



**CAMILA DA SILVA DOURADO**

**MINERAÇÃO DE DADOS CLIMÁTICOS PARA ANÁLISE DE  
EVENTOS EXTREMOS DE PRECIPITAÇÃO**

**CAMPINAS**

**2013**





UNIVERSIDADE ESTADUAL DE CAMPINAS  
FACULDADE DE ENGENHARIA AGRÍCOLA

**CAMILA DA SILVA DOURADO**

**MINERAÇÃO DE DADOS CLIMÁTICOS PARA ANÁLISE DE  
EVENTOS EXTREMOS DE PRECIPITAÇÃO**

**Orientador: Prof. Dr. Stanley Robson de Medeiros Oliveira**

**Co-orientadora: Dra. Ana Maria Heuminski de Avila**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Agrícola da Faculdade de Engenharia Agrícola da Universidade Estadual de Campinas, para obtenção do Título de Mestre em Engenharia Agrícola, na área de concentração de Planejamento e Desenvolvimento Rural Sustentável.

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO  
DEFENDIDA PELA ALUNA CAMILA DA SILVA DOURADO  
E ORIENTADA PELO PROF.DR. STANLEY ROBSON DE M. OLIVEIRA.

Assinatura do Orientador

A handwritten signature in blue ink, which appears to be 'Stanley Robson de Medeiros Oliveira', is written over a horizontal line.

CAMPINAS

FEVEREIRO DE 2013

FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DA ÁREA DE ENGENHARIA E ARQUITETURA - BAE - UNICAMP

D74m	<p>Dourado, Camila da Silva, 1982-</p> <p>Mineração de dados climáticos para análise de eventos extremos de precipitação / Camila da Silva Dourado. -- Campinas, SP: [s.n.], 2013.</p> <p>Orientador: Stanley Robson de Medeiros Oliveira Coorientador: Ana Maria Heuminski de Avila. Dissertação de Mestrado - Universidade Estadual de Campinas, Faculdade de Engenharia Agrícola.</p> <p>1. Precipitação (Meteorologia). 2. Zonas climáticas. 3. Chuvas. 4. Processamento sequencial (Computação). I. Oliveira, Stanley Robson de Medeiros. II. Avila, Ana Maria Heuminski de. III. Universidade Estadual de Campinas. Faculdade de Engenharia Agrícola. IV. Título.</p>
------	---

Título em Inglês: Mining climatic data for analysis of extreme precipitation events

Palavras-chave em Inglês: Precipitation (Meteorology), Zones climate, Rainfall,  
Sequential processing (Computer science)

Área de concentração: Planejamento e Desenvolvimento Rural Sustentável

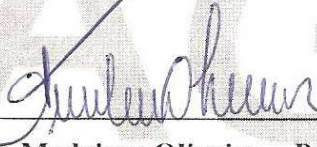
Titulação: Mestra em Engenharia Agrícola

Banca examinadora: Carlos Alberto Alves Meira, Gabriel Constatino Blain

Data da defesa: 26-02-2013

Programa de Pós Graduação: Engenharia Agrícola

Este exemplar corresponde à redação final da **Dissertação de Mestrado** defendida por **Camila da Silva Dourado**, aprovada pela Comissão Julgadora em 26 de fevereiro de 2013, na Faculdade de Engenharia Agrícola da Universidade Estadual de Campinas.



---

**Prof. Dr. Stanley Robson de Medeiros Oliveira – Presidente e Orientador**  
Feagri/Unicamp



---

**Dr. Gabriel Constantino Blain**  
IAC



---

**Dr. Carlos Alberto Alves Meira – Membro Titular**  
Embrapa/CNPq

Faculdade de  
**Engenharia Agrícola**  
Unicamp

## DEDICO

A Deus, *Mestre* dos mestres, *Autor* da minha vida e da minha história, meu *Condutor* e *Consolador* - em todos os tempos.

Aos meus modelos de toda uma vida, formadores de minha personalidade, meus amados pais, **Ranulfo** (*in memoriam*) e **Joanita**: levo o olhar e as palavras de vocês, que ecoam como uma fortaleza e me sustentam, induzindo ao prosseguimento nesta caminhada.

À minha avó, pelo cuidado incondicional e amor incessante.

Aos meus, queridos irmãos: **Priscilla**, **Raphael** e **Gabriel**, por toda compreensão da distância vivenciada e amenizada em cada reencontro.

Ao meu querido sobrinho **Levi**, pela alegria contagiante e doce num olhar confortante.

## **AGRADECIMENTOS**

À Universidade Estadual de Campinas, em especial à Faculdade de Engenharia Agrícola pela oportunidade dada.

Ao CNPq pelo incentivo financeiro.

Ao CEPAGRI pela infraestrutura cedida e conhecimento compartilhado.

À Agência Nacional das Águas pelo fornecimento dos dados pluviométricos utilizados neste trabalho.

Ao meu orientador, Prof. Dr. *Stanley Robson de Medeiros Oliveira*, pela orientação inestimável, pela confiança, dedicação, atenção e todo aprendizado vivenciado dia a dia.

À minha coorientadora, Dr<sup>a</sup>. *Ana Maria Heuminski de Avila*, por compartilhar seus conhecimentos, pelo profissionalismo, cuidado, dedicação e pela amizade e carinho ofertados.

Aos pesquisadores, funcionários e colegas do Cepagri – *Prof. Jurandir, Prof. Hilton, Edilene, Claudir, Gustavo, Solange, Andrea, Susian, Juliano, Renata, Camila Lazarim, Octávia e Daniela*, pelo aprendizado coletivo, pela convivência e apoio nos momentos difíceis.

Aos colegas e amigos, *Wanderson, Flávio, Guilherme e Agmon* pelas resoluções dos enigmas computacionais surgidos no desenvolvimento do trabalho.

Aos amigos, *Gabriel, Diego e Thiago*, por todo apoio, aventuras e risos que alegraram e confortaram meu coração.

À minha amiga Prof<sup>a</sup>. *Cláudia Bloisi Vaz Sampaio*, pelo apoio em minha vida profissional, pelas inúmeras palavras de incentivo e pela torcida contagiante.

Aos funcionários da secretaria da Pós-graduação da FEAGRI: *Célia, Marta, Fabio e Sidney*, pela atenção e presteza. Em especial, ao coordenador Prof. Dr. *Zigomar M. de Souza*.

## RESUMO

O conhecimento das condições climáticas, identificando regiões com maiores riscos de ocorrências de eventos extremos, que possam impactar os diversos setores socioeconômicos e ambientais, tornou-se um grande desafio. No Brasil as maiores ocorrências de eventos extremos estão relacionadas aos fenômenos hidrológicos. Em particular, o estado da Bahia apresenta alta variabilidade temporal e espacial no clima, desde áreas consideradas áridas ou com risco de aridização (ao Norte) a regiões com clima úmido na faixa litorânea. O estado tem sido alvo nesses últimos anos de diferentes eventos extremos de chuva, com enchentes em algumas áreas e secas severas em outras. Neste contexto, o objetivo deste trabalho foi utilizar técnicas de mineração de dados para analisar a frequência das ocorrências dos eventos extremos de precipitação durante o período de 1981 a 2010 no estado da Bahia, com o propósito de subsidiar a tomada de decisão referente a ações preventivas e mitigadoras dos impactos socioeconômico e ambientais. Para isto, foram utilizados dados climáticos de precipitação fornecidos pelo Sistema de Informações Hidrológicas da Agência Nacional de Águas. Aplicando-se a tarefa de agrupamento (clusterização), por meio do algoritmo *k-means*, as séries históricas de dados climáticos foram agrupadas em cinco zonas pluviometricamente homogêneas. Posteriormente, foram realizadas análises em diferentes escalas temporais (anual, mensal e diária) identificando através da Técnica dos Quantis limiares superiores e inferiores de intensidade de chuva em cada região homogênea, para cada escala temporal. Na escala mensal, foram identificados padrões sequenciais das ocorrências dos eventos extremos positivos e negativos ao longo dos trinta anos. Os resultados reforçam a potencialidade da técnica de mineração de dados em agrupar zonas homogêneas por similaridade pluvial, com o uso do algoritmo *k-means*. Revelam ainda, para todas as escalas temporais utilizadas, uma alta variabilidade pluviométrica. Os anos registrados com maior ocorrência de eventos extremos negativos estão na década de 90 e os anos registrados com mais eventos extremos positivos na década de 2000.

**PALAVRAS-CHAVE:** Descoberta de conhecimento; Precipitação pluviométrica; Zonas homogêneas; Fenômenos extremos; Padrões sequenciais, Técnica dos quantis.



## **ABSTRACT**

The knowledge of climate conditions, identifying areas with the greatest risk of occurrence of extreme events, that may impact the various socioeconomic and environmental sectors, has become a major challenge. In Brazil the largest occurrences of extreme events are related to hydrological phenomena. In particular, the state of Bahia presents a high temporal and spatial variability of climate, from areas considered arid or with risk to become arid – (in the North) to regions with humid along the coast. The state has been targeted of different extreme rainfall events recently, with floods in some areas and severe droughts in others. In this context, the aim of this study was to use data mining techniques to analyze the frequency of occurrences of extreme precipitation events during the period from 1981 to 2010 in the state of Bahia, in order to support decision making regarding the preventive and mitigative environmental and socioeconomic impacts. To accomplish that, it was used climate data of precipitation supplied by the Hydrological Information System of the National Water Agency. By applying the task of grouping (clustering) by means of the k-means algorithm, the time series of climate data were grouped into five homogeneous rainfall zones. Subsequently, analyzes were performed on different time scales (annually, monthly and daily) identifying by quantile methods the upper and lower thresholds of rainfall intensity in each homogeneous region, for each time scale. At the monthly scale, sequential patterns of occurrences of extreme positive and negative events were identified over the thirty years. The results reinforce the potential of the data mining technique to group homogeneous zones by similarity of rain, using the k-means algorithm. They also reveal, for all time scales used, a high rainfall variability. The years with the highest recorded extreme negative events are in the 90's and those registered with more extreme positive events are in the 2000's.

**KEYWORDS:** Knowledge discovery; Rainfall; homogeneous zones; Extreme phenomena; Sequential patterns, Quantile methods.

## LISTA DE FIGURAS

Figura 1: Mapa do Relevo do Estado da Bahia (Fonte: SEI, 2010). .....	18
Figura 2: Mapa da Distribuição dos Ecossistemas da Bahia (Fonte: SEI, 2007). .....	19
Figura 3: Distribuição dos setores e atividades econômicas do Estado da Bahia (Fonte: SEI, 2010). .....	21
Figura 4: Mapa do Clima da Bahia. (Fonte: Adaptado do Atlas Nacional do Clima do Brasil do IBGE, 2010). .....	24
Figura 5: Áreas Pertencentes ao Semiárido Brasileiro (Fonte: BRASIL, 2005). .....	26
Figura 6: Área de distribuição dos principais regimes de chuva sobre o Nordeste. Fonte: adaptado de Kousky, 1979. ....	28
Figura 7: Processo do KDD, adaptado de FAYYAD <i>et al.</i> (1996). .....	35
Figura 8: Fases da metodologia CRISP-DM (CHAPMAN <i>et al.</i> , 2000). .....	36
Figura 9: Tarefas de mineração de dados de acordo com REZENDE <i>et al.</i> (2005). .....	38
Figura 10: Mapa de Localização da Área de Estudo: Estado da Bahia (Fonte: IBGE, 2011). ....	46
Figura 11: Distribuição espacial das 92 estações meteorológicas do estado da Bahia. ....	48
Figura 12: Mapa Pluviométrico da Bahia (1981 a 2010). .....	58
Figura 13: Regionalização pluviométrica do estado da Bahia. ....	59

## LISTA DE GRÁFICOS

Gráfico 1: Produtividade agrícola do Estado da Bahia referente ao período de 1990 a 2003. Fonte: SEI, 2010. ....	22
Gráfico 2: Sazonalidade da Precipitação de 1981 a 2010 com destaque para o ano de 1981. ....	54
Gráfico 3: (a), (b), (c), (d), (e) - Distribuição do acumulado anual de precipitação para o período de 1981 a 2010 nos cinco <i>clusters</i> . ....	62
Gráfico 4: Precipitação média mensal dos <i>clusters</i> referente ao período de 1981 a 2010. ....	63
Gráfico 5: Sazonalidade da precipitação dos anos de 1985 e 1993 – <i>Cluster 0</i> . ....	70
Gráfico 6: Sazonalidade da precipitação dos anos de 1985 e 1993 – <i>Cluster 1</i> . ....	71
Gráfico 7: Sazonalidade da precipitação dos anos de 1985 e 1993 – <i>Cluster 2</i> . ....	72
Gráfico 8: Sazonalidade da precipitação dos anos de 1985 e 1993 – <i>Cluster 3</i> . ....	72
Gráfico 9: Sazonalidade da precipitação dos anos de 1985 e 1990 – <i>Cluster 4</i> . ....	73
Gráfico 10: Distribuição de Eventos de Chuva por década, na Bahia. Tem-se para MS = Muito Seco; S = Seco; N = Normal; C = Chuvoso; MC = Muito Chuvoso. ....	74
Gráfico 11: Produtividade agrícola do estado da Bahia. Fonte: SEI, 2010. ....	77
Gráfico 12: (a), (b), (c), (d), (e) – Distribuição da precipitação pluviométrica acumulada mensal para o período de 1981 – 2010. ....	78
Gráfico 13: Acumulado de chuva mensal para os meses classificados como muito secos no <i>cluster 0</i> . ....	80
Gráfico 14: Acumulado de chuva mensal para os meses classificados como muito chuvosos - <i>cluster 0</i> . ....	81
Gráfico 15: Acumulado de chuva mensal para os meses classificados como muito secos - <i>cluster 1</i> . ....	81
Gráfico 16: Acumulado de chuva mensal para os meses classificados como muito chuvosos - <i>cluster 1</i> . ....	82
Gráfico 17: Acumulado de chuva mensal para os meses classificados como muito secos - <i>cluster 2</i> . ....	83
Gráfico 18: Acumulado de chuva mensal para os meses classificados como muito chuvosos - <i>cluster 2</i> . ....	83

Gráfico 19: Acumulado de chuva mensal para os meses classificados como muito secos - <i>cluster 3</i> . .....	84
Gráfico 20: Acumulado de chuva mensal para os meses classificados como muito chuvosos - <i>cluster 3</i> . .....	85
Gráfico 21: Acumulado de chuva mensal para os meses classificados como muito secos - <i>cluster 4</i> . .....	85
Gráfico 22: Acumulado de chuva mensal para os meses classificados como muito chuvosos - <i>cluster 4</i> . .....	86
Gráfico 23: Distribuição do total de dias secos e número de sequências com dias secos consecutivos- 1981 a 2010 – <i>Cluster 0</i> .....	93
Gráfico 24: Distribuição do total de dias secos e número de sequências com dias secos consecutivos - 1981 a 2010 – <i>Cluster 1</i> .....	93
Gráfico 25: Distribuição do total de dias secos e número de sequências com dias secos consecutivos - 1981 a 2010 – <i>Cluster 2</i> .....	94
Gráfico 26: Distribuição do total de dias secos e número de sequências com dias secos consecutivos - 1981 a 2010 – <i>Cluster 3</i> .....	94
Gráfico 27: Distribuição do total de dias secos e número de sequências com dias secos consecutivos - 1981 a 2010 – <i>Cluster 4</i> .....	94
Gráfico 28: Distribuição do número máximo de dias secos consecutivos em cada ano (1981 – 2010) para o <i>cluster 0</i> . .....	96
Gráfico 29: Distribuição do número máximo de dias secos consecutivos em cada ano (1981 – 2010) para o <i>cluster 1</i> . .....	96
Gráfico 30: Distribuição do número máximo de dias secos consecutivos em cada ano (1981 – 2010) para o <i>cluster 2</i> . .....	96
Gráfico 31: Distribuição do número máximo de dias secos consecutivos em cada ano (1981 – 2010) para o <i>cluster 3</i> . .....	97
Gráfico 32: Distribuição do número máximo de dias secos consecutivos em cada ano (1981 – 2010) para o <i>cluster 4</i> . .....	97

## LISTA DE TABELAS

Tabela 1: Mudanças observadas em temperatura e extremos de precipitação referentes ao período de 1961 a 1990 e projeções para 2071 a 2100. Fonte: Adaptado do SREX - IPCC, 2012. ....	8
Tabela 2: Características da tipologia climática da Bahia. ....	24
Tabela 3: Categorias de Precipitação usando a Técnica dos Quantis ( $p = 0,15$ e $0,85$ ). ....	52
Tabela 4: Categorias de Precipitação usando a Técnica dos Quantis ( $p = 0,05$ e $0,95$ ). ....	52
Tabela 5: Esquematização para obtenção do Índice de Sazonalidade (Acumulado da Precipitação Mensal e Média da Precipitação Anual). ....	53
Tabela 6: Esquematização para obtenção do Índice de Sazonalidade (Índice Sazonal Mensal e Sazonalidade Mensal para o período de 1981 a 2010). ....	53
Tabela 7: Esquematização do conjunto de entrada base do algoritmo <i>GSP</i> . ....	55
Tabela 8: Valores acumulados anuais de precipitações, médias, desvio-padrão, coeficiente de variação, máximo e mínimo para o período de 1981 a 2010. ....	62
Tabela 9: Contribuição da precipitação nos períodos chuvosos e secos para precipitação anual. ....	65
Tabela 10: Precipitação e Classificação dos anos pelos quantis nas regiões pluviometricamente homogêneas (mm) Categorias: MS= “muito seco”; S= “seco”; N= “normal”; C= “chuvoso” e MC= “muito chuvoso”. ....	66
Tabela 11: Limites de classificação anual para identificação de eventos extremos “Muito Seco” e “Muito Chuvoso” para todos os <i>clusters</i> . ....	67
Tabela 12: Identificação dos anos com ocorrência de eventos extremos na Bahia. ....	68

Tabela 13: Relação dos anos de ocorrência e intensidade dos fenômenos de El Niño e La Niña. ....	76
Tabela 14: Definição de limiares de chuva para o estado da Bahia. ....	79
Tabela 15: Identificação dos meses com mais ocorrências de eventos extremos para uma sequência de tamanho um. ....	87
Tabela 16: Sequências de eventos extremos que mais ocorreram na série temporal de 1981 a 2010. ....	89
Tabela 17: Definição de limiares de eventos extremos para um acumulado em 24horas/dia. ....	92

## Sumário

1. INTRODUÇÃO.....	1
1.1 ESTRUTURA DO TRABALHO.....	5
1.2 HIPÓTESE.....	5
1.3 OBJETIVO.....	6
2. REVISÃO BIBLIOGRÁFICA.....	7
2.1 Eventos Extremos.....	7
2.2 Eventos Extremos de Precipitação.....	9
2.3 Impactos dos Eventos Extremos na Agricultura.....	11
2.4 Modelos de Avaliação e Quantificação dos Eventos Extremos de Chuva.....	14
2.5 Aspectos Fisiográficos.....	17
2.5.1 Aspectos Socioeconômicos.....	19
2.5.2 Aspectos Climáticos.....	23
2.5.3 Sistemas Meteorológicos Responsáveis pelo Regime Pluviométrico do Estado da Bahia.....	26
2.6 A Descoberta do Conhecimento em Bases de Dados – <i>Data Mining</i> .....	34
2.7 Modelo do Processo de Descoberta de Conhecimento em Bases de Dados.....	35
2.8 Tarefas e Técnicas de Mineração de Dados.....	37
2.8.1 Predição.....	38
2.8.2 Descrição.....	38
2.9 Descrição das Técnicas de Mineração de Dados Utilizadas.....	39
2.9.1 Clusterização.....	39
2.9.2 Padrões Sequenciais.....	41
2.10 Aplicações de Técnicas de Mineração de Dados na Agricultura e Climatologia.....	42
3. MATERIAL E MÉTODOS.....	45
3.1 Área de Estudo.....	45
3.2 Modelo do Processo.....	46
3.2.1 Compreensão do Domínio.....	46
3.2.2 Entendimento dos Dados.....	46
3.2.2.1 Pluviometria do Estado da Bahia.....	48

3.2.3	Preparação dos Dados.....	49
3.2.4	Modelagem.....	50
3.2.4.1	Obtenção de Zonas Pluviometricamente Homogêneas .....	50
3.2.5	Avaliação dos Clusters.....	51
3.3	Análise de Eventos Extremos de Precipitação.....	51
3.3.1	Análise Anual .....	52
3.3.2	Análise Mensal.....	54
3.3.3	Análise Diária .....	55
4.	RESULTADOS E DISCUSSÃO .....	57
4.1	Distribuição da Precipitação Pluviométrica .....	57
4.2	Identificação das Zonas Pluviometricamente Homogêneas .....	59
4.3.	Análise da Precipitação nas Zonas Homogêneas .....	61
4.4	Análise de Eventos Extremos de Precipitação em Diferentes Escalas Temporais. ....	65
4.4.1	Análise Anual.....	65
4.4.2	Análise Mensal .....	78
4.4.3	Análise Diária.....	90
5.	CONCLUSÃO.....	98
	REFERÊNCIAS BIBLIOGRÁFICAS .....	100
	APÊNDICE I – Tabelas de classificação de eventos extremos em escala anual.....	113
	APÊNDICE II – Relação dos municípios e estações meteorológicas utilizadas. ....	114
	APÊNDICE III – Sazonalidade <i>Cluster 0</i> .....	117
	APÊNDICE IV – Sazonalidade <i>Cluster 1</i> .....	120
	APÊNDICE V – Sazonalidade <i>Cluster 2</i> .....	123
	APÊNDICE VI- Sazonalidade <i>Cluster 3</i> .....	126
	APÊNDICE VII – Sazonalidade <i>Cluster 4</i> .....	129



# 1. INTRODUÇÃO

O interesse pelo tema da variabilidade, mudanças e eventos extremos climáticos por diversos pesquisadores, órgãos públicos e privados vem aumentando de forma bastante acentuada nas últimas décadas. Isso se deve principalmente às inúmeras modificações no sistema terrestre, ocorridas em proporções nunca antes assistidas na história da humanidade. Em frequência e intensidade, os eventos que tornam evidentes a variabilidade e mudança no clima global têm gerado consequências à biodiversidade dos sistemas naturais, perdas econômicas e de vidas (TANAJURA, 2010).

Como resposta à crescente demanda de estudos relacionados ao tema supracitado, a Organização Mundial de Meteorologia e o Programa das Nações Unidas para o Meio Ambiente criaram o *Intergovernmental Panel on Climate Change* (IPCC, sigla em inglês). Desde então, relatórios de avaliação sobre o clima atual e projeções futuras têm sido elaborados e divulgados pela comunidade científica.

