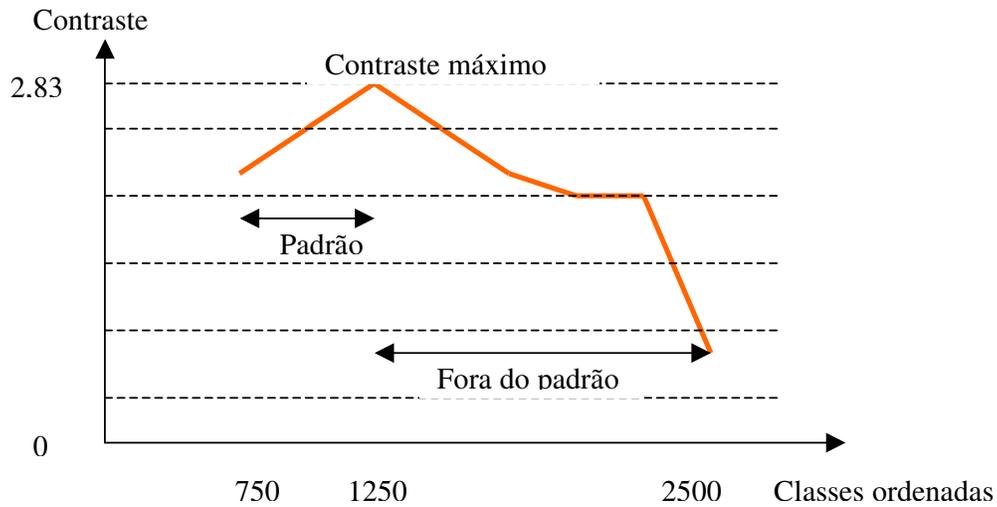
classes	n° pontos	W+	W-	contraste (C)	confiança (C/S(C))	classe binária
250	0					2
500	0					2
750	2	2,143	-0,149	2,292	2,956	2
1000	4	2,261	-0,335	2,596	4,285	2
<b>1250</b>	<b>6</b>	<b>2,269</b>	<b>-0,57</b>	<b>2,839</b>	<b>5,07</b>	<b>2</b>
1500	6	2,135	-0,563	2,697	4,821	1
1750	6	1,849	-0,544	2,393	4,282	1
2000	6	1,675	-0,529	2,204	3,946	1
2250	6	1,593	-0,521	2,114	3,787	1
2500	6	1,392	-0,497	0,889	3,386	1

**Tabela 3.4** Exemplo de pesos para dado com 10 classes ordenadas  
Fonte: adaptado de [34]

Para a transformação do tema em um mapa binário faz-se a observação do contraste e de seu grau de confiança (estimado como a razão entre o contraste e seu desvio padrão) para

estabelecer o limiar onde será separada a área anômala da não anômala, segundo o critério de contraste máximo.

Neste exemplo percebe-se que tanto o contraste como o seu grau de confiança é máximo na classe 1250, assim estabelece-se essa classe como limiar para separação das classes binárias. As classes anteriores a esta foram reclassificadas na classe binária 2 (indicando a presença da anomalia) e o restante na classe binária 1 (indicando a ausência da anomalia).



**Figura 3.4** Critério do contraste máximo  
Fonte: adaptado de [34]

Após a separação das classes binárias é feito o cálculo para a determinação dos novos pesos positivos e negativos, usados posteriormente no cálculo da probabilidade final pelo método de Pesos de Evidência.

A probabilidade condicional (posteriori) é a probabilidade calculada de acordo com as evidências adicionais disponíveis. O modelo de Pesos de Evidência expressa a probabilidade condicional de uma ocorrência ocorrer quando um padrão binário anômalo está presente em função da probabilidade prévia da ocorrência de depósitos na área e de um multiplicador vinculado à presença do padrão binário anômalo. Assim, segundo Bonham-Carter [14], temos a seguinte seqüência de equações:

$$P(D|B) = P(D) \frac{P(B|D)}{P(B)} \quad \text{onde: } P(D|B) \text{ é a probabilidade posterior}$$

$$P(D) \text{ é a probabilidade priori}$$

equação 3.19

Uma expressão semelhante relaciona a probabilidade posterior de um depósito ocorrer na ausência do padrão binário anômalo, ou seja:

$$P(D|B') = P(D) \frac{P(B'|D)}{P(B')} \quad \text{equação 3.20}$$

Como dito anteriormente a conversão da probabilidade em *log odds* ou *logits* é mais fácil de ser interpretado. Para converter a equação da probabilidade posterior para *odds*, divide-se ambos lados da equação da probabilidade posterior por  $P(D'|B)$ , conduzindo para:

$$\frac{P(D|B)}{P(D'|B)} = \frac{P(D)}{P(D'|B)} \frac{P(B|D)}{P(B)} \quad \text{equação 3.21}$$

Mas da definição de probabilidade condicional tem que:

$$P(D'|B) = \frac{P(D' \cap B)}{P(B)} = \frac{P(B|D') P(D')}{P(B)} \quad \text{equação 3.22}$$

Substituindo esta expressão por  $P(D'/B)$  no denominador no lado direito da equação acima e re-arranjando os termos temos:

$$\frac{P(D|B)}{P(D'|B)} = \frac{P(D)}{P(D')} \frac{P(B)}{P(B)} \frac{P(B|D)}{P(B|D')} \quad \text{equação 3.23}$$

O *odds* de uma ocorrência  $O(D)$  é igual a  $P(D) / (1-P(D))$ , ou  $P(D) / P(D')$ , então substituindo *odds* na equação acima e teremos a expressão desejada:

$$O(D|B) = O(D) \frac{P(B|D)}{P(B|D')} \quad \text{onde } O(D|B) \text{ é a odds condicional (posterior) de D}$$

$$O(D) \text{ é a odds priori de D}$$

$$P(B|D)/P(B|D') \text{ é a razão de suficiência LS}$$

equação 3.24

Aplicado o logaritmo natural em ambos os lados da equação acima teremos:

$$\text{logit}(D|B) = \text{logit}(D) + W+ \quad \text{equação 3.25}$$

A fórmula pode ser entendida como a chance posterior de ser encontrado algum evento na presença do padrão binário que é o somatório da chance prévia ( $\text{logit}(D)$ ) com o ponderador positivo de evidência ( $W+$ ).

Manipulações algébricas semelhantes conduzem à derivação da expressão *odds* para a probabilidade condicional de D dada à ausência do padrão de indicador binário, onde temos:

$$O(D|B') = O(D) \frac{P(B'/D)}{P(B'/D')} \quad \text{equação 3.26}$$

A forma logit da equação acima é

$$\text{logit}(D|B') = \text{logit}(D) + W^- \quad \text{equação 3.27}$$

Pode-se converter o *logits* para a probabilidade posteriori usando a equação:

$$\text{Probabilidade} = \frac{\text{Odds}}{(1 + \text{Odds})} \quad \text{equação 3.28}$$

Assim temos que a probabilidade posteriori de uma ocorrência dada à presença do padrão binário B é:

$$P(D|B) = \frac{O(D|B)}{(1 + O(D|B))} \quad \text{equação 3.29}$$

E a probabilidade posteriori de uma ocorrência dada à ausência do padrão binário B é:

$$P(D|B') = \frac{O(D|B')}{(1 + O(D|B'))} \quad \text{equação 3.30}$$

Segundo Bonham-Carter [14], quando as evidências de alguns mapas são combinadas, os pesos (ou seus equivalentes) são calculados para cada mapa independentemente e então combinados em uma simples equação.

A probabilidade condicional (posterior) de um depósito ocorrer, dada à presença de dois predicativos mapas de padrão, B1 e B2 é:

$$P(D | B1 \cap B2) = \frac{P(D \cap B1 \cap B2)}{P(B1 \cap B2)} = \frac{P(B1 \cap B2 | D) P(D)}{P(B1 \cap B2)} \quad \text{equação 3.31}$$

$$P(D | B1 \cap B2) = \frac{P(B1 \cap B2 | D) P(D)}{P(B1 \cap B2 | D) P(D) + P(B1 \cap B2 | D') P(D')} \quad \text{equação 3.32}$$

O efeito da interação entre B1 e B2 pode ser ignorada fazendo uma suposição de independência condicional, que será vista na próxima seção. A independência condicional provê uma simplificação, permitindo que cada mapa binário seja individualmente avaliado e então combinado.

A independência condicional assumida pode ser escrita como:

$$P(B1 \cap B2 | D) = P(B1 | D) P(B2 | D) \quad \text{equação 3.33}$$

Que permite a equação acima seja simplificada da seguinte forma:

$$P(D | B1 \cap B2) = P(D) \frac{P(B1 | D)}{P(B1)} \frac{P(B2 | D)}{P(B2)} \quad \text{equação 3.34}$$

Usando a formulação de *odds*, a *odds* condicional ou posteriori pode ser expressa por dois mapas de padrões determinados por:

$$O_{D|B1 \cap B2} = O(D) * LS1 * LS2 \quad \text{equação 3.35}$$

Ou com o pesos de evidência log-linear:

$$\text{Logit}(D | B1 \cap B2) = \text{logit}(D) + W^+_1 + W^+_2 \quad \text{equação 3.36}$$

O tema de resposta é o mapa de saída que expressa a probabilidade que uma unidade de área tem de conter um ponto, estimado pela combinação dos pesos de variáveis exploratórias (temas de evidência). O mapa de resposta é usualmente classificado em um pequeno número de valores e descrito como favorabilidade relativa. Se os pontos de treinamento são pontos de erosão, então o mapa de resposta mostrará para esse evento uma estimativa de potencial (também conhecido como favorabilidade). Pela equação 3.36 verifica-se que o tema de resposta ou mapa de favorabilidade é dado pela soma dos pesos adicionado ao *logit* priori.

### 3.2.2.2 Independência condicional dos dados

Segundo Meirelles et. al. [60], o termo independência condicional se refere à probabilidade de um evento não interferir no outro. Pode ser explicado como a aplicação de teste estatístico para verificar a condição de independência dos dados relativos entre si e as ocorrências, conduzindo possivelmente à rejeição ou substituição de alguns mapas de entrada.

A independência condicional estabelece a independência de dois termos, dada a presença de um terceiro. Em termos práticos, pode-se citar três eventos: o evento A representando o trovão e o evento B representando as chuvas e o evento C representando relâmpagos. Observa-se que o trovão é condicionalmente independente de chuva, ou seja, o trovão não causa chuva, dada a presença de relâmpago, ou seja:

$$P(A|B,C) = P(A|C) \quad \text{equação 3.37}$$

$$P(\text{trovãol chuva, relâmpago}) = P(\text{trovãolrelâmpago})$$

Segundo Bonham-Carter [14] o modelo Bayesiano exige a condição de independência entre as camadas de dados (mapas de entrada). Na prática, essa condição é violada de alguma forma, porém não deve ser violada num grau elevado. Alguns testes são usados para quantificar o nível máximo de aceitação da violação da condição de independência.

Os testes para verificar a condição de independência dos dados utilizados podem ser feitos de duas formas. O primeiro teste é executado entre pares de camadas de dados: são os testes estatísticos Chi-quadrado ( $\chi^2$ ) e similarmente  $G^2$ . O segundo teste é realizado para verificar o ajuste global da condição de independência, que podem ser: o teste simples proposto por Bonham-Carter [14], o teste estatístico Chi-quadrado ( $\chi^2$ ) de intervalos da probabilidade a posteori e o teste de Kolmogorov-Smirnov.

#### **Teste entre pares de dados - Teste estatístico Chi-quadrado ( $\chi^2$ )**

As evidências binárias geralmente são comparadas em pares, podendo ser utilizado como parâmetro de medida o índice Chi-quadrado ( $\chi^2$ ) e o Coeficiente de Contingência (C), tendo por objetivo testar o grau de associação entre classes de cada tema binário. Os procedimentos devem ser os mesmos para todas as análises dos pares de evidências.

Primeiramente deve-se executar uma tabulação cruzada entre as duas evidências, o que gerará uma matriz de confusão, também conhecida como tabela de contingência, de onde

serão extraídos os parâmetros necessários para o cálculo do  $\chi^2$ . As linhas da tabela de contingência são as classes de um tema e as colunas da tabela de contingência são as classes do segundo tema. Abaixo se encontra a tabela de contingência entre dois temas de evidência.

		Tema1		
		favorável	não-favorável	totais
Tema2	favorável	R1 (T1)	R2 (T2)	N(B2∩D)
	não-favorável	R3 (T3)	R4 (T4)	N(B'2∩D)
	totais	N(B1∩D)	N(B'1∩D)	N(D)

**Tabela 3.5** Tabela de contingência  
Fonte: adaptado de [2]

Segundo Agterberg e Cheng [2] os valores R1, R2, R3 e R4 são os valores observados no mapa, ou seja, os valores que representam as interseções entre as classes do primeiro mapa binário com o segundo mapa binário. Os valores T1, T2, T3 e T4 representam o número de pontos de treinamento esperados que ocorrem para uma específica sobreposição das classes do tema 1 e as classes do tema 2, caso os temas de padrões binários analisados sejam independentes. Os valores T1, T2, T3 e T4 são calculados pelo produto dos pontos total marginais dividido pelo número de pontos total, assim:

$$T1 = \frac{N(B1 \cap B2 \cap D)}{N(D)} = \frac{N(B2 \cap D) * N(B1 \cap D)}{N(D)} \quad \text{equação 3.38}$$

$$T2 = \frac{N(B'1 \cap B2 \cap D)}{N(D)} = \frac{N(B2 \cap D) * N(B'1 \cap D)}{N(D)} \quad \text{equação 3.39}$$

$$T3 = \frac{N(B1 \cap B'2 \cap D)}{N(D)} = \frac{N(B'2 \cap D) * N(B1 \cap D)}{N(D)} \quad \text{equação 3.40}$$

$$T4 = \frac{N(B'1 \cap B'2 \cap D)}{N(D)} = \frac{N(B'2 \cap D) * N(B'1 \cap D)}{N(D)} \quad \text{equação 3.41}$$

Com a tabela de contingência criada pode-se estimar o valor de Chi-quadrado ( $\chi^2$ ). Segundo Bonham-Carter [14] o Chi-quadrado é uma medida baseada na frequência observada e na esperada, através da fórmula:

$$\chi^2 = \sum \frac{(\text{Observado} - \text{Esperado})^2}{\text{Esperado}} \quad \text{equação 3.42}$$

$$\chi^2 = \frac{(R1 - T1)^2}{T1} + \frac{(R2 - T2)^2}{T2} + \frac{(R3 - T3)^2}{T3} + \frac{(R4 - T4)^2}{T4} \quad \text{equação 3.43}$$

O valor de  $\chi^2$  mostra se o desvio entre os valores observados e os esperados é pequeno ou não. Quando há independência condicional, as frequências observadas tendem a ficar perto das frequências esperadas. Neste caso o valor de  $\chi^2$  deve ser pequeno. Em outras palavras, um valor pequeno de  $\chi^2$  indica que as variáveis podem ser independentes. Por outro lado, um valor grande na estatística  $\chi^2$ , sinaliza que as diferenças entre as frequências observadas e frequências esperadas não devem ser meramente casuais, ou seja, deve haver associação entre as duas variáveis. A partir do valor de  $\chi^2$  calculado pode-se calcular também o valor do coeficiente de contingência (C) [14], onde:

$$C = \sqrt{\frac{\chi^2}{N(D) + \chi^2}} \quad \text{equação 3.44}$$

Essas operações são então repetidas para os demais temas de evidências, fazendo-se várias comparações, considerando-se todas as evidências. Segundo Moreira [68] se os resultados dos coeficientes de contingência forem valores próximos a zero isto demonstra que as evidências violaram o princípio da independência condicional em grau aceitável, não sendo necessário nenhum ajuste das mesmas.

O teste estatístico  $G^2$  também pode ser usado para verificar a independência condicional dos dados. Apesar do teste  $G^2$  ser diferente do teste estatístico Chi-quadrado eles geralmente produzem resultados similares. O teste da tabela de contingência não é indicado quando o número de ocorrências é pequeno, sendo indicado assim um teste global para a verificação de independência condicional dos dados.

### **Teste Global - Teste “Omnibus”**

Segundo Agterberg e Cheng [2] o teste global estabelece que idealmente a independência condicional de todos os mapas implica que a soma da probabilidade posteriori para todas as unidades de células na área de estudo (T) é igual ao numero total de eventos (n).

Porém, em aplicações práticas, (T) geralmente excede (n) e a diferença aumenta quando mapas adicionais são incluídos, indicando possível falta de independência condicional.

Segundo Kemp, Bonham-Carter, e Raines [53], citado por [2], pode ser assumido que se  $T > n$  é dado pela falta de independência condicional de camadas de mapa. Isto é a razão do teste global, denominado também “teste de omnibus” para condicional independência.

Um simples procedimento pode ser aplicado para determinar o número total de pontos de ocorrências preditos (T). O número de pontos preditos pelo modelo pode ser determinado pelo somatório do produto da área  $N(A)$  pela probabilidade posterior (Pf) para cada única condição do mapa (se “p” padrões binários são considerados então existem  $2^p$  únicas condições), assim:

$$T = \sum (Pf \cdot N(A)) \quad \text{equação 3.45}$$

Na prática, o número predito é sempre maior que o número observado no modelo de Pesos de Evidência. Se o total de números de pontos preditos (T) é muito maior que o total do número observado (n), isto sugere que a independência condicional está sendo violada. Se o número predito for de 10 a 15% maior que o número observado, então se deve fazer um novo teste de independência condicional dos dados [53].

Em termos práticos pode-se utilizar um exemplo adaptado de Agterberg e Cheng [2], onde imagina-se que três mapas de evidência (i, j, k) foram analisados em uma área de estudos ligeiramente maior que 37.534 Km<sup>2</sup>. Na área de estudo há cerca de 13 ocorrências de um determinado evento, ou seja,  $n=13$ . O modelo para esses 3 mapas resultam numa probabilidade posterior de 0,0453 que é cerca de 130 vezes menor que a probabilidade a priori ( $13 / 37.534$ ) = 0,00034635.

Mapa	W+	S(W+)	W-	S(W-)	C	S(C)
I	2,251	1	-1,232	0,29	3,481	1,041
J	2,037	1	-0,811	0,29	2,848	1,041
K	0,632	0,41	-0,338	0,378	0,97	0,558

**Tabela 3.6** Tabela de contingência entre dois temas  
Fonte: adaptado de [2]

Como temos três mapas de evidência, há a separação de oito ( $2^3$ ) condições únicas que são listadas na tabela 3.7.

<i>IJK</i>	Area (km <sup>2</sup> )	Pf	S(Pf)	T	S <sup>2</sup> (T)
222	1,344	0,0003	0,0005	0,0403	0,0045
212	0,9007	0,0057	0,0065	0,5134	0,3428
221	0,4351	0,0008	0,0013	0,0348	0,0032
122	0,4187	0,0111	0,0126	0,4648	0,2783
112	0,3415	0,1709	0,0907	5,8362	9,5939
211	0,2223	0,0154	0,0176	0,3423	0,1531
111	0,1771	0,3604	0,153	6,3827	7,3421
121	0,1456	0,0297	0,336	0,4324	23,9332
Sum	3,985			14,047	41,6511

**Tabela 3.7** Exemplo de pesos para três mapas  
Fonte: adaptado de [2]

A primeira coluna mostra as únicas condições do modelo, representando com o número 1 a presença e 2 a ausência do padrão binário. Na terceira e quarta coluna encontram-se os resultados das probabilidades posteriores (Pf) e o seu desvio padrão S(Pf).

A quinta e coluna representa as oito frequências preditivas dada pelo produto de cada probabilidade posterior (Pf) pela área em unidades de células de cada única condição (N(A)) resultando em oito frequências preditivas. A estimativa de (T) é gerada pelo somatório do produto das probabilidades posteriores pelas áreas em unidades de célula de cada única condição, resultando no valor aproximado de  $T = 14,05$ .

A sexta e última coluna correspondente à variância  $S^2(T)$  de cada frequência preditiva. Assim a variância total gerou um valor de correspondente a 41,6511, onde seu desvio padrão (raiz quadrada da variância) corresponde ao valor de  $S(T) = 6,45$ .

A diferença de  $T-n = 14,05 - 13 = 1,05$  não é estatisticamente significativa e a hipótese de independência condicional pode ser aceita para o modelo com esses 3 temas.

### 3.2.3 Redes Neurais

Inteligência artificial poderia ser definida como um conjunto de modelos, algoritmos, técnicas, ferramentas e aplicações, em um sistema computadorizado, que emula algumas das habilidades cognitivas do homem. A inteligência artificial abrange o conhecimento de diversas ciências, tais como, Ciência da Computação, Linguística, Matemática,

Neurofisiologia, Psicologia, que conjuntamente, objetivam o desenvolvimento de computadores e programas capazes de imitar e/ou entender aspectos do pensamento humano [59].

As Redes Neurais Artificiais têm como proposta a construção de modelos cognitivos artificiais inspirados nos modelos naturais de redes neurais. As redes neurais estão inseridas no ramo de estudo da inteligência artificial, que estuda o conjunto de paradigmas que procura obter um comportamento inteligente a partir de implementações artificiais, normalmente com a utilização de sistemas computacionais. A idéia baseada dos sistemas inteligentes é a reprodução da capacidade de auto-organização e de aprendizado do cérebro [43].

A estrutura do cérebro serviu de inspiração para grande parte dos conceitos existentes de redes neurais, que são implementadas como um conjunto, muitas vezes numeroso, de entidades simples denominadas neurônios ou nodos. Espera-se que com essa estrutura seja possível obter capacidades cerebrais, como executar operações, adaptar-se a novas circunstâncias, entre outras. Resumidamente a base fundamental do estudo das redes neurais é acreditar que uma rede neural possa ser ensinada. As redes neurais artificiais possuem a capacidade de aprender a partir de exemplos e de realizar analogias do que foi aprendido e armazenado [43].

As Redes Neurais Artificiais são modelos computacionais inspirados no raciocínio humano e que podem adquirir conhecimento a partir dos dados através de técnicas de inteligência artificial. Essas redes são compostas por elementos de processamento simples chamados de neurônios, os quais são interconectados por links associados a coeficientes numéricos (pesos) que indicam a força de cada conexão. São adequadas ao reconhecimento de padrões e a classificação de dados espaciais [87].

As Redes Neurais Artificiais apresentam um modelo matemático inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento através da experiência [4]. Uma rede neural artificial é composta por vários elementos de processamento (nodos) que computam determinadas funções matemáticas (normalmente não lineares). Tais unidades são dispostas em uma ou mais camadas interligadas por um grande número de conexões (sinapses). Na maioria dos modelos estas conexões estão associadas a pesos, os quais armazenam o conhecimento apresentado no modelo e servem para ponderar a entrada recebida por cada neurônio [60].

Uma Rede Neural é constituída por algumas camadas: uma camada de entrada onde os sinais (ou padrões) são apresentados a rede; uma ou duas camadas intermediárias (ou

escondidas ou ocultas), onde é efetuada a maior parte do processamento através das conexões ponderadas; e a uma camada de saída, onde o resultado final é concluído e apresentado [22].

Na classificação de dados espaciais há funções que podem ser utilizadas na camada intermediária que determinam os coeficientes numéricos a serem associados aos neurônios, permitindo a sua classificação. Há diversas funções que podem ser utilizadas, porém as mais utilizadas são as MLPs (Multi-Layer Perceptrons) e as funções de base radial - RBFs (Radial Basis Functions). Nas redes neurais MLP a função da camada oculta envolve operações simples de adição, multiplicação e uma transformação não-linear aos sinais de entrada. As redes neurais RBFs utilizam funções radiais para determinar a proximidade dos sinais de entrada daqueles determinados na fase de execução [87].

Em resumo pode-se considerar que em inteligência artificial é comum encontrarmos menção a aprendizado de máquinas e a capacidade de aprender pode ser considerada uma característica fundamental de um comportamento inteligente. As Redes Neurais Artificiais possuem a capacidade de aprender a partir de exemplos e de realizar interpolações (analogia) do que foi aprendido e armazenado em conexões sinápticas (memorização) [43].

### **3.2.4 Regressão Logística**

Segundo Câmara [21] um dos tipos de estudos mais comuns com dados de área é o uso de modelos de regressão. Um modelo de regressão é uma ferramenta estatística que utiliza o relacionamento existente entre duas ou mais variáveis de maneira que uma delas possa ser descrita ou o seu valor estimado a partir das demais.

Ao observar modelos de regressão nota-se que em alguns casos as variáveis dependentes são de natureza dicotômica. Em situações como estas a variável resposta apresenta apenas dois valores possíveis, como sim ou não, alto ou baixo, bom ou ruim, e a decisão por uma delas dependerá de uma série de atributos relacionados ao evento (as variáveis independentes). A melhor forma de manipular estes dados é a partir de um modelo ajustado que permitirá estimar a probabilidade de ocorrência para cada uma das respostas. A abordagem matemática utilizada para se chegar a este modelo é denominada Regressão Logística [70].

A Regressão Logística vem sendo utilizada nas mais diversas áreas. Este método, assim como a regressão linear e múltipla, estuda a relação entre uma variável resposta e uma ou mais variáveis independentes. A diferença entre estas técnicas de regressão se deve ao fato de que na Regressão Logística as variáveis dependentes estão dispostas em categorias,

enquanto na regressão linear estas variáveis são dados contínuos. Outra diferença é que na Regressão Logística a resposta é expressa por meio de uma probabilidade de ocorrência, enquanto na regressão simples obtém-se um valor numérico [70].

O método de Regressão Logística é aplicável nos casos onde a variável resposta ou dependente é discreta, assumindo dois ou mais possíveis valores (a variável possui um caráter qualitativo). Este método foi originalmente concebido para responder às necessidades das ciências biomédicas e de saúde pública, visando à modelagem de variáveis forçantes de patologias de um modo geral [3].

Em regressão logística existem três procedimentos distintos para manipular dados que são denominados Regressão Logística Binária, Ordinária, e Nominal. A escolha de qual método utilizar depende do número de categorias e das características da variável resposta [70].

### **3.3 Técnicas de avaliação espaço-temporal**

Na área de Sistemas de Informação Geográfica, um dos principais desafios para os próximos anos é transformar os sistemas atuais, essencialmente estáticos, em ferramentas capazes de prover representações realistas de processos espaço-temporais. A modelagem de grande quantidade de processos físicos, em aplicações como Geomorfologia, Estudos Climáticos, Dinâmica Populacional e Impacto Ambiental, requer que os SIGs tenham capacidade de representar os tipos de processos dinâmicos encontrados em estudos de sistemas físicos e sócio-econômicos [20].

Desta forma, além das técnicas mencionadas nos itens anteriores, procura-se destacar duas técnicas de análise de dados que atualmente vem ganhando destaque por sua importância como recursos auxiliares no apoio ao processo de tomada de decisão: a modelagem dinâmica e os autômatos celulares.

#### **3.3.1 Modelagem dinâmica**

Segundo Câmara e Pedrosa [20] o objetivo dos modelos dinâmicos no Sistema de Informação Geográfica é realizar a simulação numérica de processos dependentes do tempo, como nos modelos hidrológicos, que simulam o fluxo e transporte de água. Na definição de Burrough [18], “um modelo espacial dinâmico é uma representação matemática de um

processo do mundo real em que uma localização na superfície terrestre muda em resposta a variações em suas forças direcionadoras”.

Tipicamente, os SIGs são desenvolvidos a partir de suposições pré-estabelecidas quanto a homogeneidade, uniformidade e universalidade das propriedades de seus principais componentes, que incluem o espaço e as relações espaciais, o tempo e o modelo matemático que descreve o fenômeno. Entretanto, para modelar processos dinâmicos em SIG com o nível necessário de realismo, estas suposições rígidas têm que ser flexibilizadas de tal forma que o sistema seja capaz de representar [20]:

- O espaço como uma entidade não homogênea tanto nas suas propriedades quanto na sua estrutura.
- As vizinhanças como relações não estacionárias.
- As regras de transição como regras não universais.
- A variação do tempo como um processo regular ou irregular.
- O sistema como um ambiente aberto a influências externas.
- Retroalimentação.

Para implementar sistemas espaciais dinâmicos com as características mencionadas, alguns princípios básicos relativos aos principais elementos destes sistemas devem ser considerados. Entre estes elementos destacam-se:

- a questão da representação do espaço e do tempo,
- o modelo dinâmico a ser utilizado para a representação do fenômeno espacial
- e a abordagem computacional.

Os modelos espaciais dinâmicos são essenciais, pois descrevem a evolução de padrões espaciais de um sistema ao longo do tempo. Um modelo deve ser capaz de responder a questões que indiquem “Porquê”, “Quando” e “Onde”, e ser capaz de descrever quantitativamente um fenômeno e prever sua evolução, integrando suas escalas temporal e espacial. Os modelos podem ser classificados em empíricos e de sistemas [20].

Como exemplo de modelos empíricos pode-se citar as cadeias de Markov, modelos logísticos de Difusão e Regressão. Estes modelos utilizam equações matemáticas simples e requerem poucos dados, além de serem compatíveis com o formato de dados oriundos de

fontes de sensoriamento remoto e, como consequência, facilmente implementados em um Sistema de Informação Geográfica [20].

Os modelos de sistemas são classificados como modelos exploratórios, porque fornecem condições para que várias simulações possam ser investigadas a partir de diferentes cenários. Entretanto, estes modelos requerem um substancial conhecimento do fenômeno em estudo e acabam por se tornar sistemas altamente especializados, não podendo ser aplicados a outras classes de fenômenos [20].

Atualmente existem ambientes computacionais para modelagem dinâmica para lidar com a natureza complexa de dados e fenômenos espaciais. Um exemplo de um ambiente computacional para modelagem dinâmica é o PCRaster ([www.dpi.inpe.br](http://www.dpi.inpe.br)), que é um sistema computacional que simula a modelagem dinâmica de processos físicos.

### **3.3.2 Autômatos Celulares**

Os Autômatos Celulares são ferramentas de modelamento. É uma forma concreta para trabalhar as habilidades de modelamento, porque incorpora a noção de tempo e espaço. O espaço é visto como uma matriz homogênea dividida da mesma forma que um tabuleiro de xadrez, sendo que uma célula elementar do Autômato Celular pode assumir diferentes estados [95].

Segundo Dilão [32], os Autômatos Celulares são modelos matemáticos discretos, facilmente implementáveis em computador, e com aplicações na simulação em tempo real de sistemas físicos, biológicos e sociais, com um grande número de componentes em interação.

Os Autômatos Celulares são simulações computacionais que tentam emular a suposta maneira pela qual as leis da natureza funcionam. É um sistema dinâmico, onde o tempo e o espaço são discretos. O sistema é dividido em células, seus elementos básicos. Tais células possuem um conjunto finito de estados predefinidos e um conjunto de condições necessárias para a mudança de estados. Os estados das células são alterados conforme um conjunto de regras de transição. Tais regras de transição são baseadas no estado atual da célula e de suas vizinhas [55].

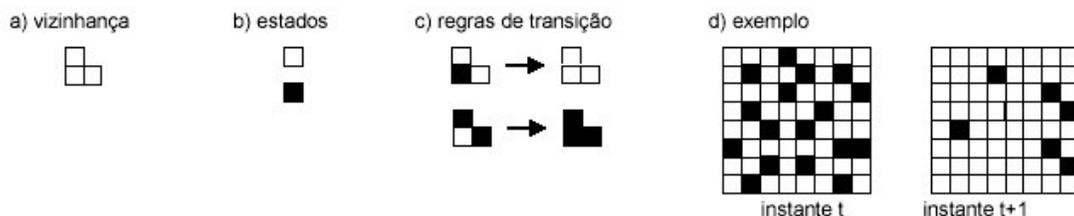
Um Autômato Celular consiste de uma grade uniforme e regular, onde cada célula pode assumir um estado discreto. Autômatos Celulares evoluem em passos de tempo discretos, em que as células são atualizadas, simultaneamente, com base no estado das células em sua vizinhança, no passo de tempo anterior, e de acordo com um conjunto de regras locais.

A vizinhança de uma célula é definida pela célula em si e todas as células imediatamente adjacentes [69].

Um exemplo que ajuda a identificar e a ilustrar como se dá o funcionamento dos autômatos celulares é apresentado num exemplo simples baseado em Câmara [23]. Neste exemplo, uma célula pode assumir dois estados (branco e preto) e sua vizinhança é definida sobre duas células adjacentes. As regras de transição especificam que:

- o estado de uma célula num instante  $t+1$  é igual ao dos seus vizinhos no instante  $t$ , se estes vizinhos tiverem os estados iguais;
- caso contrário, o estado da célula permanece o mesmo.

Pelo exemplo pode-se identificar os componentes básicos de um Autômato Celular: o espaço dividido em uma grade de células, a vizinhança, o conjunto de estados, as regras de transição e os intervalos de tempo com atualização simultânea das células. [69]



**Figura 3.5** Autômato celular  
Fonte: adaptado de [23]

Segundo Leite et. al. [55], para um sistema de Autômatos Celulares, é necessário que se tenha uma grade com geometria regular. É necessário também especificar quais serão os vizinhos de uma célula, pois as regras de transição serão aplicadas baseadas nos estados deles e da própria célula. Há algumas formas de vizinhança que seguem uma lógica ou regra e há formas que são arbitrárias. Todo autômato celular deve ter um conjunto de estados, que serão os possíveis estados em que as células poderão se encontrar. Por exemplo, o conjunto de estados pode ser um conjunto de cores (preto e branco), uma certa propriedade (ligado e desligado), etc, onde cada célula poderá encontrar-se em apenas um estado. Por fim, para implementação de um sistema Autômato Celular, a mais importante especificação do sistema são as regras de transição, pois suas especificações farão com que o sistema evolua.

O applet é uma simulação de um sistema Autômato Celular, bidimensional, com as células de formato quadrado. Um dos applet mais conhecidos é o famoso “jogo da vida” inventado por John Conway em 1970, com o objetivo explorar como regras simples poderiam dar surgimento a estruturas que simulavam a vida e comportamento. Ele acreditava que "organismos" no seu mundo artificial podiam atuar como aqueles do mundo real – movendo, crescendo, reproduzindo, evoluindo e, talvez, até mesmo pensando [55].

As regras do jogo inventado por John Conway são: se uma célula está OFF, ela fica ON se exatamente 3 dos seus vizinhos estão ON; se uma célula está ON, ela continua ON se exatamente 2 ou 3 de seus vizinhos estão ON; caso contrário, ela fica OFF.

O Autômato Celular pode ser utilizado para explorar como elementos de um sistema interagem uns com os outros, o que revela uma importante característica de que os conceitos usados na construção do autômato podem ser usados para modelar comportamentos complexos do mundo natural (químico, físico e biológico) além de fenômenos sociais (crescimento populacional, população x alimento, planejamento urbano, entre outros) [95].

Os modelos celulares têm sido utilizados em Sistemas de Informações Geográficas envolvendo diversas aplicações de meio ambiente. As motivações para o uso deste modelo de dados são influenciadas pela natureza “pixelizada” dos dados de sensoriamento remoto e pela conveniência da programação e implementação de estruturas baseadas em grades [69]. Alguns exemplos de áreas onde se utilizam os autômatos celulares são: modelamento de fluxo de trânsito veicular ou de pedestres, modelamento de fluidos (gases ou líquidos), modelamento da evolução de células ou vírus como o HIV ou o modelamento de processos de predição.

Os Autômatos Celulares podem ser utilizados para modelar numerosos sistemas físicos que se caracterizam por um grande número de componentes homogêneos e que interagem localmente entre si. Qualquer sistema real ao qual se possa analogar os conceitos de vizinhança, estados dos componentes e função de transição, é candidato a ser modelado por um Autômato Celular.

### **3.4 Considerações finais**

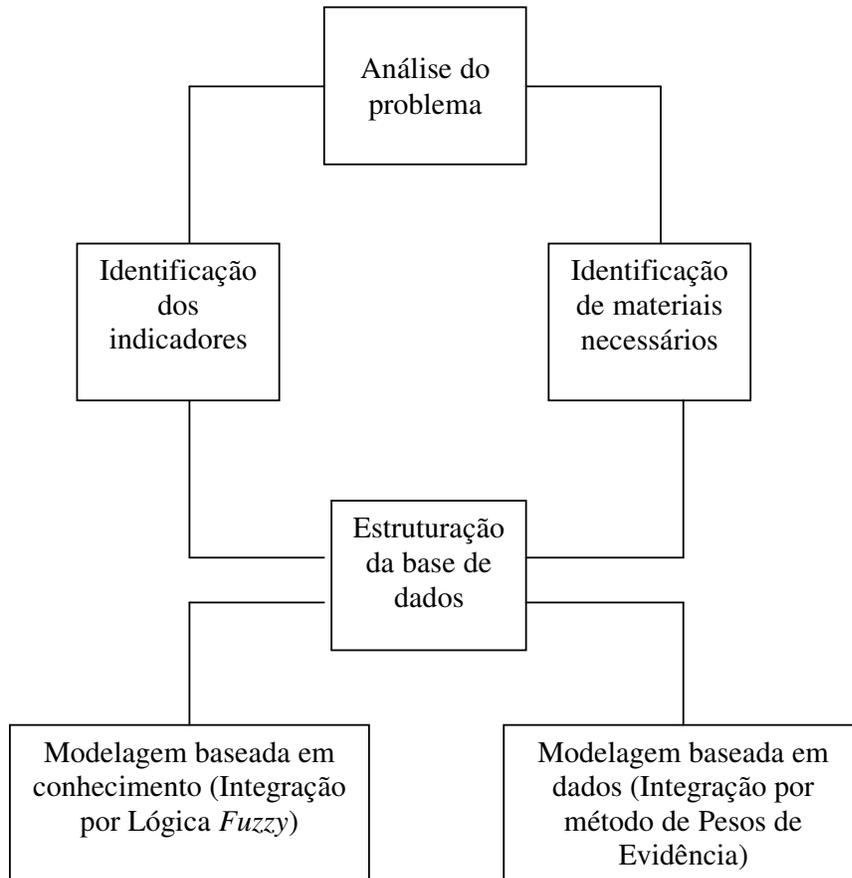
Neste capítulo foram apresentados conceitos de integração temática de dados e algumas técnicas de análise espacial para integração temática baseada em dados e baseada em conhecimento. No próximo capítulo será discutida a metodologia proposta para a utilização destas técnicas em uma área de estudo e serão apresentados os passos a serem seguidos para avaliar a erosão em bacias hidrográficas.

## CAPÍTULO 4 – METODOLOGIA PROPOSTA

Neste capítulo será apresentada a metodologia proposta para efetuar a avaliação do processo de erosão em uma área de estudo utilizando as técnicas de Lógica “*Fuzzy*” e Pesos de Evidência para o auxílio no processo de suporte à decisão.

### 4.1 Fluxograma metodológico

A metodologia foi dividida em etapas representadas no fluxograma metodológico apresentado na figura 4.1.



**Figura 4.1** Fluxograma metodológico

## **4.2 Descrição das etapas da metodologia**

### **Análise do problema**

A primeira etapa é a análise do problema, onde se busca entender e formular questionamentos que direcionem a análise voltada para o processo de erosão. Desta forma podem surgir questões como: o que causa ou o que está relacionado ao processo erosivo em uma área? O estudo do processo de erosão é um problema que envolve múltiplas variáveis/indicadores? Quais? Mais conhecimento sobre o comportamento do processo de erosão pode ser obtido através de outros indicadores, dados ou experimentos?

Nesta etapa deve-se analisar e relacionar os problemas que influenciam o processo erosivo, como por exemplo, considerar que a ação do homem, que remove a cobertura vegetal e deixa o solo exposto, está relacionado diretamente ao processo erosivo, que tem se tornado muito severo nos últimos tempos, constituindo um problema para algumas atividades econômicas. Além da ação do homem deve-se relacionar as características ambientais que contribuem para o processo erosivo e que torna difícil o processo de localização e tratamento dessas áreas.

Costa [30] ressalta a importância de se considerar essas variáveis ambientais para conhecimento do processo erosivo, em especial de variáveis como solo, vegetação, clima, etc, relacionando as características dessas variáveis com a identificação da erosão.

Dessa forma, a partir dos problemas relacionados são estabelecidas algumas premissas, estabelecendo como aspectos ambientais e antrópicos influenciam o processo erosivo. Em seguida, essas premissas irão ajudar a formular uma hipótese orientativa para nortear a análise e auxiliar as operações de identificação e predição de novos casos de erosão, a partir do conhecimento desses aspectos e de suas dinâmicas. Esse conhecimento corresponde a um dado essencial para a compreensão do processo erosivo e para o dimensionamento do impacto que ele causa em nosso ambiente, desenvolvendo sistemas e métodos para o monitoramento e análise da degradação ambiental.

### **Identificação de variáveis associadas ao processo de erosão**

A partir das considerações levantadas no item anterior podem ser identificadas algumas variáveis ambientais (que serão utilizadas como indicadores), e gerar modelagens que possam auxiliar na identificação de locais onde ocorrem processos erosivos na área de estudo. O estudo para observação de potenciais erosivos e a elaboração de um guia para

utilização de ferramentas de predição usadas neste trabalho têm o objetivo de contribuir para minimizar este problema.

Os critérios utilizados para a escolha dos indicadores devem levar em consideração os impactos que os indicadores avaliados teriam no processo erosivo de uma bacia hidrográfica considerando quaisquer fatores, positivos e negativos, que possam aumentar ou diminuir a possibilidade de ocorrência do processo erosivo.

Os indicadores devem englobar parâmetros passíveis de permitir a utilização das técnicas utilizadas na análise do problema. Os indicadores devem ainda ser simples, para que os participantes do processo decisório possam compreender seu significado e, ao mesmo tempo, coerentes, pertinentes e consistentes, de forma a fazer com que todos os aspectos relevantes sejam avaliados.

### **Identificação de materiais necessários**

Nesta etapa de metodologia busca-se identificar e reunir os materiais e dados para o desenvolvimento do trabalho. Para a avaliação do potencial erosivo em uma região podem ser utilizados dados como cartas topográficas, imagens de sensores remotos, fotos aéreas e mapas de reconhecimento que auxiliam a identificar os pontos mais críticos na área. Além destes dados é necessário reunir os dados cartográficos que representem os indicadores ambientais, como mapas temáticos de solo, conflito de uso, vegetação, que serão utilizados na análise.

Nesta etapa também devem ser apresentados os sistemas computacionais que serão utilizados para a implementação da base de dados cartográfica e os sistemas computacionais de processamento de imagens, se for o caso. Ainda são apresentados os demais equipamentos que podem ser utilizados na execução do trabalho, como GPS, estação total, entre outros.

### **Estruturação da base de dados**

Nesta etapa deve-se levar em consideração o tratamento dos dados, no que se refere à preparação dos mesmos para iniciar a análise. A base de dados em formato digital, contendo os diversos planos de informação devidamente georeferenciados, é um componente essencial e indispensável em qualquer projeto que envolva análise espacial em ambiente SIG. Isso permite que todas as informações disponíveis sobre um determinado assunto possam ser facilmente acessadas, que se estabeleçam todas as possíveis correlações sobre temas, e que se façam diversos tipos de operações matemáticas sobre os dados, desde as mais simples até as

mais sofisticadas, dependendo do que se deseje e do sistema escolhido [9]. Dessa forma, deve-se fazer um prévio tratamento dos dados, para que os mesmos sejam compatíveis com os sistemas computacionais e com as técnicas de integração utilizadas na análise.

A análise espacial baseada em conhecimento é uma análise que depende do conhecimento de um especialista que definirá as classes dos dados que têm maior influência em um determinado evento. Portanto, devem ser incluídos como parte do banco de dados as informações e pesos estabelecidos pelos especialistas. O conhecimento destes, principalmente no processo de integração baseada em conhecimento, refletirá a confiabilidade e qualidade dos resultados da análise.

### **Integração por método de Lógica “Fuzzy”**

Neste item será abordada a metodologia de modelamento da técnica de Lógica “Fuzzy” utilizada na geração de mapa para estimar áreas que possuem potencial à erosão de uma bacia hidrográfica a partir da integração de dados.

O objetivo deste modelamento é fazer uma avaliação através do mapeamento de indicadores ambientais de forma que as características dos indicadores utilizados possam ser tratadas como um modelo descritivo que poderá guiar na busca por novas ocorrências de erosão em outras áreas de estudo.

A Lógica “Fuzzy” é um método que leva em conta o grau de pertinência das variáveis envolvidas em um processo decisório ao invés de considerar valores exatos como a Lógica *Booleana*. Segundo Zadeh [97], esta é uma técnica que procura modelar a imprecisão do raciocínio possibilitando a dedução de uma resposta aproximada a uma questão baseada em um conhecimento incompleto ou impreciso.

Os sistemas de inferências baseados em Lógica “Fuzzy” possibilitam a captura do conhecimento próximo ao “modelo cognitivo” utilizado na análise de problemas. Isto significa que o processo de aquisição do conhecimento é mais fácil, mais confiável e menos sujeito a erros não identificados [76].

A integração de dados pelo método de Lógica “Fuzzy” pode ser dividida nas seguintes etapas:

### **A- Definição do conjunto de parâmetros de análise**

Nesta etapa são implementadas as bases cartográficas no ambiente SIG, através da utilização de mapas temáticos (indicadores) para a construção do modelo “*Fuzzy*”. Ainda nesta etapa é identificada a abrangência da área onde será feito o estudo de caso e a unidade de área a ser utilizada. As classes com dados ausentes (missing data) também deverão ser identificadas.

### **B- Geração de membros “*Fuzzy*”**

Os indicadores devem ter sua importância representada através dos membros “*Fuzzy*” expressando o grau de aceitação da evidência no modelo proposto. Nesta etapa são utilizados os pesos criados por especialistas ou utilizadas funções de pertinência para a geração dos membros “*Fuzzy*” de cada indicador.

Na metodologia “*Fuzzy*” é necessário que todos os indicadores estejam no formato numérico, necessitando que os mapas temáticos sejam ponderados para valores que variam entre 0 e 1 (membros “*Fuzzy*”).

### **C- Geração dos planos de informação “*Fuzzy*”**

Com base nos membros “*Fuzzy*” gerados na etapa anterior são criados nesta etapa os planos de informação “*Fuzzy*”, também chamados mapas de pertinência. Os planos de informação “*Fuzzy*” são mapas de probabilidade que representam a importância de cada classe dos mapas temáticos utilizados na análise.

### **D- Escolha do operador “*Fuzzy*”**

O próximo passo é fazer a escolha do operador “*Fuzzy*” que será utilizado na integração dos planos de informação “*Fuzzy*”. Os operadores “*Fuzzy*” (“and”, “or”, “sum”, “product” ou “gamma”) poderão gerar diferentes cenários de potencialidade à ocorrência de erosão.

### **E- Definição do grau para o operador gamma**

Esta etapa somente é realizada se na etapa anterior for escolhido o operador gamma para realizar a integração dos planos de informação “*Fuzzy*”. A fim de se alcançar diferentes cenários para a análise do processo erosivo são identificados diferentes valores para o grau utilizado pelo operador gamma.

### F- Integração dos planos de informação “Fuzzy”

Nesta etapa é feita a integração dos planos de informação “Fuzzy” através dos operadores “Fuzzy” escolhidos anteriormente. O resultado da integração dos planos de informação “Fuzzy” é um mapa de probabilidade representando o potencial de erosão na área analisada.

### G- Geração do mapa final de pertinência “Fuzzy”

Nesta etapa é gerado o mapa final de probabilidade feito após a avaliação dos parâmetros utilizados na análise. Caso sejam identificados novos parâmetros que melhorem o resultado, a análise é refeita para a geração deste mapa.

O processo de análise dos pontos de erosão através do método de Lógica “Fuzzy” pode ser resumido através do seguinte modelamento:

ETAPA	OBJETIVOS
Definição do conjunto de parâmetros de análise	Construção da base de dados espacial para análise
Geração de membros “Fuzzy”	Geração dos membros “Fuzzy” a partir da utilização de pesos criados por especialistas
Geração de planos de informação “Fuzzy”	Geração dos mapas de pertinência para cada indicador utilizado na análise
Escolha dos operadores “Fuzzy”	Identificar o operador “Fuzzy” a ser utilizado na análise
Definição do grau para o operador gamma	Identificação de diferentes valores para o grau do operador gamma gerando diferentes cenários para a análise
Integração dos planos de informação “Fuzzy”	Combinação dos planos de informação “Fuzzy” com os operadores
Geração do mapa final de pertinência “Fuzzy”	Criação do mapa final de potencial após a avaliação dos parâmetros utilizados

**Tabela 4.1** Síntese do modelamento “Fuzzy”

## **Integração por método de Pesos de Evidência**

Neste item será abordada a metodologia de modelamento da técnica estatística de inferência espacial chamada Pesos de Evidência para a geração de um mapa de favorabilidade de pontos de erosão em uma bacia hidrográfica a partir da integração de dados.

O objetivo deste modelamento é avaliar individualmente, com base na relação espacial entre as classes de cada indicador e o conjunto de pontos identificados em campo como voçorocas, o potencial de predição à localização de novos pontos de erosão, calculando os pesos e contrastes de cada uma das variáveis através do método.

Como descrito na fundamentação teórica no capítulo três o método de Pesos de Evidência é um método quantitativo para combinar evidências em suporte de uma hipótese [53], citado por [40]. A análise de Pesos de Evidência constitui um método de sobreposição e inspeção de diversos mapas temáticos, ou seja, é uma técnica que envolve processos de correlação e integração entre mapas com suporte de formulações matemáticas de probabilidades, objetivando a identificação de áreas onde um determinado fenômeno ou evento possa estar presente [40].

No método de pesos de Evidências as evidências foram combinadas a partir de alguns mapas de predição, usando a formulação da regra de Bayes. O modelamento por Pesos de Evidências inicia-se com o cálculo da probabilidade a priori de ocorrência de um determinado evento por unidade de área baseada nos pontos de treinamento e na área total a ser estudada. A análise ainda envolve algumas etapas como, por exemplo, a reclassificação de mapas de múltiplas classes para generalizar mapas com apenas duas classes, cálculo da probabilidade posteriori e geração do mapa de favorabilidade. As etapas para a integração de dados pelo método de Pesos de Evidência são detalhadas abaixo com base em Araújo [9], Bonham-Carter [14] e Câmara [21].

### **A - Definição do conjunto de parâmetros de análise**

Nesta etapa são definidos os parâmetros que irão nortear a análise por Pesos de Evidência, sendo necessário identificar a abrangência da área onde será feito o estudo de caso e a unidade de área a ser utilizada.

Deverá ser indicado ainda um conjunto de dados pontuais que servirá para definir ocorrências conhecidas do evento analisado. Esse conjunto de dados é chamado de pontos de treinamento. As classes com dados ausentes (missing data) também deverão ser identificadas.

### **B – Cálculo dos pesos para cada mapa de predição**

Nesta etapa é feito o cálculo estatístico da medida da associação espacial entre as classes dos diversos mapas, considerados como condicionantes do evento de erosão e posteriormente efetuar a conversão dessas medidas em pesos. O objetivo dessa etapa é proceder ao cálculo da medida da associação espacial entre as classes de cada mapa de evidência e um conjunto de pontos de treinamento, que podem ser representados pelas voçorocas em um estudo do processo erosivo. Os pontos de treinamento são eventos previamente conhecidos e são usados como indicadores da probabilidade a priori [9].

### **C – Generalização dos temas de evidência**

A modelagem de Pesos de Evidência pode ser executada tanto em mapas com classes múltiplas, quanto em mapas binários, mas segundo Bonham-Carter [14], para a análise sobre mapas com múltiplas classes as estimativas tornam-se instáveis e suas variâncias apresentam-se elevadas, de forma que a melhor solução para esse problema é transformar os mapas multi-classes em mapas binários onde a evidência está presente ou ausente. A reclassificação de mapas de predição multi-classes em mapas binários é chamada generalização dos temas de evidência. Ainda segundo Bonham-Carter [14], a utilização de mapas binários na modelagem, além de conferir maior robustez estatística às estimativas, torna mais fácil a interpretação dos resultados. O limite para reclassificação, tanto pode ser escolhido com base no julgamento de especialistas (subjetivamente) quanto estatisticamente.

### **D – Combinação dos temas de evidência**

A partir dos mapas binários gerados na etapa anterior é realizada uma tabulação cruzada das evidências com os pontos de treinamento (ocorrências conhecidas ou verdades de campo). Através desse procedimento é gerada uma matriz de dados onde cada pixel corresponde à interseção das classes de evidências com os pontos de treinamento. Esses valores de correlação são utilizados para a obtenção dos valores de probabilidade posteriori.

### **E – Avaliação das incertezas e geração do mapa de favorabilidade**

Nesta etapa é realizada uma avaliação das incertezas inerentes à modelagem, que normalmente estão associadas ao processo de cálculo de pesos de evidência, bem como devido à ausência de dados em algum local. São avaliados os pesos, variâncias, contrastes e a confiança para cada tema de evidência. Após a avaliação das incertezas é feita a geração do

mapa de favorabilidade a partir dos valores de probabilidade posteriori para cada única condição do tema de resposta da análise.

### **F – Verificação da independência condicional**

A modelagem por Pesos de Evidência pressupõe que exista independência condicional entre os dados, contudo é muito difícil na prática conseguir um ajuste que implique na independência total, de forma que é necessária a aplicação de testes estatísticos para verificar em que grau ocorreu à violação desse pressuposto importante. Para isso, geralmente são utilizados os testes pareado e global para a avaliação do grau de violação da independência condicional envolvendo os mapas modelados. Dependendo do resultado dessa análise, devem ser efetuados os devidos ajustes, combinando ou eliminando mapas, para que esse pressuposto importante do método seja respeitado.

A análise pelo método de Pesos de Evidência pode ser resumido pelo modelamento:

ETAPA	OBJETIVOS
Definição do conjunto de parâmetros de análise	Indicar o dado que representará a área de estudo, definir a unidade de área a ser utilizada, definir os valores que serão utilizados para a classe de dados ausentes, definir os dados usados com pontos de treinamento
Calculo dos pesos para cada mapa de predição (W+, W-, C)	Calcular estatisticamente a medida de associação espacial entre as classes dos indicadores e os pontos de treinamento
Generalização dos temas de evidências	Reclassificação dos temas em mapas binários
Combinação dos temas de evidência	Combinar os temas de evidencia binários que foram reclassificados na etapa anterior
Avaliação das incertezas e geração do mapa de favorabilidade	Avaliação das incertezas para a produção do mapa de favorabilidade (mapa potencial)
Verificação da independência condicional	Avaliar se há independência condicional entre os dados utilizados

**Tabela 4.2** Síntese do Modelamento por Pesos de Evidência

## **Considerações finais**

Neste capítulo foram detalhados todos os procedimentos da metodologia proposta para o modelamento “*Fuzzy*” e para o modelamento por Pesos de Evidência para a análise de erosão em uma área de estudos. Foram descritos todos os passos envolvidos na aplicação destes procedimentos, desde o cuidado com a base de dados espacial até a obtenção do resultado final, passando pelas diferentes etapas de análise.

Objetivando proporcionar uma visão mais aprofundada destes procedimentos, será apresentado no próximo capítulo um estudo de caso descrevendo em detalhe todos os passos da aplicação desta metodologia utilizando um sistema computacional para a análise espacial.

## **CAPÍTULO 5 - ESTUDO DE CASO**

Neste capítulo é apresentada a aplicação da metodologia proposta no capítulo anterior. Na primeira seção deste capítulo são apresentadas a localização e algumas características da área de estudos: a Bacia do Alto Taquari. Posteriormente, é apresentado o problema que se quer analisar na bacia, levantando-se a hipótese orientativa que auxiliará no desenvolvimento do trabalho.

Em seguida são feitos levantamentos sobre os indicadores ambientais e seu tratamento para a análise, bem como uma breve descrição dos materiais necessários, incluindo a descrição da ferramenta de análise espacial para este estudo de caso.

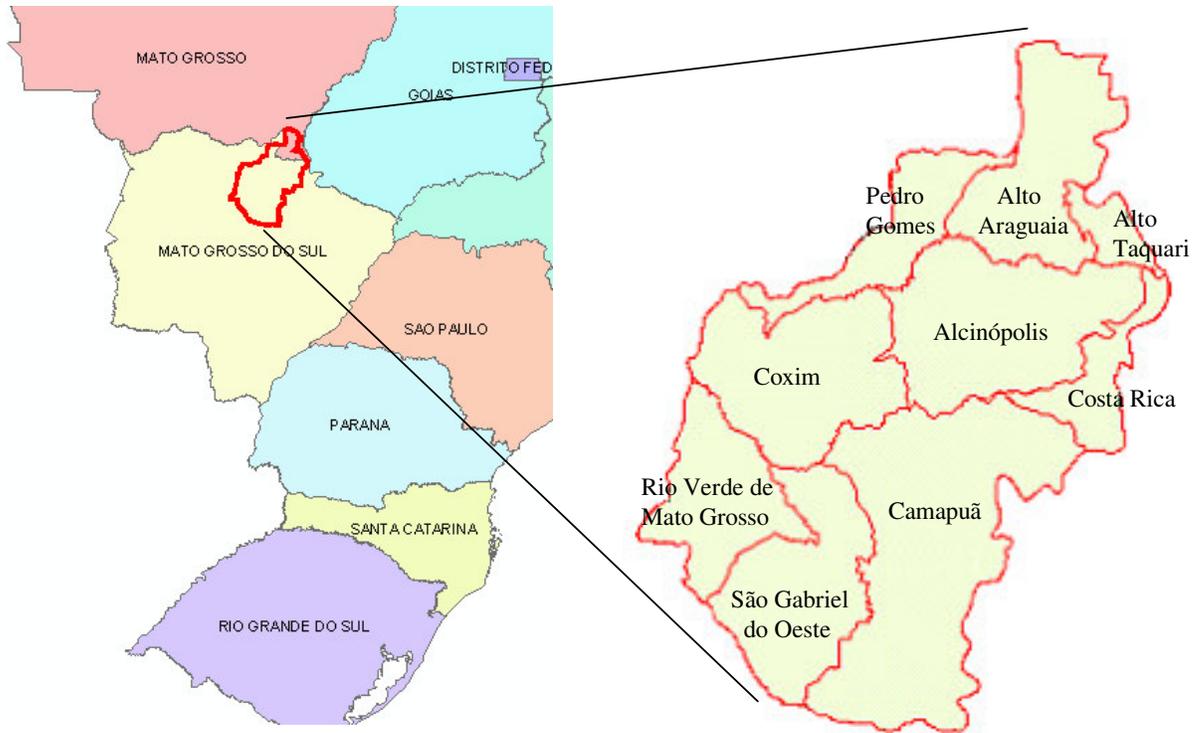
Por fim é apresentada a aplicação da integração de dados utilizando um método baseado no conhecimento (modelo de Lógica “*Fuzzy*”) e um método baseado nos dados (modelo de Pesos de Evidência), sendo apresentados também os resultados obtidos.