Ao abordar o tema de riscos de extremos climáticos, novas evidências foram divulgadas no mais recente relatório do IPCC. O Relatório Especial sobre Gestão dos Riscos de Extremos Climáticos e Desastres revela uma intensificação nesses últimos anos de eventos extremos, como secas, enchentes, ondas de calor e de frio, furacões e tempestades, afetando diferentes partes do planeta (IPCC, 2012).

Neste contexto, diversos estudos (CARVALHO, *et al.* 2002; MARENGO, *et al.*, 2007; GRIMM, *et al.*, 2009) sobre eventos extremos climáticos no Brasil se juntam a essas pesquisas uma vez que o clima do país também tem experimentado alterações, durante os últimos anos, causando perdas econômicas e de vidas. Em particular, na região semiárida, o aumento da ocorrência de secas severas em quase sua totalidade e focos de enchentes são exemplos reais que servem de advertência para o futuro.

Historicamente, a região Nordeste sempre foi submetida aos rigores das variações climáticas que caracterizam o clima semiárido. E o fenômeno da seca se repete com frequência e variada intensidade, caracterizando-se como um risco para os diversos setores econômicos e sociais, ao longo de toda uma história de ocupação e povoamento da área.

Apesar de não se tratar de um elemento novo no cenário nordestino, manifesta-se como o fator de maior impacto econômico, social e ecológico, pelos prejuízos que causa às áreas afetadas. Neste aspecto, diversos levantamentos já foram realizados por estudiosos do

assunto, no sentido de resgatar a história das secas. Tais registros históricos indicam que anos de secas e chuvas abundantes se alternam de formas erráticas, e grandes são as secas de 1710-11, 1723-27, 1736-57, 1744-45, 1777-78, 1808-09, 1824-25, 1835-37, 1844-45, 1877-79, 1982-83, 1992-93, 1995-1996, 1997-98, 2000-01, 2003 (SUDENE, 1992; CARVALHO, 1994; BARBOSA, 2000; SEI, 2010).

Além do aprofundamento no entendimento da climatologia da região, a frequência e intensidade dos eventos extremos ocorridos se tornaram alvo de grandes estudos, impulsionando com mais intrepidez a busca pelo conhecimento das características, alterações e tendências desses eventos extremos, com vistas à tomada de decisão de precaução, mitigação e adaptação para os diversos setores da sociedade, que dependem diretamente dessa variável, tais como: agricultura, pesca, produção de energia elétrica, dentre outros.

Inserido ao sul da região Nordeste, o estado da Bahia, apresenta vasta extensão de seu interior submetido aos rigores do clima semiárido. Do ponto de vista climático, essa extensa área interiorana do território baiano, mesmo participando da problemática comum do conjunto do semiárido, apresenta certas peculiaridades. Ou seja, nem todos os eventos de seca que atingem o Nordeste podem afetar igualmente a Bahia. Tal fato decorre principalmente de sua posição geográfica de transição, não apenas em relação ao norte do Nordeste, mas também as demais regiões limítrofes ao estado, que apresentam diferentes domínios climáticos (BARBOSA, 2000).

Caracterizado por uma diversidade de domínios climáticos a Bahia apresenta um quadro de alta variabilidade na precipitação pluvial, decorrente da conjunção de diferentes sistemas meteorológicos que atuam nessa região, tais como, Zona de Convergência Intertropical (ZCIT), Vórtices Ciclônicos (VCAN), Sistemas Frontais (FPA), Zona de Convergência do Atlântico Sul (ZCAS), ondas de leste, brisas marítimas/terrestres e ventos vale/montanha, bem como, de um relevo constituído por planícies, vales, serras e montanhas, (BARBOSA 2000; BRAGA *et al.*, 1998).

Além da atuação dos sistemas meteorológicos supracitados, a presença do fenômeno ENOS (Índice de Oscilação Sul) influencia na distribuição pluviométrica da região provocando seca na fase quente (El Niño) e chuvas intensas na fase fria (La Niña). A ocorrência dos fenômenos de secas em quase todo território baiano e eventos de chuvas fortes (enchentes) em alguns municípios impõem ao estado atitudes preventivas e emergenciais em

virtude dos eventos de excesso e déficit hídrico ocorridos distintamente em sua totalidade territorial.

O estado tem sido palco nos últimos anos de fortes ocorrências desses eventos extremos, como é o caso da grande seca que atingiu a região de Irecê no ano de 1993. Tal evento gerou elevadas perdas na produção de feijão, com totais percentuais de -80,5% em relação ao ano anterior (BARBOSA, 2000). Apesar desses registros, a Secretaria da Agricultura e Irrigação (SEAGRI, 2012) sinaliza de forma otimista um avanço no setor agropecuário no estado, devido à disponibilidade de terras agricultáveis, à expansão de áreas de cultivos, à tecnologia especializada e incentivos governamentais.

É conhecido que as atividades agrícolas estão sempre expostas às variações dos elementos climáticos. A precipitação e a temperatura podem apresentar grande variabilidade, o que reduz a produtividade agrícola. A agricultura moderna exige que se conheça o comportamento do clima ou, pelo menos, a tendência térmica e hídrica de uma determinada região para que os plantios e suas práticas culturais sejam bem definidas e orientadas, aumentando-se a produtividade e, conseqüentemente, os lucros econômicos.

Diante disto, surge a necessidade de sistemas que consigam analisar, com confiabilidade, informações sobre o comportamento e a frequência de ocorrência de alguns fenômenos climáticos, determinantes para os diversos setores socioeconômicos e ambientais. Particularmente, para o setor agrícola que se mostra altamente dependente de tais fenômenos, uma análise de como os eventos naturais usuais e extremos do clima têm ocorrido ao longo dos anos, torna-se uma ferramenta eficaz no processo de tomada de decisão, vislumbrando a redução de impactos severos na produção.

Como uma alternativa promissora para analisar o comportamento e a frequência de ocorrência de fenômenos climáticos, as técnicas de mineração de dados demonstram forte potencial, uma vez que estas permitem transformar informações embutidas em grandes volumes de dados em conhecimento útil, fundamentado na definição do processo KDD - Descoberta do Conhecimento em Bases de Dados (FAYYAD *et al.*, 1996).

Não existe a melhor técnica no processo da mineração de dados, já que cada uma possui vantagens e desvantagens. A escolha de uma técnica exige uma análise mais detalhada do problema em questão e a decisão de qual representação e estratégia de descoberta seja mais adequada (NONATO, 2010).

Em particular, neste trabalho, as tarefas aplicadas ocorrem em duas fases: na primeira, como o intuito de transformar as séries históricas de dados de precipitação do estado da Bahia em zonas pluviometricamente homogêneas, foi aplicada a *Clusterização* (ou análise de agrupamento). A partir da identificação das zonas pluviométricas, com o uso da Técnica dos Quantis, foram detectados e analisados a ocorrência de eventos extremos para as escalas anuais, mensais e diárias para cada região. Em particular, para escala mensal foi utilizada outra tarefa de mineração de dados, a *Associação*, visando identificar os padrões sequenciais das ocorrências dos eventos extremos da região, em cada zona homogênea.

Espera-se com isso, que os resultados obtidos sobre o comportamento da ocorrência e frequência dos eventos extremos para cada região homogênea pluviométrica, subsidiem ações preventivas e mitigadoras desses eventos extremos no Estado, e assim promova suporte à tomada de decisão para os diversos setores socioeconômicos, em particular ao setor agrícola.

## **1.1 ESTRUTURA DO TRABALHO**

Para uma melhor compreensão do trabalho, o capítulo 2 apresenta uma revisão bibliográfica da temática desse estudo. Em particular, o capítulo está subdividido em: definição dos eventos extremos; eventos extremos de precipitação; impactos dos eventos extremos na agricultura; modelos de avaliação e quantificação dos eventos extremos de precipitação; conceitos gerais e tarefas de mineração de dados, destacando-se a técnica de clusterização (agrupamento de dados) e padrões sequenciais, as quais foram utilizadas para desenvolvimento do trabalho, por fim, aborda as características fisiográficas, socioeconômicas e climáticas do estado da Bahia.

O capítulo 3 apresenta o material e os métodos utilizados, desde a aquisição dos dados, processamento e análise dos resultados para as diferentes escalas temporais (anual, mensal e diária) utilizadas no trabalho.

O capítulo 4 expõe os resultados obtidos com o desenvolvimento deste trabalho, e traz para discussão dos mesmos outros trabalhos que venham corroborar o uso das metodologias aplicadas.

Finalmente, o capítulo 5 apresenta as conclusões das atividades, revelando a novidade da pesquisa e benefícios que a mesma proporciona para comunidade científica, órgãos públicos e privados, consolidando o conhecimento sobre o tema escolhido para estudo.

## **1.2 HIPÓTESE**

É possível caracterizar padrões de ocorrência de eventos extremos de precipitação por meio de técnicas de mineração de dados, com o propósito de identificar o comportamento e a frequência das ocorrências desses eventos, em séries históricas de dados climáticos.

## **1.3 OBJETIVO**

### **Objetivo Geral**

Utilizar técnicas de mineração de dados que possibilite a extração do conhecimento em um banco de dados climáticos, para fornecer uma análise sobre o comportamento e a frequência das ocorrências dos eventos extremos de precipitação entre 1981 e 2010 no estado da Bahia, com a finalidade de auxiliar a tomada de decisão referente às ações preventivas e mitigadoras dos impactos socioeconômicos e ambientais desses eventos.

### **Objetivos Específicos**

- 1) Identificar zonas com homogeneidade pluviométrica no estado da Bahia, para o período de 1981 a 2010, tendo como ferramentas algoritmos para agrupamentos de dados;
- 2) Caracterizar o regime pluviométrico de cada zona homogênea por meio da análise de séries históricas dos dados de precipitação em escala anual, mensal e diária;
- 3) Identificar os limiares de intensidade de chuva a partir do qual um evento será considerado extremo para cada zona homogênea;
- 4) Avaliar o comportamento e a frequência nas ocorrências dos eventos extremos ao longo dos anos para cada região homogênea;
- 5) Investigar se existe relação entre os anos de ocorrência de eventos extremos, positivos ou negativos, com os anos de presença dos fenômenos El Niño e La Niña;
- 6) Identificar a estação chuvosa de cada região do estado, para definir os padrões sequenciais que caracterizem os eventos extremos positivos e negativos dos respectivos meses desse período, considerando séries históricas de 1981 a 2010.

## 2. REVISÃO BIBLIOGRÁFICA

### 2.1 Eventos Extremos

Visando a consolidação do esforço de inúmeras pesquisas científicas sobre as conexões entre a variabilidade climática e as atividades humanas, mudanças climáticas e cenários futuros de mudança no clima global, o Programa das Nações Unidas para o Meio Ambiente (PNUMA) e a Organização Meteorológica Mundial (OMM) estabeleceram em 1988 o Painel Intergovernamental de Mudanças Climáticas, cuja sigla em inglês é IPCC.

Além dos Relatórios de Avaliação das Mudanças Climáticas, o IPCC anunciou em março de 2012 o Relatório Especial sobre Gestão dos Riscos de Extremos Climáticos e Desastres (SREX, na sigla em inglês). Este relatório leva em consideração os efeitos da mudança do clima sobre eventos extremos, desastres e o gerenciamento de risco de desastres. Ele analisa como os extremos climáticos, fatores humanos e o meio ambiente interagem para influenciar os impactos dos desastres e gerenciamento de risco, e as opções de adaptação (IPCC, 2012).

Segundo o Relatório Especial (IPCC, 2012), um evento extremo é definido como a ocorrência de um valor de uma variável de condição meteorológica ou clima acima (ou abaixo) de um valor limite, perto das extremidades ('caudas') superiores (ou inferiores) da faixa de valores da variável observados. Consolidando essa definição, Cameron *et al.* (2012) afirmam que um clima em mudança provoca alterações na frequência, intensidade, extensão espacial e duração de extremos das condições meteorológicas e climáticos, podendo resultar em eventos sem precedentes.

Os eventos climáticos podem ser classificados em eventos usuais e extremos. Barbosa (2008) define eventos usuais como os episódios registrados com maior frequência possibilitando uma melhor absorção pelas sociedades e um planejamento de adaptação ao seu ritmo natural. Tais eventos, também são definidos como aqueles que não se afastam significativamente das médias, com uma frequência alta em qualquer escala temporal de ocorrência (GONÇALVES, 2003).

Há evidências de que eventos extremos, tais como secas, enchentes, ondas de calor e de frio, furacões e tempestades têm sido mais frequentes ou mais intensos, e têm impactado

tanto os sistemas humanos quanto os ecossistemas, além de provocar perdas em diversos setores econômicos e vidas em várias partes do planeta.

Com base nos números registrados pelo *Emergency Events Database* (EM - DAT) do *International Disaster Database*, nos últimos anos ocorreu um aumento exponencial da frequência e intensidade de desastres causados por eventos extremos (EM – DAT, 2012). E os resultados do SREX, para o Brasil, confirmam esse aumento com as observações dos dados de temperatura e precipitação referentes ao período de 1961 a 1990 revelando também as mudanças projetadas para o período de 2071 – 2100, conforme Tabela 1.

Região	Período	Tendências em Temp Máx.(Dias quentes e frios)	Tendências em Temp Mín.(Noites quentes e frias)	Tendências em Ondas de calor / períodos de calor	Tendências em Precipitação pesada (chuva, neve)	Tendências em Aridez e estiagem
N o r d e s t e	1961 a 1990	Aumento no nº de dias quentes	Aumento no nº de noites quentes	Evidência insuficiente	Aumento em muitas áreas, diminuição em poucas áreas	Tendências variáveis e inconsistentes
	2071 a 2100	Provável aumento de dias quentes e diminuição de dias frios	Provável aumento de quentes e diminuição de noites frias	Ondas de calor/ períodos de calor provavelmente mais frequentes e longos	Mudança leve ou nenhuma em relação ao período de 1961 - 1990	Aumento de dias

**Tabela 1:** Mudanças observadas em temperatura e extremos de precipitação referentes ao período de 1961 a 1990 e projeções para 2071 a 2100. Fonte: Adaptado do SREX - IPCC, 2012.

Avalia-se que, os desastres naturais mais comuns que ocorrem no Brasil são os relacionados com os eventos hidrológicos, as enchentes, seca, erosão e escorregamentos ou deslizamentos de terra (MARENGO, 2007; BRASIL, 2007). Desta forma, os registros do *International Disaster Database* para ocorrências de desastres naturais no Brasil, referentes ao período de 1900 a 2010, mostram que o número de pessoas atingidas por enchentes e deslizamentos chega a 14.179.949 e atingidos pela seca alcança 45.000.000 de pessoas (EM – DAT, 2012).

Barbosa (2008) destaca a possibilidade da ocorrência de uma aceleração do ciclo hidrológico, gerando intensificação de eventos extremos. Com isso, eventos como deslizamento de terra, inundações e vendavais podem tornar-se mais frequentes e intensos, principalmente nas regiões Sul e Sudeste do Brasil, enquanto o quadro de desertificação pode agravar-se na região Nordeste.



Eventos extremos, como secas ou enchentes severas alteram consideravelmente as características habituais de uma dada região, desde a física, movimentando massas e redistribuindo algumas características da paisagem, como a social, por exemplo, causando grandes transtornos sociais, como no caso dos longos períodos de estiagem no sertão.

A região Nordeste do Brasil, apesar de chover tanto quanto em muitas outras regiões do mundo, em particular na parte semiárida, é periodicamente afetada pela ocorrência de secas com perdas parciais ou totais na agropecuária, devido principalmente à irregularidade da quadra chuvosa na região, com predominância de chuvas intensas e de curta duração (Santos *et al.*, 2010). Estudos recentes indicam que esta região vem sofrendo com fortes variabilidades climáticas (SILVA, 2004; SOUZA JÚNIOR, 2006).

A ocorrência de um período de seca no ano de 1993 atingiu todos os estados do Nordeste e, de acordo com dados da Superintendência de Desenvolvimento do Nordeste, esse evento levou um total de 1.857.655 trabalhadores rurais, que perderam suas lavouras, a se alistarem nas chamadas "frentes de emergência", onde a Bahia foi o estado com maior número de trabalhadores alistados, cerca de 369.000.

## **2.2 Eventos Extremos de Precipitação**

As relações entre os eventos extremos positivos e negativos das chuvas resultam em impactos que provocam muitos problemas para zona rural e urbana. Tais impactos revelam a tamanha vulnerabilidade das sociedades frente aos fenômenos climáticos. Monteiro *et al.* (2012) afirmam que as populações urbanas, geralmente são mais atingidas quando as chuvas apresentam anomalias positivas, enquanto que as rurais quando as chuvas são escassas.

Numa abordagem conceitual sobre os eventos extremos de precipitação, Sarewitz *et al.* (2000) tratam estes fenômenos como sendo ocorrências que apresentam uma incidência rara, distanciando-se da média, variando em sua magnitude. Porém, essa perturbação ocorre por um período determinado, voltando posteriormente ao seu estado habitual.

Outro aspecto abordado por Monteiro *et al.* (2012) é que a chuva é uma variável aleatória e o valor de sua altura acumulada (em milímetros) não poderá ser previsto com uma exatidão determinística, mas na verdade ela será de natureza probabilística. Ou seja, pode-se atribuir uma probabilidade para que a altura da chuva fique compreendida entre dois limites arbitrariamente escolhidos.

Diversos são os parâmetros e aspectos relacionados à classificação desses eventos, dentre eles, mencionam-se os atributos de quantidade de chuva em um intervalo de tempo e a abrangência espacial desta chuva.

Porém, há uma grande dificuldade entre os pesquisadores em determinar limiares para os eventos extremos, afirma Monteiro *et al.* (2012). Isso se deve à dificuldade de estabelecer valores (em milímetros) confiáveis para regiões com características pluviométricas e climáticas diferenciadas em todo país e a partir de que montante pode ser considerado um evento extremo.

Sobre esta abordagem, Xavier *et al.* (1999) afirmam que a Técnica dos quantis (muito utilizada em seus trabalhos) envolve o princípio da relativização estatística, no sentido de se buscar um significado apropriado para altura pluviométrica registrada para dada localidade. Sobre isto, os autores relatam que 350mm de precipitação acumulada para um dado ano poderá ser “normal” para um local, onde a média pluviométrica seja de 480mm. No entanto, poderá cair para o nível “muito seco” para outro local onde a média seja de 1460mm, ou para o nível “muito chuvoso” se a média for de 110mm.

Essa situação foi observada nos estudos de Monteiro *et al.* (2012), onde para valores de chuva acumulada em um ano, que ultrapassaram 1200mm para região de Mossoró (interior do Rio Grande do Norte), foi indicado como um ano muito chuvoso, enquanto que os mesmos 1200mm em uma região litorânea no mesmo estado foi considerado como um ano normal. Outros trabalhos em diversas partes do Brasil encontraram situações semelhantes, visto por Xavier *et al.* (2001) para o Ceará; Cardoso *et al.* (2004) para o Sul e Sudeste; André *et al.* (2008); Farias *et al.* (2012) para Pernambuco.

De forma generalizada, o histórico dos eventos extremos de seca que assolaram a região Nordeste abrange toda extensão territorial de seus estados. Com base nas afirmações supracitadas, sobre homogeneidade pluviométrica das regiões e classificação de eventos extremos, Xavier *et al.* (1999), utilizando a série histórica para o período de 1964 a 1998, reavaliaram os anos registrados como de grandes secas para o estado do Ceará (1979 – 1983, 1993, 1997 e 1998). Verificaram que somente os registros de seca para os anos de 1982/83, 93 e 98, foram confirmados como anos de seca para todo o estado, os demais anos apresentaram discrepância quanto à classificação em algumas regiões. Os anos de destaque para seca em todo o estado estão relacionados aos anos de ocorrência de El Niño com intensidade forte.

Outra questão a ser colocada é o que se refere à análise de identificação de eventos extremos numa escala diária. Souza *et al.* (2012) questionaram sobre a partir de quantos milímetros uma chuva pode ser definida como um evento danoso. Em seus estudos para cidade de Recife, utilizaram a técnica do quantil para dados de precipitação acumulada em 24 horas, consideraram como limiares para definição dos eventos extremos, em escala diária, o quantil Q (0,5) para um Dia Seco (apresentando chuva inferior a 2mm/dia) e o quantil Q (0,95) para uma Chuva Muito Forte.

É importante ressaltar que a escolha do valor 2mm/dia para um Dia Seco, fundamenta-se na determinação de limiares inferiores para regiões de clima quente, discutida pelos autores Adejuwon e Odekunle (2006), após diversos estudos em que a precipitação total diária menor que 2mm é de pouca importância para agricultura e abastecimento de água. Estes autores justificam que valores inferiores a este limiar vão evaporar antes mesmo de infiltrar no solo, em virtude das elevadas temperaturas e taxas de evaporação e evapotranspiração.

Outra abordagem para análise de eventos extremos em escala diária foi utilizada por Dufek (2008) ao adotar índices para dias secos (cdd) e dias úmidos (cwd), demonstrando tendências positivas ou negativas para América do Sul. Foi analisado o número máximo de dias consecutivos em que a chuva diária foi inferior (cdd) ou superior (cwd) a 1mm, podendo assim definir períodos de excessiva seca e umidade, respectivamente.

### **2.3 Impactos dos Eventos Extremos na Agricultura**

A agricultura é uma atividade altamente dependente de parâmetros agrometeorológicos, como temperatura, pluviosidade, umidade do solo e radiação solar. A frequência e a severidade de eventos extremos provocam mudança na intensidade da colheita devido à alteração no número de graus-dia de crescimento, modificam o aparecimento e a severidade de pragas e doenças, e podem provocar abortamento de frutos (MORAES *et al.*, 2011).

No contexto da agricultura de sequeiro, Azevedo e Silva (2007) destacam que a estação de crescimento das culturas depende da época em que as chuvas efetivamente começam. Os períodos chuvosos com início precoce são, em geral, potencialmente superiores na produtividade das culturas em relação aos períodos chuvosos de início tardio. Diniz *et al.* (2011) também atentaram para a importância do comportamento irregular do período chuvoso

para a agricultura quando estudaram a duração da estação chuvosa em Feira de Santana, Bahia.

Monteiro (1991) ressalta que os eventos extremos de precipitação podem resultar em impactos positivos, ou seja, o aumento da sanidade em ecossistemas causado por inundações periódicas, ou a despoluição da atmosfera de um determinado local em virtude de um impacto pluvial são alguns desses impactos positivos. Do mesmo modo, para o setor agrícola, há a possibilidade da inserção de novas áreas com aptidão agrícola para produção de outras culturas.