### **5.1 Área de estudo**

O estudo foi realizado na sub-bacia hidrográfica do Alto Taquari (BAT), pertencente à Bacia do Rio Paraguai. Segundo Silva [84], a Bacia do Alto Taquari está localizada em uma área adjacente ao Pantanal, no Estado do Mato Grosso do Sul e pequena parte no Estado do Mato Grosso, com uma área total de aproximadamente 28.046 Km<sup>2</sup>. A BAT compreende a área do planalto drenada pelo rio Taquari e seus afluentes.

O rio Taquari nasce no Estado do Mato Grosso (MT), no município do Alto Taquari, a uma altitude de 860 metros, percorrendo, aproximadamente, 40 Km antes de chegar ao Estado do Mato Grosso do Sul (MS). Em seu alto e médio curso percorre os municípios de Alcinópolis, Pedro Gomes e Coxim, para em seguida, no baixo curso, adentrar na região do Pantanal, no município de Corumbá [84].

Ao nível de delimitação política, a parte da BAT a ser estudada neste trabalho compreende os municípios acima mencionados e os municípios Alto Taquari, Costa Rica, Rio Verde do Mato Grosso, Camapuã e São Gabriel do Oeste, todos com fortes evidências de degradação ambiental e de reflexos nocivos ao Pantanal [10].



**Figura 5.1** Localização da BAT

Os critérios que foram utilizados para a escolha da área levaram em consideração o alto potencial erosivo da BAT, devido ao grande desmatamento e mudança do uso do solo para introdução de atividades agropecuárias, pelos tipos predominantes de solo na bacia, que são vulneráveis à erosão, pelo tipo de relevo e pela existência e disponibilidade de bases cartográficas e temáticas, geradas por especialistas que trabalharam na área através do Plano de Conservação da Bacia do Alto Paraguai – PCBAP [75] e do projeto “SOS Taquari”, coordenado pela Embrapa Solos ([www.cnps.embrapa.br](http://www.cnps.embrapa.br)).

## 5.2 Análise do problema

O rio Taquari constitui um importante afluente do rio Paraguai, exercendo um papel importante no ecossistema do Pantanal. A bacia do rio Taquari possui um potencial de erosão naturalmente alto. Há uma série de fatores como solo arenoso, topografia acidentada e fortes chuvas de verão que fazem com que o rio Taquari carregue sedimentos em sua descida do planalto para o Pantanal, o que leva a um processo natural de assoreamento da região.

O processo de assoreamento, que está intimamente associado aos processos erosivos que ocorrem na região do Alto Taquari, termina por levar a um grande aporte de sedimentos à

região baixa do rio, fazendo com que sejam verificados muitos impactos negativos devido ao assoreamento extremo dos rios, com graves prejuízos para a pesca, navegabilidade, com a formação de enchentes, redução da área disponível para pecuária, e com impactos sobre a vida animal e vegetal [51].

Além das características ambientais, nas últimas décadas, a ação do homem acelerou o processo de erosão que deveria acontecer naturalmente. Sem orientação para trabalhar a terra de acordo com as características da região, houve a retirada da vegetação nativa impulsionada pela expansão da agricultura e pela criação de áreas de pasto através do desmatamento. Como consequência houve elevação da temperatura, devido à incidência direta do sol sobre o solo, que, aliada à umidade, provoca a decomposição da matéria orgânica do solo. Conseqüentemente, o solo tem sua estrutura desestabilizada e é facilmente levado pelas enxurradas. Esse processo é agravado pelo aumento na quantidade de chuvas na região [27].

Desta forma, foi possível verificar os problemas relacionados com o processo erosivo na Bacia do Alto Taquari. Nesse sentido, foram estabelecidas duas premissas. A primeira diz respeito às variáveis ambientais que estão relacionadas diretamente ao processo de erosão, como tipo de solo, conflito de uso do solo, erosividade da chuva, entre outras. A segunda premissa assume a intervenção humana como grande responsável pelo agravamento do processo erosivo nas bacias hidrográficas.

Considerando-se as premissas fixadas, a hipótese orientativa do desenvolvimento deste trabalho é que a partir do conhecimento sobre fatores ambientais e antrópicos que influenciam o processo erosivo e a identificação de pontos conhecidos de erosão avançada (voçorocas) em uma bacia hidrográfica pode-se, com o auxílio de um Sistema de Informação Geográfica, estimar a localização de pontos de erosão a partir do método de Lógica “*Fuzzy*”, bem como estimar a probabilidade de novos pontos de erosão (predição) a partir do método de Pesos de Evidência.

### **5.3 Escolha dos indicadores**

Nesta seção são apresentados os indicadores ambientais escolhidos para a aplicação da metodologia proposta no estudo de caso. A escolha dos indicadores para a análise do processo de erosão na BAT foi definida a partir da avaliação de quais representariam melhor o processo erosivo. A avaliação foi feita a partir de estudos e consultas à bibliografia existentes referentes ao processo erosivo em bacias hidrográficas.

A escolha de indicadores ambientais para avaliação do processo erosivo na BAT ainda envolveu a realização de análise das informações referentes às relações existentes entre os processos naturais e as atividades antrópicas da região.

Para a análise do processo erosivo na BAT foi avaliada a possibilidade de se utilizar os fatores integrantes da Equação Universal de Perda de Solo (*Universal Soil Loss Equation - USLE*) [94] para estimar o valor de perda de solo na BAT. Porém os valores quantitativos de perda de solo gerados pela USLE são considerados como estimativas para fins comparativos, fornecendo uma análise qualitativa da distribuição espacial da erosão, como pode ser observado no trabalho de Galdino et. al. [36], que faz uma estimativa de perda de solo para a BAT. Além deste fato, a equação de perda de solo é um processo multiplicativo, se um dos fatores integrantes da equação for impreciso pode ser observado uma propagação de erro que influenciará na estimativa do valor final de perda do solo.

A USLE geralmente é apropriada para estudos realizados em escalas grandes, onde se possa medir o comprimento da rampa (fator topográfico). Como a BAT apresenta uma extensão muito grande, impossibilitando medir o fator topográfico (comprimento de rampa) de maneira precisa, em vez de utilizar a equação universal de perdas de solo (que não é eficiente para escalas pequenas, ou seja, estudos regionais) procurou-se considerar a integração temática de dados baseada em técnicas avançadas de inferência geográfica, utilizando-se mapas temáticos para a avaliação e predição do processo erosivo.

Nesse sentido, para a realização da integração temática de dados foram escolhidos somente quatro indicadores para representar o processo erosivo, acreditando-se que os quatro indicadores são representativos do processo de erosão na área de estudos. Os indicadores escolhidos foram: uso do solo, solo, erosividade da chuva e conflito de uso.

Para efetuar a análise do processo de erosão na BAT optou-se ainda por trabalhar com os indicadores utilizando a unidade territorial microbacia, considerada uma unidade homogênea. Essa aproximação se deve ao fato de alguns indicadores serem calculados de acordo com a sua preponderância, ou seja, o valor representativo da microbacia é o valor da sua classe preponderante [62].

#### **5.4 Identificação de materiais necessários**

O principal objetivo nesta seção é fazer um levantamento dos dados básicos, equipamentos e sistemas computacionais que auxiliarão a diagnosticar o processo erosivo na área de estudo, fornecendo subsídios aos modelos de inferência espaciais que serão utilizados

na etapa posterior. Nesse sentido, para a realização do estudo de caso foram utilizados mapas digitais, dados cartográficos e sistemas computacionais de análise espacial conforme descritos abaixo.

#### **5.4.1 Mapas digitais e dados cartográficos**

O estudo da BAT utilizou como base as informações na escala de 1:250.000, que compõem o Plano de Conservação da Bacia do Alto Paraguai – PCBAP [75] geradas entre 1995 e 1996, complementadas por outras informações, como o mapeamento realizado por Silva [84] de algumas informações para a Bacia do Alto Taquari no ano de 2000, gerando como produtos básicos mapas temáticos referentes aos indicadores utilizados.

O Plano de Conservação da Bacia do Alto Paraguai foi uma iniciativa das Secretarias de Estado do Meio Ambiente dos estados de Mato Grosso e Mato Grosso do Sul, com a finalidade de se elaborar um trabalho técnico-científico no contexto sócio ambiental, para se definir diretrizes para a conservação ambiental da bacia. A motivação para o trabalho foi a identificação dos vários problemas ambientais na bacia, com destaque para erosões, assoreamento, pesca predatória, caça predatória, garimpos clandestinos, desmatamentos irregulares, entre outros [75].

Um dos objetivos específicos no estudo de Silva [84] em 2000 foi a identificação, o mapeamento e a organização dos indicadores ambientais úteis ao planejamento ambiental tendo como estudo de caso a BAT. Segundo Abdon [1], a complementação dos dados do PCBAP com o mapeamento de Silva [84] justifica-se em função das rápidas alterações que a BAT sofre em função de sua ocupação desordenada e por fornecer dados mais atuais, passíveis de serem comparados com os impactos ambientais identificados nesta região.

Foram realizados ainda trabalhos de campo, no âmbito do projeto “SOS Taquari”, ([www.cnps.embrapa.br](http://www.cnps.embrapa.br)) para complementar as informações mencionadas acima, quando foram tomadas coordenadas geográficas através de receptores GPS - Global Position System, localizando as voçorocas na área de estudos, servindo para consolidar a base de dados utilizada na análise.

#### **5.4.2 Sistemas computacionais**

Atualmente existem muitas ferramentas de análise espacial que permitem extrair informação para o planejamento, avaliação, monitoramento e gestão do meio ambiente. Para a realização da análise feita neste trabalho optou-se por utilizar a ferramenta ArcSDM - *Spatial*

*Data Modeller for ArcGIS* ([www.ige.unicamp.br/sdm/default\\_e.htm](http://www.ige.unicamp.br/sdm/default_e.htm)) tendo em vista que esta ferramenta possui tanto técnicas voltadas para a análise exploratória de dados dirigidas pelo conhecimento quanto as dirigidas pelos dados, além de ser de fácil utilização e de ser uma ferramenta disponibilizada gratuitamente na internet.

Os responsáveis pela implementação e desenvolvimento do ArcSDM são os pesquisadores Bonham-Carter e Laura Kemp, do Serviço Geológico do Canadá - GSC ([http://gsc.nrcan.gc.ca/index\\_e.php](http://gsc.nrcan.gc.ca/index_e.php)), em conjunto com Gary Raines, do Serviço Geológico dos Estados Unidos - USGS ([www.usgs.gov](http://www.usgs.gov)), sob patrocínio de serviços geológicos e de algumas empresas de mineração [87].

O ArcSDM é uma extensão gratuita do *software* de SIG *ArcView* ([www.esri.com](http://www.esri.com)), criado para servir como uma ferramenta para avaliação de potencial e previsão de fenômenos espaciais. Esse sistema computacional foi originalmente desenvolvido para ser utilizado juntamente com o *software* de SIG *ArcView* e a sua extensão de análise espacial *Spatial Analyst*. O *ArcView* funciona como um ambiente de espacialização, possibilitando a visualização, edição, classificação e a integração da base de dados geográficos. O *Spatial Analyst* é responsável por fazer algumas análises espaciais nos dados antes dos mesmos serem utilizados, como por exemplo, a transformação de todos os dados para o formato *raster* antes de serem utilizados no ArcSDM.

O ArcSDM fornece ferramentas para aplicação de vários métodos de análise espacial na avaliação de favorabilidade de ocorrência de um fenômeno e suas aplicações dividem-se na abordagem exploratória dos dados dirigida pelo conhecimento (“knowledge-driven”) e na abordagem dirigida pelos dados e estatística (“data-driven”). As funções de análise que compreendem essas duas abordagens são:

- Pesos de Evidência (*WofE - Weights of Evidence*);
- Regressão Logística (*LR logistic regression*);
- Lógica Difusa (“*Fuzzy*” *Logic*) e
- Redes Neurais (*Neural Network*)

Todas as funções de análise do ArcSDM pressupõem, além de uma grande base de dados, a existência de um bom controle de qualidade sobre a base de dados, o que se relaciona diretamente com a qualidade dos produtos finais. O ArcSDM atualmente encontra-se na versão 3.1 para o SIG *ArcView* 9.1, que está disponibilizada gratuitamente em: [www.ige.unicamp.br/sdm/default\\_e.htm](http://www.ige.unicamp.br/sdm/default_e.htm)

## 5.5 Estruturação da base de dados

Toda a base de informação utilizada na análise da BAT foi armazenada e analisada no SIG *ArcView*, constituindo-se no banco de dados geográficos da bacia hidrográfica do Alto Taquari.

O banco de dados geográficos utilizado na análise foi composto por informações georreferenciadas e espacializadas sobre uso do solo, erosividade da chuva, solo e conflito de uso. Além dos mapas temáticos, o banco de dados ainda contou com pontos amostrais do processo erosivo em estágio avançado (voçorocas), obtidos a partir de trabalhos de campo, e com as informações fornecidas por especialistas sobre os dados utilizados na análise.

A análise do processo de erosão na Bacia do Alto Taquari demandou o tratamento e a classificação dos quatro indicadores escolhidos, que são descritos com mais detalhes abaixo.

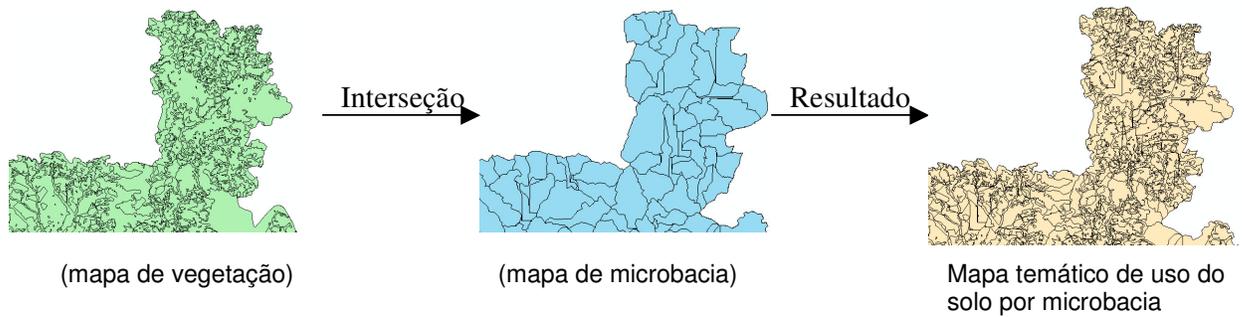
### 5.5.1 Uso do solo

Segundo Silva [84], citado por [37], as áreas ocupadas pela agropecuária correspondem a 61,9% da superfície da BAT, tendo destaque para as pastagens, que em 2000 eram cultivadas em 54,8% das terras da bacia.

As áreas da Bacia do Alto Taquari ocupadas com pastagens cultivadas são as mais castigadas pela erosão, devido ao uso inadequado do solo, com desmatamento indiscriminado nas encostas e nos topos dos morros e predomínio de solos com textura arenosa de baixa fertilidade.

As pastagens cultivadas são dominadas pelas gramíneas do gênero *Brachiaria decumbens*, que por ser pouco exigente quanto à fertilidade dos solos, têm sido amplamente implantada na BAT. Em 2000, 47,9% das pastagens da BAT eram cultivadas em Neossolos Quartzarênicos Órticos, correspondendo a 26,3% da superfície da bacia [84] [37].

O uso do solo foi obtido a partir das informações de vegetação do PCBAP e do mapeamento de uso da terra que Silva [84] realizou para a Bacia do Alto Taquari no ano de 2000. Foi feita uma representação do mapa temático de uso do solo por microbacia através da interseção entre o mapa vegetação e o mapa de microbacias da BAT. A interseção garante a representação do mapa temático por cada microbacia que é o valor representativo da sua classe preponderante.



**Figura 5.2** Seqüência para representação do mapa temático uso do solo por microbacia

Em seguida o mapa temático de uso do solo foi classificado por especialistas da Embrapa Solos com relação ao risco de gerar impacto erosivo. A classificação foi feita de acordo com a tabela abaixo, que foi desenvolvida para utilização no Sistema de Análise Espacial do Risco de Impacto Ambientais Derivados de Erosão em Bacias Hidrográficas – SAEB [62], elaborado pela Embrapa Solos.

Tipo de vegetação	Classificação
Floresta Estacional semidecidual Aluvial (mata, mata ciliar)	Fa
Floresta Estacional semidecidual Submontana (mata, mata ciliar)	Fs
Enclave Savana/Floresta Estacional semidecidual (mata)	SF
Cerradão	Sd
Savana (cerrado) florestada + arborizada	Sd+Sa
Savana (cerrado) arborizada + florestada	Sa+Sd
Savana (cerrado) arborizada (campo, cerrado, cerrado aberto)	Sa
Savana (cerrado) arborizada + gramíneo-lenhosa	Sa+Sg
Savana (cerrado) gramíneo-lenhosa (campo, campo limpo, campo sujo caronal, campina e campo alagado)	Sg
Áreas desmatadas	USO
Não classificado	N/C

**Tabela 5.1** Classificação do indicador uso do solo  
Fonte: adaptado de [62]

Finalmente o mapa de uso do solo, que se encontrava no formato vetorial, foi convertido para o formato *raster* (imagem formada por *pixels*), utilizando a extensão do SIG *ArcView* chamada “*Spatial Analyst*” ([www.esri.com](http://www.esri.com)), gerando o mapa temático final de uso do solo que foi utilizado nas análises espaciais.

### 5.5.2 Erosividade da chuva

Segundo Soriano et. al. [86] a precipitação total anual média na BAT é estimada em 1.506 mm. A análise da distribuição mensal das chuvas na BAT evidenciou que o período mais chuvoso, ou seja, quando pode ocorrer maior erosão, inicia-se em novembro e estende-se até março, quando ocorre 71,6% do índice total anual pluviométrico na BAT. Os meses mais chuvosos são dezembro e janeiro, com 509 mm, ou seja, 33,8% do volume total anual [37].

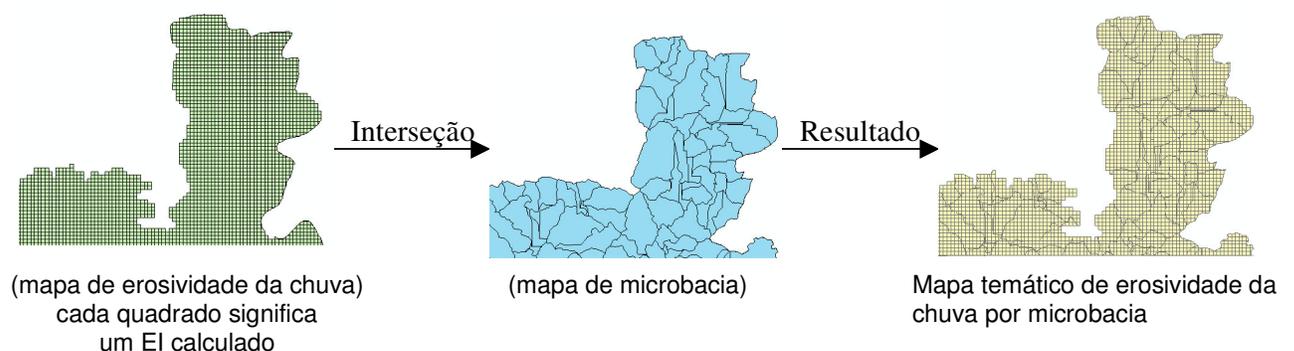
A erosividade da chuva usualmente pode ser estimada pelo produto da energia cinética total pela intensidade da chuva máxima em trinta minutos. Porém, há uma indisponibilidade de registros de intensidade máxima de chuva, obtidos através de pluviógrafos, dificultando o cálculo de precipitação pluvial [62] [17] [86].

Assim, Lombardi Neto ajustou uma regressão linear simples entre o índice médio mensal de erosão (EI) e um índice pluviométrico chamado coeficiente de precipitação, para gerar o cálculo dos índices mensais de erosividade. Utilizando o método de Lombardi Neto (1977), citado por [86], através da fórmula abaixo, foi gerado o índice de erosividade das chuvas [62].

$$EI = 6,872 (p^2/P)^{0,85}$$

onde EI = média mensal do índice de erosão (Ton mm/ha.h);  
p = precipitação média mensal (mm);  
P= precipitação média anual (mm).

Foi realizado o processo de interseção entre o mapa de erosividade da chuva e o mapa de microbacias da BAT para a geração da representação do mapa temático de erosividade da chuva por microbacia.



**Figura 5.3** Seqüência para representação do mapa temático de erosividade da chuva por microbacia

Em seguida cada classe de erosividade da chuva foi classificada com relação ao seu risco de gerar impacto erosivo, levando-se em consideração os valores calculados dos índices mensais de erosividade, sendo geradas cinco novas classes de erosividade da chuva.

Ei Ano	Classificação
1 - 250	Muito Baixa
250 - 500	Baixa
500 - 750	Média
750 - 1000	Alta
1000 - 2000	Muito Alta

**Tabela 5.2** Classificação do indicador erosividade da chuva  
Fonte: adaptado de [62]

Por fim também foi feita a conversão do mapa temático de erosividade da chuva do formato vetorial para o formato *raster*, utilizando a extensão “*Spatial Analyst*”.

### 5.5.3 Solo

Segundo [37], a utilização de mapas digitais de solos do Plano de Conservação da Bacia do Alto Paraguai, obtidos dos levantamentos de campo na escala de 1:250.000, geraram informações específicas de solos para a Bacia do Alto Taquari.

Segundo Silva [84], o solo é dominado por Latossolos, Neossolos Quartzarênicos órticos, Argissolo e Neossolos Litólicos. Estes tipos de solo são mais sujeitos a erosão e recobrem 79,19% da superfície da BAT. Os Neossolos Quartzarênicos órticos ocupam quase a metade de toda a superfície da BAT.

CLASSE DE SOLO	ÁREA	
	km <sup>2</sup>	%
Neossolo Quartzarênico órtico	13113,7	46,09
Neossolo Quartzarênico Hidromórfico	524,4	1,84
Gleissolo Háptico	56,6	0,2
Latossolo Vermelho Acriférico	154,5	0,54
Latossolo Vermelho	4204,3	14,78
Latossolo Vermelho-Amarelo	973,3	3,42
Neossolo Litólico	3793,7	13,33
Argissolo Vermelho-Amarelo	5624	19,77

**Tabela 5.3** Tipos de solos da BAT  
Fonte: adaptado de [37]

As informações de solo também foram obtidas a partir do mapeamento feito pelo PCBAP. Foi realizado o processo de interseção entre o mapa de solo e o mapa de microbacias da BAT para a geração da representação do mapa temático de solo por microbacia.



**Figura 5.4** Seqüência para representação do mapa temático de solo por microbacia

Em seguida, cada classe de solo foi classificada por especialistas da Embrapa Solos com relação ao seu risco de gerar impacto erosivo e levando-se em consideração a sua classe, textura, horizonte A, relevo, principais limitações e aptidão agrícola sendo geradas cerca de cinquenta novas classes de solo. Abaixo se encontram as primeiras classificações do indicador solo.

Classe	textura	horizonte A	Relevo	Principais Limitações	Aptidão agrícola	Classificação
LVd	m. arg.	mod.	pl, sond	f, h	1bC	LVd1
PVd	m. arg. e arg	mod.	pl, sond	f, h	1bC	
LVd	arg e m. arg	mod. e Chern.	Sond, pl	f, h, m	1bC	LVd2
PVe	méd/arg.	mod.	Sond	f, h, m	2abc	
LVd	méd.	mod.	Sond	f, h, m	2(b)c	
LVd	arg	mod.	PI	f, h	1bC	LVd3

**Tabela 5.4** Classificação do indicador solo  
Fonte: adaptado de [62]

Por fim, também foi feita a conversão do mapa temático de solo do formato vetorial para o formato *raster*, utilizando a extensão “*Spatial Analyst*”.

#### 5.5.4 Conflito de uso

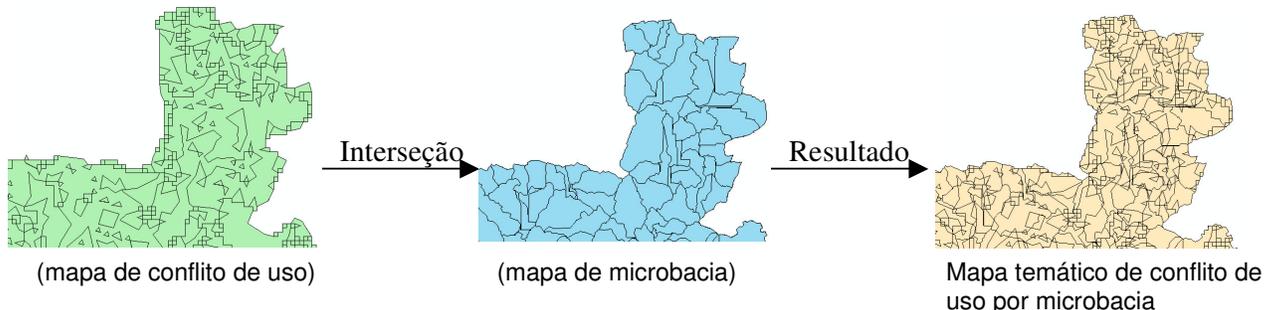
Segundo Galdino et. al. [37], a análise da ocorrência das principais classes de uso/coertura vegetal evidenciou que em 1994, a maior parte da BAT, estava recoberta por

pastagens cultivadas (1.536.529 ha). A vegetação nativa recobria 1.062.945 ha e a soja era cultivada em 198.830 ha.

TIPO DE USO	ÁREA	
	km <sup>2</sup>	%
Agricultura (Soja e Milho)	1989,83	7,09
Pecuária (Pastagem cultivada)	15365,29	54,79
Vegetação nativa	10629,45	37,9
Outros (Urbanização e corpos d'água)	61,62	0,22

**Tabela 5.5** Tipos de uso da terra da BAT  
Fonte: adaptado de [37]

Uma área é considerada como não tendo conflito de uso quando a utilização da terra se dá de acordo com sua capacidade de uso ou aptidão agrícola. Para determinar a existência ou não de conflito pode-se confrontar o mapa de aptidão agrícola das terras, feito com base nas unidades ambientais naturais (informações de solo), com o mapa de uso e cobertura vegetal [62]. Após ser gerado o mapa de conflito de uso a partir da aptidão agrícola e do mapa de uso foi realizado o processo de interseção para a geração da representação do mapa temático de conflito de uso por microbacia.



**Figura 5.5** Seqüência para representação do mapa temático de conflito de uso por microbacia

Em seguida cada classe de conflito de uso foi classificada com relação ao seu risco de gerar impacto erosivo levando-se em consideração um índice de conflito de uso gerado, sendo criadas seis novas classes como apresentadas abaixo.

Valores de Conflito	Classificação
20	Baixo Conflito
40	Médio Conflito
60	Alto Conflito
100	Alto Conflito
200	Alto Conflito
300	Alto Conflito

**Tabela 5.6** Classificação do indicador conflito de uso  
Fonte: adaptado de [62]

Por fim, também foi feita a conversão do mapa temático de conflito de uso do formato vetorial para o formato *raster*, utilizando-se a extensão “*Spatial Analyst*” do ArcView ([www.esri.com](http://www.esri.com)).

### 5.5.5 Pontos de localização de voçorocas

A Bacia do Alto Taquari possui várias áreas de erosão identificadas como voçorocas, devido a maior intensidade do processo erosivo em algumas áreas. Uma lista de cerca de vinte e dois pontos com coordenadas, foram observados em campo nos municípios de Camapuã e Alcinópolis.

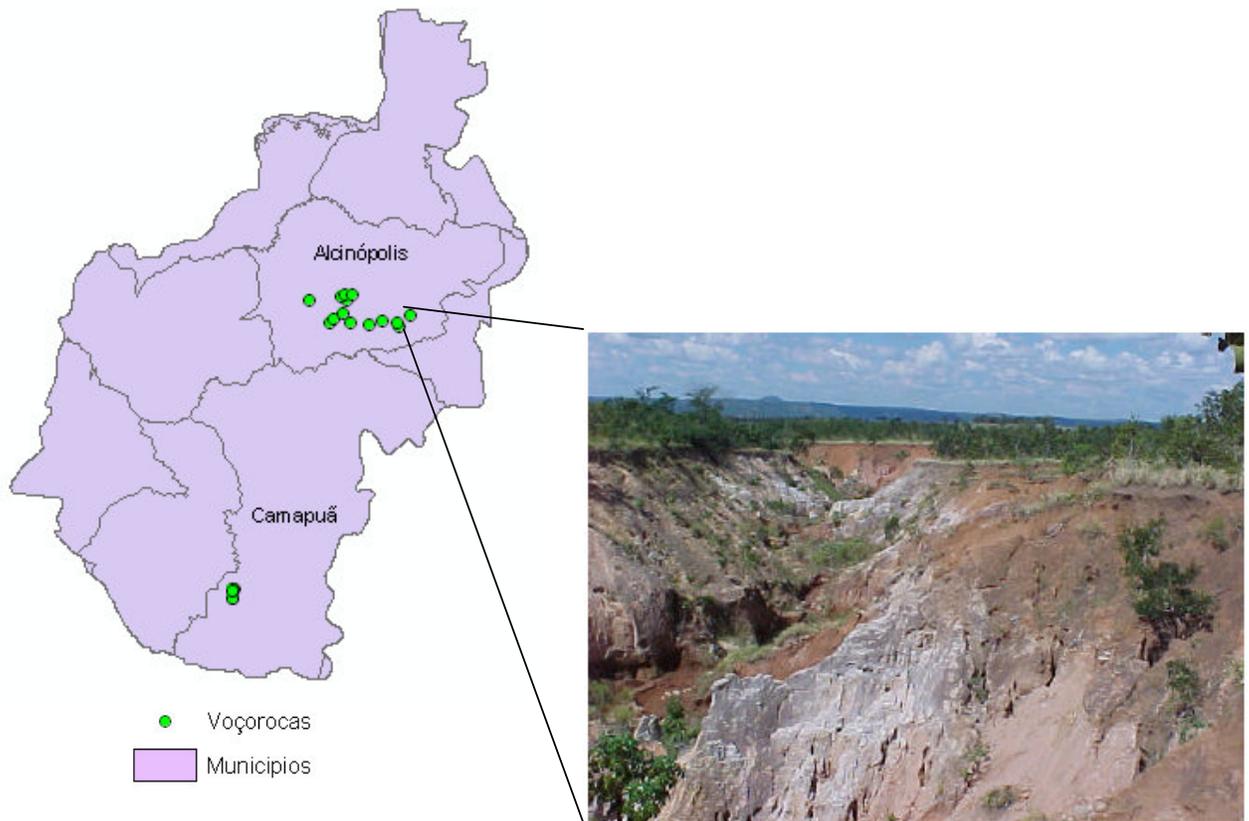
As coordenadas desses pontos foram coletadas a partir de um receptor GPS - Global Position System ([www.garmin.com](http://www.garmin.com)) no âmbito do projeto “SOS Taquari”, a fim de servir como amostras para fins de avaliação da correlação espacial existente entre as mesmas e as classes dos indicadores investigados. Abaixo encontram-se as tabelas com as coordenadas obtidas:

lat (S)	long (W)	Município
18° 21' 31.7"	53° 52' 4.4"	Alcinópolis
18° 21' 31.7"	53° 47' 31.2"	Alcinópolis
18° 21' 38.2"	53° 43' 28.9"	Alcinópolis
18° 21' 46.8"	53° 36' 45.0"	Alcinópolis
18° 21' 2.5"	53° 37' 11.3"	Alcinópolis
18° 21' 0.0"	53° 40' 36.1"	Alcinópolis
18° 20' 49.2"	53° 51' 31.0"	Alcinópolis
18° 19' 30.0"	53° 49' 14.9"	Alcinópolis
18° 19' 27.1"	53° 34' 16.3"	Alcinópolis
18° 16' 44.4"	53° 56' 57.8"	Alcinópolis
18° 16' 23.2"	53° 48' 39.6"	Alcinópolis
18° 15' 46.8"	53° 49' 45.5"	Alcinópolis
18° 15' 31.0"	53° 48' 49.3"	Alcinópolis
18° 15' 32.0"	53° 47' 7.8"	Alcinópolis

lat (S)	long (W)	município
19° 20' 37.7	54° 12' 55.6	Camapuã
19° 20' 36.6	54° 13' 00.4	Camapuã
19° 18' 36.7	54° 12' 49.0	Camapuã
19° 18' 33.9	54° 12' 55.4	Camapuã
19° 18' 37.9	54° 12' 59.3	Camapuã
19° 18' 38.8	54° 12' 59.1	Camapuã
19° 18' 57.5	54° 12' 55.2	Camapuã
19° 18' 57	54° 12' 56.5	Camapuã

**Tabela 5.7** Coordenadas dos pontos de localização de voçorocas na BAT

As coordenadas das voçorocas identificadas em campo na região da BAT foram usadas como pontos de treinamento. O sistema computacional utilizado para fazer as análises espaciais no estudo de caso necessita de pontos de treinamento, já que os mesmos definem a ocorrência de um evento na área de estudos.



**Figura 5.6** Espacialização de áreas de voçorocamento na BAT  
 Fonte: [www.cnps.embrapa.br](http://www.cnps.embrapa.br)

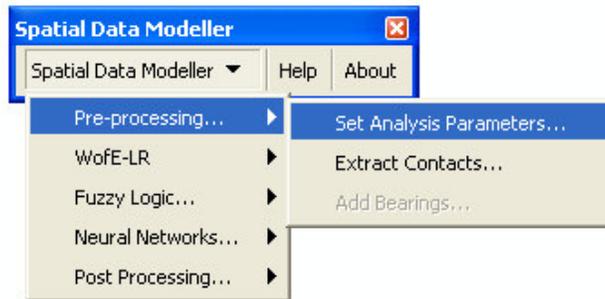
## 5.6 Aplicação da metodologia de classificação “Fuzzy”

Nesta seção serão apresentados os procedimentos utilizados para implementar a integração de dados temáticos pelo método baseado na utilização de Lógica “Fuzzy”, baseada na utilização do sistema computacional ArcSDM, a fim de se estudar o fenômeno de erosão na BAT.

### 5.6.1 Definição do conjunto de parâmetros de análise

Nesta etapa foram adicionados no Sistema de Informação Geográfica *ArcView* todos os dados utilizados na análise. Foram adicionados os quatro indicadores ambientais descritos no item 5.5 e o arquivo vetorial de pontos representando a localização das voçorocas. Foi necessário salvar o projeto no SIG *ArcView* no qual estão inseridos os temas que serão utilizados antes de usar qualquer função do ArcSDM.

Em seguida foi selecionada a opção *Spatial Data Modeller* > *Pre-processing* > *Set Analysis Parameters*. O “*Set Analysis Parameters*” auxilia a selecionar e editar os parâmetros utilizados na análise.



**Figura 5.7** Iniciando “*Set Analysis Parameters*”

Na caixa de diálogo chamada “*Set analysis parameters...-Layers*” foi selecionada a opção “*Fuzzy Logic*”, indicando-se, posteriormente, quatro parâmetros necessários para o início das análises: o “*Study Area Grid*”, o “*Unit Área*”, o “*Training Points*” e o “*Missing Data*”.

O “*Study Area Grid*” é definido como a abrangência geográfica da área de estudo analisada. Para a aplicação da metodologia “*Fuzzy*” torna-se necessário que o *Study Area Grid*, bem como todos os indicadores (também chamados de evidências) estejam no formato *raster*. O ArcSDM estabelece que a área de estudo não deve ter um tamanho de célula maior que o menor tamanho de célula de qualquer outra evidência a ser usada na análise.

No campo “*Unit Area*” é necessário definir a unidade de área que o ArcSDM usará como a unidade de área para medição durante o cálculo das análises. Para a análise da BAT foi colocada a unidade de um quilômetro quadrado.

O campo “*Training Sites*” corresponde aos pontos de treinamento, que são um conjunto de dados pontuais que definem a ocorrência de um evento, neste caso pontos de erosão. Para o estudo de caso da BAT, cada ponto de treinamento foi considerado uma ocorrência de voçoroca identificada em campo. Para “*Training Points*” foram utilizadas as coordenadas listadas na tabela 5.7.

O “*Missing data*” é um conceito usado no ArcSDM para identificar dados dentro da área de estudos que serão tratados como “*no data*” ou para identificar áreas de “*no data*” que estão dentro da área de estudos. No caso em questão será usado o valor -99.

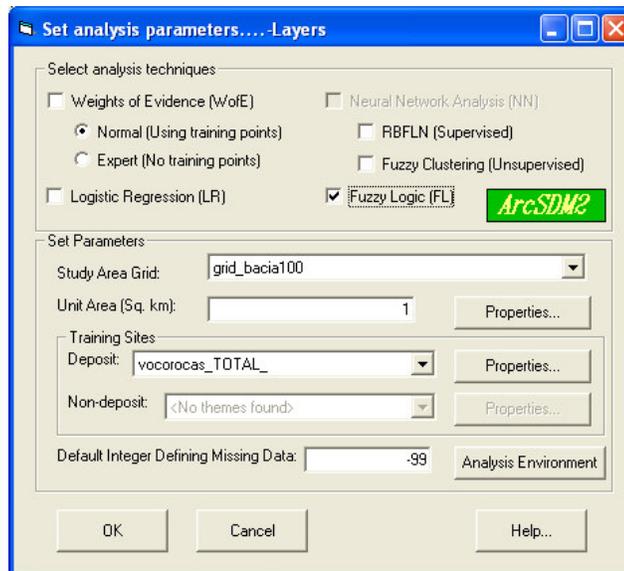


Figura 5.8 Caixa de diálogo “Set Analysis Parameters”

### 5.6.2 Geração dos membros “Fuzzy”

Os valores de pertinência “Fuzzy” refletem a importância relativa de cada mapa, bem como a importância relativa de cada classe do mapa. Os valores são selecionados baseados no julgamento subjetivo sobre a importância relativa dos mapas e suas várias classes. Os membros “Fuzzy” na análise da BAT foram definidos na opção: *Spatial Data Modeller* > “Fuzzy” Logic > *Define Fuzzy Membership*.

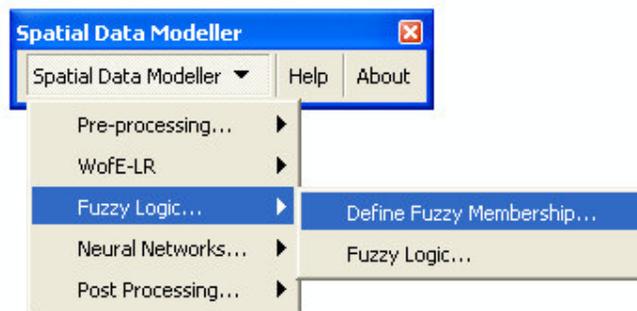


Figura 5.9 Iniciando “Define Fuzzy membership”

Na caixa de diálogo “Calculate Response Theme” foi indicado em “Select raster” o dado de uso do solo que foi trabalhado primeiramente. Quando o *raster* é escolhido o

ArcSDM abre automaticamente uma caixa de diálogo que pergunta se o dado é categórico ou ordenado.

Os dados ordenados indicam um *raster* para o qual o atributo é algum tipo de medida, já os dados categóricos indicam um *raster* cujos atributos são informações geralmente textuais e não representam medidas.

De acordo com o tipo de dado utilizado pode-se optar por dois métodos para a geração dos valores de pertinência “*Fuzzy*” no ArcSDM. O primeiro método é usado com dados categóricos e envolve essencialmente a edição de uma tabela. Esse método é chamado “*Free data Fuzzification Functions*”. O segundo método usa funções matemáticas para transformar o dado de entrada em valores de pertinência “*Fuzzy*” no intervalo entre 0 e 1. Esse método é utilizado para dados ordenados e é chamado “*Ordered Data Fuzzification Functions*”.

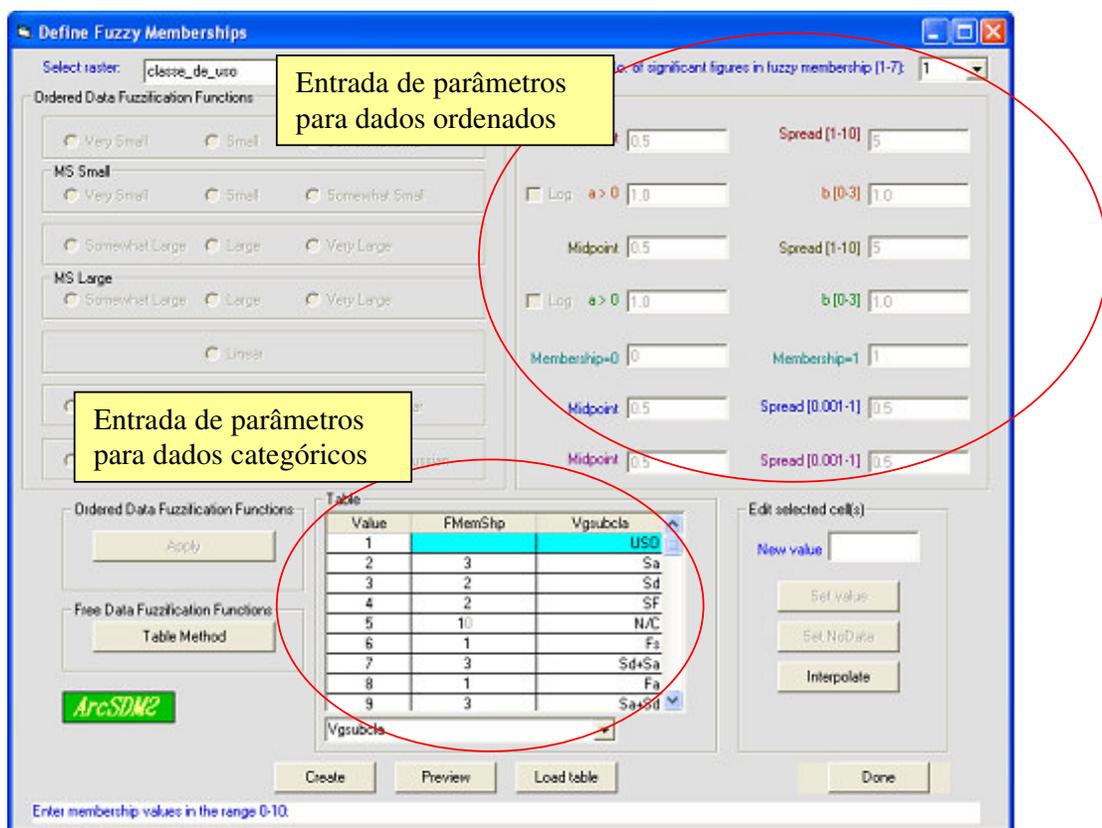


Figura 5.10 Caixa de diálogo “*Define Fuzzy membership*”

### 5.6.3 Geração dos planos de informação “*Fuzzy*”

Nesta etapa foi feito o processo de geração dos planos de informação “*Fuzzy*” para os dados da BAT a partir da *fuzzificação* (transformação de um conjunto de dados em um conjunto contínuo baseado nas funções de pertinência “*Fuzzy*” definidas anteriormente) dos

indicadores ambientais da BAT, que foram tratados como dados categóricos ou dados ordenados como apresentado a seguir.

### 5.6.3.1 Planos de informação “Fuzzy” para os indicadores categóricos

Para o estudo de caso da BAT foi utilizada a *fuzzificação* de dados categóricos para o indicador uso do solo e o indicador solo. Os pesos atribuídos a cada indicador foram levantados por especialistas através de *workshops* realizados na Embrapa-Solos.

Para dados categóricos, o método para proceder a entrada dos pesos propostos por especialistas é feito interativamente, selecionando-se a linha que se quer colocar o valor, entrando em seguida com o peso no campo “New Value” e clicando em “Set Value”. O novo valor aparece então na coluna de pertinência “Fuzzy”, chamada “FmemShp”, para a linha selecionada, conforme pode ser visto abaixo.

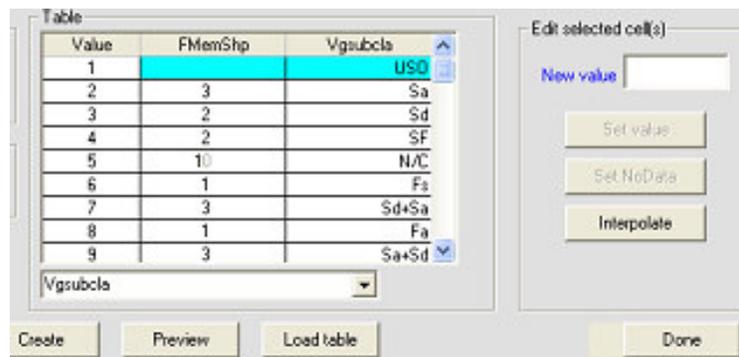


Figura 5.11 Entrada de parâmetros para dados categóricos

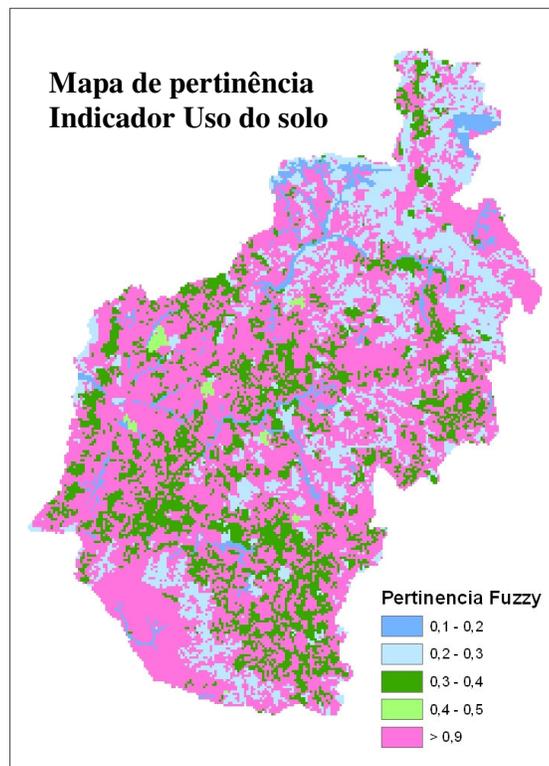
Assim, o ArcSDM pondera os mapas temáticos em valores que variam entre 0 e 1 (valores de pertinência “Fuzzy”) a partir desses pesos. A definição dos pesos é uma das mais importantes etapas, pois os pesos indicarão quais classes são consideradas mais relevantes para o surgimento do evento (no caso erosão), sendo posteriormente representadas pelos membros “Fuzzy” que irão expressar a importância da classe no modelo. Para o indicador uso do solo foram utilizados os pesos definidos por especialistas conforme a tabela abaixo:

Tipo de Vegetação	Classificação	Peso
Floresta Estacional Semidecidual Aluvial(Mata, Mata Ciliar)	Fa	1
Floresta Estacional Semidecidual Submontana(Mata,Mata Ciliar)	Fs	1
Encrave Savana/Floresta Estacional Semidecidual(Mata)	SF	2
Cerradão	Sd	2
Savana(Cerrado) Florestada + Arborizada	Sd+Sa	3
Savana(Cerrado) Arborizada + Florestada	Sa + Sd	3

Savana(Cerrado) Arborizada(Campo Cerrado,Cerrado,Cerrado Aberto)	Sa	3
Savana(Cerrado) Arborizada + Gramineo-Lenhosa	Sa + Sg	4
Savana(Cerrado) Gramineo-Lenhosa (Campo,Campo Limpo,Campo SujoCaronal, Campina e Campo Alagado)	Sg	4
Áreas Desmatada	USO	10
Não Classificado	N/C	10

**Tabela 5.8** Valores dos pesos para geração dos membros “Fuzzy” de uso do solo  
 Fonte: adaptado de [62]

Finalmente, foi criado o plano de informação “Fuzzy” para o indicador uso do solo através do botão “Create”, gerando o mapa de saída com os membros “Fuzzy” variando de 0 a 1 (também chamado mapa de pertinência). Abaixo é apresentado o mapa de pertinência contendo cinco classes definidas pelo método “Natural Breaks” do SIG ArcView.



**Figura 5.12** Mapa de pertinência do indicador uso do solo

Os mapas de pertinência indicam que os valores mais próximos de “1” pertencem à classe que tem maior influência no processo erosivo, enquanto os valores próximos de “0” indicam os que pertencem à classe que tem menor influência. Observa-se que a maior parte da bacia encontra-se com o uso do solo classificado como áreas desmatadas, cujo peso indicado

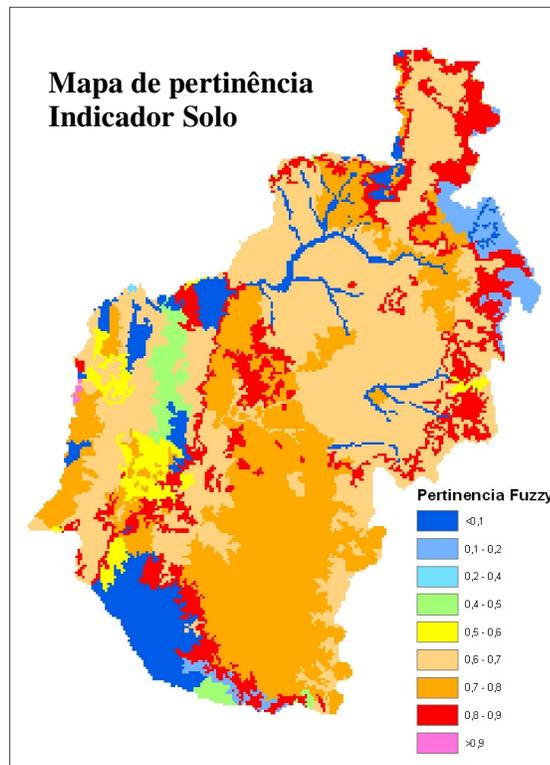
pelos especialistas da Embrapa Solos é alto, ou seja, o valor de pertinência “*Fuzzy*” para as regiões que cobrem esta classe também é alto.

O indicador solo também foi tratado como um dado categórico, sendo utilizada a tabela de pesos apresentada abaixo para a geração dos membros “*Fuzzy*”.

Classificação	Peso						
LVd1	1	LVd12	6	PVAe2	7	RQo8	8
LVd2	1	LVd13	6	GXbd	1	RQo9	8
LVd3	1	LVd14	2	RQg1	1	RQo10	8
LVd4	1	LVdf	2	RQg2	1	RQo11	6
LVd5	1	LVdf1	7	RQg3	1	RQo12	8
LVd6	1	LVdf2	1	RQo1	7	RQo13	8
LVd7	4	PVAd1	7	RQo2	7	RLd1	9
LVd8	1	PVAd2	7	RQo3	7	RLd2	9
LVd9	5	PVAd3	1	RQo4	7	RLd3	9
LVd10	1	PVAd4	1	RQo5	7	RLd4	9
LVd11	5	PVAd5	1	RQo6	7	RLd5	10
		PVAe1	8	RQo7	7	RLe	9
						RLq	9

**Tabela 5.9** Valores dos pesos para geração dos membros “*Fuzzy*” de solo  
Fonte: adaptado de [62]

Em seguida também foi criado o plano de informação *Fuzzy* para o indicador solo, gerando o mapa de pertinência apresentado abaixo, que foi classificado em nove classes.



**Figura 5.13** Mapa de pertinência do indicador solo

Observa-se que as áreas com maiores valores de pertinência “*Fuzzy*” para o indicador solo encontram-se distribuídas ao norte e a leste da bacia, podendo também ser encontrada em algumas regiões na parte central da bacia.

### 5.6.3.2 Planos de informação “*Fuzzy*” para os indicadores ordenados

Foi utilizada a *fuzzificação* de dados ordenados para os indicadores erosividade da chuva e conflito de uso. Para esses indicadores também foram estabelecidos pesos que foram atribuídos por especialistas através de *workshops* realizados na Embrapa Solos.

Para a utilização desse método de geração dos membros de pertinência “*Fuzzy*” deve ser definido primeiramente o arquivo *raster* referente ao tema que será usado na caixa de diálogo “*Select raster*”, conforme pode ser visto na figura 5.10. Em seguida deve ser escolhido quantos números significativos terá o dado de entrada, a partir da opção “*No. Of significant figures em fuzzy membership*”. Se a quantidade de números significativos for “um” (1) os valores de entrada para a geração dos membros “*Fuzzy*” deverão ser inteiros entre 0 e 10. Se a quantidade de números significativos for “dois” (2) os valores de entrada poderão ser inteiros entre 0 e 100 e assim por diante. O dado de saída, quando o mapa de pertinência “*Fuzzy*” for criado, terá sempre valores do tipo “*floating*” (ponto flutuante) entre 0 e 1.

Para dados ordenados, deve-se escolher ainda uma função de *fuzzificação* e os parâmetros desejados para a função. Após a entrada dos parâmetros é feito o mapeamento dos valores de entrada do *raster*, clicando no botão “*Apply*”, calculando em seguida os valores de membros “*Fuzzy*” que aparecerão na coluna de pertinência *Fuzzy* “*FmemShp*”.

As funções de *fuzzificação* que podem ser utilizadas são: *Small, MS small, Large, MS large, Linear, Near e Gaussian*. Para os indicadores erosividade da chuva e conflito de uso foi utilizada a função *Linear*.

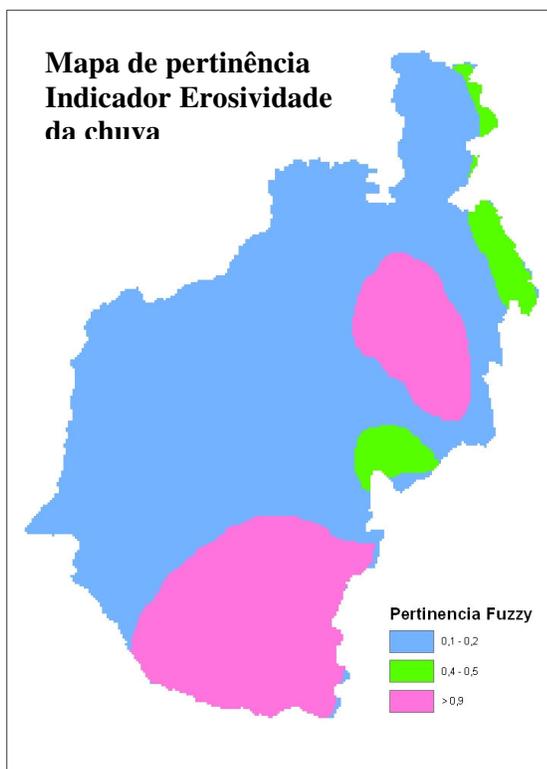
Segundo Moreira [66], a função *Linear* é definida por dois segmentos de reta inclinados que se encontram em um ponto central, cujo valor de pertinência é 1. Nas bordas o valor é 0,5 e a inclinação das retas define a zona de transição “*Fuzzy*”.

Apesar de ser utilizada a função de *fuzzificação Linear*, os valores atribuídos a cada classe dos indicadores solo e conflito de uso foram alterados conforme os pesos indicados pelos especialistas da Embrapa. Abaixo segue a tabela com os valores dos pesos para o indicador erosividade da chuva.

Ei-Ano	Classificação	Peso
1 - 250	Muito baixa	2,5
250 - 500	Baixa	5
500 - 750	Média	5
750 - 1000	Alta	10
1000 - 2000	Muito Alta	10

**Tabela 5.10** Valores dos pesos para geração dos membros “Fuzzy” de erosividade da chuva  
 Fonte: adaptado de [62]

Em seguida, também foi criado o plano de informação “Fuzzy” para o indicador erosividade da chuva, gerando o mapa de pertinência apresentado abaixo, que foi classificado em apenas três classes definidas pelo método “Natural Breaks” do SIG ArcView.



**Figura 5.14** Mapa de pertinência do indicador erosividade da chuva

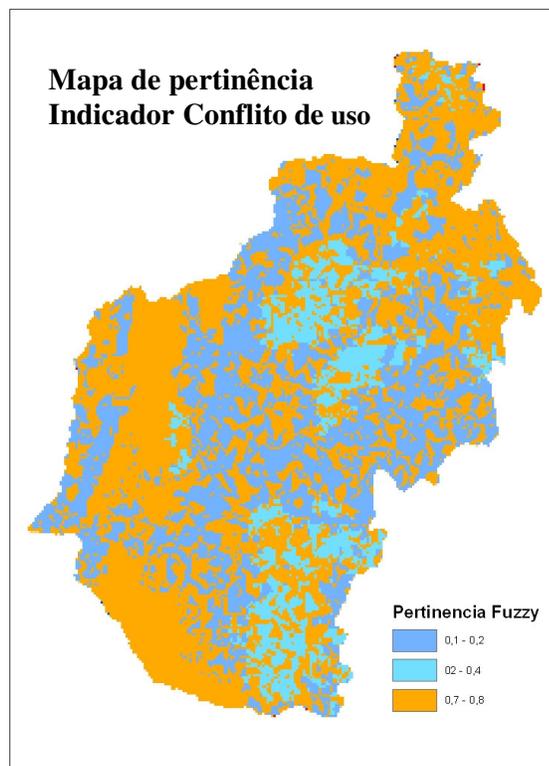
Observa-se que as áreas com maiores valores de pertinência “Fuzzy” para o indicador erosividade da chuva encontram-se na parte leste e ao sul da bacia, correspondendo à área cuja média mensal do índice de erosão (Ei) está acima de 750 Ton mm/há.h.

O indicador conflito de uso também foi tratado como um dado ordenado, sendo utilizada a tabela de pesos apresentada abaixo.

Valores de conflito	Classificação	Peso
20	Baixo Conflito	2
40	Médio Conflito	4
60	Alto Conflito	8
100	Alto Conflito	8
200	Alto Conflito	8
300	Alto Conflito	8

**Tabela 5.11** Valores dos pesos para geração dos membros “Fuzzy” de conflito de uso  
Fonte: adaptado de [62]

Em seguida também foi criado o plano de informação “Fuzzy” para o indicador conflito de uso, gerando o mapa de pertinência mostrado abaixo, que foi classificado em três classes definidas pelo método “Natural Breaks” do SIG ArcView.



**Figura 5.15** Mapa de pertinência do indicador conflito de uso

Observa-se que as áreas com maiores valores de pertinência “Fuzzy” para o indicador conflito de uso encontram-se onde o conflito de uso foi classificado como baixo, sendo dado maior peso para essas classes pelos especialistas.

#### **5.6.4 Escolha do operador “Fuzzy”**

Para a integração dos planos de informação “Fuzzy” visando a obtenção do mapa final de potencialidade à ocorrência de erosão na BAT os indicadores foram combinados através do operador “Fuzzy” gamma ( $\gamma$ ). Uma breve descrição do operador “Fuzzy” gamma pode ser encontrada no item 3.2.1.4 do terceiro capítulo.