Apesar disto, em diversas partes do planeta, inúmeros são os registros dos impactos decorrentes dos episódios de eventos extremos de chuva relacionados aos setores socioeconômicos e ambientais. Segundo a FAO (agência da ONU para agricultura e alimentação), as chuvas e inundações no Paquistão destruíram 73% das lavouras do país e 67% dos estoques alimentares na populosa província do Sindh, mataram quase 78 mil cabeças de gado, e que pelo menos 5 milhões de animais estão ameaçados pela falta de ração, abrigo e cuidados veterinários (FAO, 2012).

Aqui no Brasil, 73 municípios no estado do Rio Grande do Sul foram afetados pelas chuvas de novembro de 2009 e janeiro de 2010, situados em especial no extremo noroeste, regiões Sul e Central daquele estado. O prejuízo econômico total que chegou aproximadamente a R\$ 595 milhões. Entre-Ijuís foi o município que teve as maiores perdas, relacionadas principalmente à agricultura, as quais representaram mais de 80% de seu PIB (SAITO, 2011).

Além dos impactos devido aos excessos das chuvas, grandes são os prejuízos no setor agrícola, gerados por causa da escassez hídrica. Não é rara na história do Nordeste brasileiro, em particular a região semiárida, a sucessão de anos seguidos de seca. No entanto, a limitação hídrica ocorre anualmente devido ao longo período seco que leva à desperenização dos rios e riachos endógenos. A reduzida capacidade de absorção de água da chuva no solo é dificultada em virtude do relevo alterado e dos solos rasos e pedregosos (SILVA, 2007).

De acordo com Silva (2007), no ano de 2000, a capacidade de acumulação total dos reservatórios existentes na região Nordeste era estimada em 86 bilhões de metros cúbicos de água. As águas represadas em açudes do Departamento Nacional de Obras Contra as Secas (DNOCS) propiciavam a exploração de 82 mil hectares de áreas de vazantes e da produção de

17 mil toneladas de pescado, o aproveitamento com irrigação de cerca de 107 mil hectares, sendo 67 mil hectares em áreas privadas e 30 mil hectares em áreas públicas, e o abastecimento humano de 1,7 milhão de pessoas em cidades do Semiárido brasileiro. Mesmo assim, a irrupção de estiagens prolongadas provoca crise de abastecimento de água para a população.

Andrade *et al.* (2012) descrevem o impacto da seca ocorrida no ano de 1992/93 para o município de Itaetê na Bahia gerando morte de animais, e perda de toda a safra de mamona, feijão e milho da região. Para o mesmo ano, Barbosa (2000) relata os grandes prejuízos ocorridos em todo Estado, com perdas totais de safras. No mesmo trabalho, o autor destaca o município de Irecê, famoso pelo seu potencial agropecuário, tendo recebido o título de “Cidade do Feijão”, pelas grandes safras colhidas em 1980 e 1990. Tal município teve toda sua produção de feijão desestruturada levando ao endividamento dos produtores rurais.

Registrados nos estudos de Diniz *et al.* (2008), que utilizaram uma série com dados de precipitação de dez anos (1996 a 2005), outros anos acometidos pelas secas trouxeram prejuízos para o setor agrícola no município de Feira de Santana na Bahia. Dentre eles, encontram-se os anos de 1998, 2001, 2002 e 2003, sendo o ano de 1998 o mais seco. Os autores relacionam os eventos de seca nesses anos aos de ocorrência de El Niño, reforçando a estreita relação da presença deste fenômeno e as severas secas que marcaram a região Nordeste.

Com base nas afirmações supracitadas sobre a relação direta dos impactos de eventos extremos na agricultura, segundo registros da Superintendência de Estudos Econômicos e Sociais da Bahia (SEI, 2010), a produção agrícola do estado da Bahia nos anos de 1990, 1993, 1996, 1998, 2001, 2003 e 2009 sofreu uma queda na safra agrícola (tanto na produção de grãos quanto no setor de fruticultura) devido a condicionantes climáticos.

Entretanto, a Bahia apresenta vários aspectos positivos para o cultivo de diversas culturas, devido à fertilidade de seus solos e principalmente disponibilidade energética, muito embora ofereça restrições relacionadas às disponibilidades hídricas (SILVA *et al.*, 2007). Tais vantagens vinculadas ao avanço de tecnologias, disponibilidade de áreas agricultáveis e incentivos agrícolas pelo governo têm proporcionado uma expansão da agropecuária do território baiano (SEAGRI, 2012).

## 2.4 Modelos de Avaliação e Quantificação dos Eventos Extremos de Chuva

A distribuição das chuvas, sua quantificação e monitoramento têm sido motivos da intensificação de inúmeras pesquisas científicas. Em particular, para os trabalhos de avaliação e quantificação dos eventos extremos de precipitação em diferentes regiões do Brasil, são utilizadas principalmente técnicas e modelos estatísticos. Nestes casos, destacam-se os trabalhos de Severo e Gan (2004), Gan (1982) para região Sul; Muza (2005, 2006), com a região Sul e Sudeste; Xavier *et al.* (1999, 2003) e Santos *et al.* (2011) na região Nordeste.

Outras pesquisas que reforçam o uso dessas técnicas, com enfoque para os eventos de excessos de chuvas (ESPÍRITO SANTO, 2002; GRIMM, 1998, 2009; TEIXEIRA, 2005; TEDESCHI, 2008) e eventos de déficit de chuvas (REPELLI *et al.*, 1991). Da mesma forma, a utilização de outros índices como: *Palmer Drought Severity Index* (PSDI), Índice de Anomalia de Precipitação, Índices de Porcentagem Normal, Índice de Precipitação Normalizada e Métodos dos Quantis, são aplicados para o estudo desses fenômenos (BLAIN, 2005; GÓIS, 2005; BRUNINI *et al.*, 2002a; PINKAYAN, 1966; XAVIER, 1984; XAVIER *et al.*, 2007).

Para quantificar a intensidade, déficit e excessos das chuvas, diversos índices já foram desenvolvidos por especialistas do assunto, cabendo ao pesquisador definir aquele que mais se aproxime das necessidades da análise. Altamirano (2010) menciona em seus estudos o Índice de Anomalias de Precipitação (IAP), que é calculado simplesmente como a diferença entre a média de precipitação para cada série e o tempo de escala selecionado, dividido pelo desvio padrão. O mesmo autor, numa abordagem sobre o Índice de Porcentagem Normal (IPN), relata que um dos principais inconvenientes do uso deste índice é que a distribuição da precipitação média muitas vezes não é o mesmo que a mediana da precipitação, o que torna difícil comparar os valores do índice para diferentes locais ou escalas temporais.

Originalmente proposto por Palmer (1965), o *Palmer Drought Severity Index* (PSDI) é uma medida de seca meteorológica. Este índice aparece com frequência em muitos dos estudos de períodos secos ou úmidos. A aplicação dele implica em um acréscimo de outros atributos, além da precipitação, tais como: temperatura, evapotranspiração, escoamento e condições de água do solo. O PSDI é baseado na equação do balanço hídrico do modelo de solo de duas camadas e varia aproximadamente entre - 4 e + 4. Cada conjunto definido de valores corresponde a uma classe particular de eventos.

Melo (2010), utilizando o PDSI, avaliou os fenômenos de seca e a existência de flutuações não-estacionárias em regiões do Nordeste. Altamirano (2010) apresenta como limitações seu uso no monitoramento dos eventos de seca em diferentes regiões. Neste sentido, o autor cita a sensibilidade da capacidade de água do solo (o qual necessita ter o conhecimento das propriedades do solo local), a estimativa da evapotranspiração potencial utilizando o método de Thornthwaite ou o uso de constantes para avaliar o coeficiente de ponderação que são arbitrariamente baseados em estudo de Palmer para Iowa Central e Ocidental de Kansas nos Estados Unidos.

Desenvolvido no Colorado, EUA, por McKee *et al.* (1993), o *Standardized Precipitation Index* (SPI) pode ser comparado entre regiões com climas distintos e usado em uma localidade com mais de 30 anos de chuvas (HAYES *et al.*, 1999). Está relacionado ao número de desvios padrão que a precipitação acumulada para uma determinada série de tempo (1 mês, 3 meses, 6 meses, 12 meses, etc.) se afasta da média climatológica para uma variável aleatória com distribuição normal.

Os estudos de precipitação realizados no Brasil com o uso do SPI concentram-se em eventos de déficit de chuvas (seca), não relacionando aos eventos de excesso de chuvas. Sobre isto, pode-se citar os Estados de: Minas Gerais (SILVA *et al.*, 2005; GÓIS, 2005), São Paulo (BRUNINI *et al.*, 2002b; BLAIN, 2005; BLAIN e BRUNINI, 2007a, 2007b), Paraíba (KUMAR e SOUSA, 2002; ALENCAR *et al.*, 2006) e Tocantins, na Bacia do rio São Francisco (MORAIS *et al.*, 2008).

Numa análise comparativa da aplicação do SPI em relação aos outros índices, tem-se como vantagem o fato de sua determinação ser feita apenas com base em dados de precipitação, de forma que a avaliação da seca ou chuva se torna possível, mesmo que outras variáveis hidrometeorológicas não sejam disponíveis.

Bastante utilizada para definição de limiares de eventos extremos, a técnica dos quantis está presente em diversos estudos, trazendo resultados mais confiáveis. A técnica envolve a relativização estatística, no sentido de se procurar interpretar de forma adequada o significado de um total pluviométrico. Sendo imune a eventual assimetria na função densidade de probabilidade que descreve o fenômeno aleatório (XAVIER *et al.*, 2002).

Sobre este fenômeno, Monteiro *et al.* (2012) afirmam que a chuva é uma variável aleatória e o valor de sua altura acumulada (em milímetros) não poderá ser previsto com uma

exatidão determinística, mas na verdade ela será de natureza probabilística. Ou seja, pode-se atribuir uma probabilidade para que a altura da chuva fique compreendida entre dois limites arbitrariamente escolhidos.

No cenário de busca por ferramentas e metodologias mais promissoras para determinação e análise de eventos extremos, o método dos quantis é um instrumento importante nos estudos desses “eventos raros”. Xavier *et al.* (2007) afirmam que, no que concerne a ocorrências de chuvas intensas, secas enchentes catastróficas, além de eventos identificados na categoria de desastres naturais, incluindo os de origem antrópica, este se torna um método bastante eficaz.

Por definição, “Quantis (“*quantiles*”, “*fractiles*”) são medidas de separação para distribuições de probabilidade ou para suas amostras. Um quantil de ordem  $p$  (definido para  $0 < p < 1$ ) é um valor numérico que secciona a distribuição em duas partes, com probabilidades  $p$  (à esquerda deste quantil “teórico”) e  $1-p$  (à direita). Em termos amostrais, permitindo separar uma amostra em duas massas de observações numéricas, com  $100xp\%$  dos elementos localizados à esquerda do quantil “amostral” e os demais  $100x(1-p)\%$  à direita”(XAVIER *et al.*, 2007).

O precursor da aplicação da Técnica dos Quantis para análises de eventos extremos foi Pinkayan em 1966, que categorizou em cinco classes os eventos de chuva conforme limiar estabelecido pela ordem quantílica, a saber: Muito Seco, Seco, Normal, Chuvoso, Muito Chuvoso. Estas classes são representadas pelos seus respectivos quantis  $Q(0,15)$ ,  $Q(0,35)$ ,  $Q(0,65)$ ,  $Q(0,85)$ . Adotando estas mesmas ordens quantílicas, a técnica do Quantil tem sido amplamente divulgada nos trabalhos de Xavier (1984, 1999, 2002, 2007) para o monitoramento de períodos secos e chuvosos para o estado do Ceará.

Tal análise necessita ser regionalizada para se conseguir levar em conta as disparidades, no que diz respeito à disponibilidade hídrica, em termos exclusivamente do aporte pluvial da região. Constatado por Monteiro *et al.* (2012), em estudos no estado Rio Grande do Norte, utilizando a técnica do quantis, verificou-se que os valores normais e extremos para chuva variam em determinadas áreas do Estado, sendo necessária a homogeneização das zonas pluviométricas para detecção de limiares mais confiáveis para os dados de precipitação.



Além do intervalo quantílico de 85%, outros percentis são utilizados para estudos de eventos extremos de chuvas, como exemplo os trabalhos de May (2004) em que os percentis de 99%, 99,5% e 99,75% foram utilizados para analisar a variabilidade dos eventos extremos de precipitação diária, durante a moção de verão na Índia. Muza *et al.* (2009) adotaram o percentil de 95% para avaliar os eventos extremos de chuva no Sudeste da América do Sul. Grimm e Tedeschi (2009) fizeram uso do percentil 90% para identificar os eventos extremos na região da América do Sul e analisar a influência dos fenômenos ENSO na frequência desses eventos.

Numa comparação da Técnica dos Quantis com o desempenho do SPI para o monitoramento das chuvas no Alto Sertão da Paraíba, Lima *et al.* (2006) adotando o intervalo quantílico de 15 – 85%, de acordo com as classificações de Xavier (2002), mostraram que o método SPI tende a produzir estações chuvosas normais em demasia, apontando uma maior confiabilidade para o uso da técnica de quantis para medir a qualidade da estação chuvosa ou do ano hidrológico para microrregião do Alto do Sertão da Paraíba.

## **2.5 Aspectos Fisiográficos**

O território da Bahia está situado na fachada atlântica do Brasil, com relevo caracterizado pela presença de planícies, planaltos, depressões e as formas tabulares e planas (chapadas, chapadões e tabuleiros). Apresenta como relevo predominante a depressão. Na região litorânea estão situadas as planícies, onde a altitude não ultrapassa os 200 metros, propiciando a formação de praias, dunas, restingas e até pântanos (Figura 1).

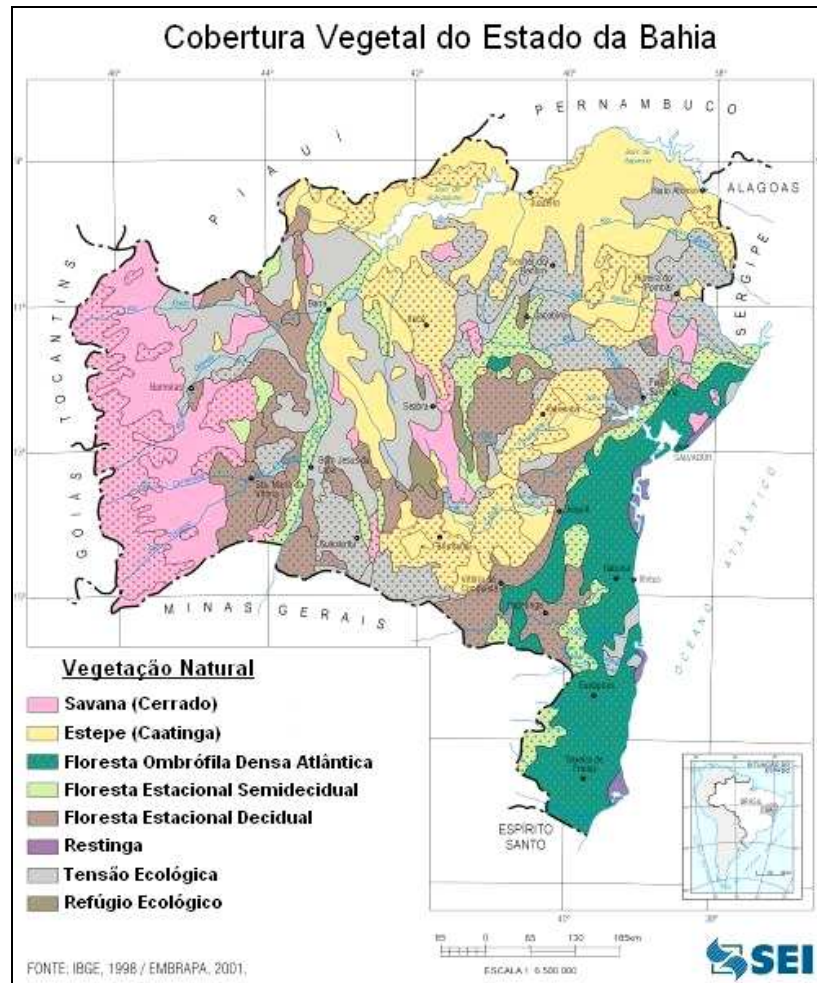
Com características bem acentuadas, os planaltos ocupam quase todo o estado, apresentando uma série de patamares, por onde cruzam rios vindos da Chapada Diamantina, da serra do Espinhaço, que nasce no centro de Minas Gerais, indo até o norte do estado. E a própria Chapada Diamantina, de formato tabular, apresenta as áreas de maiores altitudes do território baiano, com 800 a 1200m acima do nível do mar. O planalto semiárido, localizado no sertão brasileiro, é caracterizado por baixas altitudes.



**Figura 1:** Mapa do Relevo do Estado da Bahia (Fonte: SEI, 2010).

O estado detém um imenso potencial ambiental representado principalmente pelos biomas do Cerrado, da Caatinga e da Mata Atlântica, além dos ecossistemas de restingas, lagunas e manguezais. O Cerrado destaca-se por sua rica biodiversidade e seu potencial aquífero, com destaque para a Bacia do Rio São Francisco. Nestes espaços concentram-se atualmente as mais importantes atividades agrícolas do Estado.

A Caatinga predomina na maior porção do território apresentando algumas “ilhas” de prosperidade, diversidade de paisagens e riqueza biológica ímpar. A Mata Atlântica representa cerca de 6% da área original, abrigando remanescentes da segunda maior floresta tropical úmida do Brasil, que originalmente estendia-se do Rio Grande do Norte ao Rio Grande do Sul, partindo do litoral e adentrando o território (Figura 2).



**Figura 2:** Mapa da Distribuição dos Ecossistemas da Bahia (Fonte: SEI, 2007).

### 2.5.1 Aspectos Socioeconômicos

As principais atividades econômicas estão distribuídas em espaços específicos do território baiano. Segundo a SEI (2008), na Região Metropolitana de Salvador, a leste, prevalecem os setores de serviços, assim como as indústrias petroquímica e automotiva. Na região Oeste estão os cerrados, responsáveis pela maior produção baiana de grãos, especialmente a soja. Ao Norte, no submédio São Francisco, está a produção de fruticultura irrigada e, nos municípios do entorno, se encontra o maior rebanho caprino do país.

Ao Sul do estado da Bahia concentram-se a maior produção pecuária bovina, a produção de café – no Planalto de Conquista – a cacauicultura, a silvicultura e a produção industrial de celulose, ao longo do litoral. Apresenta na região do Recôncavo, grande

potencial na produção de *citrus*, mandioca, e fumo (com caráter de exportação). Na região central do estado, de maneira mais atomizada, observam-se, principalmente, atividades de mineração e fruticultura diferenciada com cultivares de clima temperado devido às elevadas altitudes da região da Chapada Diamantina.

No aspecto disponibilidade de terras para uso agrícola, a Bahia detém uma vasta extensão territorial não ocupada para esta atividade (Figura 3). Segundo dados da Secretaria da Agricultura da Bahia (SEAGRI, 2012), o estado possui uma extensão territorial de 56 milhões de hectares, dos quais 32 milhões são áreas agricultáveis, as lavouras ocupam 4 milhões e as pastagens 15 milhões de hectares, restando ainda 13 milhões disponíveis para o uso agrícola.

Os relatórios da SEAGRI (2012) revelam um crescimento e uma atenção maior para expansão agrícola no estado nestes últimos anos. O resultado de todo esse empenho tem sido a consolidação do agronegócio, que se expande com sucesso em direção às áreas irrigadas do Vale do São Francisco, da região Oeste e do extremo Sul do estado. Os mesmos relatórios apontam novas perspectivas de expansão da fruticultura tropical devido às áreas disponíveis para irrigação e à possibilidade de obtenção de duas ou mais safras por ano.

Atualmente uma produção diferenciada se destaca no território baiano, além da produção de frutas tropicais. O clima propiciado pela altitude da Chapada Diamantina tem atraído produtores para o cultivo de frutas como ameixa, morango, caqui e uva sem semente, que são típicas de regiões temperadas, diversificando ainda mais a atividade no estado.