A escolha do grau que é usado para gamma garante a tendência com que terão os valores de saída da análise, gerando diferentes cenários que vão do mais favorável ao mais desfavorável. Segundo Câmara [22] se o valor  $\gamma$  adotado for maior que 0,8 os valores resultantes tendem a se tornar mais próximos da soma algébrica. Porém se o valor  $\gamma$  for menor que 0,35 os valores resultantes tendem a se tornar mais próximos do produto algébrico, ou seja, a escolha do gamma pode variar de “pessimista”, com  $\gamma$  abaixo de 0,35 a “otimista”, com  $\gamma$  acima de 0,8 [61].

#### **5.6.5 Definição do grau para o operador gamma**

Nesta parte do processamento, foi feita uma seleção empírica de diferentes valores de gamma, procurando valores que estivessem distribuídos de forma a gerar diferentes cenários para tentar auxiliar no processo de análise. Foram selecionados os valores de gamma 0,70, 0,75, 0,80, 0,85, 0,90 e 0,95, gerando-se um total de 6 cenários diferentes, para tentar identificar o que mais se aproximou de outros trabalhos de pesquisa que também avaliaram processos de erosão e de perda de solo na BAT, como por exemplo os resultados encontrados pelo mapeamento das perdas de solo na Bacia do Alto Taquari realizado pela Embrapa em 1994 [36].

O cenário que mais se aproximou deste trabalho foi o de grau 0,95 sugerindo que este resultado estaria mais próximo dos encontrados por outros trabalhos de pesquisa. Para se chegar a valores mais aproximados dentro do contexto do operador utilizado é necessário realizar ajustes na geração dos membros “Fuzzy” de cada tema e nos valores do grau do operador gamma, junto com especialistas de cada indicador utilizado.

### 5.6.6 Integração dos planos de informação “Fuzzy”

Nesta etapa do processamento “Fuzzy” foi feita a combinação dos planos de informação (mapas de pertinência “Fuzzy”) de todos os indicadores. Depois de definir as funções de *fuzzificação* e gerar os mapas de pertinência “Fuzzy” foi selecionada a opção “Fuzzy Logic” no menu ArcSDM para a combinação.

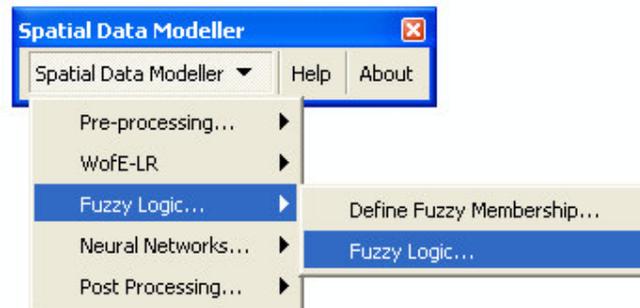


Figura 5.16 Inicando “Fuzzy Logic”

Desta forma, na caixa de diálogo “Fuzzy logic Model Tables”, foi selecionada a opção “Make and add a new table” para criar uma tabela *dBase* que armazena cada operação (combinações usando os operadores “Fuzzy”) do modelo. Essa tabela pode ser salva em um diretório e visualizada posteriormente. O nome padrão para a tabela é “FLmdl” (*Fuzzy Logic Model*) seguido de um número sequencial. Assim que a tabela é salva, é fornecida a caixa de diálogo chamada “Fuzzy Logic”, onde pode ser feita a seleção dos indicadores e operadores utilizados na análise.

Na caixa de diálogo “Fuzzy Logic” selecionou-se os quatro indicadores e em seguida escolheu-se a função gamma. O botão “Calculate” abre a janela “Gamma Operator” que permite definir o grau utilizado para o operador gamma.

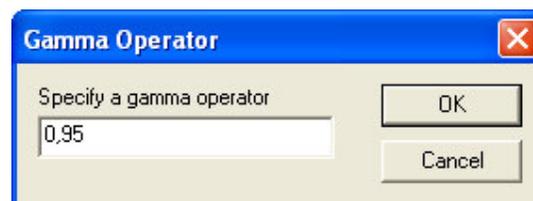


Figura 5.17 Caixa de diálogo “Gamma Operator”

### 5.6.7 Geração do mapa final de pertinência “Fuzzy”

Após o processamento foi aberta a caixa de diálogo “*Result of Fuzzy GAU operation*” onde foi definido um diretório e um nome de saída para o arquivo *raster* criado. O arquivo *raster* criado é o mapa de potencial “Fuzzy” que pode ser visto abaixo.

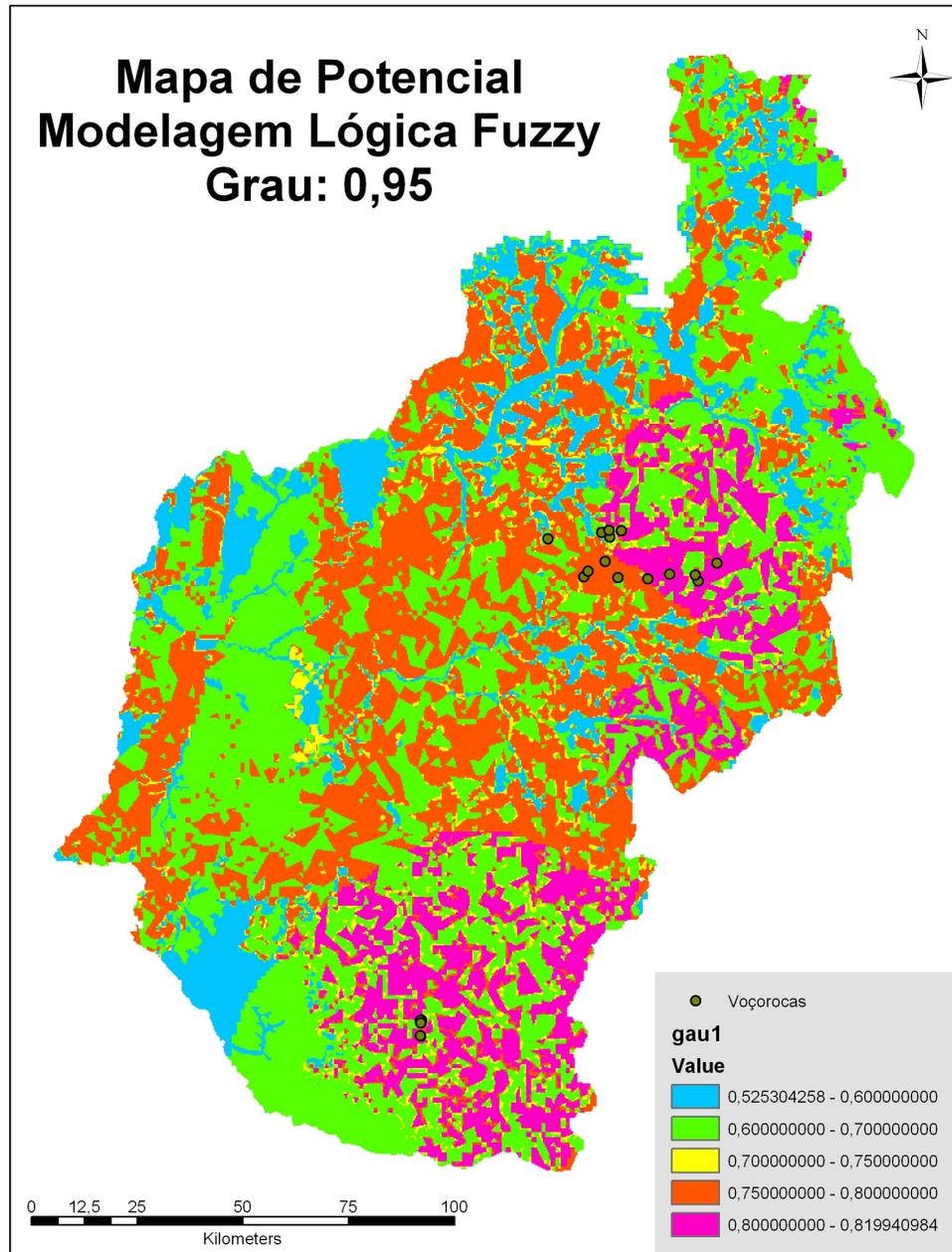


Figura 5.18 Mapa de potencial final “Fuzzy”

No mapa final de pertinência “*Fuzzy*” para os pontos de erosão na BAT, foi realizado um fatiamento criando-se cinco classes definidas pelo método “*Natural Breaks*” do SIG *ArcView*, para facilitar a visualização das áreas com maior potencial à erosão. Percebe-se que as áreas em tons de vermelho e rosa apresentam maior potencial para a ocorrência de erosão (possuem valores “*Fuzzy*” mais próximos de 1) e ocupam grande parte da Bacia do Alto Taquari.

## **5.7 Aplicação do método de predição por Pesos de Evidência**

O modelo de Pesos de Evidência é um método estatístico que envolve a estimativa de um mapa de favorabilidade a partir de um conjunto de variáveis preditivas. Neste trabalho o método foi desenvolvido em uma aplicação em meio ambiente, no qual as evidências consistem de um conjunto de indicadores ambientais (mapas de evidências) e de pontos de ocorrências de voçorocas, citados no item 5.3 deste capítulo.

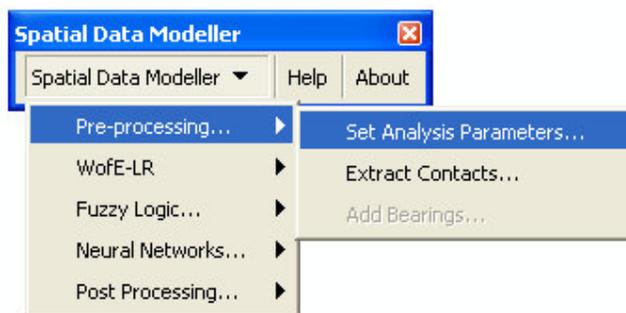
Este método de análise é baseado nos conceitos de probabilidade bayesiana, que foi definido por Araújo [9] como um método estocástico que está apoiado principalmente na idéia de probabilidade “a priori” e “a posteriori”, tendo como característica a transformação de uma série de dados determinísticos em probabilidades.

Assim, nesta seção serão apresentados os procedimentos utilizados para implementação prática da metodologia por Pesos de Evidência descrita no quarto capítulo sobre a área de estudo.

### **5.7.1 Definição do conjunto de parâmetros de análise**

Foram utilizados para a análise de Pesos de Evidência os mesmos dados utilizados no modelamento por Lógica “*Fuzzy*”. Nesta etapa, também foram adicionados no SIG *ArcView* os quatro indicadores ambientais (temas de evidência) e o arquivo vetorial de pontos representando as voçorocas.

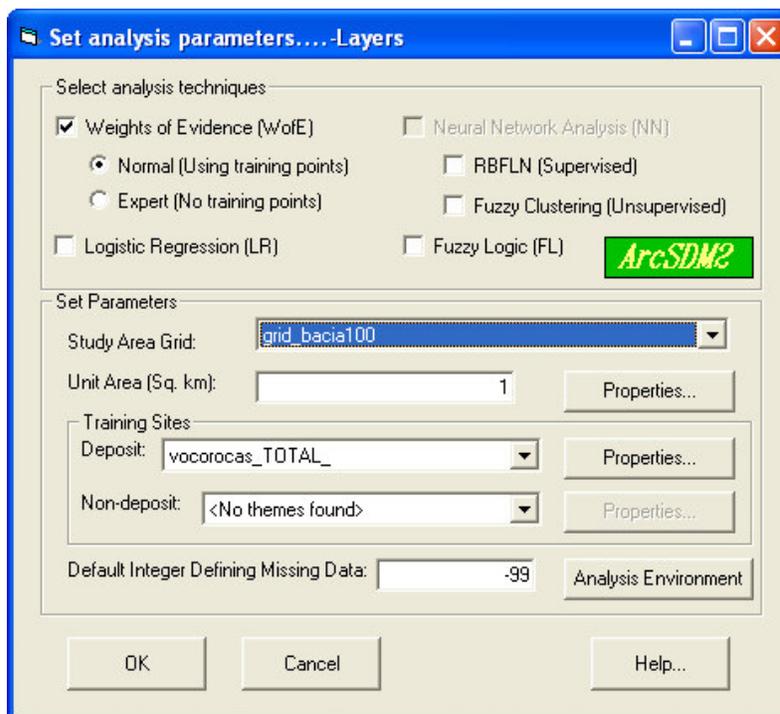
Em seguida foi selecionada a opção *Spatial Data Modeler > Pré-processing > Set Analysis Parameters*. Como dito anteriormente o “*Set Analysis Parameters*” auxilia a selecionar e editar os parâmetros utilizados na análise.



**Figura 5.19** Iniciando “Set Analysis Parameters”

Na caixa de diálogo “Set analysis parameters...- Layers” foram selecionadas as opções abaixo, explicadas anteriormente no item 5.6.1 deste capítulo.

- Foi selecionada a opção “Weights of Evidence (WofE)”, que indica a técnica de análise que foi utilizada. A opção “Normal (Using training points)” indica que na análise foram usados pontos de treinamentos;
- Em “Study Area Grid” foi indicado o *raster* que representava a área de estudos;
- Para “Unit Area (Sq. Km)” foi indicado o valor um (1 km);
- Para “Training Sites” foi indicado o arquivo que representa os pontos de voçoroca coletados em campo, que para a análise de Pesos de Evidência correspondem aos pontos de treinamento.



**Figura 5.20** Caixa de diálogo “Set Analysis Parameters”

## 5.7.2 Cálculo dos pesos para cada mapa de predição

O próximo passo foi calcular os pesos de evidência para cada mapa de predição. No modelamento pelo método de Pesos de Evidências um conjunto de estatísticas, incluindo área, número de pontos de treinamento, pesos positivos ( $W^+$ ), pesos negativos ( $W^-$ ), contraste e seus desvios padrões são calculados para cada classe dos temas e escritos em um arquivo “*dBase*”. Uma breve descrição sobre os pesos pode ser encontrada no item 3.2.2.1 do terceiro capítulo deste trabalho.

O arquivo “*dBase*” criado pode ser utilizado posteriormente no projeto, podendo ser re-visualizado e usado para auxiliar no processo de generalização (reclassificação dos temas de evidência) para o posterior cálculo do mapa de probabilidade.

Para o cálculo dos pesos foi selecionada a opção *Spatial Data Modeller* > *WofE-LR* > *Calculate Theme Weights*, conforme mostrado na figura abaixo.

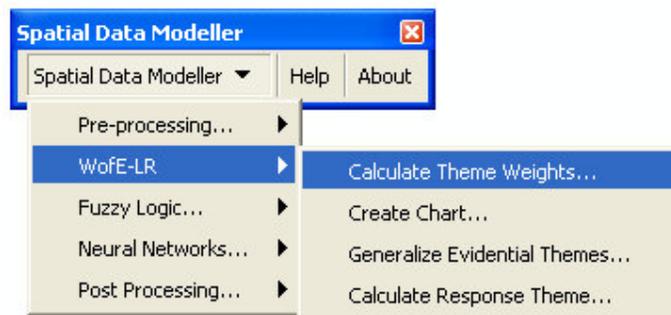
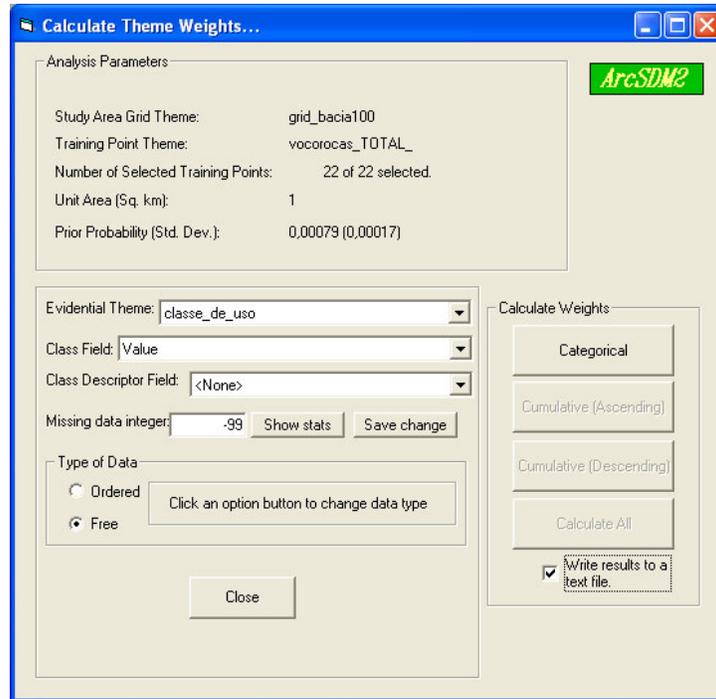


Figura 5.21 Iniciando “*Calculate Theme Weights*”

A caixa de diálogo “*Calculate Theme Weights...*” indica como os pesos foram calculados. Esta caixa de diálogo também mostra o valor da probabilidade “a priori” calculado pelo aplicativo, além dos itens listados abaixo.

- “*Study Area Grid Theme*” que indica o raster que representa a área de estudo analisada
- “*Training Point Theme*” que indica o arquivo que representa os pontos de treinamento
- “*Number of Selected Training Points*” que indica o número total de pontos de treinamento - Somente o número de pontos selecionados na área de estudo são reportados como o total de número de pontos de treinamento. Este número pode ser menor que o número de pontos reportado na caixa de diálogo “*Set Analysis parameters...*”.

- “*Unit Área (Sq. Km)*” que indica a unidade de área em Km<sup>2</sup>
- “*Prior Probability (Std. Dev.)*” que indica a ”Probabilidade “a priori” (com seu desvio padrão seguido em parênteses)



**Figura 5.22** Caixa de diálogo “*Calculate Theme Weights*”

A opção “*Evidential Theme*” (tema de evidência) lista todos os *rasters* disponíveis para o cálculo dos pesos. O primeiro tema utilizado para o cálculo dos pesos foi o tema “uso do solo”. A opção “*Class field*” seleciona a coluna da tabela de atributos do tema de evidência que contém as classes para o cálculo dos pesos. Neste campo foi indicado o campo “Value” do tema “uso de solo”, que apresenta as classes deste tema.

O campo “*Class Description Field*” seleciona uma coluna da tabela de atributos do tema indicado que contenha uma descrição de texto das classes. Para o tema “uso do solo” foi selecionada a opção “*None*”, pois não se quis usar nenhuma descrição para as classes.

O campo “*Missing Data*” especifica um inteiro usado para definir áreas onde os dados são desconhecidos. O valor padrão é -99, que é o mesmo valor especificado na caixa de diálogo “*Set Analysis parameters...*”.

Em seguida foi pedido para especificar se o dado seria tratado como um dado ordenado (*Ordered*) ou não ordenado (*Free*). Esta escolha afeta os pesos acumulados que são calculados pelos dados. Esta opção foi armazenada na tabela chamada “*Layers\_Data Type of*

*Evidential Themes*”, criada pelo ArcSDM, para que possa ser lembrada e para não precisar ser especificada novamente se o dado é do tipo “ordenado” ou “não ordenado” numa nova análise. O tema “uso do solo” foi tratado como um dado do tipo não ordenado (*Free*).

Para o cálculo das estatísticas há quatro opções de como os pesos podem ser calculados. A opção “*Categorical*” calcula os pesos para cada classe do tema de evidência. A opção acumulativa ascendente (*Cumulative (Ascending)*) calcula os pesos acumulados da primeira classe até a última estabelecendo que quanto mais distante da evidência (ou quanto menor concentração de um evento) menor a possibilidade de presença das ocorrências. A opção acumulativa decrescente (*Cumulative (Descending)*) calcula os pesos acumulados da primeira classe até a última classe estabelecendo que quanto mais próximo da evidência (ou quanto maior a concentração de um evento) maior a possibilidade de presença das ocorrências. A última opção chamada “*Calculate All*” calcula todas as três opções acima.

Para o tema “uso do solo” e para todos os outros temas foi utilizada a opção “*Categorical*” e ainda habilitada a opção “*Write results to a text file*”, que escreve os pesos em uma tabela *dBase*. O nome da tabela criada é o nome do tema de evidência acrescido do sufixo “.CT” (indicando que a opção “*Categorical*” foi utilizada).

A tabela de pesos criada contém as seguintes colunas que são descritas abaixo.

- **Class** - Números ou textos que identificam as classes do tema especificado.
- **Area Sq km** - Área de cada classe dada em km<sup>2</sup>.
- **Area Units** - Área de cada classe dada em unidades de células.
- **Points** - O número de pontos de treinamento (neste caso voçorocas) que estão na mesma classe
- **(W+)** - O valor do peso positivo (W+) calculado
- **S(W+)** - O desvio padrão de W+
- **(W-)** - O valor do peso negativo (W-) calculado
- **S(W-)** - O desvio padrão de W-
- **(C)** - O valor do contraste, dado pela diferença entre W+ e W-
- **S(C)** - Desvio padrão do contraste
- **Stud(C)** - O valor de “Student” do contraste, dado pela razão entre o contraste e seu desvio padrão

Class	Area Sq km	Area Units	#Points	W+	s(W+)	W-	s(W-)	C	s(C)	stud(C)
1	16107,5300	16108	14	0,0946	0,2674	-0,1467	0,3537	0,2413	0,4434	0,5442
2	5037,7500	5038	6	0,4099	0,4085	-0,1188	0,2501	0,5287	0,4790	1,1039
3	2307,1900	2307	0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
4	2858,0200	2858	1	-0,8158	1,0002	0,0619	0,2183	-0,8777	1,0237	-0,8574
5	5,1500	5	0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
6	290,8300	291	0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
7	32,8700	33	0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
8	808,0300	808	1	0,4484	1,0006	-0,0171	0,2183	0,4655	1,0242	0,4545
9	232,9000	233	0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
10	140,9100	141	0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
11	0,6600	1	0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

**Tabela 5.12** Cálculo dos pesos de evidência para o indicador uso do solo

Observa-se que os pesos não foram calculados para as classes que não tinham nenhum ponto de treinamento, ou seja, quando a coluna “Points” é zero, os pesos e contrastes associados a esses registros têm o valor zero.

Em seguida foi feito o cálculo dos pesos para os temas de evidência “solo”, “erosividade da chuva” e “conflito de uso” utilizados na análise.

Class	Area Sq km	Area Units	#Points	W+	s(W+)	W-	s(W-)	C	s(C)	stud(C)
1	3157,0500	3157	0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
2	682,0800	682	0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
3	6,1700	6	0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
4	773,8300	774	2	1,1859	0,7080	-0,0671	0,2237	1,2531	0,7425	1,6876
5	2829,0800	2829	2	-0,1121	0,7074	0,0119	0,2237	-0,1240	0,7419	-0,1671
6	753,9000	754	0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
7	438,1100	438	0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
8	204,1200	204	0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
9	1,2300	1	0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
10	130,5400	131	0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

**Tabela 5.13** Cálculo dos ponderadores de evidência para as dez primeiras classes do indicador solo

Class	Area Sq km	Area Units	#Points	W+	s(W+)	W-	s(W-)	C	s(C)	stud(C)
1	18901,2200	18901	7	-0,7703	0,3780	0,7798	0,2584	-1,5500	0,4579	-3,3849
2	1236,1600	1236	0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
3	7362,4900	7362	15	0,9364	0,2585	-0,8341	0,3780	1,7706	0,4579	3,8663
4	8,8000	9	0	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

**Tabela 5.14** Cálculo dos ponderadores de evidência para o indicador erosividade da chuva

Class	Area Sq km	Area Units	#Points	W+	s(W+)	W-	s(W-)	C	s(C)	stud(C)
1	9537,0500	9537	4	-0,6359	0,5001	0,2200	0,2358	-0,8559	0,5529	-1,5479
2	15262,9000	15263	8	-0,4129	0,3536	0,3454	0,2674	-0,7583	0,4434	-1,7104
3	2983,1900	2983	10	1,4456	0,3168	-0,4929	0,2887	1,9385	0,4286	4,5226

**Tabela 5.15** Cálculo dos ponderadores de evidência para o indicador conflito de uso

### 5.7.3 Generalização dos temas de evidências

Devido ao fato de mapas com duas classes serem mais fáceis de interpretar e pelo fato de conferir maior confiança estatística à modelagem de Pesos de Evidência foi feita uma reclassificação dos temas de evidência usadas na análise.

O processo de reclassificação dos mapas de evidência é obtida no ArcSDM através da operação denominada “generalização”. As ferramentas do ArcSDM usam um método onde primeiramente é selecionado o tema de evidência e o campo que contém as classes para serem generalizadas. Assim, utilizando-se uma das ferramentas de generalização, é possível especificar interativamente como a classe atual será generalizada e que novas classes serão criadas. Uma vez especificadas, as novas classes são guardadas em um arquivo denominado “GEN” que é relacionado ao tema de evidência para ser posteriormente utilizado na análise.

O nome convencional do arquivo gerado é formado pelo nome “GEN” mais o nome da tabela de pesos do tema de evidência, por exemplo, o nome da tabela de generalização do tema de evidência “uso do solo” passou a se chamar GEN\_uso\_do\_solo\_CT, uma vez que a sua tabela de pesos tinha o nome uso\_do\_solo\_CT.

Freqüentemente um tema de evidência é generalizado em duas classes, criando um mapa binário no qual cada célula ou está dentro ou fora do padrão. Porém, em alguns casos, reduzir um tema de evidência em um padrão binário pode não ser o caminho mais apropriado para controlar os dados. Um mapa geológico, por exemplo, que pode possuir muitas classes poderia ser tratado mais adequadamente com três ou quatro classes no mapa.

Para a generalização foi selecionada a opção *Spatial Data Modeller*> *WofE-LR*> *Generalize Evidential Themes*, conforme mostrado na figura abaixo.

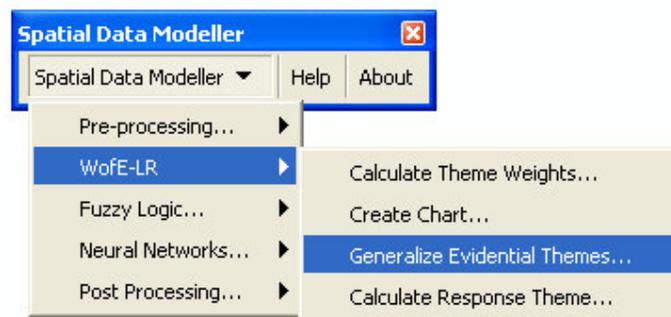
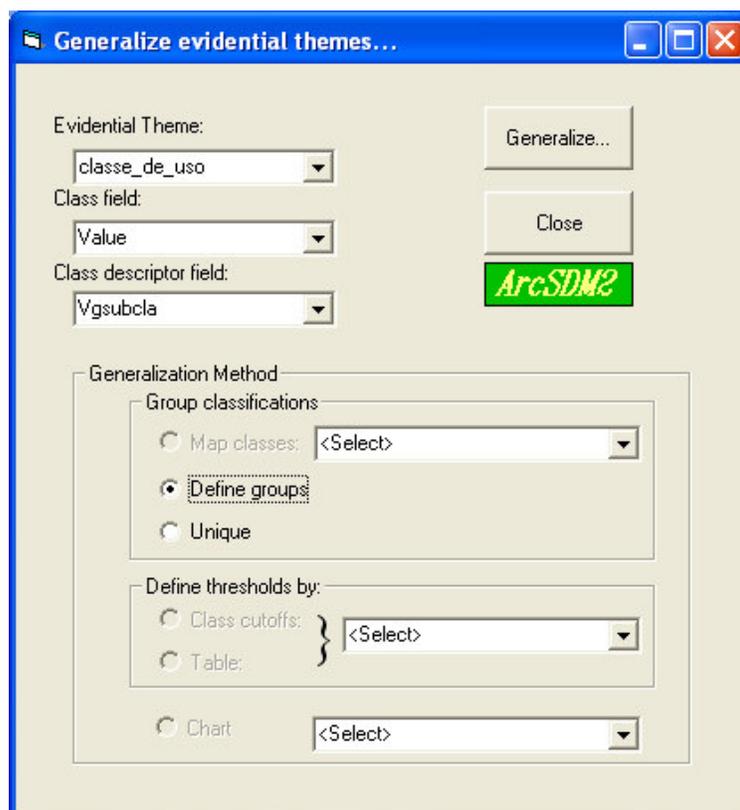


Figura 5.23 Iniciando “Generalize evidential themes...”

Esta opção apresenta a caixa de diálogo “*Generalize evidential themes...*”.



**Figura 5.24** Caixa de diálogo “*Generalize evidential themes...*”

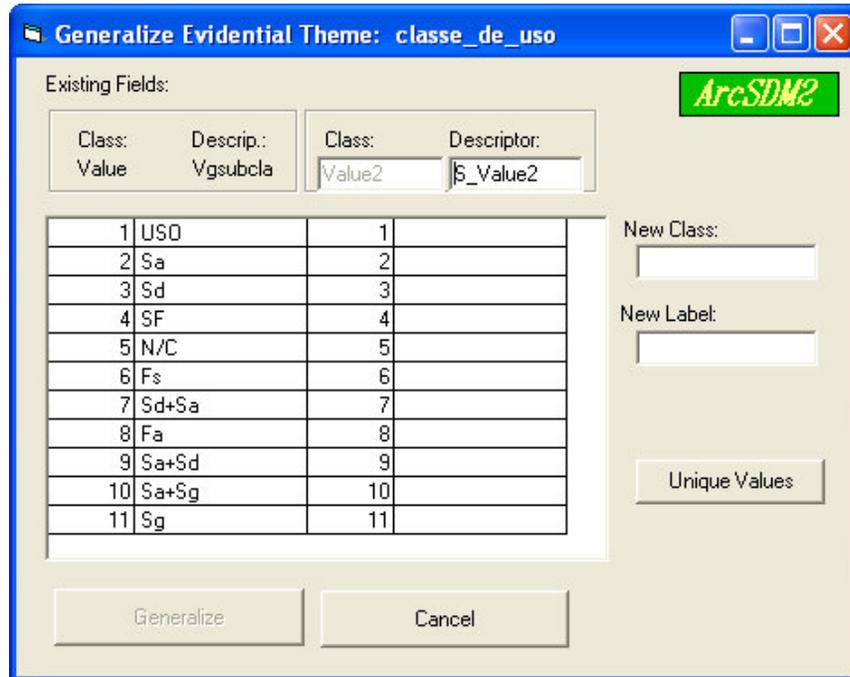
O campo “*Evidential Theme*” mostra todos os temas de evidências (*rasters*) do projeto que podem ser generalizados. O primeiro tema escolhido para ser generalizado foi “uso do solo”.