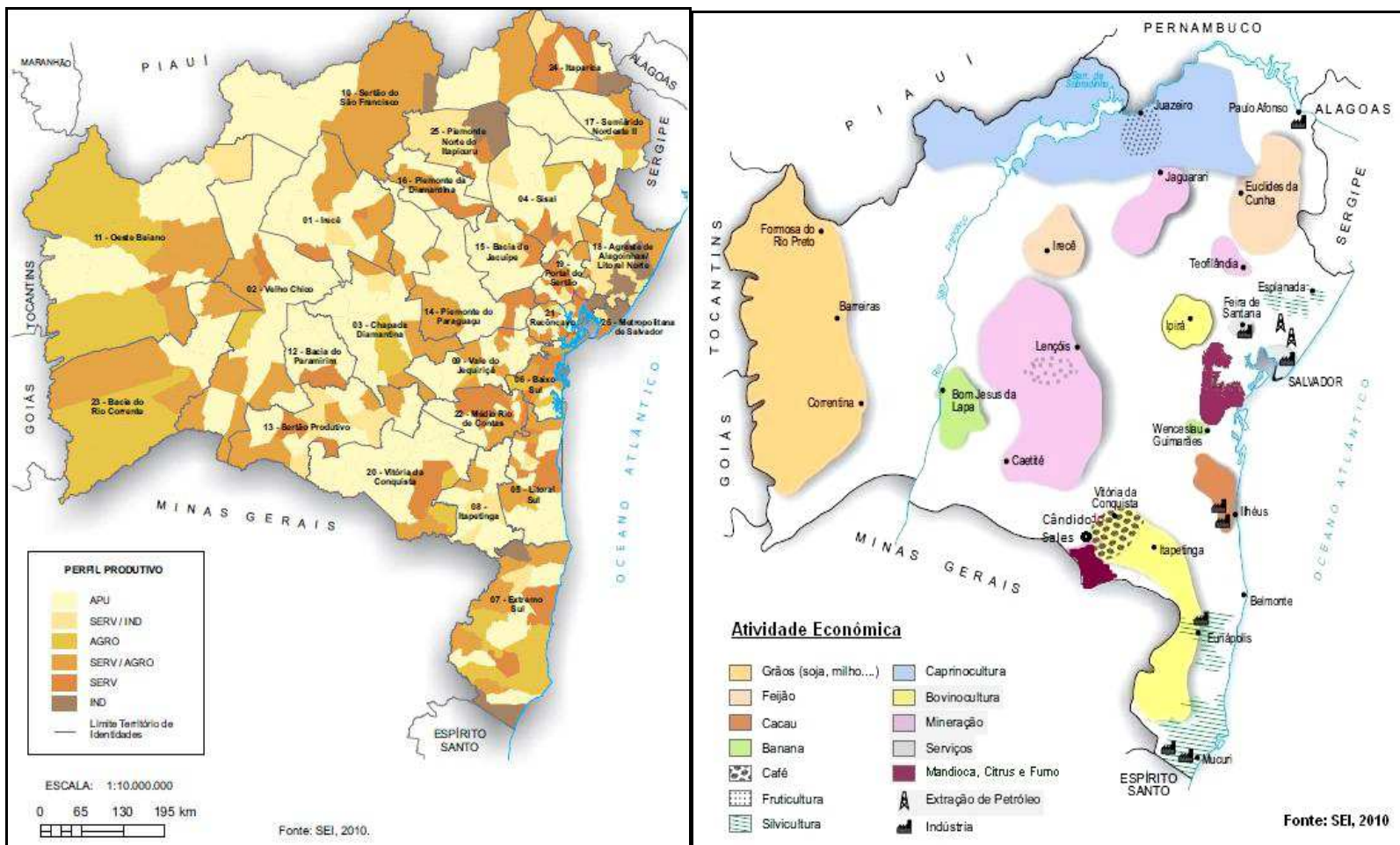
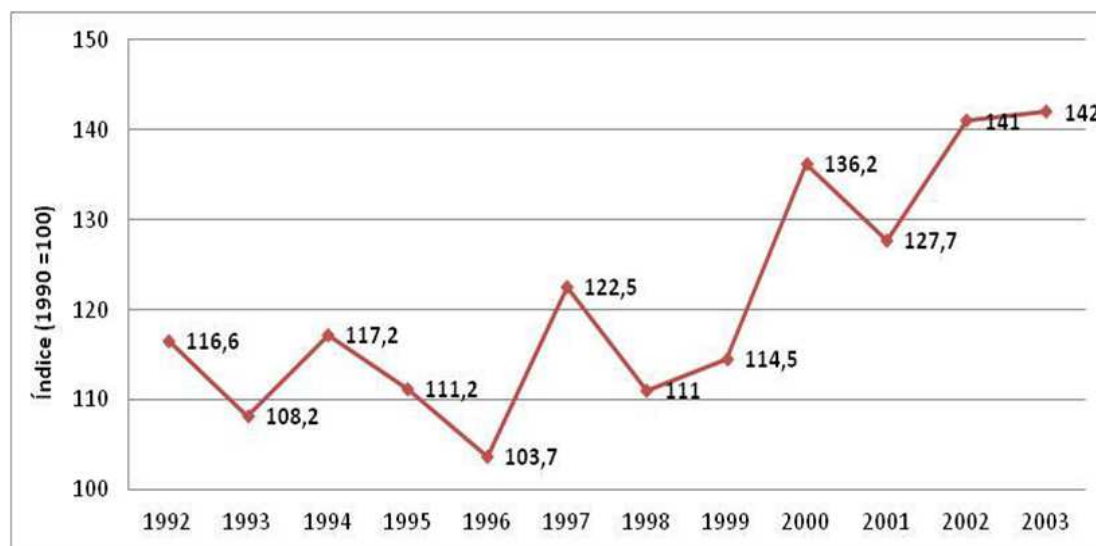


Figura 3: Distribuição dos setores e atividades econômicas do Estado da Bahia (Fonte: SEI, 2010).

Segundo Rocha (2011), o tipo climático da área potencializa o desenvolvimento de determinadas atividades econômicas, que por sua vez impulsiona a dinâmica socioeconômica da região. Por outro lado, a variabilidade da distribuição espaço-temporal das chuvas é uma das responsáveis pelo dinamismo decrescente das atividades agropecuárias. Sobre isto, registros da SEI (2010) revelam a estreita relação dos anos com queda na produtividade agrícola da Bahia e os anos que o estado sofreu com as adversidades climáticas, a seca.

Neste aspecto, de riscos e insucessos agrícolas, foi confeccionado pela Superintendência de Estudos e Informações da Bahia um quadro destacando os anos que apresentaram queda de produtividade agrícola e os principais produtos atingidos. Anuncia-se uma queda de -7,3% em 1993, -5,1% em 1995, -6,7% em 1996, -9,4 em 1998 e -6,2% em 2001 (Gráfico 1), destacando as perdas mais expressivas nas safras do feijão, mandioca, cana-de-açúcar e cacau, que estão entre os dez principais produtos no ranking de produção agrícola baiano.



**Gráfico 1:** Produtividade agrícola do Estado da Bahia referente ao período de 1990 a 2003. Fonte: SEI, 2010.

## 2.5.2 Aspectos Climáticos

O estado da Bahia, do ponto de vista climático, apresenta algumas características peculiares devido à sua posição geográfica de transição, onde as regiões limítrofes ao estado apresentam diferentes domínios climáticos: o Nordeste semiárido, o Sudeste úmido e o Centro-Oeste com períodos secos e úmidos bem definidos. Desta forma, sugere-se uma ampliação da diversidade de domínios climáticos e regimes pluviais sobre a região do estado da Bahia.

Dos elementos climáticos, a pluviosidade se destaca com estreitas correlações entre o regime fluvial e as formações vegetais. A posição geográfica concorre para a configuração das características naturais, pois a influência da massa líquida na costa contribui para a ocorrência de umidade elevada e chuvas constantes. Ao avançar para o interior (sentido Leste-Oeste) os efeitos da continentalidade e da topografia contribuem para diminuição desses fatores (SEI, 2008).

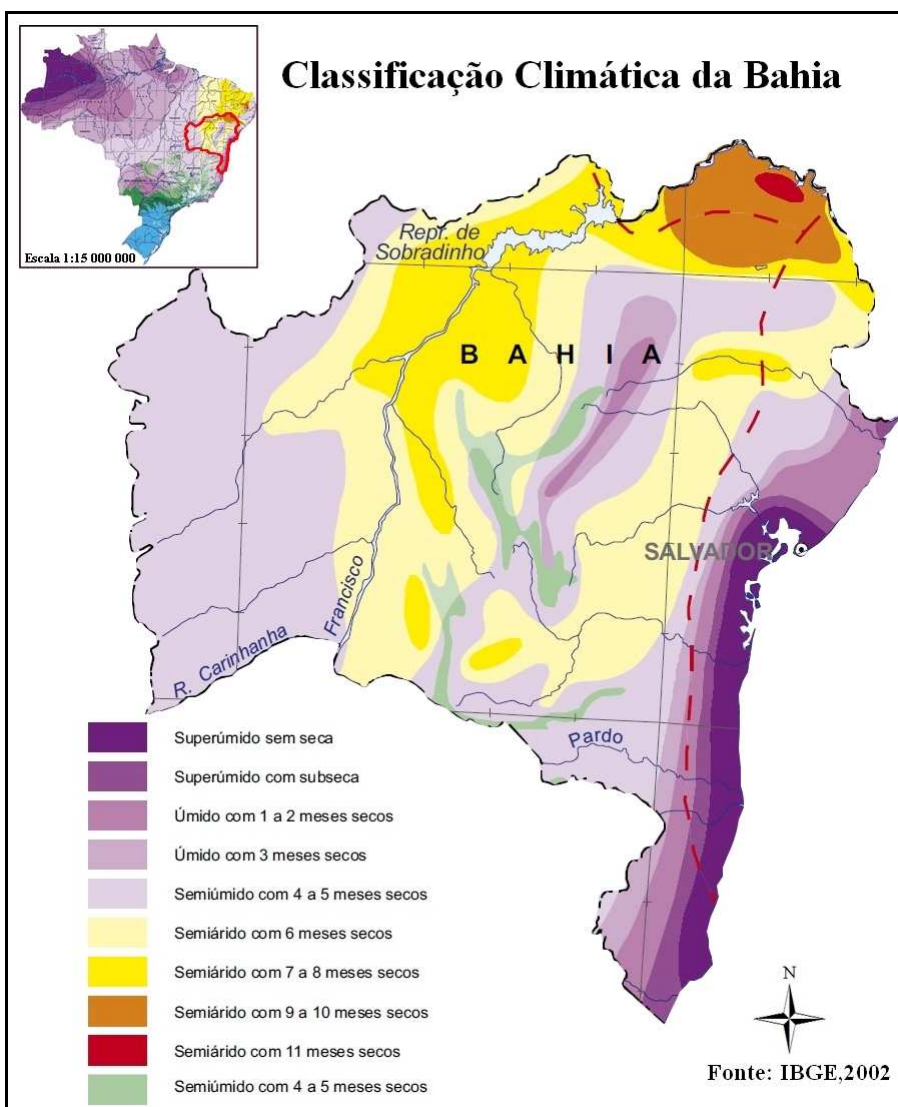
A Chapada Diamantina representa com maior expressividade o efeito orográfico da Bahia, afetando as circulações no sentido do litoral para o interior, que expressam o contraste dos parâmetros de umidade e vegetação de suas vertentes. Enquanto a vertente oriental retém mais umidade, a ocidental traz as características do semiárido (BRAGA, 1998).

O estado apresenta cinco tipos climáticos, distribuídos nos sentidos Norte-Sul e Leste-Oeste, conforme a classificação de Thornthwaite e Mather (1955), em: árido, semiárido; subúmido a seco, úmido, úmido a subúmido, conforme Tabela 2.

Segundo o Instituto Brasileiro de Geografia e Estatística (2012), o território baiano é constituído por 10 subtipos de climáticos regionais. Estes podem ser agrupados em dois grandes grupos, o tipo úmido e tipo semiárido, ambos apresentam médias de temperatura acima de 18° C, caracterizando-se como tipo clima quente. Destaca-se no centro e centro Sul do estado um clima tipo subquente com temperaturas médias entre 15° a 18° C, isso ocorre devido às elevadas altitudes dessa região (Figura 4).

**Tabela 2:** Características da tipologia climática da Bahia.

Clima	T.Méd.(°C)	Prec. Méd.(mm)	Def. Hídr.(mm)	Exc.Hídr.(mm)	Índice de Aridez
ÁRIDO	23,5 a 27,1	300 a 500	834 a 1401	0	67 % a 79%
SEMIÁRIDO	20,7 a 26,8	500 a 800	341 a 1060	0 a 90	30 a 66 %
SUBÚMIDO A SECO	18,8 a 25,4	800 a 1100	29 a 547	0 a 237	2 a 43 %
ÚMIDO A SUBÚMIDO	19,4 a 24,6	1100 a 2000	3 a 426	22 a 424	0 a 35
ÚMIDO	18,4 a 25,3	> 2000	0 a 345	200 a 1409	0 a 20



**Figura 4:** Mapa do Clima da Bahia. (Fonte: Adaptado do Atlas Nacional do Clima do Brasil do IBGE, 2010).



Como em todo o Nordeste, existe uma discrepância quanto à tendência geral de decréscimo do nível da pluviosidade na medida em que se adentra pelo continente (RADAMBRASIL, 1983). Os totais médios anuais da precipitação pluvial variam de 300 a 2000mm, podendo ainda ultrapassar valores de 2000mm em algumas regiões de clima úmido. Conforme a determinação das regiões características, do ponto de vista da precipitação mostrada por Araújo e Rodrigues (2000), a Bahia apresenta diferentes períodos de ocorrência das chuvas sobre as mesmas e/ou diferentes áreas do estado.

O território baiano apresenta vasta extensão de seu interior submetido aos fortes rigores do clima semiárido. De acordo com a Nova Delimitação da Região do Semiárido Brasileiro (BRASIL, 2005a; BRASIL, 2005b), instituída em Portaria nº 89, em 16 de março de 2005, foi acrescentado ao estado da Bahia mais oito municípios pertencentes ao Polígono das Secas, totalizando 265 municípios com abrangência de 23,4% da região semiárida e 63,9% da área do estado, ou seja, uma área de 564.273,0 km<sup>2</sup>.

Esta nova delimitação das áreas pertencentes ao semiárido, realizada pelo Ministério Nacional da Integração (Figura 5), utilizou como critérios técnicos de classificação três itens: 1º) Precipitação pluviométrica média anual inferior a 800 mm; 2º) Índice de aridez de até 0,5 calculado pelo balanço hídrico que relaciona as precipitações e a evapotranspiração potencial, no período entre 1961 e 1990; e 3º) Risco de seca maior que 60%, tomando-se por base o período entre 1970 e 1990.



**Figura 5:** Áreas Pertencentes ao Semiárido Brasileiro (Fonte: BRASIL, 2005).

Azevedo *et al.* (1998) afirmam que, do ponto de vista climático, a região Nordeste do Brasil é considerada semiárida por apresentar substanciais variações temporal e espacial da precipitação pluviométrica, e elevadas temperaturas ao longo do ano. Apesar das elevadas temperaturas registradas durante todo o ano, as amplitudes térmicas máximas são em torno de 6°C (FERREIRA, 2005).

### 2.5.3 Sistemas Meteorológicos Responsáveis pelo Regime Pluviométrico do Estado da Bahia

Segundo Molion *et al.* (2002), as precipitações pluviais estão diretamente relacionadas com a convecção local. Esta é caracterizada por movimentos ascendentes de ar úmido, resultantes da ocorrência de pressões atmosféricas mais baixas junto à superfície

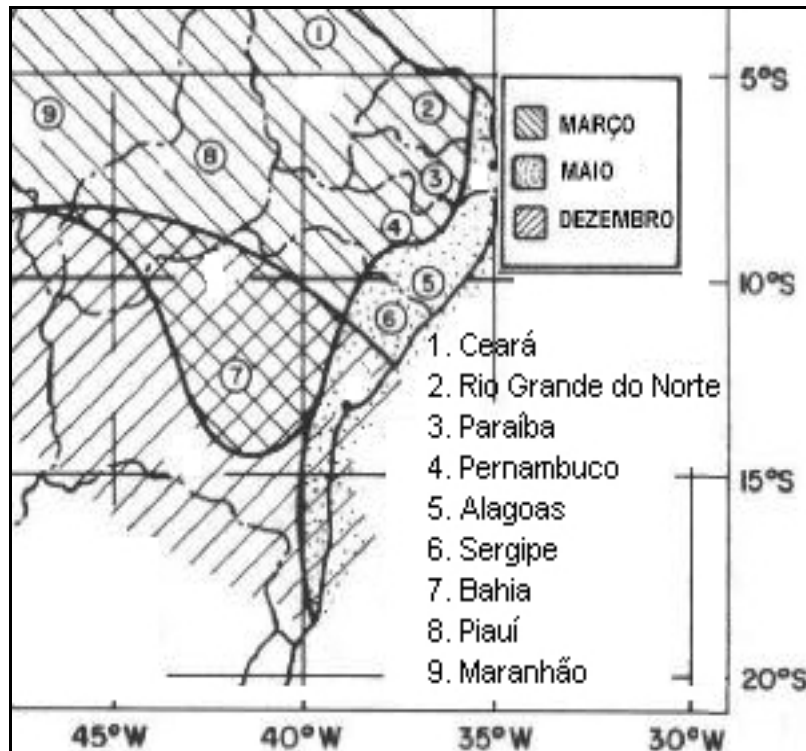
terrestre, seja em consequência do aquecimento do ar em contato com essa superfície, seja pela ação de fenômenos transientes, de caráter puramente dinâmico, como sistemas frontais, ou frentes frias, e perturbações ondulatórias no campo dos ventos.

A convecção tropical é essencialmente controlada - intensificada ou inibida - pela circulação geral da atmosfera, fenômenos de escala global, resultantes da interação complexa entre a superfície do planeta, particularmente a distribuição de continentes e oceanos com fornecimento desigual de energia solar, topografia e cobertura vegetal.

A variabilidade interanual das chuvas no estado da Bahia é influenciada principalmente pela Zona de Convergência Intertropical (ZCIT), por Sistemas Frontais (SFs) associados à Zona de Convergência do Atlântico Sul (ZCAS) e por efeitos locais como orografia e vegetação. A atuação dos Vórtices Ciclônicos de Ar Superior (VCAS), de acordo com a posição do centro desse sistema, também exercem grande influência no volume pluviométrico da região (DINIZ, 2008).

A análise da distribuição espacial e temporal das precipitações pluviais elaboradas para o estado da Bahia de acordo com os estudos do Centro de Planejamento da Bahia e do Centro de Estatística e Informação (CEPLAB, 1976; CEI, 1991), evidenciaram a distribuição decrescente dos totais médios anuais, das extremidades leste (Planície Litorânea) e Oeste (Chapadões Ocidentais) para o interior semiárido, confirmando a ação periférica desses sistemas dinâmicos que possibilitam a existência de três regimes pluviais diferentes e superpostos sobre o território baiano.

A máxima precipitação em março – abril, no Norte e no centro do Nordeste, deve-se à influência da ZCIT. Para as áreas do sul do Nordeste, a precipitação máxima ocorre de novembro a março, com um pico em dezembro, é ocasionada pela incursão de Sistemas Frontais e seus remanescentes entre 5°S e 18°S, que integram com a convecção local. E sob a influência da ZCAS essa área se expande para as latitudes entre 15°S e 20°S. Em algumas situações, um VCAS pode estar associado a um SF estacionário no leste do Nordeste. Na região litorânea, em consequência de influências de brisas que advectam nebulosidades provoca maior concentração de chuvas nessa área, com totais anuais superiores a 1500mm, conforme Figura 6 (CAVALCANTI *et al.*, 2009).



**Figura 6:** Área de distribuição dos principais regimes de chuva sobre o Nordeste. Fonte: adaptado de Kousky, 1979.

#### a) Zona de Convergência Intertropical

A Zona de Convergência Intertropical (ZCIT) pode ser definida como uma banda de nuvens que circunda a faixa equatorial do globo terrestre, formada principalmente pela confluência dos ventos alísios do hemisfério norte com os ventos alísios do hemisfério sul, em baixos níveis (o choque entre eles faz com que o ar quente e úmido ascenda e provoque a formação das nuvens), baixas pressões, altas temperaturas da superfície do mar, intensa atividade convectiva e precipitação (FERREIRA, 2005).

Barbosa (2000) e Cavalcanti *et al.* (2009) destacam a ZCIT como o fator mais importante na determinação de quão abundante ou deficiente serão as chuvas no norte e centro do Nordeste do Brasil. Normalmente a ZCIT migra sazonalmente de sua posição mais ao norte, aproximadamente de 10°N a 14°N em agosto-outubro, tornando este o período mais seco do norte do Nordeste. E ao atingir sua posição mais ao Sul, em torno de 2°S a 4°S nos meses de março a abril, no final do outono austral, justifica-se ser este o principal mecanismo dinâmico responsável pelas chuvas entre fevereiro a abril.

Esse deslocamento da ZCIT está relacionado aos padrões da Temperatura da Superfície do Mar (TSM) sobre essa bacia do oceano Atlântico Tropical. A ZCIT é mais significativa sobre os oceanos, e por isso, a Temperatura da Superfície do Mar (TSM) é um dos fatores determinantes na sua posição e intensidade.

Os estudos de Barbosa (2000) e relatórios da SEI (2010) descrevem que sob o domínio da Zona de Convergência Intertropical (ZCIT), encontra-se a parte mais setentrional da Bahia – Juazeiro, Remanso, Paulo Afonso, região do Raso da Catarina. Essa área está submetida ao regime de precipitações com estação chuvosa de verão-outono e máximos de chuvas em dezembro/janeiro a abril, quando a ZCIT, na porção ocidental do Atlântico Tropical, situa-se em sua posição mais ao sul, próxima à costa setentrional nordestina.

É a área que mais se integra ao domínio semiárido do sertão nordestino, vinculadas pelo mesmo sistema de circulação. É, também, aquela de maior escassez e irregularidade, com períodos chuvosos muito curtos, onde as chuvas só acontecem, quando existem condições atmosféricas e oceânicas para a ZCIT atingir a região.

Na Bahia, essa área registra as mais baixas médias pluviométricas do Estado. Os autores Kousky (1979) e Kousky e Moura (1981) chamam a atenção para a estreita ligação entre a atuação desse sistema de circulação e as secas nordestinas, afirmando que as grandes secas ocorrem quando a ZCIT permanece mais ao norte do equador e, portanto não alcançam o Nordeste. Neste caso, o anticiclone subtropical encontra-se fortalecido e os ventos alísios de sudeste (SE) dominam a região.

Apenas o extremo nordeste da área beneficia-se das precipitações litorâneas, podendo ocorrer um máximo secundário no outono-inverno. Nessa época (março-abril), verifica-se a simultaneidade de atuação dos dois sistemas básicos para a produção de chuvas no Nordeste e na Bahia, em particular – ZCIT e FPA.

## **b) Frente Fria**

Ferreira *et al.* (2005) retratam que um importante mecanismo causador de chuvas no Nordeste do Brasil está ligado à penetração de frentes frias até às latitudes tropicais entre os meses de novembro e janeiro.

As frentes frias são bandas de nuvens organizadas que se formam na região de confluência entre uma massa de ar frio (mais densa) com uma massa de ar quente (menos

densa). A massa de ar frio penetra por baixo da quente, como uma cunha, e faz com que o ar quente e úmido suba, forme as nuvens e, conseqüentemente, as chuvas.

Na Bahia, Aouad (1983) através de estudos desenvolvidos na Secretaria de Planejamento, Ciência e Tecnologia (SEPLANTEC-BA), atestou a participação da Frente Polar Atlântica (FPA) na produção de chuvas no estado, correlacionando os índices pluviométricos com a atuação das correntes atmosféricas.

Constatou-se a existência de estreita correlação entre o número de passagens das frentes frias e as chuvas na Bahia, durante todo o ano. Os meses secos estavam relacionados à frequência dos ventos alísios de Sudeste. Nas pesquisas de Molion (1985) há confirmação da importância dos sistemas frontais na produção de chuvas no Nordeste, com a ressalva sobre o fato das configurações da circulação em latitudes subtropicais estejam favoráveis. Nesses mesmos estudos.

O Centro Norte, Centro Sul e Sudoeste do semiárido baiano integram-se ao regime pluviométrico do extremo oeste do estado, onde as chuvas ocorrem mais intensivamente na primavera-verão (outubro/novembro a abril) com o mês mais chuvoso em dezembro. Tal regime pluvial, que atinge o semiárido baiano de forma periférica, liga-se principalmente aos mecanismos de ação da Frente Polar Atlântica – FPA- que neste período atuam no interior, filiados ao regime climático da região Sudeste do país (BARBOSA, 2000).

Na faixa costeira, o grande volume e a relativa regularidade pluvial que caracterizam essa área ligam-se à propagação das frentes frias, durante todo o ano. As chuvas, no litoral, são mais intensas no período de outono-inverno (março a julho), ocasião em que esses sistemas frontais podem estabelecer interação com a ZCIT, favorecendo a pluviosidade e explicando parte da ocorrência dos eventos chuvosos extremos que frequentemente atingem o litoral, nesta época do ano. Tais eventos são mais intensos quando ocorre o estacionamento da FPA (BARBOSA, 2000).