No campo “*Class field*” indica-se a coluna da tabela de atributo que possui as classes para serem generalizadas. O campo “*Class Descriptor field*” mostra qualquer campo do tipo “*string*” encontrado na tabela de atributos do tema de evidência que descreva a classe selecionada.

O ArcSDM possui alguns métodos de generalização que estão localizadas em “*Generalization Method*” na opção “*Group classification*”. Os métodos são “*map class*”, “*define groups*” e “*unique*” que são descritos a seguir.

\* **Map Class** – esse método se refere ao mapeamento classe por classe para generalização dos temas de evidência. Essa opção permite definir novas classes individualmente para cada classe do tema de evidência.

A opção “*Map Class*” faz aparecer a seguinte caixa de diálogo:



**Figura 5.25** Caixa de diálogo do método “*Map Class*” de generalização

A primeira coluna apresenta as classes encontradas no tema de evidência e que serão reclassificadas (coluna especificada anteriormente na caixa de diálogo “*class field*”). A segunda coluna se refere à descrição da classe (coluna especificada anteriormente na caixa de diálogo “*class descriptor field*”).

A terceira coluna se refere à coluna que contém os novos valores da generalização que serão editadas (inicialmente contém o mesmo valor da primeira coluna). E a última coluna conterá o texto de descrição das novas classes que forem adicionadas.

\* **Define groups** – esse método permite realizar-se uma reclassificação de um grupo de classes para uma nova classe. Os grupos de classes são definidos aplicando-se uma sentença (“*query*”) para os atributos. Esse foi o método escolhido para a generalização dos temas de evidência do estudo de caso na Bacia do Alto Taquari. A opção “*Define groups*” aciona a caixa de diálogo “*Group class*” apresentada a seguir.

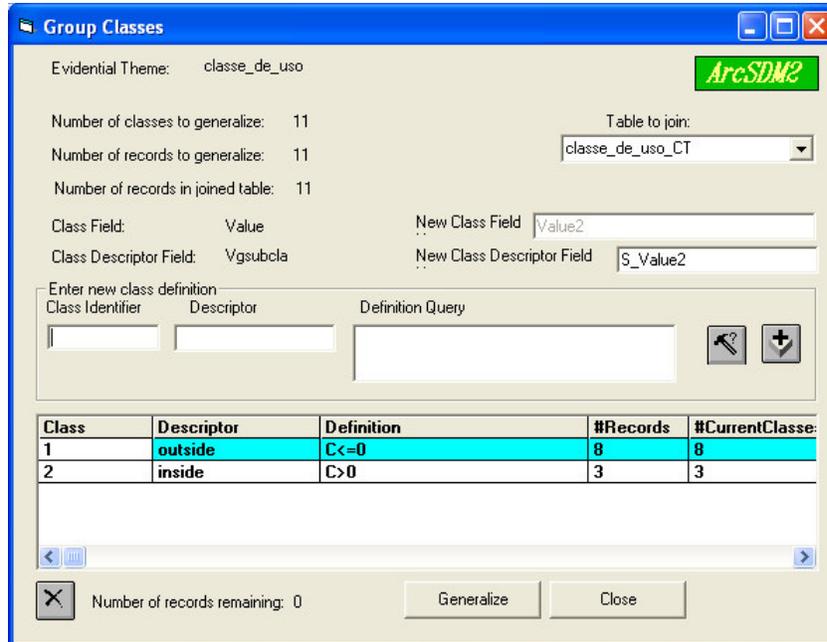


Figura 5.26 Caixa de diálogo “Group Classes”

Deve-se primeiramente indicar no campo “Table to join” a tabela com os pesos de evidência calculados para cada tema de evidência. Logo após ser indicada a tabela de pesos a mesma será aberta automaticamente pelo programa, a fim de ajudar na visualização e para facilitar a construção de sentenças para separar as novas classes. Os agrupamentos das classes podem ser baseados nos pesos, contrastes ou demais elementos encontrados na tabela de pesos associada.

O campo “New Class Field” corresponde à coluna que será gerada para armazenar as novas classes criadas. O valor de “New class field” é fixado como “Value2” e não pode ser alterado. O ArcSDM assume que o novo valor de generalização no arquivo “GEN” é sempre o campo “Value2”. O campo “New Class Descriptor Field” se refere à coluna de descrição das novas classes que serão criadas.

No canto superior da caixa de diálogo o campo “Number of records to generalize” reporta o número de registros na tabela de atributos dos temas de evidências e o “Number of classes to generalize” reporta o número de classes dos temas de evidências, que contém o mesmo (ou menor) número de registros.

O campo “Class Identifier” é o primeiro controle para a entrada da informação sobre a nova generalização. O campo aceita somente números inteiros. Não se pode fornecer nenhum valor neste campo antes de ser especificado a tabela de pesos. Como a intenção é generalizar as classes dos temas de evidência em um padrão binário foram criados dois valores para esse

campo, onde o valor um (1) indicou a ausência do padrão e o valor dois (2) indicou a presença do padrão. Embora seja mais fácil de interpretar uma convenção do tipo 2 = presença e 1 = ausência, qualquer valor inteiro pode ser usado para identificar as classes.

O campo “*Descriptor*” é usado para descrever a nova classe que se está criando. Não são exigidos dados nesta linha para criar uma generalização, porém utilizou-se como descrição a palavra “*outside*” para descrever a ausência do padrão e a palavra “*inside*” para descrever a presença do padrão.

Finalmente, no campo “*Definition Query*” é possível fornecer a sentença que se deseja usar para separar as classes em padrões binários. Para os temas de evidência utilizados na análise da BAT foi usada a sentença que estabelecia que se o contraste fosse menor ou igual a zero ( $C \leq 0$ ) o padrão era ausente e se o contraste fosse maior que zero ( $C > 0$ ) o padrão era presente.

Depois que for especificada uma sentença para definir o grupo de classes para a generalização, a mesma é verificada clicando no botão “*Compose definition query*” (simbolizado por um martelo) e em seguida o botão “*Add Generalization*” (simbolizado por um sinal de soma) fica verde e é habilitado. Ao clicar neste botão a sentença é adicionada em uma lista na caixa de diálogo.

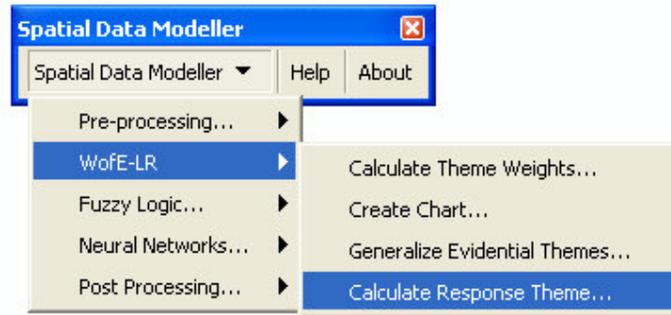
A coluna “*Records*” indica o número de registros na tabela do tema de evidência que são selecionados. Quando a sentença é verificada e adicionada o número de registros é subtraído do número de registros remanescentes.

O botão “*Generalize*” aplica a reclassificação baseada nas novas classes. Cada sentença é processada em seqüência. A sentença é aplicada nos atributos da tabela de pesos. As classes especificadas e as descrições dessas classes escrevem novos campos na tabela, conforme indicado nas sentenças. O botão de generalização somente fica ativo se o número de registros remanescentes for zero, ou seja, se todos os registros estiverem sendo contemplados pelas sentenças criadas.

\* **Unique** – O método “*Unique*” pode ser usado para definir uma generalização de valores únicos num tema, ou seja, fazer a generalização de todas as classes em uma única.

#### 5.7.4 Combinação dos temas de evidência

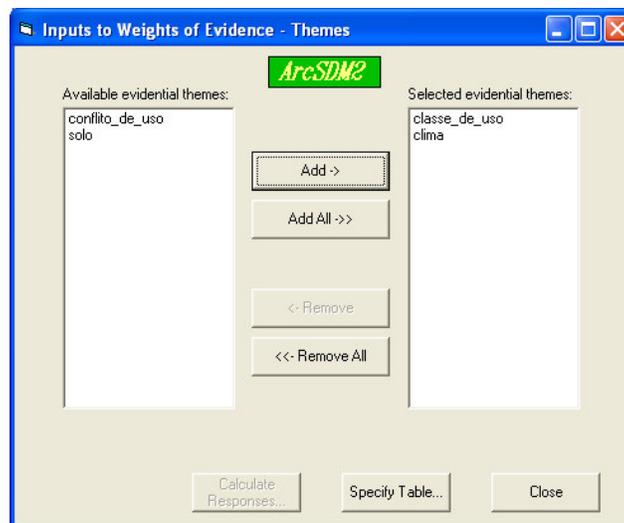
Para combinar o conjunto de temas de evidência em um *raster* de condição única foi selecionada a opção *Spatial Data Modeller > Wofe-LR > Calculate Response Theme*, conforme mostrado abaixo.



**Figura 5.27** Iniciando “*Calculate Response Theme*”

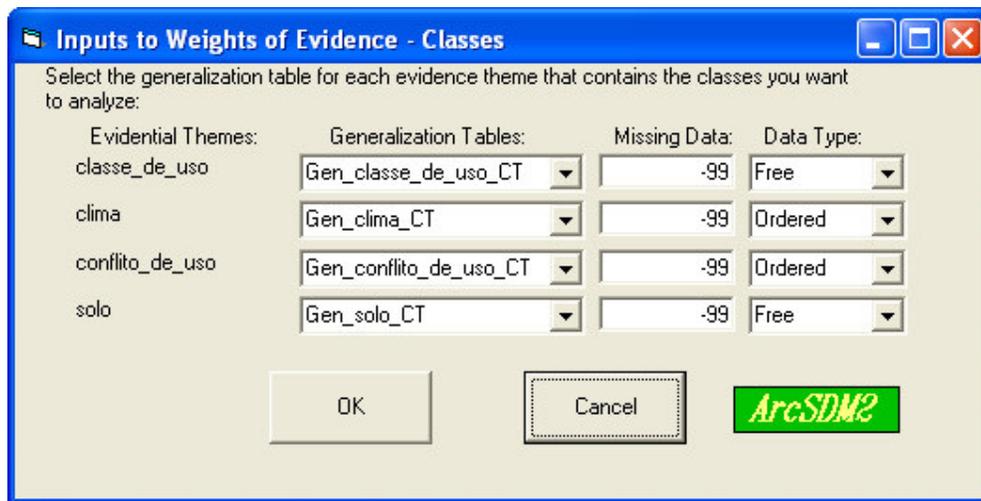
A caixa de diálogo “*Input to Weights of Evidence –Themes*” é utilizada para combinar o conjunto de temas de evidências em um *raster* de única condição pressupondo a independência condicional dos dados (uma breve teoria sobre independência condicional dos dados encontra-se na seção 3.2.2.2). Será apresentada a avaliação da independência condicional para este estudo de caso adiante, na seção 5.7.6.

Assim, foram selecionados os quatro temas de evidência e em seguida adicionados através do botão “*Add*” para indicar que eles seriam os temas utilizados para gerar o mapa de favorabilidade.



**Figura 5.28** Caixa de diálogo “*Inputs to Weights of Evidence – Themes*”

Depois de selecionados os temas de evidência foi habilitado o botão “*Specify Table*”, que abriu a caixa de diálogo “*Inputs to Weights of Evidence-Classes*”. Nesta caixa de diálogo foram selecionadas as tabelas de generalização que contém as classes reclassificadas de cada tema, conforme a próxima figura.



**Figura 5.29** Caixa de diálogo “*Inputs to Weights of Evidence - Classes*”

Em seguida, voltando na caixa de diálogo “*Inputs to Weights of Evidence – Themes*”, foi habilitado o botão “*Calculate Responses*”. Essa opção faz os temas de evidência serem combinados gerando um raster de única condição e algumas tabelas que relacionam as estatísticas do cálculo.

### 5.7.5 Avaliação das incertezas e geração do mapa de favorabilidade

Foram calculados pesos, variâncias e contrastes para cada tema de evidência e guardados em três tabelas: uma tabela de pesos (chamada WOE), uma tabela de variância (chamada WOEVAR) e uma terceira tabela (chamada WOFE) que representa as condições únicas. As tabelas são descritas a seguir.

#### \* Tabela de pesos (WOE)

O nome padrão para a tabela de pesos (WOE) é “SDMUC\_woe.dbf”. Essa tabela significa um resumo dos cálculos dos ponderadores de evidência (W+ e W-) para cada tema, ou seja, um resumo das estatísticas resultantes da aplicação da modelagem por Pesos de Evidência na área de estudo (BAT). A tabela WOE gerada na aplicação da BAT foi:

Evidence_N	Evidence_T	Class_Fiel	W-99	W1	W2	Contrast*	Confidence
uso	SDMTEMP3	Gen_uso_do_solo_CT	0,0000	-1,5348	0,1904	1,7252	1,6853
clima	SDMTEMP4	Gen_clima_CT	0,0000	-0,8339	0,9359	1,7698	3,8647
conflito	SDMTEMP5	Gen_conflito_de_uso_CT	0,0000	-0,4926	1,4436	1,9362	4,5173
solo	SDMTEMP6	Gen_solo_CT	0,0000	-1,4053	1,1512	2,5565	4,6235
	vocorocas_TOTAL	22	27817,6600	0,0008	0,0000	0,0000	0,0000

**Tabela 5.16** Tabela de pesos (WOE)

A coluna “*Evidence\_N*” representa o nome do tema de evidência e a coluna “*Evidence\_T*” representa o nome do raster temporário que foi criado para a generalização dos temas de evidências e usado como entrada para combinar e produzir o *raster* de única condição. Este nome aparece na tabela de atributo do *raster* de condição única. A coluna “*Class\_fiel*” é a coluna que registra o nome da tabela de generalização na qual os pesos foram calculados.

As colunas “W1” e “W2” representam os peso positivo (W+) e o peso negativo (W-), respectivamente, para cada tema de evidência de entrada. Embora essa tabela possa acomodar um número grande de temas de evidência, é recomendado que temas de evidência com múltiplas classes estejam limitadas em pequeno número (não mais que cinco) para facilitar a interpretação.

A coluna “*Contrast\**” registra o contraste (C) baseado na diferença entre os pesos calculados. A coluna “*Confidence*” registra a confiança (correlação espacial entre a área anômala e as ocorrências) que é dada pela razão entre o contraste e seu desvio padrão. Essa razão serve como um teste informal da hipótese de que o contraste seja igual a zero, e, portanto não estatisticamente significativo, pois nesse caso os pesos de evidência teriam valores iguais a zero. Segundo Bonham-Carter et. al. [16] se os valores da confiança são maiores que 1,96 rejeita-se a hipótese nula, ou seja, o dado é estatisticamente significativo. O valor de 1,96 representa um valor na distribuição normal e equivale a taxa de um número A (representando uma média) e um outro número B (representando duas vezes o desvio padrão de A). Assim, o valor de 1,96 representa uma significância estatística de 95% de probabilidade.

Na última linha da tabela de pesos (WOE) alguns parâmetros são escritos. O nome dos pontos de treinamento são escritos na última linha da coluna “*Evidence\_T*”. O número total de pontos de treinamento são escritos na coluna onde se encontram os nomes das tabelas de generalização. A área total de estudos é escrita na coluna “W-99” e a probabilidade a priori (número total de pontos de treinamento divididos pela área total de estudo) é escrita na coluna “W1”.

**\* Tabela de variância (WOEVAR)**

O nome padrão para a tabela de variância (WOEVAR) é “SDMUC\_woevar.dbf”. Esta tabela tem a mesma estrutura da tabela de pesos com três exceções: são reportados os valores de variância para cada tema de evidência, o contraste e confiança não são reportados e os pontos de treinamento e a área de estudos não são reportados na última linha

Value	Evidence_N	Evidence_T	Class_Fiel	V-99	V1	V2
1	classe	SDMTEMP3	Gen_classe_de_uso_CT	0,0000	1,0002	0,0477
2	clima	SDMTEMP4	Gen_clima_CT	0,0000	0,1429	0,0668
3	conflito	SDMTEMP5	Gen_conflito_de_uso_CT	0,0000	0,0834	0,1003
4	solo	SDMTEMP6	Gen_solo_CT	0,0000	0,2500	0,0557

**Tabela 5.17** Tabela de variância (WOEVAR)

**\* Tabela de pesos de evidência (WOFE)**

O ArcSDM calcula uma série de estatísticas para cada condição única e as escreve em uma tabela, que é a tabela de pesos de evidência (WOFE). A tabela possui o nome padrão “SDMUC\_wofe.dbf”. A tabela WOFE gerada na aplicação da BAT teve cinquenta e cinco classes. As primeiras vinte classes da tabela criada para os temas de evidência encontram-se abaixo.

ID	Post_Prob	Post_Logit	Sum_Wts	Uncrty	Msnrg_Data	Tot_Uncrty
1	0,00079086	-7,14159272	0,00000000	0,00016848	0,02810973	0,02811023
2	0,00034367	-7,97549272	-0,83390000	0,00014910	0,01853466	0,01853526
3	0,00021002	-8,46809272	-1,32650000	0,00010946	0,01449038	0,01449079
4	0,00006233	-9,68299272	-2,54140000	0,00004703	0,00000000	0,00004703
5	0,00010200	-9,19039272	-2,04880000	0,00007111	0,01009914	0,01009939
6	0,00001110	-11,40819272	-4,26660000	0,00001370	0,00000000	0,00001370
7	0,00001817	-10,91559272	-3,77400000	0,00002180	0,00426286	0,00426291
8	0,00080289	-7,12649272	0,01510000	0,00049137	0,00000000	0,00049137
9	0,00014312	-8,85169272	-1,71010000	0,00016488	0,00000000	0,00016488
10	0,00002556	-10,57429272	-3,43270000	0,00003002	0,00505603	0,00505612
11	0,00553952	-5,19029272	1,95130000	0,00344931	0,00000000	0,00344931
12	0,00099131	-6,91549272	0,22610000	0,00114833	0,00000000	0,00114833
13	0,00043194	-7,74679272	-0,60520000	0,00033061	0,00000000	0,00033061
14	0,00007697	-9,47199272	-2,33040000	0,00009548	0,00000000	0,00009548
15	0,00004527	-10,00289272	-2,86130000	0,00005105	0,00672788	0,00672808
16	0,00023481	-8,35649272	-1,21490000	0,00013752	0,01532152	0,01532214
17	0,00004184	-10,08169272	-2,94010000	0,00004762	0,00646798	0,00646816
18	0,00014349	-8,84909272	-1,70750000	0,00009370	0,01197781	0,01197818
19	0,00048340	-7,63419272	-0,49260000	0,00017344	0,02198023	0,02198092
20	0,00007408	-9,51029272	-2,36870000	0,00008076	0,00860661	0,00860698

**Tabela 5.18** Tabela de pesos de evidência (WOFE)

A coluna “ID” indica o identificador de cada classe gerada. A coluna “Post\_prob” indica a probabilidade posteriori calculada. A coluna “Sum\_Wts” é a soma dos pesos para cada classe de tema de evidência ocorrendo na única condição. A coluna “Uncrty” indica a incerteza devido ao cálculo dos pesos (desvio padrão). A coluna “Msnrg\_Data” indica a

incerteza devido aos dados *missing* (desvio padrão) e a coluna “*Tot\_Uncrty*” é a incerteza total dada pela combinação das incertezas dos pesos e dos dados *missing* (desvio padrão).

Após salvar as tabelas foi gerado o mapa *raster* de condição única onde as unidades de células (pixels) com a mesma probabilidade posteriori formam classes que pertencem a 2<sup>p</sup> possíveis condições únicas, onde “p” significa a quantidade de padrões binários considerados na análise. Por exemplo, a combinação de 3 temas binários produzem um mapa com 8 únicas condições.

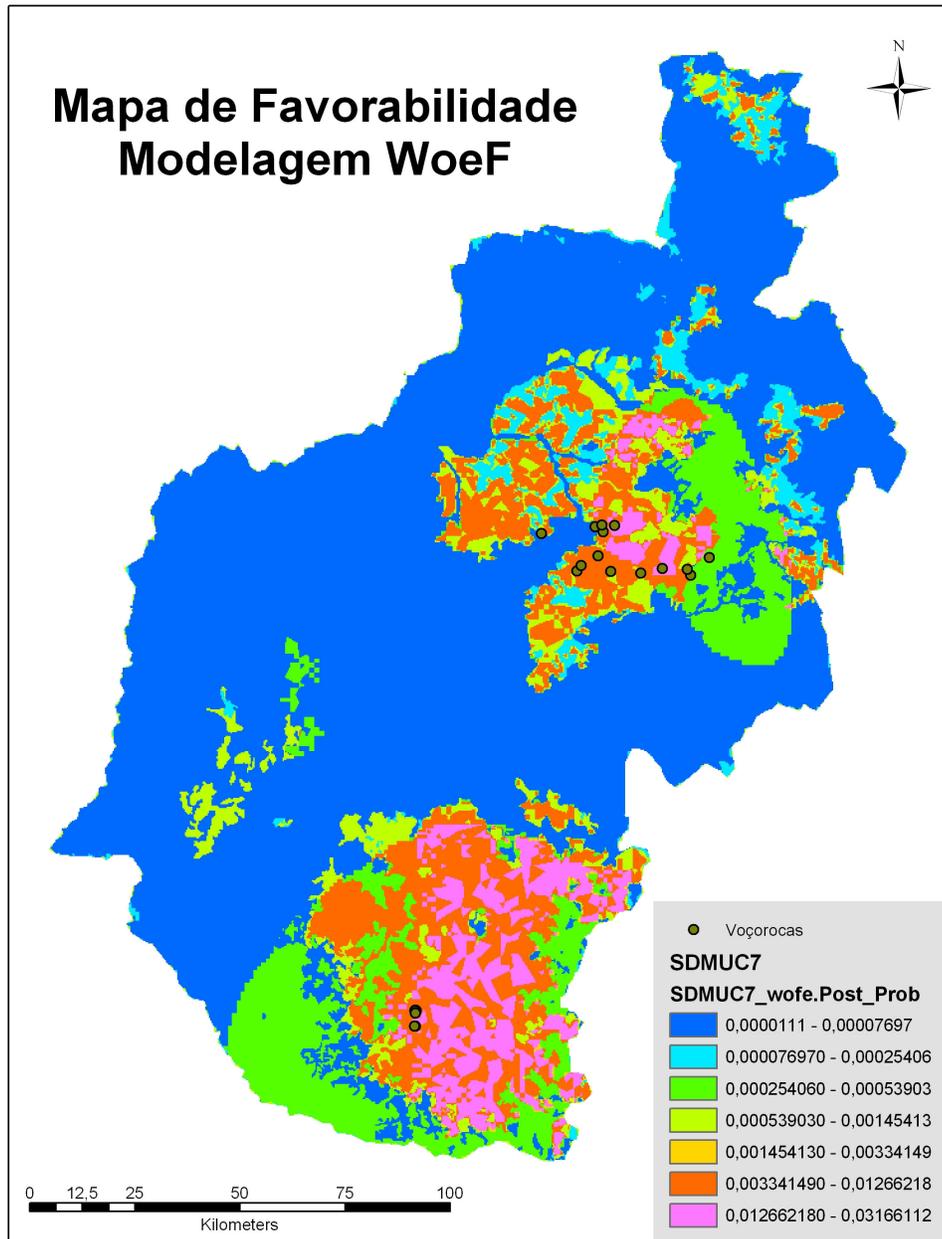
A tabela do *raster* de condição única possui valores de classes que são simbolizados com o número 2 para a presença do padrão e 1 para a ausência do mesmo. O número de condições únicas aumenta nitidamente se algum dos temas de evidência é multi-classes. No estudo de caso da BAT foram geradas cinquenta e cinco condições únicas. As vinte primeiras encontram-se listadas abaixo.

ObjectID	Value	Count	Sdmtemp3	Sdmtemp4	Sdmtemp5	Sdmtemp6
0	1	7080	-99	-99	-99	-99
1	2	549	-99	1	-99	-99
2	3	615	-99	1	1	-99
3	4	1313660	2	1	1	1
4	5	1176	2	1	-99	1
5	6	350972	1	1	1	1
6	7	476	1	1	-99	1
7	8	104652	2	1	1	2
8	9	82526	1	1	1	2
9	10	19477	1	-99	1	1
10	11	125916	2	1	2	2
11	12	12241	1	1	2	2
12	13	15068	2	1	2	1
13	14	1066	1	1	2	1
14	15	24	1	1	1	-99
15	16	898	2	-99	-99	1
16	17	210	1	-99	-99	1
17	18	9650	2	-99	1	1
18	19	155	-99	-99	1	-99
19	20	12	1	1	-99	-99
20	21	1	2	-99	-99	-99

**Tabela 5.19** Tabela de condição única para BAT

A tabela do *raster* de condição única é automaticamente conectada a tabela de pesos WOFÉ. O *raster* de condição única é então simbolizado em um de seus atributos para criar o tema de resposta. O produto final do modelo de Pesos de Evidência é um *raster* de condição

única simbolizado de acordo com a probabilidade “a posteriori” calculada. No estudo de caso da Bacia do Alto Taquari (BAT) o mapa *raster* de condição única foi simbolizado baseado na coluna “*Post\_Prob*” que possui os valores de probabilidade “a posteriori” calculados. Foram usadas sete classes definidas pelo método “*Natural Breaks*” do SIG *ArcView* para a simbolização, resultando no mapa apresentado abaixo.



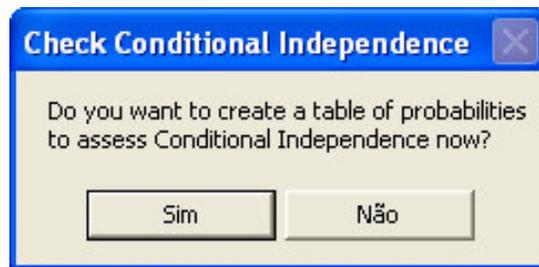
**Figura 5.30** Mapa de favorabilidade da modelagem Peso de Evidência (Wofe)

Observa-se que as áreas em tons de vermelho e rosa são mais favoráveis à ocorrência de novos pontos de erosão.

### 5.7.6 Verificação da independência condicional

O último passo é a observação da independência condicional dos dados. A independência condicional estabelece a independência de dois termos dada à presença de um terceiro. O modelo de Pesos de Evidência exige a condição de independência entre as camadas de dados e temas considerados na modelagem do problema, no caso erosão. Na prática, essa condição é violada de alguma forma. Alguns testes são usados para quantificar o nível máximo de aceitação da violação da condição de independência [14].

Uma vez que o mapa de favorabilidade está completo, o ArcSDM reporta duas medidas de independência condicional. A primeira medida é uma taxa de independência condicional (CI), que pode ser usada como uma avaliação global da independência condicional entre os dados, e a segunda medida é o teste Agterberg-Cheng de independência condicional.



**Figura 5.31** Caixa de diálogo “*Check Condicional Independence*”

Para o cálculo das probabilidades e a verificação da independência condicional dos dados, o ARCSHM gera os valores de (T), de (n) e de (T-n). Esses termos são explicados no item 3.2.2.2 do terceiro capítulo deste trabalho. Os valores destes termos são gerados a partir de duas tabelas obtidas no processamento do modelo: a tabela SDMUC\_TBL e a tabela SDMUC\_CI.

A tabela SDMUC\_TBL traz informações a respeito da área, medida em metros quadrados, e o número de pontos de treinamento que ocorre em cada única condição.

A tabela SDMUC\_CI reporta duas colunas. A coluna “*PrdPntFreq*” é o produto gerado entre o campo “*Post\_Prob*” da tabela SDMUC\_Wofe (que corresponde a probabilidade posterior) e a coluna “*Area\_sqm*” da tabela SDMUC\_TBL (área de cada condição única). A coluna “*VarPntFreq*” traz a variância de cada única condição. Os quinze primeiros registro destas tabelas encontram-se listados a seguir.