### **c) Vórtice Ciclônico de Altos Níveis**

Os Vórtices Ciclônicos de Altos Níveis (VCANs) que penetram na região Nordeste do Brasil formam-se no oceano Atlântico, principalmente entre os meses de novembro e março, e sua trajetória normalmente é de leste para oeste, com maior frequência entre os

meses de janeiro e fevereiro, conforme Gan e Kousky (1982). O tempo de vida desses sistemas varia em média, entre 7 a 10 dias.

Os VCANs são um conjunto de nuvens que, observado pelas imagens de satélite, têm a forma aproximada de um círculo girando no sentido horário. Na sua periferia há formação de nuvens causadoras de chuva e no centro há movimentos de ar de cima para baixo (subsidiência), aumentando a pressão e inibindo a formação de nuvens. Segundo Silva (2005), é nítida a associação da ausência de chuvas ao norte do estado da Bahia, devido à presença do centro do sistema diretamente sobre esta região.

#### **d) Zonas de Convergência do Atlântico Sul**

Durante a estação compreendida entre primavera-verão no hemisfério Sul, os sistemas frontais se posicionam preferencialmente sobre a parte central do continente sul-americano, tendo seu eixo no sentido NW-SE, de inclinação variável, criando uma zona de convergência de umidade que, posteriormente, foi denominada zona de convergência do Atlântico Sul (ZCAS). O deslocamento da ZCAS para 12°S - 15°S, e sua permanência com atividade intermitente, causa a estação chuvosa (novembro a março) do sul do Nordeste (MOLION *et al.*, 2002).

A Zona de Convergência do Atlântico Sul (ZCAS) é definida como uma região de alta variabilidade convectiva posicionada a leste da Cordilheira dos Andes com orientação noroeste-sudeste estendida desde o sul da Amazônia, regiões Centro-Oeste e Sudeste do Brasil, Centro Sul da Bahia e prolonga-se até o Atlântico Sudoeste (KODAMA, 1992; CARVALHO *et al.*, 2004).

A variabilidade espacial e temporal da ZCAS tem papel fundamental para a distribuição dos extremos de chuvas principalmente na região Sudeste (CARVALHO *et al.*, 2004). A frequência e a intensidade desses eventos são fatores importantes no total de precipitação em determinada estação do ano, fazendo com que ela seja considerada mais seca ou mais chuvosa que a normal climatológica.

#### **e) Fenômeno ENOS (El Niño de Oscilação Sul)**

É um fenômeno de grande escala que afeta a circulação atmosférica da região. Possui uma componente de natureza oceânica e outra atmosférica. A componente oceânica atua

provocando o aquecimento (El Niño) e o resfriamento (La Niña) anômalos das águas superficiais do Oceano Pacífico Equatorial Central e Oriental. Esses fenômenos naturais produzem fortes mudanças na dinâmica geral da atmosfera, alterando o comportamento climático (BERLATO *et al.*, 2003).

O impacto causado pelo fenômeno El Niño-Oscilação Sul (ENOS), um exemplo de perturbação climática de escala global, pode ser sentido principalmente pela modificação no regime e no total de precipitação que, dependendo da intensidade do evento, pode resultar em secas severas ou chuvas intensas, interferindo, de forma expressiva, nas atividades humanas.

De maneira geral os que eventos El Niño-Oscilação Sul (ENOS) impactam o tempo e o clima global, principalmente nos trópicos. Nobre e Molion (1988) sugeriram que, na fase quente dos eventos ENOS (El Niño), o ramo ascendente da Circulação Hadley-Walker, usualmente sobre a Amazônia, seja deslocado para sobre as águas anormalmente quentes do Pacífico Este ou Central, produzindo centros ciclônicos nos altos níveis sobre o norte/nordeste da América do Sul e uma forte subsidência sobre essa região e sobre o Atlântico tropical. Essa subsidência enfraqueceria a ZCIT e a convecção sobre o Nordeste, provocando a diminuição das chuvas.

Silva *et al.* (2012), com base em dados diários de precipitação, dados da reanálise do NCEP (*National Centers for Environmental Prediction*, sigla em inglês) e dados de Temperatura de Superfície do Mar (TSM), observou que as anomalias positiva e negativa de precipitação sobre o sul do Nordeste são associadas, respectivamente, às fases positiva e negativa do fenômeno ENSO.

Xavier *et al.* (2003) e Santos e Brito (2007) afirmam que a precipitação sobre o NEB é bastante sensível a extremos de TSM no Oceano Pacífico Equatorial associado ao evento El Niño Oscilação Sul (ENOS). Corrobora tal afirmação os resultados de Patrocínio (2008) revelando que os eventos mais expressivos de El Niño coincidem com as secas mais prolongadas no Nordeste brasileiro.

Nas pesquisas de Molion (1985) foram observadas que há anos que aparecem bloqueios no escoamento atmosférico, devido às anomalias atmosféricas-oceânicas de escala global – ENOS, fenômenos esses que impedem o avanço das frentes frias para o Nordeste, instalando-se a seca, na parte sul da região, que corresponde ao semiárido baiano e norte de Minas Gerais.



Observa-se que as precipitações na faixa litorânea também se reduzem, embora outros mecanismos ligados às brisas possam produzir chuvas nesta área. Algumas áreas do Sertão (semiárido) nordestino, essa diminuição pode alcançar até 80% do total médio do período chuvoso.

Recentes estudos a respeito da seca de dezembro de 2011 a março de 2012 na região oeste da Bahia, realizados por Santos *et al.* (2012), demonstraram relações entre a seca recentemente ocorrida no interior da Bahia e os padrões de teleconexão El Niño – Oscilação Sul (ENOS) e Oscilação Antártica (AAO). Verificaram também que sempre que o ENOS ou a AAO impõem anomalias positivas (negativas) de precipitação na Amazônia, o interior no Nordeste, em especial o interior da Bahia apresenta anomalias negativas (positivas).

Ao estudar a série pluviométrica de 1996 a 2005 do município de Feira de Santana na Bahia, Diniz *et al.* (2011) verificou que os anos identificados como secos foram os anos de ocorrência do El Niño com intensidade forte, o que levou a sérios prejuízos agrícolas principalmente para safras do milho, feijão e mandioca.

Os episódios de La Niña tendem a causar menos danos do que os de El Niño. Como consequência da La Niña, as frentes frias que atingem o centro-sul do Brasil têm sua passagem mais rápida que o normal e com mais força. Como as frentes têm mais força a passagem pela região sul e sudeste ocorre de forma mais rápida que o normal, consequentemente ocorre uma redução nos índices pluviométricos e a frente alcança o Nordeste do Brasil mais facilmente. Desta forma a região nordeste, principalmente o sertão e o litoral baiano e alagoano, são afetados por um aumento de chuvas o que pode ser positivo para a região semiárida, mas causa grandes prejuízos à agricultura.

A ocorrência de chuva acima da média pode ocorrer se também for formado um Dipolo Térmico do Atlântico favorável, ou seja, com temperatura da superfície do mar acima da média no Atlântico Tropical Sul e abaixo da média no Atlântico Tropical Norte. Em geral, a circulação atmosférica tende a apresentar características de anos normais na presença de La Niña, mas a distribuição de chuva, de fevereiro a maio, no semiárido do Nordeste pode se caracterizar por uma elevada irregularidade espacial e temporal mesmo em anos de La Niña.

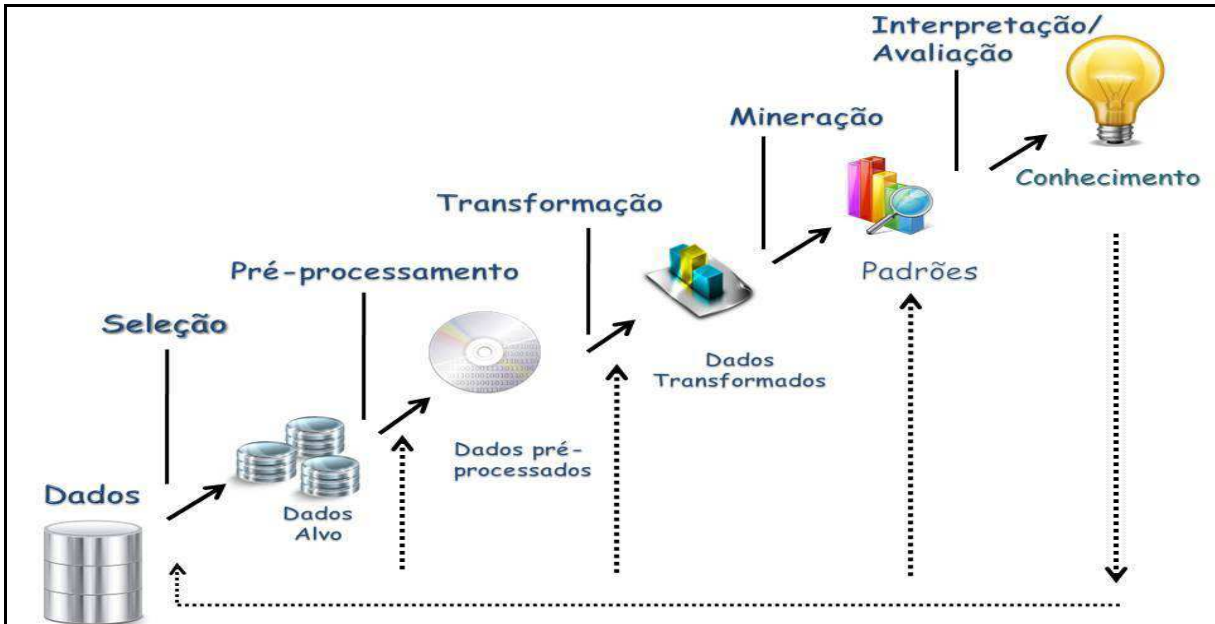
## 2.6 A Descoberta do Conhecimento em Bases de Dados – *Data Mining*

O constante crescimento do volume de dados nas mais diversas áreas do conhecimento tem impulsionado e exigido da comunidade acadêmica o desenvolvimento de técnicas que possibilitem a extração de informações desses dados. Neste contexto, a Descoberta do Conhecimento em Bases de Dados (*Knowledge Discovery in Databases - KDD*) se apresenta como uma tecnologia detentora de ferramentas eficientes para transformar informações contidas em grandes volumes de dados em conhecimento útil.

A Descoberta de Conhecimento em Bases de Dados é definida como um processo de identificação de padrões válidos, não conhecidos, potencialmente úteis e interpretáveis em dados armazenados (FAYYAD *et al.*, 1996). Uma das etapas do processo de KDD é a aplicação das técnicas de mineração de dados (*data mining*, em inglês) que em muitas vezes é usada como sinônimo do KDD (HAN *et al.*, 2011).

De acordo com Rezende *et al.* (2005), mineração de dados é a interação entre o especialista do domínio, o analista e o usuário final. Cabe ao especialista do domínio possuir amplo conhecimento sobre o assunto da aplicação e deve fornecer suporte para o desenvolvimento do processo. O analista é o especialista no processo de “extração do conhecimento” e responsável por sua execução. Ele deve conhecer profundamente as etapas que compõem o processo. O usuário final é representado por aqueles que utilizam o conhecimento extraído no processo, para auxiliá-lo na tomada de decisão.

A Figura 7 demonstra as etapas de Descoberta do Conhecimento em Bases de Dados – KDD (FAYYAD *et al.*, 1996).

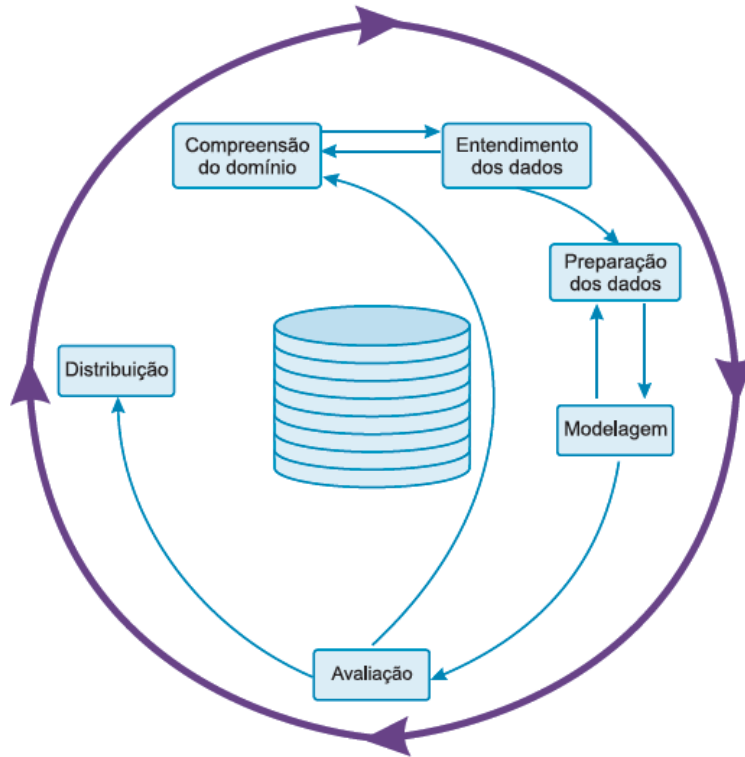


**Figura 7:** Processo do KDD, adaptado de FAYYAD *et al.* (1996).

O processo de KDD consiste primeiramente na compreensão do domínio, ou seja, é a etapa de identificação da meta da descoberta do conhecimento pelo ponto de vista do usuário; em seguida a seleção dos dados, etapa de aquisição dos dados a serem utilizados; partindo para o pré-processamento, onde é feita uma limpeza de ruídos, inconsistência e tratamento de dados faltantes; posteriormente, é aplicada a transformação dos dados, de suma importância para montagem da forma e disposição dos dados a serem aplicados os algoritmos de mineração de dados. E finalmente, a etapa de interpretação e avaliação, onde é extraído o conhecimento a partir dos resultados gerados pelos algoritmos.

## 2.7 Modelo do Processo de Descoberta de Conhecimento em Bases de Dados

Chapman *et al.* (2000) criaram em 1996 o modelo de processo CRISP-DM (*Cross-Industry Standard Process for Data Mining*) com o intuito de padronizar o ciclo de vida de um projeto de mineração de dados. Este ciclo é dividido em seis fases, a saber: compreensão do domínio, entendimento dos dados, preparação dos dados, modelagem, avaliação e distribuição, conforme Figura 8.



**Figura 8:** Fases da metodologia CRISP-DM (CHAPMAN *et al.*, 2000).

Uma descrição sucinta de cada fase do processo CRISP-DM é apresentada a seguir:

- **Compreensão do domínio:** visa obter conhecimento sobre seus objetivos e requisitos para defini-lo e convertê-lo em um problema de mineração de dados, traçando-se, então, um plano preliminar designado para alcançar seus objetivos.
- **Entendimento dos dados:** tem início com a coleção e seleção de dados inicial seguindo com atividades de exploração de dados, para se familiarizar, identificar problemas de qualidade, formalizar as primeiras hipóteses e identificar possíveis subconjuntos que possam abrigar informações ocultas sobre esses dados.
- **Preparação dos dados:** esta fase implica em todas as atividades necessárias para a construção do conjunto de dados final, sobre o qual serão aplicadas as técnicas de modelagem. As tarefas incluem limpeza de dados, seleção e transformação dos dados. É nessa etapa que dados inconsistentes são eliminados.
- **Modelagem:** nessa fase são selecionadas e aplicadas diversas técnicas de mineração de dados, e seus parâmetros são calibrados. Por possibilitar o uso de várias técnicas para o

mesmo problema, em muitos casos existe a necessidade de voltar para a fase de preparação de dados, para obtenção de formatos específicos.

- **Avaliação:** nesta etapa, tem-se o modelo (ou modelos) com boa qualidade numa perspectiva de análise de dados. Os resultados são comparados e interpretados conforme a área de aplicação. Ressalta-se a importância de reavaliar as etapas anteriores do processo para se ter a certeza de que o modelo atende às necessidades e aos objetivos do projeto.
- **Distribuição:** na última fase do processo, tem-se um modelo construído. Porém, isso não implica na finalização do projeto. O conhecimento obtido deve ser documentado, organizado e apresentado para os usuários, para que estes possam saber quais ações devem ser realizadas para aproveitar os modelos criados.

## 2.8 Tarefas e Técnicas de Mineração de Dados

Amo (2004) relata a importância da distinção entre os conceitos de técnica e tarefa da mineração de dados. A mesma autora define uma tarefa de mineração de dados como a especificação do que se pretende buscar nos dados, que tipo de regularidades ou categoria de padrões tem-se interesse em encontrar, ou que tipo de padrões revelaria algo novo. E a técnica de mineração consiste na especificação de métodos que garantam como descobrir os padrões de maior interesse.

Na prática, os dois objetivos principais da mineração de dados são a predição e a descrição. Dentre essas tarefas, dependendo do objetivo procurado, dividem-se as principais categorias, a saber: classificação, regressão, associação e clusterização.

Esses objetivos podem ser alcançados por meio de vários tipos de tarefas. A escolha de uma ou mais tarefas depende do problema em questão. As tarefas tradicionais de mineração de dados estão representadas na Figura 9 e são brevemente descritas a seguir (HAN *et al.*, 2011; FAYYAD *et al.*, 1996).



**Figura 9:** Tarefas de mineração de dados de acordo com REZENDE *et al.* (2005).

### 2.8.1 Predição

Segundo Rezende *et al.* (2011), a predição consiste em examinar atributos de um conjunto de entidades e, baseado nos valores destes, assinalar valores e atributos de uma nova entidade que se quer caracterizar. A predição usa atributos para predizer o desconhecido ou os valores futuros de outra variável (também conhecida como atributo meta). Os dois principais tipos de tarefas para predição são classificação e regressão. A mesma autora define as tarefas de Regressão e Classificação em:

**Regressão:** consiste em encontrar uma função que mapeie um item de dados para uma variável de predição de valor numérico contínuo.

**Classificação:** busca modelos (funções) que descrevam e distingam classes ou conceitos, com o intuito de utilizá-los para predizer classes de objetos que ainda não foram classificados.

### 2.8.2 Descrição

A descrição tem por objetivo descrever o que está acontecendo em uma base de dados, no intuito de aumentar o entendimento sobre as pessoas, produtos ou processos que produziram os dados (BERRY *et al.*, 1997). Atividades de descrição consistem na identificação de comportamentos intrínsecos do conjunto de dados, sendo que estes dados não possuem uma classe especificada. Algumas das tarefas de descrição são clusterização e regras de associação.

**Associação:** é a descoberta de regras de associação indicando condições de atributo valor que ocorrem frequentemente em um conjunto de dados. A regra é apresentada com forma padrão  $X \rightarrow Y$ , onde  $X$  e  $Y$  são conjuntos de itens frequentes e  $X \cap Y = \emptyset$  (AGRAWAL, 1993).

**Agrupamento (*Clustering*):** consiste em agrupar os dados em classes ou *clusters*, onde os objetos de uma classe tenham alta similaridade entre si e sejam ao máximo diferente entre os objetos da outra classe (HAN *et al.*, 2011).

**Sumarização:** correspondem às formas para encontrar uma descrição compacta para um subconjunto de dados.

## 2.9 Descrição das Técnicas de Mineração de Dados Utilizadas

### 2.9.1 Clusterização

Análise de *clusters* (agrupamentos) – diferentemente da classificação, em que os dados estão previamente rotulados, a análise de *clusters* trabalha sobre dados em que as classes não estão definidas. A técnica consiste em identificar novos agrupamentos, que contenham características similares e agrupar os registros, ou seja, particionar (segmentar) uma dada população de objetos ou itens em conjuntos (KUMAR *et al.*, 2009).

No processo de clusterização existem vários métodos descritos na literatura. Sua escolha dependerá do tipo de dado a ser analisado, assim como do propósito e da aplicação da análise. Dentre eles, há um destaque para três métodos: particionamento, hierárquico e os fundamentados em densidade.

#### 2.9.1.1 Principais Métodos de Clusterização

##### a) Método de Particionamento

Os métodos de particionamento trabalham com um número fixo de grupos. Dados  $n$  objetos para serem agrupados em  $k$  grupos, os métodos de particionamento tentam encontrar as  $k$  melhores partições para os  $n$  objetos. Segundo Ng e Han (2002), é muito comum encontrar casos onde os  $k$  grupos encontrados pelo método de particionamento são de melhor qualidade (i.e., mais similares) que os grupos encontrados nos métodos hierárquicos. Por isso,

os métodos de particionamento têm recebido maior atenção por especialistas da área de análise de grupos. Além disso, muitos métodos de particionamento baseados no *k-means* e no *k-medoid* têm sido desenvolvidos.

No que se refere à metodologia de particionamento, de acordo com Rezende *et al.* (2005), os objetos são alocados nos clusters de acordo com a medida de similaridade adotada, onde se busca a maximização da distância entre clusters e a minimização da distância entre os objetos do mesmo grupo. Na execução deste método, utiliza-se o algoritmo *k-means* como técnica de agrupamento por meio da métrica da distância euclidiana.

Este método é sensível a ruídos e *outliers*, pois em sua execução cada elemento pertencente ao conjunto de dados é inserido em um *cluster*. Desta forma, é possível que um número pequeno de tais dados (*outliers*) influencie o valor do centróide do *cluster* ao qual ele pertence.

#### **b) Método Hierárquico**

Para Han *et al.* (2011), os métodos hierárquicos são classificados em aglomerativos ou divisivos, e estes são baseados no modo de conformação da decomposição hierárquica. Difere-se do método de particionamento, no sentido dos algoritmos não necessitarem de uma definição prévia do número de cluster. Esses dois métodos, aglomerativos e divisivos, são chamados hierárquicos, pois criam uma relação de hierarquia entre os grupos formados. A partir da visualização da hierarquia desses grupos, o usuário pode decidir com quais grupos deseja trabalhar.

Dados  $n$  objetos para serem agrupados, nos métodos aglomerativos começamos o algoritmo com  $n$  grupos, cada qual formado por um objeto. A cada passo do algoritmo, dois grupos semelhantes são aglomerados transformando-se em um novo grupo. Este processo é repetido até que exista apenas um único grupo contendo todos os  $n$  objetos.

Nos métodos divisivos, dados  $n$  objetos, o algoritmo começa com um único grupo contendo todos os  $n$  objetos. A cada passo do algoritmo, os grupos formados são divididos de acordo com a similaridade dos objetos contidos neles. Este processo é repetido até que  $n$  grupos com apenas um objeto sejam formados.



### c) Método Baseado em Densidade

É um método baseado na noção de densidade, uma vez que considera como cluster regiões densas de objetos no espaço, separadas por regiões vizinhas de baixa densidade (ESTER *et al.*, 1996). Um cluster é definido com um conjunto de pontos densamente conectados, enquanto ruídos representam conjuntos de objetos que não estão contidos em nenhum cluster.

Ao mencionar os métodos baseados por densidade, tem-se o algoritmo DBSCAN, que de acordo com Han *et al.* (2011), é um algoritmo típico da disposição e crescimento dos clusters com o limiar de densidade. Este algoritmo é sensível a ruídos, *outliers*, podendo identificá-los e separá-los dos grupos formados.

#### 2.9.2 Padrões Sequenciais

As análises de sequências que dependem do tempo podem ser de dois tipos: séries temporais, onde os objetos são valores num domínio contínuo; ou sequências temporais, em que os objetos obedecem a um determinado alfabeto. Análises de dados em séries temporais geralmente buscam por modelos globais para o estudo de tendências, como por exemplo, o estudo das mudanças climáticas e do efeito estufa. No entanto, as operações de mineração em sequências temporais geralmente buscam descobrir padrões sequenciais que se repetem numa mesma sequência ou em várias sequências (DAS *et al.*, 1998).

De acordo com Agrawal e Srikant (1995), a tarefa da mineração de padrões sequenciais é encontrar subsequências frequentes em um banco de dados de sequências. Esta é uma atividade desafiadora, já que a busca pode exigir o exame de um número combinatoriamente explosivo de padrões de subsequências possíveis. Essa técnica possui diversas aplicações, incluindo a análise do comportamento de clientes, padrões de acesso a Web, análise do processo de experimentos científicos, comportamentos de eventos extremos climáticos, tratamento de doenças, teste de medicamentos, análise de DNA, etc.

Formalmente, um padrão sequencial é uma expressão da forma  $\langle I_1, \dots, I_n \rangle$ , em que cada  $I_i$  é um conjunto de itens. A ordem em que estão alinhados os conjuntos reflete a cronologia com que aconteceram os fatos representados por eles. Encontrar padrões previsíveis em um período de tempo significa que um comportamento particular em um dado momento pode ter como consequência outro comportamento ou sequência de

comportamentos dentro de um mesmo período de tempo. Por exemplo, um extremo registrado em um município brasileiro, em um determinado período, poderá se repetir em outro período de tempo.

Os principais algoritmos utilizados na busca de padrões sequenciais são: AprioriAll, GSP, PSP, SPADE, FreeSpan, Pre\_xSpan, DepthProject, entre outros. O algoritmo *Generalized Sequential Patterns* (GSP), segundo Furtado (2005), é um algoritmo iterativo que encontra, a cada iteração, o conjunto dos padrões sequenciais frequentes de comprimento  $k$ , denotado por  $L_k$ . A respeito das interações realizadas pelo GSP, Srikant e Agrawal (1996) descrevem que estas ocorrem em duas fases: a fase da geração de candidatos e a fase do cálculo do suporte (frequência de ocorrência).

A primeira fase é a geração de candidatos que, por sua vez, é subdividida na etapa da junção e na etapa da poda. Nesta primeira iteração, o algoritmo GSP encontra o conjunto de todos os itens frequentes no banco de dados de tamanho 1 (1-sequências). Tal conjunto é utilizado na segunda iteração, para o cálculo das sequências de tamanho 2 (2-sequências) candidatas. Uma 2-sequência candidata pode conter um único *itemset* (conjunto frequente) com 2 itens frequentes, ou dois *itemsets*, cada um contendo um item frequente. Por exemplo, se os itens 1 e 2 são frequentes, na segunda iteração serão geradas as 2-sequências candidatas  $\langle(1),(2)\rangle$ ,  $\langle(2),(1)\rangle$  e  $\langle(1,2)\rangle$ .

Na segunda fase, o banco de dados é percorrido e o suporte das 2-sequências candidatas é calculado. O GSP elimina as sequências candidatas com suporte inferior ao suporte mínimo. A partir da terceira iteração, o conjunto das  $k$ -sequências candidatas é gerado através da combinação de  $(k-1)$ -sequências frequentes, em seguida as  $k$ -sequências candidatas que não podem ser frequentes são podadas e então o banco de dados é percorrido para cálculo do suporte das sequências candidatas restantes.

## **2.10 Aplicações de Técnicas de Mineração de Dados na Agricultura e Climatologia**

O setor agrícola apresenta intenso desenvolvimento de tecnologias para captura e armazenamento de dados agronômicos. Diversos segmentos como agrometeorologia, agricultura de precisão e mudanças climáticas têm gerado conjunto de dados cada vez maiores e mais detalhados. Surge a necessidade de transformar esses dados em conhecimento e

informação útil para auxílio em tomadas de decisão estratégicas. Devido a esse aumento no volume de informações, a partir do ano 2000 encontra-se uma maior quantidade de trabalhos científicos na tentativa de esclarecer as indagações dos cientistas e validar suas hipóteses.

Boschi *et al.* (2011), com o intuito de identificar zonas pluviometricamente homogêneas no estado Rio Grande do Sul para análises das ocorrências dos fenômenos de veranicos nesse estado, utilizaram as técnicas de *agrupamento* (clusterização) sob um banco de dados meteorológicos. Rios (2011) também utilizou métodos de clusterização para definir zonas de manejo por produtividade no âmbito da agricultura de precisão, possibilitando um maior entendimento das características particulares das frações do solo.

Fechine *et al.* (2008) utilizaram o método hierárquico aglomerativo para o agrupamento de zonas homogeneamente pluviométricas no âmbito da bacia hidrográfica do rio Brígida, semiárido do estado de Pernambuco. Em contribuição para agricultura de precisão, Silva *et al.* (2012) utilizaram algoritmos de agrupamento que consideram atributos georreferenciados para separar grupos pela sua distância geográfica, e atributos não georreferenciados para agrupar os dados de uma lavoura pelo tipo da cultura.

Com a aplicação da tarefa de *classificação*, através da indução por árvore de decisão, Meira *et al.* (2009) desenvolveu modelos de alerta da ferrugem-do-cafeeiro para lavouras com alta carga pendente de frutos, a partir de dados meteorológicos e do espaçamento entre plantas. Paul e Munkvold (2004) também utilizaram este tipo de modelagem para predizer categorias de severidade da cercosporiose do milho em estágio avançado do cultivo. E Molineros *et al.* (2005), por meio das árvores de decisão, modelaram epidemias de giberela do trigo, procurando predizer se a severidade da doença seria maior ou igual a 10%.

Souza *et al.* (2010) discutiram a aplicação da árvore de decisão na análise dos atributos do solo e da produtividade da cultura da cana-de-açúcar, em experimento montado no estado de São Paulo. Neste estudo, a árvore de decisão foi induzida no programa SAS Enterprise Miner, sendo utilizado algoritmo baseado na redução de entropia permitindo descobrir ou inferir os valores determinantes ou de correlação com a produtividade de cana-de-açúcar.

Liu *et al.* (2001) aplicaram as tarefas de classificação visando a predição de chuvas. Para isto, utilizaram o classificador Naïve-Bayes e outros algoritmos para seleção de atributos em dados meteorológicos reais da cidade de Hong-kong. Os autores encontraram ótimos

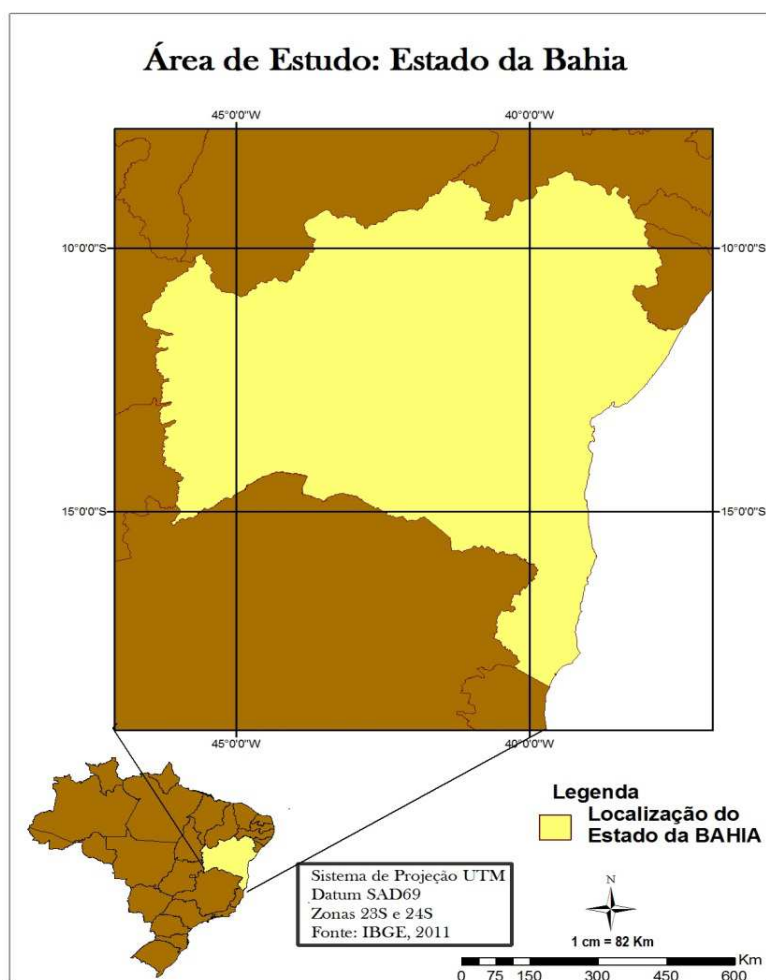
resultados com o uso desse classificador, alcançando uma acurácia de 90% após a seleção de atributos.

Com o uso das tarefas de *associação*, Dhanya e Kumar (2009) montaram um conjunto de regras que mostra fortes relações entre índices climáticos com pressão do nível do mar, Oscilação do Atlântico Norte, El Nino, temperatura da superfície do mar e os eventos extremos de chuva. Outra abordagem da tarefa de associação foi adotada por Vendrusculo *et al.* (2009) que utilizaram algoritmos de padrões sequenciais para identificar os padrões de ocorrência dos períodos de cheia e estiagem do Rio Paraguai, utilizaram o algoritmo *Generalize Sequential Patterns* (GSP), no banco de dados referentes à altura do rio, bem como a técnica SAX para redução da dimensionalidade e representação simbólica dos dados.

### 3. MATERIAL E MÉTODOS

#### 3.1 Área de Estudo

A área de estudo corresponde ao estado da Bahia, com limítrofe ao Norte o rio São Francisco, no município de Curaçá, divisa com Pernambuco, entre as coordenadas  $-8^{\circ} 32' 00''$  de latitude e  $-39^{\circ} 22' 49''$  de longitude. Ao sul, o limite extremo é a Barra do Riacho Doce, no município de Mucuri, divisa com o Espírito Santo, lat.  $-18^{\circ} 20' 07''$  e long.  $-39^{\circ} 39' 48''$ . A Leste, o ponto extremo é a Barra do Rio Real, no município de Jandaíra, divisa com o Oceano Atlântico, lat.  $-11^{\circ} 27' 07''$  e long.  $-37^{\circ} 20' 37''$ . O ponto extremo do Oeste é o divisor de águas, no município de Formosa do Rio Preto, divisa com o Tocantins, lat.  $-11^{\circ} 17' 21''$  e long.  $-46^{\circ} 36' 54''$  (SEI, 2008), conforme Figura 10.



**Figura 10:** Mapa de Localização da Área de Estudo: Estado da Bahia (Fonte: IBGE, 2011).

## **3.2 Modelo do Processo**

A metodologia aplicada para execução e consolidação do processo de descoberta do conhecimento em base de dados, teve como suporte o modelo CRIS-PDM (*Cross Industry Standard Process for Data Mining*) descrita por Chapman *et al.* (2000). A escolha do CRISP-DM se deu porque essa metodologia é amplamente adotada em projetos de mineração de dados, tanto na academia quanto na indústria, e por prover uma ferramenta de suporte rápida e barata para definir os procedimentos adotados na fase de mineração de dados.

Neste trabalho, a metodologia foi dividida em cinco etapas fundamentais, a saber: Etapa 1 – compreensão do domínio; Etapa 2 – entendimento dos dados; Etapa 3 – preparação dos dados; Etapa 4 – modelagem; Etapa 5 – avaliação dos resultados obtidos. A última etapa encontra-se no capítulo 4 (Resultados e Discussão). A seguir, apresenta-se o detalhamento das etapas metodológicas.

### **3.2.1 Compreensão do domínio**

Fase inicial do projeto que permitiu a compreensão do assunto para as questões que envolveram a definição do problema e o plano para alcançar os objetivos. Nesta fase, foram desenvolvidas atividades de revisão de literatura sobre o tema proposto, consultas a especialistas das áreas de climatologia, meteorologia, agrometeorologias, economia agrícola, estatística e mineração de dados.

### **3.2.2 Entendimento dos dados**

Por meio de consultas à base de dados de séries históricas pluviométricas da Agência Nacional das Águas, disponível no Sistema de Informações Hidrológicas ([www.hidroweb.ana.gov.br](http://www.hidroweb.ana.gov.br)) (ANA, 2011), foram selecionados e extraídos dados históricos de totais diários de precipitação referentes ao estado da Bahia. Além dos dados pluviométricos para confecção do banco de dados, foram adicionados: o nome do município onde a estação está localizada, posição geográfica (latitude e longitude).

Considerando a não homogeneidade na distribuição das chuvas no estado da Bahia, optou-se em trabalhar com séries históricas homogêneas e contínuas. Neste sentido, a seleção

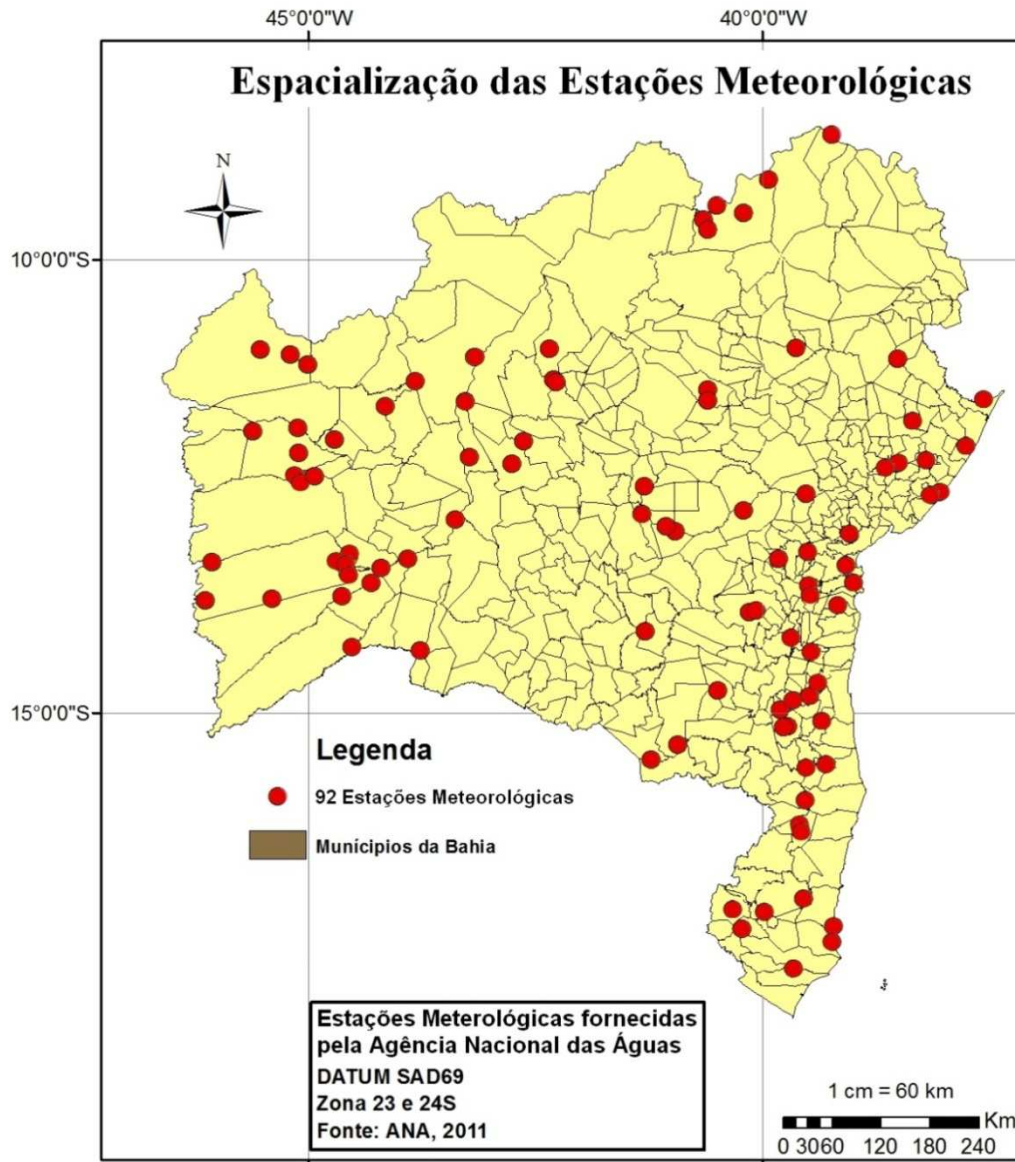
das estações meteorológicas utilizadas neste trabalho fundamenta-se nos quesitos: disponibilidade e qualidade dos dados.

O quesito qualidade dos dados se refere aos dados consistidos da ANA, ou seja, esta Agência adota a metodologia proposta pela Agência Nacional de Energia Elétrica (ANEEL), fundamentada no modelo matemático de Holanda e Oliveira (1979). Estes dados passaram por duas críticas para formulação de sua consistência. A primeira delas é a comparação dos dados em escala mensal de uma estação com mais outras duas estações vizinhas, com período de cobertura e coeficiente de correlação semelhante. Os dados obtidos nesse procedimento passam, em seguida, por uma análise de intervalo de aceitação das estações de apoio.

A segunda crítica segue basicamente o mesmo procedimento, utilizando-se a nova série gerada com os valores propostos pela primeira etapa. Com a obtenção da consistência em escala mensal, é feita a compatibilização com os dados diários, onde são eliminados aqueles correspondentes a períodos rejeitados. Desta forma, estes dados apresentam um maior grau de confiabilidade, que é fundamental para a realização de estudos científicos, nas mais diversas áreas, em especial nas áreas de socioeconomia. Dentre estes estudos, destacam-se aqueles relacionados às atividades agrícolas.

Outro critério adotado foi a delimitação de uma série histórica de dados meteorológicos, que deveria conter pelo menos 30 anos (1981 – 2010), respaldando desta forma as inferências e destaques com relação ao clima do estado da Bahia, segundo padronização da Organização Meteorológica Mundial (WMO, sigla em inglês). Sendo assim, foram selecionadas 92 estações meteorológicas que cobrem praticamente todo o estado, conforme Figura 11 e Tabela 17, disponível no Apêndice II.

A implantação dos dados espaciais foi feita por meio do programa Arc-Map versão 9.3 (ESRI, 2009), um sistema de informações geográficas (SIG) vetorial, que permite georreferenciamento e cruzamento de diferentes planos de informação. O sistema de coordenadas adotado foi o de Latitude e Longitude, Datum SAD69, e sistema de projeção *Universal Transversa de Mercator*. As informações das geoespaciais foram fornecidas pela ANA (2011) e o *shape* referente aos limites municipais do estado da Bahia pelo IBGE (2012).



**Figura 11:** Distribuição espacial das 92 estações meteorológicas do estado da Bahia.

### 3.2.2.1 Pluviometria do Estado da Bahia

A confecção do mapa pluviométrico da Bahia referente à série histórica de 30 anos (1981 a 2010) foi realizada a partir dos dados das médias anuais de precipitação pluviométrica. Foi construída uma planilha composta de município, código da estação meteorológica, latitude, longitude, altitude e precipitação média anual, de acordo com a Tabela 17 (Apêndice II). Estas informações estavam em formato Excel e posteriormente



foram importadas para o Arc MAP, onde foram transformadas para *shapefile*, para aplicação do método geostático IDW (*Inverse Distance Weighted* – Ponderação do Inverso da Distância).

De acordo com Tieppo *et al.* (2010) e Marcuzzo *et al.* (2011), uma das técnicas de interpolação mais usadas para pontos espalhados espacialmente é o IDW. A interpolação por IDW determina os valores dos pontos usando uma combinação linear ponderada dos pontos amostrados. O peso de cada ponto é o inverso de uma função da distância. Para elaboração do cálculo da interpolação do valor de um ponto através do método do IDW, Marcuzzo *et al.* (2011) utilizaram a seguinte equação matemática:

$$Z(x) = \frac{\sum_{i=1}^n \omega_i Z(x_i)}{\sum_{i=1}^n \omega_i} \quad (1)$$

Em que,  $Z(x)$  é o valor do ponto que se deseja interpolar;  $n$  é a quantidade de pontos próximos utilizados na interpolação do ponto  $x$ ;  $Z(x_i)$  é o valor do ponto  $x_i$ ; e  $\omega_i$  é o peso do valor de  $x_i$  sobre o ponto  $x$ . Para se determinar  $\omega_i$ , utiliza-se a seguinte equação matemática:

$$\omega_i = \frac{1}{h(x, x_i)^p} \quad (2)$$

Sendo,  $h(x, x_i)$  a distância entre o ponto  $x$  e o ponto  $x_i$ ; e  $p$  o parâmetro de potência, geralmente igual a dois. Parâmetros de potência maiores enfatizam pontos mais próximos, tornando o resultado menos suave. Parâmetros de potência menores enfatizam pontos mais distantes, tornando o resultado mais suave, porém menos preciso.

### 3.2.3 Preparação dos dados

A fase de preparação dos dados teve início com a construção de uma planilha em formato Excel, convertendo-se os dados pluviométricos dos arquivos em formato Microsoft Access para um único arquivo. Desta forma, foi obtida uma planilha de dados com 92 linhas (registros), condizentes às estações meteorológicas selecionadas e 10.957 colunas (atributos), relacionadas aos valores diários de precipitação para um período de 30 anos.

A partir da confecção desta planilha, com dados de precipitação em escala diária, foi construído outro banco de dados com valores de precipitação acumulados em escala mensal.

A disposição dos dias nas colunas e das estações nas linhas foi organizada com o intuito de subsidiar a geração de agrupamentos de estações meteorológicas conforme a similaridade de comportamento das chuvas. Para obtenção dos clusters foram utilizados somente os dados de precipitação pluviométrica, desprezando-se a princípio as informações de posicionamento geográfico das estações do território baiano.

É importante ressaltar que a transformação da escala temporal (diária para mensal) foi necessária devido à particularidade de alta variabilidade da variável estudada, a chuva, visto que a técnica de agrupamento busca padrões de comportamento similares entre os objetos para formação dos *clusters*.