Value	Area_sqm	TrngPoints
1	70800000,00	0
2	5490000,00	0
3	6150000,00	0
4	13136600000,00	1
5	11760000,00	0
6	3509720000,00	1
7	4760000,00	0
8	1046520000,00	0
9	825260000,00	0
10	194770000,00	0
11	1259160000,00	5
12	122410000,00	0
13	150680000,00	0
14	10660000,00	0
15	240000,00	0

Value	PrdPntFreq	VarPntFreq
1	0,0560	3,9609
2	0,0019	0,0104
3	0,0013	0,0079
4	0,8188	0,3817
5	0,0012	0,0141
6	0,0390	0,0023
7	0,0001	0,0004
8	0,8402	0,2644
9	0,1181	0,0185
10	0,0050	0,9698
11	6,9751	18,8637
12	0,1213	0,0198
13	0,0651	0,0025
14	0,0008	0,0000
15	0,0000	0,0000

**Tabela 5.20** Tabelas SDMUC\_TBL e SDMUC\_CI

O somatório da coluna “*PrdPntFreq*” corresponde ao valor do número esperado de pontos de treinamento (T), sendo calculado o valor de 63,604 para a análise da BAT. O somatório da coluna “*VarPntFreq*” corresponde a variância, sendo calculado o valor de 600,98 aproximadamente. Assim foi possível calcular o desvio padrão de T, a partir da raiz quadrada da variância (correspondendo a 24,516). Como o valor observado de pontos de treinamento (n) é 22, o valor calculado para (T-n) foi de 41,604 (= 63,604 – 22 ).

A primeira medida reportada pelo ArcSDM é a taxa de independência condicional. A taxa é calculada dividindo-se o número atual de pontos de treinamento (n) pelo número de pontos esperados do modelo (T). Esta taxa geralmente tem valor entre 0 e 1. O valor de 1 (que nunca ocorre na prática) indica independência condicional entre os temas de evidência usados no modelo. Valores muito diferentes de 1 indicam um problema na independência condicional. O valor calculado para a taxa de independência condicional no presente estudo foi 0,346 ( $n/T = 22 / 63,604$ ).

A segunda medida reportada pelo ArcSDM se refere ao teste de Agterberg-Cheng de independência condicional. O teste de Agterberg-Cheng reporta a probabilidade do modelo não ser condicionalmente independente, baseado na razão de T-n sobre o desvio padrão de T. Assim o valor da razão de T-n/SDT foi de 1,697 (= 41,604 / 24,516). Uma grande probabilidade indica que o tema de resposta tem dependência condicional significante.

O ArcSDM reporta um resumo da independência condicional dos dados:

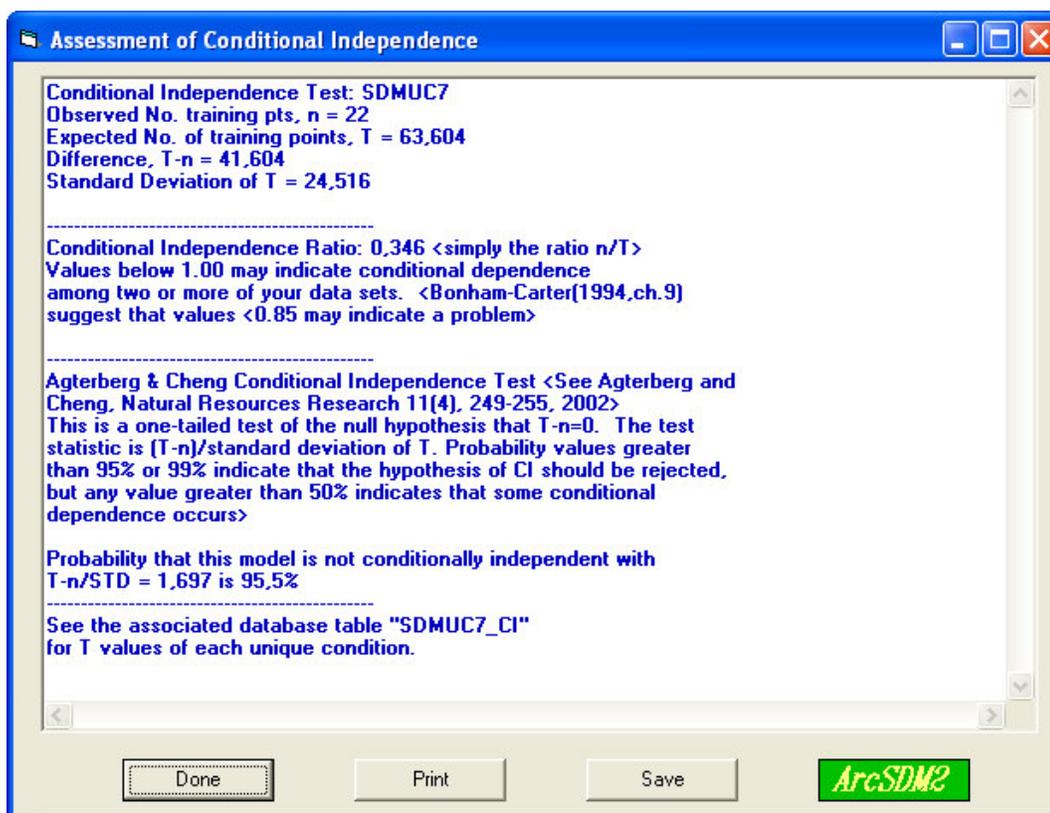


Figura 5.32 Caixa de diálogo “Assessment of Condicional Independence”

## 5.8 Considerações finais

Neste capítulo foi apresentada a aplicação da metodologia proposta na dissertação e apresentada no capítulo quatro, bem como os resultados alcançados para o estudo de caso na Bacia do Alto Taquari.

Para uma avaliação mais abrangente a respeito da metodologia, no próximo capítulo será feita uma breve discussão a respeito dos dados utilizados na análise bem como uma discussão dos resultados obtidos, procurando-se levantar os pontos positivos e negativos de cada uma das técnicas utilizadas no trabalho.

Será realizada ainda uma análise comparativa com resultados obtidos por outros trabalhos na mesma área de estudos procurando avaliar se os resultados encontrados neste trabalho estão em consonância com os demais. Em seguida serão feitas as considerações finais a respeito da aplicabilidade destes métodos para subsidiar estudos de erosão em uma bacia e sugestões para trabalhos futuros visando aprimorar os procedimentos propostos neste trabalho.

## CAPÍTULO 6 – RESULTADOS E DISCUSSÃO

### 6.1 Resultados

Os mapas de evidências (indicadores) utilizados neste trabalho demonstraram boa aplicação na modelagem das áreas de erosão na bacia no Alto Taquari, através dos Sistemas de Informações Geográficas, pelo fato de serem considerados representativos do processo erosivo e por poderem definir condições favoráveis para a ocorrência do processo de erosão.

Fazendo-se uma análise visual nos dois mapas gerados pelos modelos de Lógica “Fuzzy” e pelo modelo de Pesos de Evidência, figura 6.1, foi observada a existência de várias áreas que possuem potencial elevado de erosão na bacia que coincidem com áreas de erosão identificadas em campo, significando um bom ajuste entre o modelo de dados (estatístico) e a realidade.

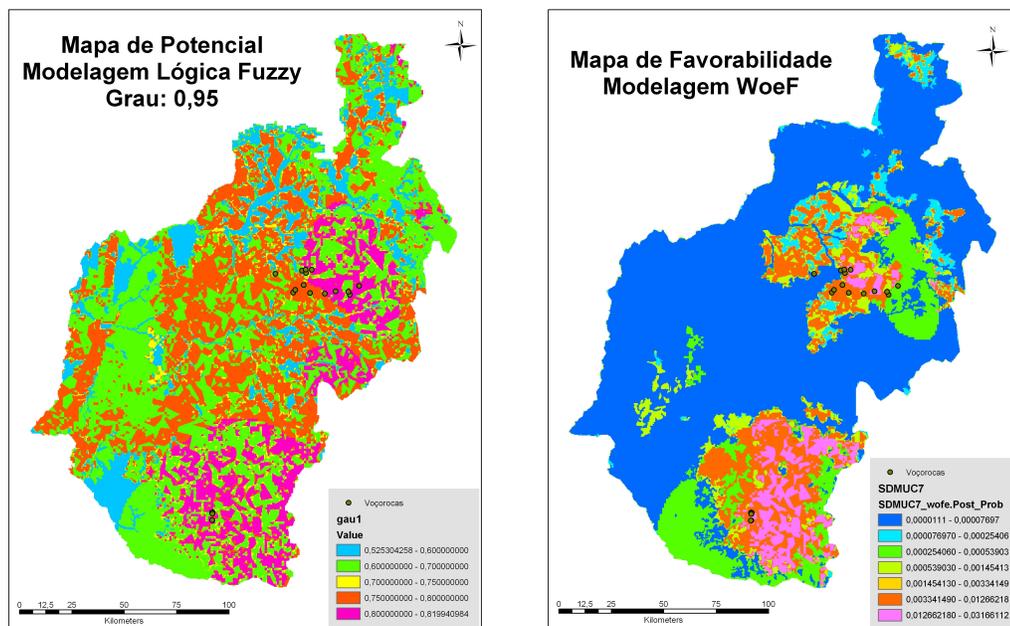


Figura 6.1 Mapas gerados no estudo de caso da BAT

Foi percebido na análise de Pesos de Evidência que nas tabelas de cálculo dos pesos para os temas de evidência (tabela 5.12 até tabela 5.15), com exceção do indicador conflito de uso, todos os outros indicadores possuem classes onde os pesos e seus respectivos contrastes não foram calculados. Isto ocorre devido à inexistência de pontos de treinamentos (voçorocas) nessas áreas. O ArcSDM só calcula os pesos de evidências para aquelas classes que possuam pontos de treinamento.

A tabela 6.1 apresenta um resumo dos pesos calculados pelo método de Pesos de Evidência para cada indicador ambiental investigado como potencialmente preditivo ao processo de erosão, bem como os valores de contraste e confiança para os mesmos.

Mapas de evidência	W+	W-	contraste	confiança
Erosividade da chuva	-0,8339	0,9359	1,7698	3,8647
Conflito de uso	-0,4926	1,4436	1,9362	4,5173
Solo	-1,4053	1,1512	2,5565	4,6235
Uso do solo	-1,5348	0,1904	1,7252	1,6853
Área total: 27817,66 Km <sup>2</sup> . Probabilidade prévia: 0,00079				

**Tabela 6.1** Tabela resumo dos pesos positivos e negativos para cada indicador

O elevado valor do contraste obtido para o mapa de “solo” comprovou a forte associação espacial entre os pontos de localização das voçorocas e esse tema de evidência. A segunda maior medida de contraste foi obtida para o mapa de “conflito de uso”, embora esse tema tenha apresentado um baixo valor de ponderador positivo de evidência (W+) e alto valor de ponderador negativo de evidência (W-), significando que a ausência do padrão binário é mais importante que sua presença, ou seja, onde não devemos procurar ocorrências de voçorocas ou pontos de erosão. Os mapas de “erosividade da chuva” e “uso do solo” apresentaram valores de contraste também altos, demonstrando que podem ser usados também na estimativa da ocorrência de novos pontos de erosão.

Foi observado também o valor da confiança para cada um dos temas de evidência. Como dito anteriormente, segundo Bonham-Carter et. al. [16], a confiança (razão do contraste pelo seu desvio padrão) serve como um teste informal para verificar a importância dos dados utilizados. Se os valores da confiança são maiores que 1,96 o dado é considerado estatisticamente significativo com respeito aos pontos de treinamento utilizados. Baseada nesta teoria, pode-se observar que quase todos os valores de confiança calculados para os temas de evidência são estatisticamente significantes, pois apresentam confiança maior que 1,96. Somente o tema uso do solo apresentou a confiança ligeiramente menor que esse limite.

No mapa gerado pela modelagem de Pesos de Evidência no sistema computacional ArcSDM foi observado que grande parte da área de estudo com alta favorabilidade ao surgimento de novos pontos de erosão coincide com as ocorrências conhecidas, situados na região central e na parte sul da bacia, mostrando que foram bem modelados pelo método de Pesos de Evidência.

Nota-se que, segundo esta modelagem, a região oeste da bacia apresenta pontos de baixa probabilidade de ocorrência de pontos de erosão, e na modelagem “*Fuzzy*”, nesta mesma parte da bacia, há pontos com grandes índices de potencial. Isso pode significar ausência de informações relevantes nesta área.

Comparando-se os valores de probabilidade posteriori calculado pelo método de Pesos de Evidência com o valor da probabilidade prévia (probabilidade a priori = 0,00079), observa-se que existem áreas onde a favorabilidade gira em torno de 0,03 (cerca de 40 vezes mais favoráveis que a probabilidade a priori da área).

Os valores do cálculo da independência condicional dos dados mostram que os testes de independência condicional não foram satisfatórios. No cálculo da taxa de independência condicional foi verificado um valor de 0,346, o que pode indicar um problema com os dados utilizados uma vez que valores menores que 1 podem indicar dependência dos dados. Acredita-se que a quantidade reduzida de pontos de treinamento tenha interferido no cálculo da independência condicional dos dados, uma vez que foram utilizados somente 22 pontos (pontos de voçorocas identificadas no campo). Os pontos de treinamento utilizados também não estão distribuídos pela bacia. Os pontos se encontram concentrados em duas áreas da bacia (parte sul e parte central da bacia), podendo também ter interferido no cálculo da independência condicional dos dados.

No segundo cálculo de independência condicional dos dados pelo método de Agterberg e Cheng foi encontrada uma probabilidade de 95% com uma taxa do valor de  $T-n/SDT = 1,697$ . Probabilidades com valores maiores que 95% ou 99% indicam que a hipótese de independência condicional pode ser rejeitada.

Esta situação não é totalmente inesperada, pois os dados ambientais podem ter alguma forma de inter-relação, como por exemplo, o dado de uso do solo com o dado de solo, ou uso do solo com erosividade da chuva. Para testar essas possibilidades e identificar se existe e qual dos dados apresentam dependência, poderia ser utilizado o teste estatístico Chi-quadrado ( $\chi^2$ ), que compara pares os dados. Porém este teste não se encontra no sistema utilizado.

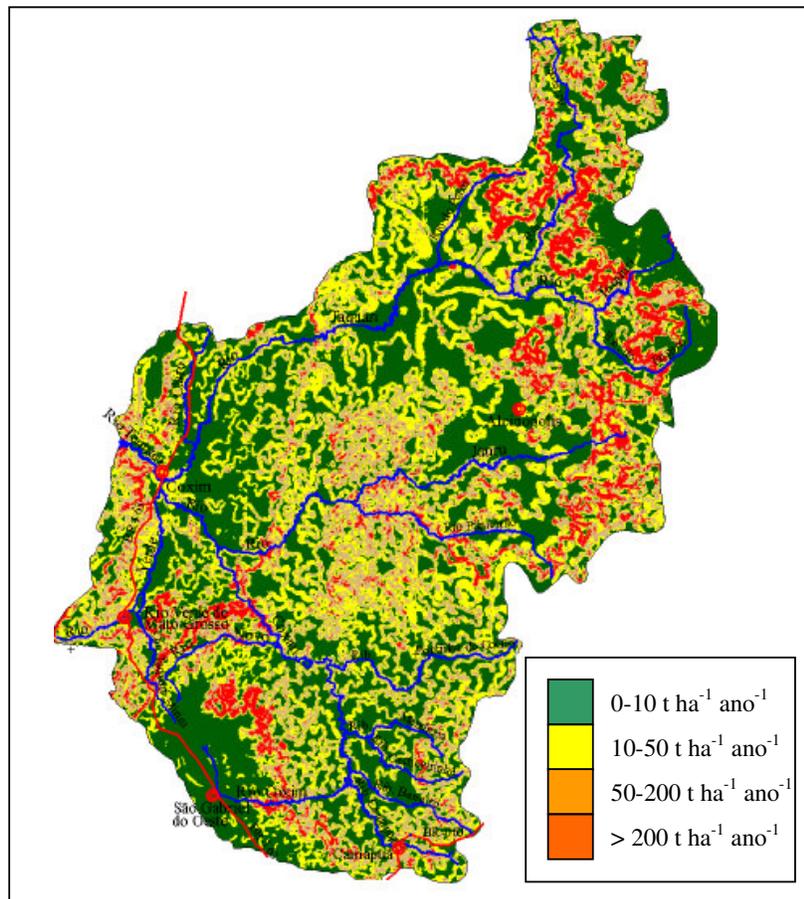
## **6.2 Análise comparativa com outros resultados**

### **6.2.1 Mapa de perda de solo**

De acordo com o Boletim de desenvolvimento e Pesquisa da Embrapa [36], que contém um mapeamento das perdas de solo na Bacia do Alto Taquari, foi determinada a ocorrência de diferentes graus de perda de solo na bacia.

Segundo o boletim, a perda média de solo na BAT em 1994 foi estimada a partir da equação universal de perda de solo (*Universal Soil Loss Equation - USLE*), chegando-se a um valor de 70,39 t ha<sup>-1</sup> ano<sup>-1</sup>, que corresponde a um grau de erosão alto (50 a 200 t ha<sup>-1</sup> ano<sup>-1</sup>) [36].

O grau de erosão foi classificado em quatro níveis onde os valores de perda de solo menores que 10 t ha<sup>-1</sup> ano<sup>-1</sup> foram classificados como ligeiro, os valores entre 10 até 50 t ha<sup>-1</sup> ano<sup>-1</sup> foram classificados como moderado, os valores de 50 até 200 t ha<sup>-1</sup> ano<sup>-1</sup> foram classificados como alto e os acima de 200 t ha<sup>-1</sup> ano<sup>-1</sup> foram classificados como muito alto.



**Figura 6.2** Mapa de distribuição de perdas de solos na BAT em 1994  
Fonte: adaptado de [36]

Segundo Galdino et. al. [36] os valores quantitativos de perda de solo gerados pela USLE ou outros modelos de simulação, devem ser considerados como estimativas para fins comparativos.

Pode-se perceber que as áreas identificadas no mapa de erosão do boletim da Embrapa [36] estão coerentes com as áreas de alta probabilidade de erosão na bacia determinadas pela

modelagem “*Fuzzy*”. Pode-se perceber ainda uma tendência ao crescimento da erosão na parte sul da bacia quando se compara o mapa da Embrapa com o mapa gerado pela análise “*Fuzzy*”.

Dentre as conclusões destacadas pelo Boletim de desenvolvimento e Pesquisa da Embrapa [36] destaca-se que a perda de solo nas áreas ocupadas por pastagens foi superior ao dobro das ocorridas em áreas cultivadas com soja e o elevado índice de perda de solo em áreas ocupadas com vegetação nativa, foi decorrente principalmente do relevo mais acentuado. A incorporação dessas áreas ao sistema agropecuário representa grande risco de intensificação dos processos erosivos na BAT, fato que ocorre com maior intensidade nos dias de hoje fazendo agravar ainda mais o processo de erosão na bacia [36].

É importante destacar que a geração do mapa de perda de solo na BAT publicada neste boletim leva em consideração a utilização de indicadores diferentes dos utilizados na análise feita no sistema computacional ArcSDM desta dissertação. Porém, os mapas resultantes dos dois trabalhos podem ser utilizados para fins comparativos.

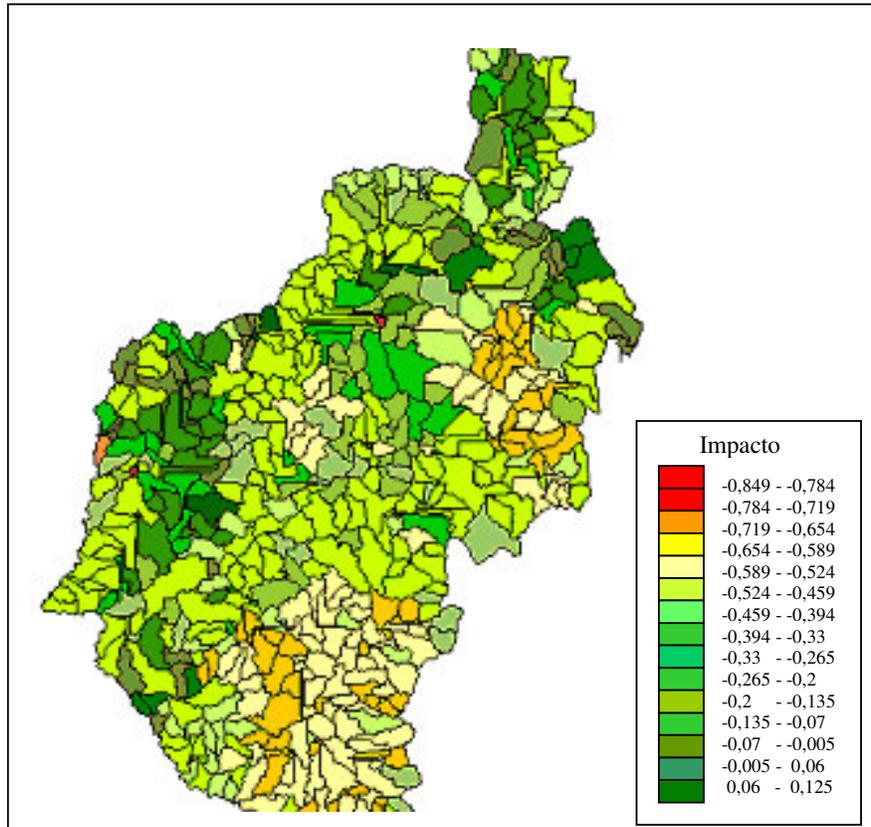
Deve-se também levar em consideração o fator temporal dos dados. O boletim de pesquisa da Embrapa traz como resultados dados que foram publicados em 1994, como dito anteriormente. Já os dados utilizados neste trabalho levam em consideração dados mais recentes, atualizados em 2000 por Silva [84]. Deve-se levar em consideração que ao longo do tempo podem ocorrer variações climáticas, de vegetação, e de outros fatores que podem contribuir para alteração da análise.

### **6.2.2 Mapa de impacto na BAT a partir do SAEB (Sistema de Análise do Risco de Impactos Ambientais Derivados da Erosão em Bacias Hidrográficas)**

Outro mapa utilizado para comparação dos resultados obtido neste trabalho foi o mapa desenvolvido pelo Sistema de Análise do Risco de Impactos Ambientais Derivados da Erosão em Bacias Hidrográficas (SAEB) [62]. O SAEB tinha por objetivo realizar uma análise para auxiliar tomadas de decisão com relação à priorização de áreas que sofrem impactos erosivos, tendo como estudo de caso a Bacia do Alto Taquari.

O SAEB utilizou o sistema computacional EMDS (Environment Manager Decision Support – Suporte à decisão para Gestão ambiental) e se baseou no desenvolvimento de uma estrutura hierárquica de indicadores, que foram organizados segundo o modelo estrutural conhecido como “Pressão-Estado-Resposta (P-E-R)”, desenvolvido pela OECD (Organization for Economic Co-operation and Development) [67].

Os indicadores ambientais de Pressão utilizados por este sistema foram densidade de estradas e conflito de uso, os indicadores de Estado utilizados foram mata ciliar, densidade de drenagem, solos, cobertura vegetal e clima e o indicador ambiental de Resposta utilizado foi práticas de conservação do solo. O resultado deste estudo foi um mapa de impacto na Bacia do Alto Taquari que pode ser observado na figura 6.3.



**Figura 6.3** Mapa de impacto na BAT a partir do SAEB  
Fonte: adaptado de [62]

Pode-se perceber que as áreas identificadas no mapa de impacto gerado pelo SAEB coincidem com as áreas de alta probabilidade de erosão na bacia determinadas pela modelagem “*Fuzzy*”. Apesar dos indicadores utilizados não terem sido os mesmos é possível fazer uma análise visual e verificar que existem áreas de erosão coincidentes.

### 6.2.3 Utilização de pontos de controle

Para a avaliação dos mapas de potencialidade gerados pensou-se em utilizar alguns dos pontos de treinamento conhecidos na BAT (voçorocas), para realizar ao final da modelagem, uma comparação através da sobreposição dos mapas gerados nas modelagens

(que indicam as áreas mais favoráveis a ter erosão) com as verdades de campo através dos pontos de treinamento que não foram utilizados.

A intenção era verificar a correlação das ocorrências de pontos de erosão com as classes dos mapas de potencialidade definidos como de alto potencial, avaliando visualmente o grau de precisão que os mapas gerados tiveram.

Devido ao número reduzido de pontos de treinamentos (22 pontos de voçorocas utilizados) comparado a extensão da área a ser analisada (cerca de 28000 Km<sup>2</sup>), tal verificação não pode ser feita, uma vez que a retirada de um ou mais pontos para o teste iria prejudicar na execução do modelamento, contribuindo para que o valor de independência condicional dos dados piorasse.

## **CAPÍTULO 7 – CONCLUSÕES E RECOMENDAÇÕES**

### **7.1 Conclusões sobre a análise**

Os resultados gerados pela Lógica “*Fuzzy*” e pela técnica de Pesos de Evidência se mostraram apropriados como subsídios à avaliação do processo erosivo em bacias hidrográficas.

A Lógica “*Fuzzy*” é uma eficiente ferramenta espacial de suporte para o processo de tomada de decisão. Este trabalho examinou a incorporação dessa tecnologia na aplicação de avaliação de processo de erosão em área ambiental baseada num ambiente SIG. Foi observado como podem ser adotados os conceitos da Lógica “*Fuzzy*” para a representação e análise de dados geográficos cuja incerteza é uma das características.

A Lógica “*Fuzzy*” é uma técnica que mostra maior liberdade para se definir os pesos e valores de pertinência para as classes dos temas de evidência utilizados em uma análise. Através desses valores de pertinência é possível identificar quais classes dos temas têm maior influência na ocorrência de um determinado evento.

Uma das vantagens da utilização da Lógica “*Fuzzy*” é a geração de diferentes cenários a partir das variações dos valores de pertinência e da possibilidade de utilização de diferentes operadores na integração temática. Isso faz com que o modelo de Lógica “*Fuzzy*” permita a utilização de mais de um modelo exploratório de dados e ajude no processo de tomada de decisão.

A utilização da Lógica “*Fuzzy*” neste trabalho teve por objetivo capturar e operar com a diversidade, a incerteza e as verdades parciais, a partir dos membros “*Fuzzy*”, em vez de operar somente com a opção de totalmente verdadeira ou totalmente falso, como é o caso da lógica clássica.

No mapa gerado a partir do uso de Lógica “*Fuzzy*” no sistema computacional ArcSDM foi observado que os pontos de erosão previamente conhecidos (voçorocas identificadas em campo) coincidem com as áreas de maior potencial de surgimento de novos locais de erosão, como pode ser verificado na figura 5.18.

A técnica de Pesos de Evidência utilizou estatística para a predição de novos locais de erosão na Bacia do Alto Taquari, onde algumas ocorrências representativas de erosão eram conhecidas. Esta técnica, originalmente utilizada em aplicações não-espaciais de diagnósticos médicos, e atualmente muito utilizada no mapeamento de favorabilidade mineral em ambientes geológicos, se mostrou satisfatória para estudos de impactos ambientais como a avaliação de processos erosivos, uma vez que conseguiu realizar a integração temática de

dados ambientais, gerando o mapa de favorabilidade para a área de estudos. Pode-se concluir que a técnica de Pesos de Evidência pode ser utilizada em qualquer área de conhecimento onde se tenha variáveis espaciais e um conjunto de fenômenos conhecidos na área de estudos.

A técnica de Peso de Evidência traz vantagens e desvantagens. Como vantagem pode-se citar que este método é dirigido pelos dados, logo evita a estimativa subjetiva dos pesos de evidência. Outra vantagem é o fato das medidas de incerteza também poderem ser mapeadas. Além disso, o processo de reclassificação binária otimiza o contraste dando maior percepção no relacionamento espacial dos temas de evidência.

Por outro lado, tem a desvantagem de ser uma técnica que requer um conjunto de pontos de treinamento suficientemente grande para o modelamento e que faz a suposição de independência condicional dos dados. Além disso, a técnica de Pesos de Evidência geralmente faz o uso de padrões binários, o que pode fazer com que algumas informações das evidências sejam perdidas depois da reclassificação de dados (generalização), uma vez que este processo pode ser feito subjetivamente ou através de métodos estatísticos.

Uma das limitações para a utilização dos Sistemas de Informação Geográfica no Brasil, não só na área de meio ambiente como em outras áreas, diz respeito às bases de dados que geralmente não são atualizadas, ou se encontram em formato analógico ou simplesmente não existem. A utilização de um Sistema de Informação Geográfica pressupõe a utilização de mapas digitais sendo fundamental manter as bases de dados sempre atualizadas.

Nesse sentido, foi observada uma dificuldade na atualização dos dados do indicador erosividade da chuva, devido à pouca disponibilidade de dados de chuva, que são obtidos através de estações meteorológicas, dificultando assim o cálculo do valor da erosividade da chuva na região de estudo.

## **7.2 Recomendações**

Os métodos e procedimentos utilizados neste trabalho são úteis nas aplicações envolvendo a análise ambiental, tendo em vista a necessidade de se quantificar e simular a importância relativa de variáveis ambientais na análise de processos de erosão.

A utilização das técnicas do aplicativo ArcSDM serviu de grande ajuda para as modelagens feitas neste trabalho. O objetivo principal da utilização destas técnicas foi fazer com que as informações espaciais fossem integradas da melhor maneira, de tal forma que fosse possível fazer um constante aprimoramento dos modelos utilizados. A utilização destas técnicas visa subsidiar a identificação de áreas favoráveis ao processo de erosão.

Outro aspecto a considerar é a importante utilização do Sistema de Informação Geográfica como ferramenta de apoio ao processo de tomada de decisão que permitiu a visualização e a realização das análises espaciais de maneira rápida e objetiva.

Com relação ao método de Lógica “*Fuzzy*”, pode-se recomendar testar a criação dos membros “*Fuzzy*” a partir das funções de “*fuzzyficação*” que o ArcSDM possui (*Small, Large, Linear, Gaussian*), verificando assim se há algum aprimoramento do modelo.