Uma etapa essencial na aplicação das técnicas de clusterização quando se utiliza atributos com diferentes ordens de grandeza é a normalização dos dados. No entanto, para o presente trabalho esta etapa não foi necessária, uma vez que o único atributo considerado foi a precipitação pluviométrica mensal.

Para criar um único arquivo com as informações meteorológicas contidas em 92 arquivos (postos pluviométricos), seguindo a disposição vetorial supracitada, foi desenvolvido um *script* em linguagem de programação *Python*. Em seguida, a fase de transformação das escalas temporais foi realizada por meio do *software* estatístico R, uma adaptada para cálculos e gráficos matemáticos e estatísticos.

### **3.2.4 Modelagem**

#### **3.2.4.1 Obtenção de Zonas Pluviometricamente Homogêneas**

Nesta etapa, o programa computacional utilizado para as análises dos dados foi o WEKA versão 3.6.3. O WEKA é um software livre que está disponível sob licença GNU (*General Public License*), composto de uma coleção de algoritmos nas áreas de aprendizado de máquina e mineração de dados (HALL *et al.*, 2009; WITTEN *et al.*, 2011).

A técnica de modelagem escolhida para seleção das zonas homogêneas foi a clusterização, por meio do algoritmo *k-means* que forneceu o agrupamento dos dados de acordo com a métrica de distâncias euclidiana. Esse agrupamento ocorre com a especificação de uma medida de distância de acordo com o algoritmo, conforme a representatividade dos dados, ou quantitativa ou qualitativa (REZENDE *et al.*, 2005).

Os *clusters* obtidos foram espacializados em ambiente Arc MAP versão 9.3 (ESRI, 2009), uma ferramenta de sistema de informações geográficas (SIG) vetorial, que permite georreferenciamento e cruzamento de diferentes planos de informação. Desta forma, foi possível uma melhor visualização dos resultados com a confecção do mapa de distribuição das zonas pluviometricamente homogêneas do estado da Bahia.

Foram utilizadas informações geográficas extraídas da base de dados do sistema hidrológico web, fornecidas pela ANA (2011). As informações originais estavam em formato *Access*, sendo transformadas para formato Excel (conforme Tabela 17 disponível no Apêndice II) e, posteriormente, importadas para o Arc MAP, onde foram transformadas para *shapefile* compondo o banco de dados que serviu de base para o desenvolvimento do trabalho.

### 3.2.5 Avaliação dos Clusters

A definição do número de clusters baseou-se nos resultados obtidos para série histórica do período de 1981 a 2010, por meio das análises da distribuição pluviométrica, mapas com a espacialização das estações em seus respectivos, literatura especializada da região (Molion *et al.*, 2002 e Cavalcanti *et al.*, 2009), e principalmente, consultas a especialistas da área de climatologia e meteorologia.

## 3.3 Análise de Eventos Extremos de Precipitação

Foi aplicada a Técnica dos quantis para estabelecer os limiares de quantidade de chuva, com o intuito de classificar a ocorrência de eventos extremos em diferentes escalas temporais (anual, mensal e diária). Este método envolve o princípio da relativização estatística, no sentido de se procurar interpretar de forma adequada o verdadeiro significado de um total pluviométrico, conforme previamente citado.

Conforme Xavier e Xavier (1987), a técnica define que uma variável  $X$  representa os valores de precipitação, onde  $Q_p$  é o limite do intervalo do quantil para variável  $X$ , sendo  $p$  a probabilidade de ocorrência, tal que  $X$  seja menor ou igual que  $Q_p$ . Desta forma, a escolha do intervalo associado a um quantil deve satisfazer a seguinte relação:  $\text{Prob.}(X \leq Q_p) = p$ , sendo ( $0 < p < 1$ ). Neste trabalho foi utilizada a metodologia de Xavier (1984, 1999, 2002, 2007), desenvolvida inicialmente por Pinkayan (1966).

Para as análises das escalas anual e mensal foram adotadas as ordens quantílicas  $p = 0,15$  e  $0,85$  (Tabela 3). E para as análises da escala temporal diária foi utilizada a ordem

quantílica  $p = 0,05$  e  $p = 0,95$ . Ressalta-se que para essas análises, na cauda superior (0,95) foram considerados apenas os dias com chuva, ou seja, acima de 2mm de chuva, conforme Adejuwon *et al.* (2006), de acordo com a Tabela 4. Tais valores de  $p$  caracterizam as caudas da distribuição de probabilidade, onde a cauda inferior corresponde a  $p = 0,15$  e  $p = 0,05$ ; e a cauda superior se refere a  $p = 0,85$  e  $p = 0,95$ .

**Tabela 3:** Categorias de Precipitação usando a Técnica dos Quantis ( $p = 0,15$  e  $0,85$ ).

<b>Classificação</b>	<b>Limite de Precipitação</b>
Muito Seco (MS)	$X_i \leq Q_{0,15}$
Seco (S)	$Q_{0,15} < X_i \leq Q_{0,35}$
Normal (N)	$Q_{0,35} < X_i < Q_{0,65}$
Chuvoso ( C )	$Q_{0,65} \leq X_i < Q_{0,85}$
Muito Chuvoso (MC)	$X_i \geq Q_{0,85}$

**Tabela 4:** Categorias de Precipitação usando a Técnica dos Quantis ( $p = 0,05$  e  $0,95$ ).

<b>Classificação</b>	<b>Limite de Precipitação</b>
Extremamente Seco	$P < Q_{0,05}$
Extremamente Chuvoso	$P > Q_{0,95}$

### 3.3.1 Análise Anual

Para a investigação na escala anual, buscando o conhecimento do comportamento e das tendências da precipitação, foi calculada a média de todas as séries pluviométricas para cada zona homogênea. Essa escala propiciou o conhecimento da tendência da precipitação em cada *cluster* no período estudado, permitindo estimar o comportamento geral e dessemelhanças espaciais.

Tomando como base as tabelas resultantes do cálculo da média, foram elaborados gráficos de representação para analisar as características das séries e sua variação interanual para o período de 1981 a 2010. Os resultados de valores médios de precipitação anual foram utilizados para o procedimento de identificação dos limiares superiores e inferiores, por meio da técnica dos quantis. Foi adotada como ordem quantílica os valores de  $p = 0,15$  e  $p = 0,85$ .

Para corroborar os resultados dos anos classificados como “MUITO SECO” e “MUITO CHUVOSO”, foi realizado o cálculo do Índice de Sazonalidade para série temporal em estudo. Para o cálculo da sazonalidade, os valores da Tabela 6 são obtidos pela divisão de seus valores correspondentes, na Tabela 5, pela média anual. Por exemplo, para o ano de 1981, mês de janeiro, o valor 1,07 (na Tabela 6) é calculado pela razão  $35,58/33,25$ . O mesmo procedimento é usado para o cálculo dos demais valores da Tabela 6. Finalmente, os índices de sazonalidade são obtidos pela média dos valores das colunas da Tabela 6. Na prática, o índice é útil para comparar os valores médios de um determinado ano com relação ao padrão da componente sazonal (ciclo sazonal), que é escrito em termos percentuais (NEVES *et al.*, 2007).

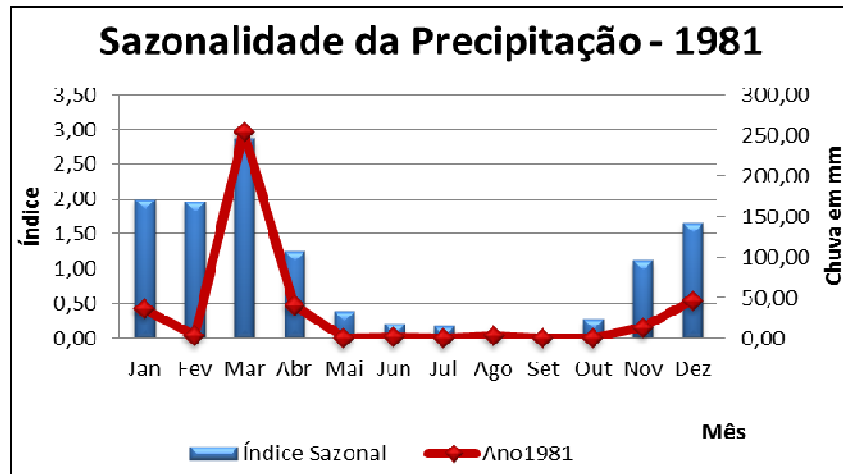
O índice de sazonalidade fornece informação sobre o comportamento pluviométrico do ano em foco em relação à sazonalidade da precipitação ao longo da série histórica (1981 a 2010). O Gráfico 2 ilustra a comparação dos valores médios de precipitação do ano de 1981 com relação à sazonalidade. Esta informação permite uma análise pluviométrica mais completa da classificação de um ano como “MUITO CHUVOSO” ou “MUITO SECO”.

**Tabela 5:** Esquematização para obtenção do Índice de Sazonalidade (Acumulado da Precipitação Mensal e Média da Precipitação Anual).

Data	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez	MedAnual
1981	35,58	3,32	255,02	40,73	0,25	1,88	0,42	2,78	0,28	0,00	12,83	45,95	33,25
1982	18,86	27,20	54,52	47,12	3,30	4,88	3,13	2,10	2,08	0,83	0,00	50,28	17,86
1983	56,93	103,15	85,43	2,52	2,77	0,33	3,70	2,35	0,02	0,88	44,08	23,77	27,16
1984	19,75	10,32	163,02	80,72	14,83	3,35	4,12	1,08	14,68	8,82	30,68	5,47	29,74
1985	348,97	127,08	159,50	153,15	44,58	57,72	5,98	8,25	0,02	5,67	27,05	210,38	95,70
...													
2010	45,73	55,48	29,90	78,85	7,05	10,78	16,20	0,82	1,28	63,74	0,10	114,18	35,34

**Tabela 6:** Esquematização para obtenção do Índice de Sazonalidade (Índice Sazonal Mensal e Sazonalidade Mensal para o período de 1981 a 2010).

Data	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
1981	1,07	0,10	7,67	1,22	0,01	0,06	0,01	0,08	0,01	0,00	0,39	1,38
1982	1,06	1,52	3,05	2,64	0,18	0,27	0,18	0,12	0,12	0,05	0,00	2,82
1983	2,10	3,80	3,15	0,09	0,10	0,01	0,14	0,09	0,00	0,03	1,62	0,88
1984	0,66	0,35	5,48	2,71	0,50	0,11	0,14	0,04	0,49	0,30	1,03	0,18
1985	3,65	1,33	1,67	1,60	0,47	0,60	0,06	0,09	0,00	0,06	0,28	2,20
...												
2010	1,29	1,57	0,85	2,23	0,20	0,31	0,46	0,02	0,04	1,80	0,00	3,23
Sazonalidade	2,00	1,97	2,87	1,25	0,39	0,20	0,17	0,06	0,05	0,27	1,12	1,65



**Gráfico 2:** Sazonalidade da Precipitação de 1981 a 2010 com destaque para o ano de 1981.

Foram utilizadas informações de ocorrência e intensidade dos fenômenos El Niño de Oscilação Sul para explicar as flutuações climáticas ocorridas durante a série trabalhada e associação com as ocorrências de eventos extremos desse período. As informações sobre El Niño de Oscilação Sul foram disponibilizadas pela *National Oceanic and Atmospheric Administration* – NOAA, sigla em inglês (NOAA, 2012).

### 3.3.2 Análise Mensal

A análise da escala mensal teve como característica abordar apenas os meses pertencentes à “estação chuvosa” de cada *cluster*. Desta forma, a aplicação das técnicas de identificação de extremos possibilitou a visualização de forma pontual do impacto da ocorrência de um evento extremo no período crucial para as atividades agrícolas.

A identificação das “estações chuvosas” de cada *cluster* foi baseada na literatura especializada (Kousky, 1979; Molion *et al.*, 2002; Cavalcanti *et al.*, 2009), nos gráficos de distribuição pluviométrica mensal – 1981 a 2010, e em de reuniões com especialistas da área de meteorologia.

A técnica utilizada para identificação dos limiares de intensidade de chuva (extremos positivos e extremos negativos) foi o Quantil obedecendo à ordem quantílica de  $p=0,85$  para cauda superior e  $p=0,15$  para cauda inferior (XAVIER, 2001). A aplicação desta técnica permitiu o conhecimento dos meses com seus respectivos anos que apresentaram fenômenos “muito secos” e “muito chuvosos”.

Uma abordagem interessante executada na análise da escala mensal foi a utilização da tarefa de associação, por meio do uso do algoritmo *Generalized Sequential Patterns* (GSP) para geração de padrões sequenciais. O algoritmo está disponível no Ambiente Weka 3.6.3. A ordem em que os padrões estão alinhados reflete a cronologia com que aconteceram os fatos representados por eles. (AGRAWAL e SRIKANT, 1995).

A geração de padrões sequenciais para a ocorrência de eventos extremos foi possível após a identificação destes eventos para cada *cluster*, ou seja, a série de padrões referentes aos eventos extremos também foi específica para cada *cluster*. O arquivo de entrada para processamento dos dados no ambiente WEKA foi composto apenas dos meses e anos que ocorreram os eventos positivos (EP) e ou negativos (EN). Por exemplo, a expressão “ano81” se refere ao ano de 1981, enquanto “(EN m2)” indica que no mês de fevereiro ocorreu um evento extremo negativo, de acordo com a primeira linha da Tabela 7.

**Tabela 7:** Esquematização do conjunto de entrada base do algoritmo *GSP*.

<b>Ano</b>	<b>Evento</b>
ano81	(EN m2)
ano81	(EP m3)
ano83	(EN m4)
ano84	(EN m2)
.	.
.	.
.	.
ano10	(EP m7)

### 3.3.3 Análise Diária

Considerando que o conjunto original das séries históricas, utilizado neste trabalho, contém dados de precipitação diária, foi possível analisar os impactos dos eventos extremos diários ao longo dos trinta anos. Como a identificação das regiões pluviometricamente homogêneas foi realizada com os dados em escala mensal, teve-se conhecimento de quais estações meteorológicas pertenciam a cada *cluster*. Desta forma, o centróide de cada *cluster* foi recalculado para sintetizar o volume pluviométrico diário das estações, por *cluster*.

A próxima etapa foi definir os limiares para identificação e quantificação de eventos extremos e classificar a precipitação diária acumulada em 24 horas, por meio da técnica do Quantil. Para a realização desta análise, a ordem quantílica utilizada foi 95% para os “dias de

chuva”, sendo considerado um dia de chuva aquele cuja precipitação for maior que 2mm, ou seja, abaixo desse valor tem-se um dia seco, pois os valores inferiores a este limiar são de pouca importância para a agricultura (ADEJUWON *et al.*, 2006).

Para abordagem dos dias considerados com ocorrência de eventos extremos negativos foi aplicada uma contagem de dias secos consecutivos. Esta análise permite o conhecimento dos anos com mais dias secos e dos anos com mais sequência de dias secos consecutivos, proporcionando uma nova discussão sobre o comportamento dos eventos extremos de seca no estado da Bahia.



## **4. RESULTADOS E DISCUSSÃO**

### **4.1 Distribuição da Precipitação Pluviométrica**

A representação geoespacial da distribuição pluviométrica para o período de 1981 a 2010 no estado da Bahia, por meio do método geostático IDW (*Inverse Distance Weighted – Ponderação do Inverso da Distância*), permitiu uma melhor visualização da disparidade do volume pluvial existente no território baiano, revelando para região menos chuvosa índices médios anuais de 438mm e para mais chuvosa 2033mm, conforme Figura 12.

A descrição dos fenômenos que caracterizam tal cenário pluviométrico, bem como o comportamento das chuvas durante a série histórica, será discutida na próxima seção de forma mais detalhada.

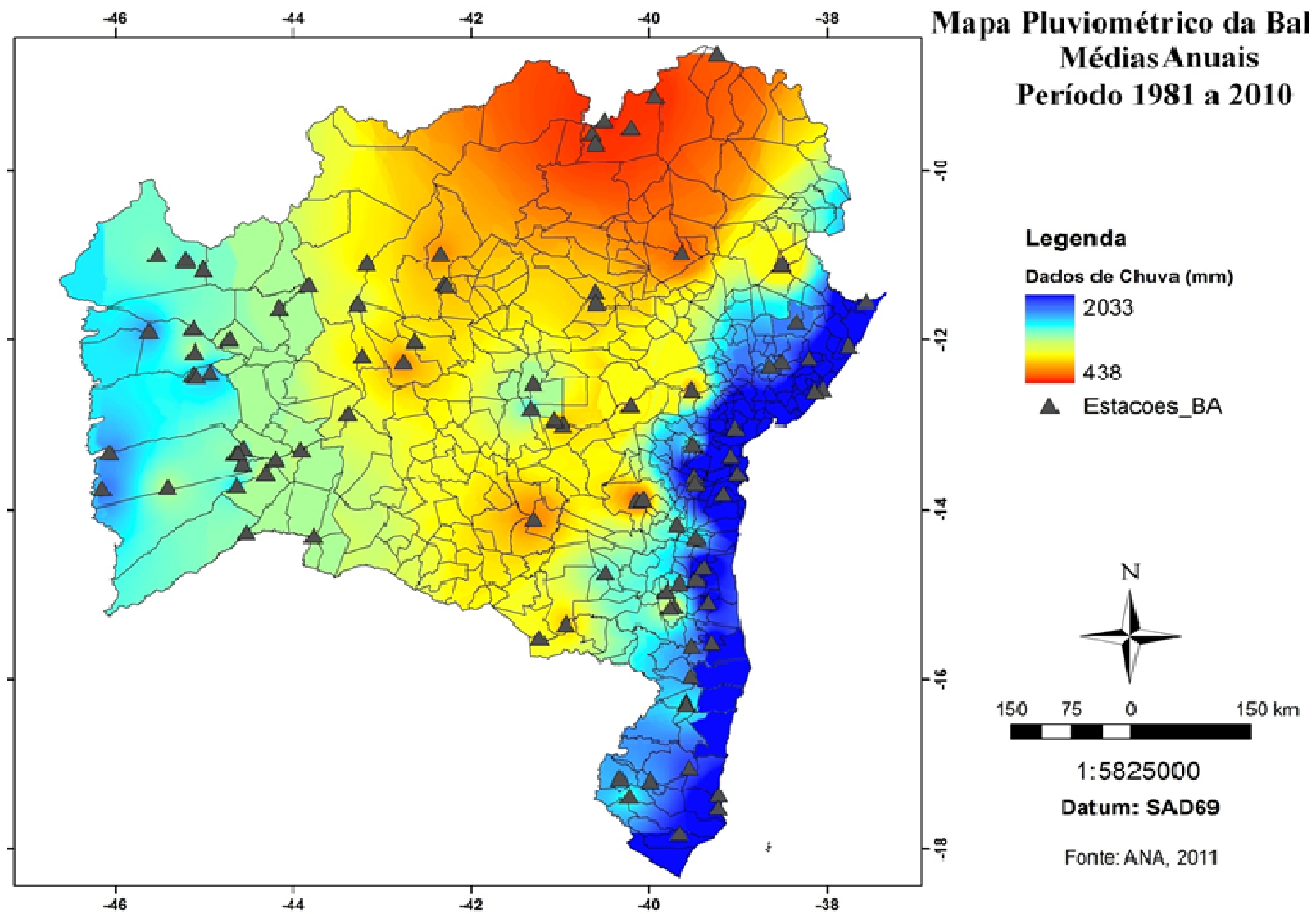
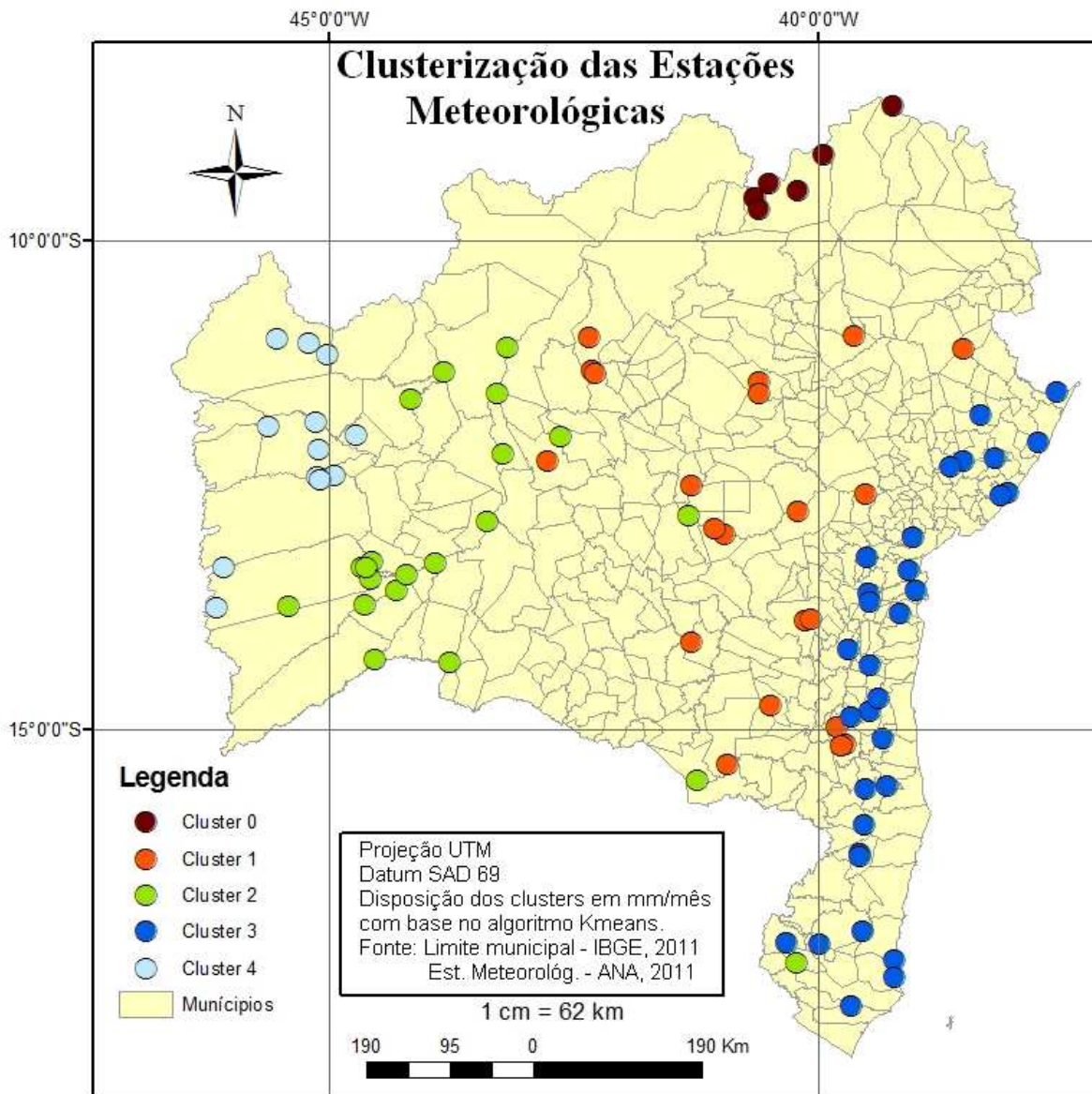


Figura 12: Mapa Pluviométrico da Bahia (1981 a 2010).