Com relação ao método de Pesos de Evidência, apesar do valor da confiança calculada ter indicado que as variáveis usadas no modelamento são bem relacionadas ao surgimento de erosão, recomenda-se a avaliação ou teste de outras variáveis ambientais e não ambientais, tais como, estradas, drenagens, etc, que venham a aprimorar o modelo de busca por novos pontos de erosão. Segundo as conclusões do Boletim de Desenvolvimento e Pesquisa da Embrapa [36], a topografia pode ser uma interessante variável a ser considerada em análises envolvendo processos erosivos, já que foi presenciado um elevado índice de perda de solo em áreas com relevo mais acentuado.

Recomenda-se, também, que a metodologia empregada no presente trabalho seja aplicada em outras bacias hidrográficas com maior quantidade de pontos de treinamento (pontos de erosão conhecidos), para validar as técnicas utilizadas. Sugere-se ainda que técnicas estatísticas de inferência espacial, como os modelos de Redes Neurais e os modelos de Regressão Logística, sejam aplicados ao conjunto de indicadores na BAT, para fins de investigação de funções estatísticas de predição de novos pontos de erosão e comparação com as técnicas aplicadas nesta dissertação. Essas técnicas encontram-se presentes na ferramenta ArcSDM, porém não foram exploradas.

### **7.3 Considerações finais**

Como consideração final pode-se afirmar que a metodologia aqui apresentada permite auxiliar a busca por novas áreas de erosão associadas aos indicadores utilizados neste trabalho. Deste modo, conclui-se ser efetivamente possível utilizar procedimentos como estes para acompanhar processos erosivos em bacias semelhantes à estudada e avaliar indiretamente as ações dos homens associados aos processos de erosão, de maneira a viabilizar a adoção de medidas de preservação ambiental e orientar a tomada de decisão. Este trabalho não teve como objetivo gerar um mapa de vulnerabilidade ou de áreas de erosão na Bacia do Alto Taquari, pois a qualidade dos mapas gerados depende de informações mais precisas de campo e de disponibilidade de dados secundários. A contribuição desta dissertação é sugerir métodos e ferramentas que possam subsidiar este tipo de questão em bacias hidrográficas no nosso país.

## REFERÊNCIAS

- [1] Abdon, M. M. (2004). “Os impactos ambientais no meio físico - erosão e assoreamento na bacia hidrográfica do rio Taquari, MS, em decorrência da pecuária”. Tese de Doutorado. Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, SP, 297p.
- [2] Agterberg, F. P.; Cheng, Q. (2002). “Conditional Independence Test for Weights-of-Evidence Modeling”. *Natural Resources Research*, Dezembro, Vol. 11, nº4.
- [3] Almeida, C. M. (2003). “Modelagem da Dinâmica Espacial como uma Ferramenta Auxiliar ao Planejamento: Simulação de Mudanças de Uso da Terra em Áreas Urbanas para as Cidades de Bauru e Piracicaba (SP), Brasil”. Tese de Doutorado. Curso da Pós-Graduação em Sensoriamento Remoto, Instituto Nacional de pesquisas espaciais (INPE), SP, 297p.
- [4] Altran, A. B.; Lopes, M. L. M.; Minussi, C. R.; Villareal, F. (2005). “Um Estudo das Funções de Base Radial Aplicadas à Previsão de Cargas Elétricas via Redes Neurais Artificiais. In: XXVIII CNMAC - Congresso Nacional de Matemática Aplicada e Computacional, São Paulo. Anais do XXVIII CNMAC.
- [5] Álvares, M. T. P.; Pimenta, M. T. (1998). “Erosão hídrica e transporte sólido em pequenas bacias hidrográficas”. 4º Congresso da água. FIL, Lisboa - 23 a 27 de Março de 1998.  
Disponível em:  
[http://snirh.inag.pt/snirh/estudos\\_proj/portugues/docs/desertificacaofichas.html](http://snirh.inag.pt/snirh/estudos_proj/portugues/docs/desertificacaofichas.html)  
Acesso em: 14/01/2007.
- [6] Alves, A.; Souza, F. J.; Márques, M. (2005). “Avaliação do potencial à erosão dos solos uma análise comparativa entre Lógica Fuzzy e o Método USLE”. Anais XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, Brasil, 16-21 abril, p. 2011-2018.
- [7] Amendola, M.; Souza, A. L.; Barros, L. C. (2005). “Manual do uso da teoria dos conjuntos Fuzzy no MATLAB 6.5”. Universidade estadual de Campinas - UNICAMP  
Disponível em : [www.ime.unicamp.br/~laeciocb/MANUAL\\_2005.pdf](http://www.ime.unicamp.br/~laeciocb/MANUAL_2005.pdf)  
Acesso em: 23/01/2007.
- [8] An, P.; Moon, W. M.; Rencz, A. (1991). “Application of fuzzy set theory for integration of geological, geophysical, and remote sensing data”. *Canadian Journal of Exploration Geophysics*, v. 27, p. 1-11.

- [9] Araújo, P. C. (2004). “Análise da suscetibilidade a escorregamentos: uma abordagem probabilística”. Tese de Doutorado. Curso da Pós-Graduação em Geociências e Meio Ambiente – área de concentração em Geociências e Meio Ambiente, Universidade Estadual Paulista, 172p.
- [10] Araújo, P. H. M., Nehme, C. C. (acesso em 2005). “Uso de modelos de incerteza no controle ambiental do processo erosivo na Bacia do Alto Taquari”. Disponível em: <http://www.geoinfo.info/geoinfo2001/papers/112pedro.pdf>  
Acesso em: 14/04/2005.
- [11] Araújo, Q. R.; Paiva, A. (acesso em 2007). “Manejo e conservação do solo e da água na região Cacaueira da Bahia aspectos”. Disponível em: <http://www.ceplac.gov.br/radar/semfaz/soloeagua.htm>  
Acesso em: 09/01/2007.
- [12] Bauchspiess, A.; Gosmann, H. L.; Filho, F. M. L. (2002). “Controle fuzzy para sistema de nível de líquidos”. XIV Congresso Brasileiro de automática. Natal - RN
- [13] Boas, C. L. V. (acesso em 2007). “Análise da aplicação de métodos multicritérios de apoio à decisão (MMAD) na gestão de recursos hídricos”. Disponível em: [www.cprm.gov.br/rehi/simposio/go](http://www.cprm.gov.br/rehi/simposio/go)  
Acesso em: 11/01/2007.
- [14] Bonham-Carter, G. F. (1994). “Geographic Information Systems for Geoscientists – Modelling with GIS”. New York, Elsevier Science Inc., 398p.
- [15] Bonham-Carter, G. F.; Agterberg, F. P.; Wright, D. F. (1988). “Integration of geological datasets for gold exploration in Nova Scotia: Photogrammetric Engineering and Remote Sensings”. v. 54, nº 11, p. 1585-1592.
- [16] Bonham-Carter, G. F.; Agterberg, F. P.; Wright, D. F. (1989). “Weights of evidence modeling: a new approach to mapping mineral potential”. Statistical applications in the Earth Science. Ottawa: Geological Survey of Canada, Paper 89-9, p. 171-183.
- [17] Bueno, M. C. D. (2003). “Utilização de redes de dependência e lógica nebulosa em estudos de avaliação ambiental”. Dissertação de Mestrado em Engenharia de computação - Área de concentração Geomática. Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 121p.
- [18] Burrough, P. A.; Mcdonnell, R. A. (1998). “Principles of Geographical Information Systems”. New York, Oxford University Press Inc., 333p.

- [19] Cabral, P. C. B. (2001). “Sistemas espaciais de apoio à decisão – O Sistema de apoio ao Licenciamento da Direcção Regional do Ambiente do Alentejo”. Dissertação de Mestrado em Sistemas de Informações Geográficas. Universidade técnica de Lisboa – Instituto Superior Técnico. 96p.
- [20] Câmara, G.; Pedrosa, B. M.. (2003) “Modelagem dinâmica: conceitos básicos e exemplos de sistemas”. In: Câmara, G.; Monteiro, A. M. V. “Introdução a modelagem dinâmica espacial”. Simpósio Brasileiro de Sensoriamento Remoto 2003. Cap. 1, p. 1-41.
- Disponível em: <http://www.dpi.inpe.br/cursos/tutoriais/modelagem>  
Acesso em: 07/01/2007.
- [21] Câmara, G.; Carvalho, M. S.; Cruz, O. G.; Correa, V. (2001). “Análise de Dados de Áreas”. In: Câmara, G.; Monteiro, A. M.; Fuks, S. “Análise Espacial de Dados Geográficos” São José dos Campos, Instituto Nacional de pesquisas espaciais (INPE). Cap. 5, p. 1-44.
- Disponível em: <http://www.dpi.inpe.br/gilberto/livro/analise/index.html>  
Acesso em: 14/08/2005.
- [22] Câmara, G.; Moreira, F. R.; Barbosa, C.; Filho, R. A.; Bonisch, S. (2001). “Técnicas de Inferência Geográfica”. In: Câmara, G.; Davis, C.; Monteiro, A. M. V. “Introdução à Ciência da Geoinformação” São José dos Campos, Instituto Nacional de pesquisas espaciais (INPE). Cap. 9, p. 2-48.
- Disponível em: <http://www.dpi.inpe.br/gilberto/livro/introd/>  
Acesso em: 25/08/2005.
- [23] Câmara, A. S. (1996). “Spatial Simulation Modelling” In: “Spatial Analytical Perspectives on GIS” M. Fisher. London, Taylor & Francis: 213-218.
- [24] Castro, L. M. A.; Baptista, M. B.; Netto, O. M. C. (2004). “Análise multicritério para a avaliação de sistemas de drenagem urbana – Proposição de indicadores e de sistemática de estudo”. Revista Brasileira de Recursos Hídricos, Porto Alegre, v. 9, n. 4, p. 05-19.
- [25] Chaves, E. O. C.; Falsarella, O. M. (2003). “Sistemas de Informação e Sistemas de Apoio à Decisão”.
- Disponível em: <http://www.chaves.com.br/TEXTSELF/COMPUT/sad.htm>  
Acesso em: 10/02/2007.
- [26] Cheng, Q. (2004). “Application of Weights of Evidence Method for Assessment of Flowing Wells in the Greater Toronto Area, Canadá”. Natural Resources Research, Vol.

13, No. 2, Junho.

- [27] Collischonn, W.; Tucci, C. E. M. (2002). “Aspectos da hidrologia da bacia do alto rio Taquari” In: II Simpósio de Recursos Hídricos do Centro Oeste. Campo Grande, MS. Disponível em: <http://galileu.iph.ufrgs.br/collischonn>  
Acesso em: 13/03/2006.
- [28] Coolbaugh, M. F.; Bedell, R. L. (2004). “A Simplification of Weights of Evidence using a Density Function and Fuzzy Distributions, using Geothermal Systems in Nevada as an example”. Geological Association of Canada Special Paper “GIS applications in the Earth Sciences”, in press.  
Disponível em: <http://www.unr.edu/geothermal>  
Acesso em: 15/06/2005.
- [29] Corrêa, O. (acesso em 2005). “O fenômeno das voçorocas”. Embrapa.  
Disponível em: <http://www.cnpqc.embrapa.br/informa/dezembro2001/vocorocas.html>  
Acesso em: 25/08/2005.
- [30] Costa, G. A. O. P. (2003). “Uma metodologia para subsidiar a avaliação da degradação ambiental a partir de dados NOAA AVHRR”. Dissertação de Mestrado em Engenharia de computação - Área de concentração Geomática. Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 152p.
- [31] Croitoru, A.; Doytsher, Y. (acesso em 2005). “Fuzzy” Fuzzy Logic for spatial decision making: Accounting for data uncertainty  
Disponível em:  
[http://www.spatial.maine.edu/~croitoru/Publications/Croitoru\\_Doytsher\\_SEASC.pdf](http://www.spatial.maine.edu/~croitoru/Publications/Croitoru_Doytsher_SEASC.pdf)  
Acesso em: 25/08/2005.
- [32] Dilão, R. (1993). “Autômatos Celulares, Máquinas de Turing ou a Natureza como Máquina de Cálculo”. Colóquio Ciências, 12, p. 3-20.  
Disponível em: <https://sd.ist.utl.pt/Awareness/general.html>  
Acesso em: 12/01/2007.
- [33] Fiori, A. P., Fiori, C. O.; Disperati, L.; Conceição A.; Kozciak, S.; Guedes, J. A.; Ciali, A. (2001). “O processo erosivo na Bacia do Alto Paraguai”. Boletim Paranaense de Geociências, n. 49, p. 63-78. Editora da UFPR.
- [34] Franca-Rocha, W., Bonham-Carter, G. F; Misi, A. (2000). “Estimando a favorabilidade para recursos minerais na Bacia de Irecê pelos métodos lógica nebulosa e pesos das evidências”. Siteintibus, Feria de Santana, nº 22, p. 27-54.

- [35] Freitas, H. (2000). “As tendências em Sistemas de Informação com base em recentes congressos”. Porto Alegre/RS: No. 13, 20 p.  
Disponível em: <http://read.adm.ufrgs.br>  
Acesso em: 12/01/2007.
- [36] Galdino, S.; Risso A.; Soriano, B. M. A.; Vieira, L. M.; Padovani C. R.; Pott A.; Melo E. C.; Junior N. A. (2003). “Perdas de solo na Bacia do Alto Taquari”. Corumbá – MS, Embrapa - Boletim de Pesquisa e desenvolvimento 44, 40 p.
- [37] Galdino, S.; Vieira L. M.; Soriano, B. M. A. (2003). “Erosão na Bacia do Alto Taquari”. Corumbá – MS, Embrapa - Documentos 52, 47p.
- [38] Galdino, S. (2003). “Projeto implementação de práticas de gerenciamento integrado de Bacia hidrográfica para o Pantanal e Bacia do Alto Paraguai”. Processos Erosivos na Bacia do Alto Taquari – Relatório Final – Embrapa, 32 p.
- [39] Galdino, S.; Soriano, B. M. A.; Vieira, L. M.; Padovani C. R.; (2002). “Mudanças do regime hidrológico da bacia do Rio Taquari – Pantanal”. Corumbá – MS, Embrapa - Boletim de Pesquisa, 26, 24p.
- [40] Galvão, W. S.; Meneses, P. R. (2005). “Identificação de variáveis geoambientais preditivas ao processo de locação de estações fluviométricas com o apoio do Modelo de Pesos de Evidência”. Anais XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, Brasil, 16-21 abril, p. 2169-2176.
- [41] Gartner, I. R. (2001). “Avaliação ambiental de projetos em bancos de desenvolvimento nacionais e multilaterais: evidências e propostas”. Editora Universa, Brasília.
- [42] Gasoto, D. G., Volpi, N. M. P. (2006). “Analytic Hierarchy Process na Avaliação de Projetos”. Mini-Simpósio MS03 - Pesquisa Operacional, Paraná, Brasil, 07-10 agosto, Universidade estadual de Londrina.
- [43] Gibilini, Daniel (2006). “Aplicação de técnicas de inteligência artificial na alocação dinâmica de canais em redes sem fio”. Dissertação de Mestrado em Engenharia. Escola Politécnica, Universidade de São Paulo, 140p.
- [44] Guerra; A. J. T. (2005) “Experimentos e monitoramentos em erosão dos solos”. Revista do Departamento de Geografia, 16. p. 32-37.
- [45] Guerra; A. J. T. (2005) “Processos Erosivos nas Encostas” In: Guerra; A. J. T.; Cunha, S. B. “Geomorfologia: Uma Atualização de Bases e Conceitos”. Ed. Bertrand Brasil, Rio de Janeiro, 6ª ed. p. 149-209.

- [46] Haridasan, M. (acesso em 2007). “Conseqüências de alterações em ecossistemas naturais”. Universidade de Brasília - Departamento de Ecologia.  
Disponível em: <http://www.unb.br/ib/ecl/docentes/Haridasan/Aula-Gama.pdf>  
Acesso em: 12/01/2007.
- [47] Infanti Jr, N.; Fornasari Filho, N. (1998). “Processos de dinâmica Superficial”. In: Oliveira, A. M. S.; Brito, S. N. A. (ed.). Geologia de Engenharia. São Paulo: Associação Brasileira de Geologia de Engenharia. Cap. 9, p. 131-152.
- [48] Jafelice, R. S. M.; Barros, L. C.; Bassanezi, R. C. (2005). “Teoria dos Conjuntos Fuzzy com Aplicações”. Sociedade Brasileira de Matemática Aplicada e Computacional - SBMAC, São Carlos - SP, 69 p.
- [49] Jenkins, A.; Peters, N.E; Rodhe, A. (1994). “Hidrology”. In Moldan, B.; Cerny, J. (Cord.) Journal Biogeochemistry of small catchments: a tool for environmental research. (scope 51). Ed. John Wiley & Sons, Chichester, p. 31-54
- [50] Jesus, A. A. (2006). “Ensaio de uso do sistema monitor SIGINDES no núcleo de desertificação de Gilbués – Piauí”. Monografia de graduação do curso de Agronomia. Faculdade de Agronomia e Medicina Veterinária. 53p.
- [51] Jonathan, M. (2005). “Classificação do uso e cobertura do solo em escala regional a partir de seqüências temporais de dados MODIS”. Dissertação de Mestrado em Engenharia de computação - Área de concentração Geomática. Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 147p.
- [52] Jonathan, M.; Meirelles, M. S. P.; Berroir, J. P.; Herlin, I. (2006) “Considerations for regional scale long-term monitoring of land use with MODIS data and its application for the high Taquari basin, MS/MT, Brazil”. Anais 1º Simpósio de Geotecnologias no Pantanal, Campo Grande, Brasil, 11-15 novembro 2006, Embrapa Informática Agropecuária/INPE, p.457-466.  
Disponível em: <http://mtc-m17.sid.inpe.br/rep-/sid.inpe.br/mtc-m17@80/2006/12.12.12.41>. Acesso em: 15/12/2006.
- [53] Kemp, L.D.; Bonham-Carter, G. F.; Raines, G. L.; Looney, C.; Yu, H. (2002). “ArcSDM and DataXplore User Guide. Spatial Data Modeller Extension for ArcView and Spatial Analyst”. Geological Survey of Canada, User Guide.  
Disponível em: [http://ntserv.gis.nrcan.gc.ca/sdm/default\\_e.htm](http://ntserv.gis.nrcan.gc.ca/sdm/default_e.htm)  
Acesso em: 01/05/2005.
- [54] Kozciak, S. (2002). “Evolução do processo erosivo da bacia dos rios Arica-Açu e Arica-

- Mirim, MT, através de análise multitemporal”. Boletim Paranaense de Geociências, n. 51, p. 91-114. Editora UFPR.
- [55] Leite, I. O. B.; Lins, J. C. S.; Cerqueira, M. G. de C. (acesso em 2007). “Autômatos celulares”. Centro de informática. UFPE  
Disponível em: <http://www.di.ufpe.br/~iobl/monografia/index.html>  
Acesso em: 12/01/2007.
- [56] Malczewski, J. (1999). “Gis and Multicriteria Decision Analysis”. New York: John Wiley & Sons, Incorporated
- [57] Manfro, P. R. (1998). “Aplicação do analytic hierarchy process (AHP) como instrumento de apoio à decisão no gerenciamento costeiro integrado”. Dissertação de Mestrado - Área de concentração Inteligência Aplicada. Universidade Federal de Santa Catarina.
- [58] Maranhão, M. R. A.; Souza, F. J. (2005). “Modelo de seleção de áreas para atualização do mapeamento sistemático baseado em lógica nebulosa”. Dissertação de Mestrado em Engenharia de computação - Área de concentração Geomática. Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 54 p.
- [59] Medeiros, J. S. (1999). “Banco de dados geográficos e redes neurais artificiais: tecnologias de apoio a gestão de território”. Tese de Doutorado – Geografia física. Universidade de São Paulo - USP. 221 p.
- [60] Meirelles, M. S. P.; Moreira, M.; Coelho, F. B. N.; (2004). “Uso de Inteligência artificial em agricultura de precisão: Redes Bayesianas e Neurais”. In: Machado, P. L. O. A.; Bernardi, A. C. C., Silva, C. A. “Agricultura de precisão para o manejo de fertilidade do solo em Sistema Plantio Direto”. Rio de Janeiro, Embrapa Solos. Cap. 10, p. 175-195.
- [61] Meirelles, M. S. P. (1997). “Análise Integrada do Ambiente Através de Geoprocessamento – Uma Proposta Metodológica para Elaboração de Zoneamentos”. Tese de doutorado. UFRJ, IGEO, Rio de Janeiro. 192 p.
- [62] Meirelles, M. S. P.; Dias, T. C. S.; Bueno, M. C. D.; Coutinho, H. L. C. (2006) “SAEB - Sistema de análise espacial do risco de impactos ambientais derivados da erosão em bacias hidrográficas”. Boletim técnico – Embrapa Solos
- [63] Meirelles, M. S. P.; Bueno, M. C. D.; Dias, T. C. S.; Coutinho, H. L. C. (2005) “Sistema de suporte a decisão para avaliação do risco de impactos ambientais em bacias hidrográficas por redes de dependência e lógica fuzzy”. Anais XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, Brasil, 16-21 abril 2005, Instituto Nacional de

- pesquisas espaciais (INPE), p. 2259-2266.
- [64] Meirelles, M. S. P.; Câmara, G.; Almeida, C. M. (2007) “Geomática: Modelos e Aplicações Ambientais”. 1º Ed. Brasília: Embrapa. 599 p.
- [65] Moldan, B.; Bilharz, S. (1997). “Sustainability indicators: report of the project on indicators of sustainable development”. Ed John Wiley & Sons. Chichester, Reino Unido. 415p.
- [66] Moreira, F. R. S. (2002). “Uso da avaliação de técnicas de integração e análise espacial de dados em pesquisa mineral aplicadas ao planalto de poços de caldas”. São José dos Campos, Instituto Nacional de pesquisas espaciais (INPE).
- [67] Organization for Economic Co-operation and Development - OECD (1993). “Core set of indicators for environmental performance reviews: a synthesis report by the Group on the State of the Environment”. Environmental Monographs. Paris, nº 83, 39 p.
- [68] Oliveira, L. F. C.; Martinez, M. A.; Pruski, F. F.; Griebeler, N. P.; Oliveira, G. C. (2005). “Rotina computacional para a determinação da velocidade de sedimentação das partículas do solo em suspensão no escoamento superficial”. Eng. Agríc., Jan./Abr. 2005, vol.25, nº 1, p.126-136. ISSN 0100-6916
- [69] Pedrosa, B. M. (2003). “Ambiente Computacional para Modelagem Espacial Dinâmica”. Tese de Doutorado – Computação aplicada. São José dos Campos, São Paulo, Instituto Nacional de Pesquisas Espaciais (INPE), SP, 111p.
- [70] Penha, R. N. (2002). “Um estudo sobre regressão logística binária”. Universidade Federal de Itajubá – Instituto de engenharia mecânica - Engenharia de Produção – 10 p.
- [71] Pereira, S. E. M.; Meirelles, M. S. P.. (2006). “Uso de sistema de suporte à decisão espacial como subsídio ao planejamento territorial: aplicação ao zoneamento ecológico-econômico”. I Encontro UERJ de Geomática. Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro.
- [72] Pereira, M.C.G. (1998) “Fatores de erosão em vertentes na Formação Barreiras – Estudo de caso em Salvador-Ba”. Dissertação de Mestrado. Instituto de Geociências da Universidade Federal da Bahia - UFBA, Pós-graduação em Geoquímica e Meio ambiente. Salvador – Bahia 115 p.
- [73] Pina, M. F.; Carvalho, M. S.; Santos, S. M. (2000). “Conceitos básicos de sistemas de informação geográfica e cartografia aplicados à saúde”. Brasília, Editora Organização Panamericana da Saúde / Ministério da Saúde. 122 p.
- [74] Portugal, G. (1992) “Erosão, assoreamento e desertificação”.

Disponível em: <http://www.gpca.com.br/Gil/art002.htm>

Acesso em: 14/04/2007.

- [75] Ross, J. L. S. (2006). “PCBAP - Plano de conservação da bacia do alto Paraguai e o zoneamento ecológicoeconômico para o Brasil”. Anais 1º Simpósio de Geotecnologias no Pantanal, Campo Grande, Brasil, 11-15 novembro 2006, Embrapa Informática Agropecuária/INPE, p.667-674.
- [76] Ruhoff, A. L.; Souza, B. S. P.; Giotto, E.; Pereira, R. S. (2005). “Lógica Fuzzy e Zoneamento ambiental da Bacia do Arroio Grande”. Anais XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, Brasil, 16-21 abril, p. 2355-2362.
- [77] Rattenbury, M. S.; Partington, G. A. (acesso em 2005). “Prospectivity models and GIS data for the exploration of epithermal gold mineralisation in New Zealand”. Epithermal Au Prospectivity Report; Institute of Geological & Nuclear Sciences Ltd  
Disponível em:  
<http://www.kenex.co.nz/pdfs/Partington%20RattenburyProspectivityReport.pdf>  
Acesso em: 23/07/2005.
- [78] Saaty, T. L. (1990) “How to Make a Decision: The Analytic Hierarchy Process” European Journal of Operational Research 48, p. 9-26.
- [79] Saaty, T. L. (1980). “The analytic hierarchy process”. New York: McGraw-Hill.
- [80] Sandri, S.; Correa, C. (1999). “Lógica Nebulosa”. V Escola de Redes Neurais, Promoção: Conselho Nacional de Redes Neurais. pp. c073-c090, 19 de julho, 1999 - ITA, São José dos Campos - SP
- [81] Schmidt, A. M. A. (1995). “Processo de apoio à tomada de decisão – Abordagens: AHP e MACBETH”. Dissertação de Mestrado em engenharia – especialidade em engenharia de produção. UFSC, Florianópolis.  
Disponível em: <http://www.eps.ufsc.br/disserta/engait95.html>  
Acesso em: 10/01/2007.
- [82] Schmoldt, D. L.; Peterson, D. L (1997). “Using the Analytic Hierarchy Process for Decision-Making in Ecosystem Management”. USDA Forest Service - Research & Development  
Disponível em: <http://www.treesearch.fs.fed.us/pubs/79>  
Acesso em: 10/01/2007.
- [83] Shaw, I. S.; Simões, M. G., (2001). “Controle e Modelagem Fuzzy”. Editora Edgard Blücher, São Paulo. 165 p.

- [84] Silva, J. S. V. da (2003). “Análise multivariada em zoneamento para planejamento ambiental Estudo de caso: bacia hidrográfica do alto rio Taquari MS/MT”. Dissertação de Doutorado em Engenharia Agrícola - Área de concentração em Planejamento e Desenvolvimento Rural Sustentável. Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas, Campinas, 338 p.
- [85] Soares, S. R. (2003). “Análise multicritério como instrumento de gestão ambiental”. Dissertação de Mestrado. UFSC, Florianópolis.  
Disponível em: <http://www.ens.ufsc.br/~soares>  
Acesso em: 10/01/2007.
- [86] Soriano, B. M. A.; Clarke, R. T.; Galdino, S. (2001). “Evolução da erosividade das chuvas na Bacia do Alto Taquari”. Corumbá – MS; Boletim de Pesquisa e Desenvolvimento 25– Embrapa.
- [87] Souza Filho, C. R.; Crosta, A. P. (2003). “Geotecnologias aplicadas à geologia”. Revista Brasileira de Geociências. Volume 33 (2ºSuplemento). 1-4 p. Publicação:junho de 2003.
- [88] Stefanakis, E.; Sellis, T. (1999). “Enhancing a Database Management System for GIS with Fuzzy Set Methodologies”. Proceedings of the 19th International Cartographic Conference, Ottawa, Canada, August.
- [89] Star, J.; Estes, J. (1991). “Geographic Information Systems: an Introduction”. Englewoods Cliffs, New Jersey.
- [90] Thiart, C.; Bonham-carter, G. F.; Agterberg, F. P. (acesso em 2005) “Conditional independence in weights-of-evidence: application of an improved test”.  
Disponível em: [http://www.iamg.org/meetings/Proceedings\\_2003/papers/](http://www.iamg.org/meetings/Proceedings_2003/papers/)  
Acesso em: 07/03/2005.
- [91] Tunstall, D. (1992). “Developing environmental indicators: definitions, framework and issues”. Workshop on global environmental indicators. Washington D. C., Estados Unidos.
- [92] Turban E., Liebowitz, J. (1990). “Managing Expert Systems”. Idea Group Publishing
- [93] Wang, H.; Cai, G.; Cheng, Q. (2002). “Data integration using weights of evidence model: applications in mapping mineral resource potentials”. Symposium on Geospatial Theory, Processing and Applicatios, Symposium sur la théorie, les traitements et les applications des données Géospaciales, Ottawa.
- [94] Wischmeier, W.H.; Smith, D.D. (1978). “Predicting rainfall erosion losses – a guide to conservation planning”. Washington: USDA/ARS. 58p.(Agricultural Handbook n° 573).

- [95] Xavier, A. A. B.; Borges A. T. (2002). “Ambiente de modelamento e simulação no ensino de ciências”. Anais da III Semana da Pós-Graduação da UFMG, Belo Horizonte, Brasil.
- [96] Zadeh, L. A. (1965). “Fuzzy sets” In: Fuzzy Sets, Information and Control. nº 8, p. 338 – 353.
- [97] Zadeh, L. A. (1988). “Fuzzy Logic”. IEEE Computer Mag. University of California, p. 88–93, Berkeley, apr.
- [98] Zambon, K. L.; Carneiro, A. A. F. M.; Silva, A. N. R.; Negri, J. C. (2005). “Análise de decisão multicritério na localização de usinas termoelétricas utilizando SIG”. Pesquisa Operacional, v.25, n.2, p.183-199, Maio a Agosto de 2005.