## 4.2 Identificação das Zonas Pluviométricamente Homogêneas

A aplicação das técnicas de agrupamento de mineração de dados permitiu a identificação das diferentes regiões pluviométricas do estado da Bahia, de acordo com a similaridade do volume pluvial, conforme ilustrado na Figura 13. Informações mais detalhadas sobre as estações meteorológicas podem ser encontradas na Tabela 17 (disponível no Apêndice). Para esta etapa do trabalho, foi adotado como padrão do estudo a escala mensal do banco de dados.



**Figura 13:** Regionalização pluviométrica do estado da Bahia.

Foram testados valores para  $k$  (nº de grupos) = 2, 3, 4 e 5 para definição do número de zonas homogêneas. Optou-se pela divisão de cinco zonas pluviometricamente homogêneas, uma vez que este número está em consonância com pareceres de especialistas da área e registros na literatura referentes à distribuição, orografia, maritimidade e sistemas pluviométricos que atingem o estado (BRAGA, 1998; CAVALCANTI, *et al.* 2009; TANAJURA *et al.*, 2010).

Os resultados mostram que o algoritmo *k-means* gerou os *clusters* de forma coerente com a distribuição pluviométrica espaço temporal do estado da Bahia. Resultados semelhantes foram encontrados ao comparar com o mapa pluviométrico do Estado referente ao período de 1981 a 2010, obtidos por meio da técnica de interpolação IDW (*Inverse Distance Weighted – Ponderação do Inverso da Distância*) conforme Figura 12.

Os métodos de agrupamentos foram utilizados em vários trabalhos, em diferentes partes do Brasil, tais como: Alagoas (Souza *et al.*, 1992), Paraíba (Silva *et al.*, 1996), Rio Grande do Norte (Lopes *et al.*, 1996), Regiões Sul e Sudeste (Cardoso *et al.*, 2004), Rio Grande do Sul (Boschi *et al.*, 2011). Neste contexto, destaca-se o trabalho de André *et al.* (2008), que utilizaram o *k-means* em dados meteorológicos do estado do Rio de Janeiro para identificar regiões pluviometricamente homogêneas, sob forte influência dos fatores maritimidade e orografia.

Observa-se que a disposição dos *clusters* no território baiano evidencia a distribuição decrescente nos volumes de chuva, das extremidades leste (Planície Litorânea) e oeste (Chapadões Ocidentais) para o interior do estado da Bahia. Desta forma, confirma a ação periférica dos sistemas dinâmicos ZCIT, FPA, ZCAS e VCAN que possibilitam a existência de regimes pluviais diferentes e superpostos sobre o território baiano, verificados da mesma forma por Barbosa (2000).

O *cluster* 0, representado por 6 estações meteorológicas, situa-se ao norte do estado. Os municípios deste grupo pertencem à região do semiárido nordestino, submetidos aos fortes rigores do clima árido, acometidos por longos períodos de estiagem, elevadas temperaturas, altas taxas de evapotranspiração e alta disponibilidade energética de luz solar. Sobre este *cluster*, atuam a Zona de Convergência Inter Tropical e os Vórtices Ciclônicos (MOLION *et al.*, 2002).

O *cluster* 1, representado por 21 estações meteorológicas, abrange a região mais setentrional, expandindo-se por toda parte central de norte ao sul do estado. Ocupa a maior área do território baiano, onde estão localizados os municípios do semiárido pertencentes ao Polígono das Secas. São responsáveis pela distribuição pluviométrica deste grupo os VCAN, ao norte, e a FPA nas demais áreas.

O *cluster* 2, composto por 21 estações meteorológicas, delimita a região de transição do semiárido para o cerrado do oeste baiano. Os municípios pertencentes a este grupo não são contíguos geograficamente como nos demais *clusters*, isto ocorre devido à expressividade do efeito orográfico da região da Chapada Diamantina e da presença dos mecanismos de ação da FPA filiados ao regime climático da região Sudeste do país (BARBOSA, 2000; BRAGA *et al.*, 1998).

O *cluster* 3, composto por 32 estações meteorológicas, é representado pelos municípios da costa litorânea do Estado, apresentando grande volume de chuva e relativa regularidade pluvial durante todo o ano. Sobre este *cluster* atuam os sistemas meteorológicos de FPA e as brisas marítimas. Representa a zona com maior índice pluviométrico, tal característica atribui a esta região a condição de baixo risco de seca, segundo os estudos de Risco de Seca do estado da Bahia realizado pela SEI (2000).

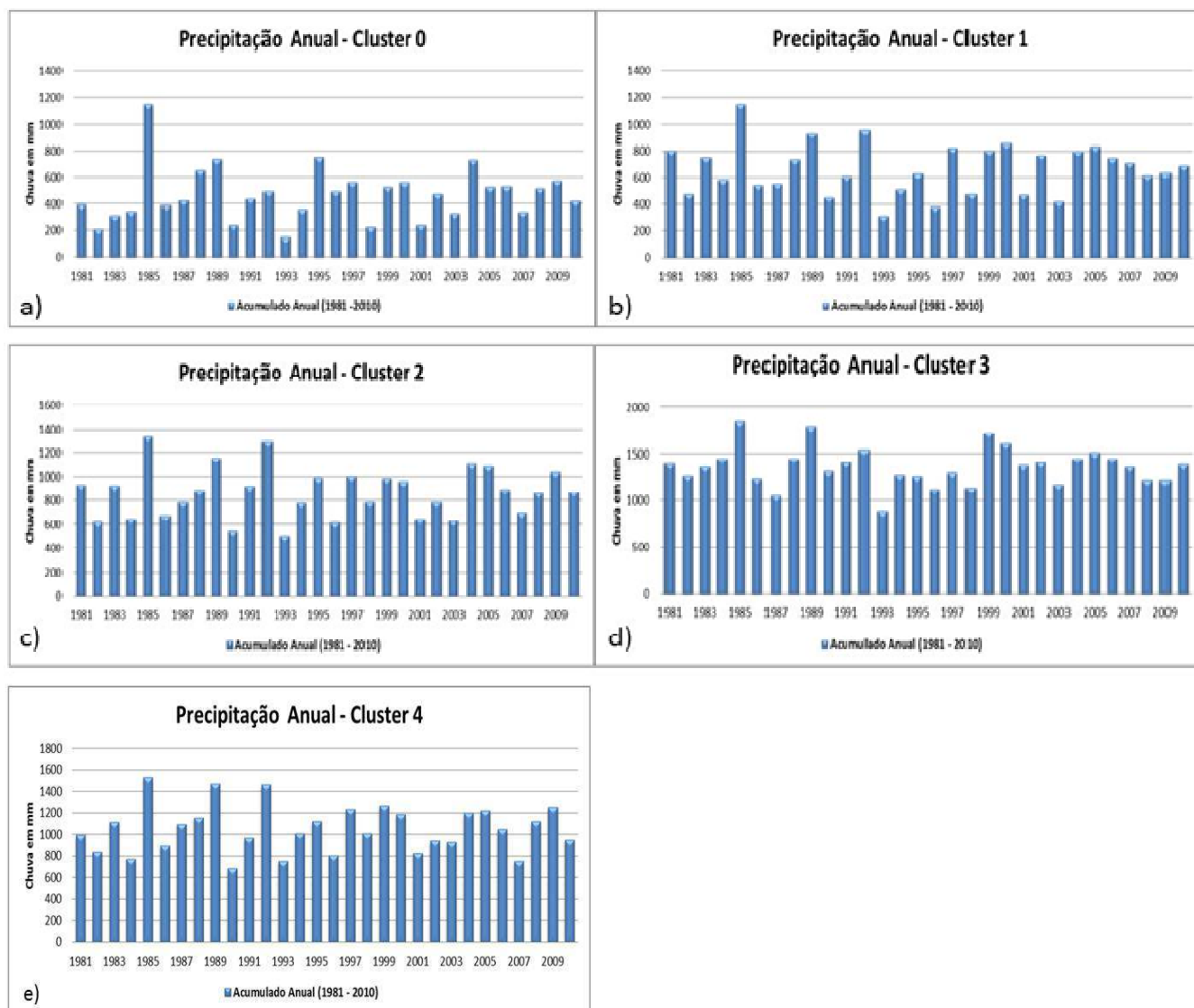
Por fim, o *cluster* 4, representado por 12 estações meteorológicas, corresponde à parte ocidental do oeste da Bahia, com períodos chuvosos e secos bem definidos ao longo do ano. É a região do cerrado baiano cuja concentração das atividades agropecuárias proporciona aos municípios desse grupo o maior PIB agrícola do estado da Bahia. Possui o segundo maior volume pluviométrico e, sobre este grupo, atuam os Sistemas Frontais e a Zona de Convergência do Atlântico Sul.

### **4.3. Análise da Precipitação nas Zonas Homogêneas**

A Tabela 8 apresenta os valores referentes ao acumulado anual das precipitações médias para os cinco *clusters*, no período de 1981 a 2010, médias, desvios-padrão, coeficientes de variação (CV), mediana, máximo e mínimo. Nela verifica-se a alta heterogeneidade na distribuição espacial das chuvas, com valores mínimos variando entre 160mm a 899mm e médios de 477mm a 1377mm para diferentes regiões, que também podem ser observados no Gráfico 3.

**Tabela 8:** Valores acumulados anuais de precipitações, médias, desvio-padrão, coeficiente de variação, máximo e mínimo para o período de 1981 a 2010.

Cluster	Média	DesvPad	CV (%)	Mediana	Máximo	Mínimo
0	477	201	42	465	1148	160
1	672	189	28	664	1147	323
2	873	210	24	886	1337	512
3	1377	209	15	1378	1861	900
4	1059	219	21	1034	1531	695

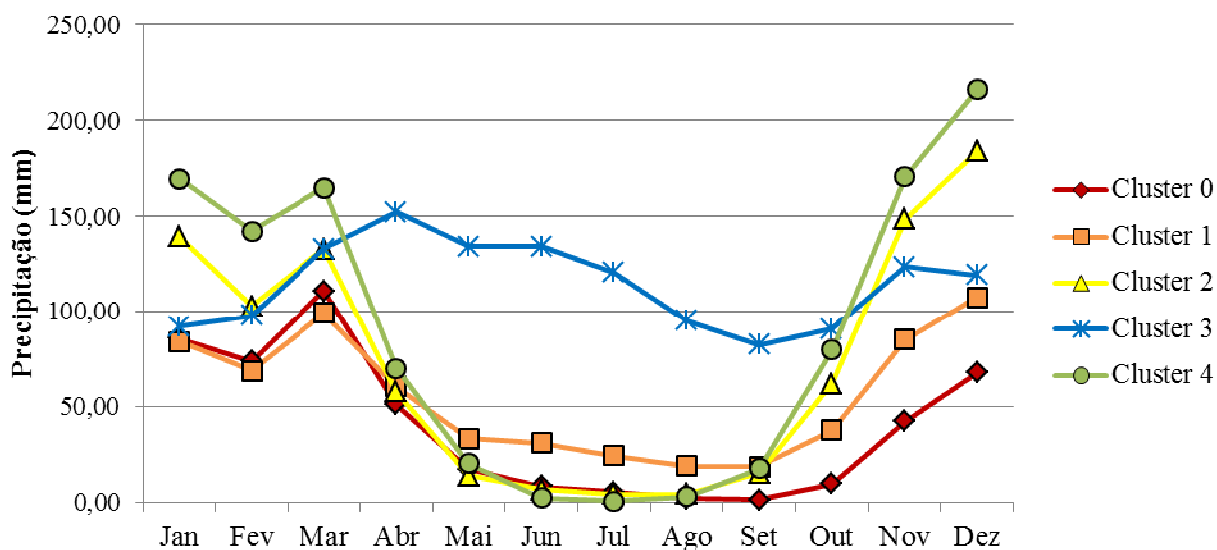


**Gráfico 3:** (a), (b), (c), (d), (e) - Distribuição do acumulado anual de precipitação para o período de 1981 a 2010 nos cinco clusters.

Verifica-se na Tabela 8 que os *clusters* 0 e 1 apresentam os menores volumes pluviométricos e os maiores coeficientes de variação 42% e 28%, respectivamente, o que evidencia as características climáticas das regiões semiáridas, vistas também por Molion *et al.* (2002) e Ferreira *et al.* (2005), ao estudar a variabilidade pluviométrica no semiárido do Nordeste brasileiro.

Já o *cluster* 3 apresenta o menor coeficiente de variação (15%), representando a região de maior altura e regularidade pluviométrica durante toda a série histórica. Observação semelhante encontra-se em Barbosa (2000). Nos *clusters* 2 e 4 o coeficiente de variação corresponde a 24% e 21%, respectivamente, indicando uma baixa variação no volume pluviométrico. O *cluster* 2 corresponde à região mais próxima do semiárido, o que representa comportamento de variabilidade mais elevado em relação ao *cluster* 4, que faz parte de uma região de cerrado e o segundo maior em níveis pluviométricos anuais.

O Gráfico 4 mostra as distribuições mensais da precipitação média de cada *cluster* determinada pelo método do *k-means*. Desta forma, é possível comparar as regiões avaliando onde há maior ou menor variabilidade da média entre os *clusters*.



**Gráfico 4:** Precipitação média mensal dos *clusters* referente ao período de 1981 a 2010.

Observa-se que durante a estação chuvosa para os *clusters* 0, 1, 2 e 4 há uma considerável dispersão entre os valores mensais. Entretanto, no período seco, apenas o *cluster* 1 apresenta variações consideráveis em relação aos demais. Verifica-se que a estação chuvosa do *cluster* 3 é composta por outros meses em relação aos demais *clusters*.

A estação chuvosa do *cluster* 0 compreende aos meses de dezembro a abril, com volume máximo no mês de março (Gráfico 4). Isso ocorre devido à influência da ZCIT, que atinge a região nesse período, observado também por Kousky (1979) e Cavalcanti (2009). A precipitação na estação chuvosa deste *cluster* é 90% do total anual, caracterizando uma região de estações bem definidas ao longo do ano.

Compartilha desta mesma característica os *clusters* 2 e 4, que apresentam no período chuvoso, um total da precipitação anual de 95% e 96%, respectivamente. Apresentam semelhança entre os meses que compreendem o período chuvoso, outubro a abril, com volumes máximos nos meses de novembro a dezembro (Gráfico 2). O *cluster* 4 representa o grupo com o maior volume pluvial no período chuvoso em relação aos demais, alcançando um total superior a 1000mm.

Observa-se que os volumes pluviais do período chuvoso do *cluster* 1 correspondem a 81% da precipitação total. Os sistemas meteorológicos FPA, ZCAS e VCANs são responsáveis pela produção de chuvas nessa região e contribuem para ocorrência de chuvas entre os meses de novembro a abril, com máxima no mês de dezembro.

A precipitação dos períodos chuvosos e secos do *cluster* 3 é bem elevada, com destaque para volumes acima de 459mm na estação seca, revelando regularidade na distribuição pluviométrica durante todo o ano. Apresenta 66% do montante pluviométrico no período chuvoso em relação ao total das chuvas anuais. Os meses que integram a estação chuvosa desta região são março a julho, novembro e dezembro, com volumes máximos nos meses de dezembro e abril.

A Tabela 9 evidencia a contribuição da precipitação para as estações chuvosa e seca, em cada *cluster*. Nela verifica-se que os *clusters* 0, 2 e 4 apresentam estações bem definidas em relação as demais, com precipitação média do período chuvoso acima de 800mm (*clusters* 2 e 4); 431mm (*cluster* 0) para estação chuvosa e acima de 44mm para estação seca (*clusters* 0, 2 e 4).



**Tabela 9:** Contribuição da precipitação nos períodos chuvosos e secos para precipitação anual.

	<b>Cluster 0</b>	<b>Cluster 1</b>	<b>Cluster 2</b>	<b>Cluster 3</b>	<b>Cluster 4</b>
<b>Prec. Anual</b>	477	672	873	1377	1059
<b>Per. Chuvoso</b>	432	544	828	918	1014
<b>Per. Seco</b>	45	127	45	459	45
<b>% Per. Chuvoso</b>	90	81	95	66	96
<b>% Per. Seco</b>	10	19	5	34	4

#### **4.4 Análise de Eventos Extremos de Precipitação em Diferentes Escalas Temporais.**

##### **4.4.1 Análise Anual**

A aplicação da Técnica dos quantis, considerando os dados de precipitação anual, resultou na classificação e identificação dos anos com ocorrência de eventos extremos de chuva. A classificação desses anos, juntamente com o total pluviométrico anual de cada região pluviometricamente homogênea, pode ser visualizada na Tabela 10.

**Tabela 10:** Precipitação e Classificação dos anos pelos quantis nas regiões pluviometricamente homogêneas (mm) Categorias: MS= “muito seco”; S= “seco”; N= “normal”; C= “chuvoso” e MC= “muito chuvoso”.

Ano	Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4
1981	399 N	810 C	942, C	1398 N	999 N
1982	214 MS	481 S	629 MS	1279 S	845 S
1983	326 S	751 C	932 N	1372 N	1122 N
1984	357 S	591 S	658 S	1460 C	779 MS
1985	1148 MC	1147 MC	1337 MC	1861 MC	1531 MC
1986	396 S	55 S	684 S	1241 S	904 S
1987	433 N	560 S	795 N	1064 MS	1102 N
1988	659 MC	736 N	891 N	1458 C	1156 C
1989	740 MC	935 MC	1163 MC	1811 MC	1471 MC
1990	245 MS	458 MS	550 MS	1322 N	695 MS
1991	447 N	614 N	931 N	1415 C	974 N
1992	501 N	958 MC	1298 MC	1551 MC	1465 MC
1993	160 MS	323 MS	512 MS	900 MS	757 MS
1994	369 S	521 S	790 S	1288 S	1016 N
1995	753 MC	633 N	997 C	1272 S	1128 C
1996	501 N	385 MS	622 MS	1132 MS	813 MS
1997	568 C	827 C	1011 C	1310 N	1238 C
1998	234 MS	482 S	797 N	1150 MS	1012 N
1999	535 C	809 C	992 C	1735 MC	1272 MC
2000	568 C	857 MC	970 C	1624 MC	1193 C
2001	245 MS	476 MS	655 S	1384 N	831 S
2002	484 N	771 C	794 S	1414 N	946 S
2003	341 S	423 MS	641 MS	1188 MS	937 S
2004	735 MC	805 C	1110 MC	1462 C	1205 C
2005	535 C	838 MC	1091 MC	1526 C	1223 C
2006	540 C	747 N	898 N	1461 C	1053 N
2007	347 S	713 N	709 S	1371N	761 MS
2008	526 N	619 N	869 N	1227 S	1122 N
2009	582 C	640 N	1045 C	1232 S	1256 MC
2010	424 N	688 N	880 N	1392 N	956 S

Observa-se que a ocorrência de anos muito secos e muito chuvosos (extremos) não apresenta similaridade em todos os *clusters*. Em outras palavras, um ano considerado seco para uma região pode ser classificado chuvoso em outra região. Resultados semelhantes foram encontrados por Souza *et al.* (1992), Xavier *et al.* (1999) e Monteiro *et al.* (2012), ao

identificar eventos extremos no estado do Alagoas, Ceará e Rio Grande do Norte, respectivamente.

Os autores supracitados afirmam que os limiares de chuva variam entre as regiões pluviométricas. Tal afirmação pode ser verificada na tabela de classificação dos limites de precipitação para as categorias correspondente aos eventos extremos “muito secos” e “muito chuvosos”, para todos os *clusters* do estado da Bahia (Tabela 11), utilizando os dados pluviométricos anuais referentes ao período de 1981 a 2010.

Em particular, verifica-se que o menor limiar para ser considerado um ano “muito seco” encontra-se no *cluster* 0 (região ao norte do estado, submetido aos rigores do clima semiárido) com limiar de 273mm, e o maior limiar para a mesma categoria encontra-se no *cluster* 3 (região da costa litorânea, o qual detém os maiores volumes pluviométricos do estado) com limiar de 1202mm.

**Tabela 11:** Limites de classificação anual para identificação de eventos extremos “Muito Seco” e “Muito Chuvoso” para todos os *clusters*.

Clusters	Limiares de Eventos Extremos	
	Muito Seco Q (0,15)	Muito Chuvoso Q (0,85)
0	$P \leq 273$	$P \geq 633$
1	$P \leq 478$	$P \geq 834$
2	$P \leq 646$	$P \geq 1075$
3	$P \leq 1202$	$P \geq 1542$
4	$P \leq 820$	$P \geq 1250$

Observa-se que o limite (1202mm) definido pela ordem quantílica para classificação de um ano como “muito seco” para o *cluster* 3 (representante da faixa litorânea) é superior em 569mm do limite (633mm) para um ano ser considerado “muito chuvoso” para o *cluster* 0 (representante da região Norte do estado), correspondendo a um percentual de 90% de diferença entre os limites de chuva dos *clusters* para classificações extremas.

A identificação dos anos com ocorrência dos eventos extremos (muito seco e muito chuvoso) para todos os *clusters* é apresentada na Tabela 12. Estudos relacionados ao histórico das secas para região Nordeste e o estado da Bahia corroboram os resultados encontrados com a técnica do Quantil para essa região (CARVALHO, 1994; BARBOSA, 2000; SEI, 2010).

Verifica-se em todos os clusters que existe uma predominância de ocorrência de anos “muito secos” na década de 90. Os anos de destaque dessa década são 1990, 1993, 1996 e 1998, os quais também marcaram a história das secas ocorridas na região Nordeste (BARBOSA, 2000). Quanto à classificação de anos muito chuvosos, ocorre uma variação entre as décadas, prevalecendo a década de 80, com os anos 1985 e 1989 com ocorrência em todo estado da Bahia.

Constata-se uma alta variabilidade entre os extremos de precipitação no estado durante os trinta anos estudados. Isto pode ser visto na Tabela 12, ao analisar os anos que antecedem ou sucedem os anos considerados como “muito secos”. Pode-se constatar essa simultaneidade nos anos 1990, 1993, 1996, 1998, 2001 e 2003 (muito secos) em relação aos anos 1989, 1992, 2000 e 2004 (muito chuvosos).

**Tabela 12:** Identificação dos anos com ocorrência de eventos extremos na Bahia.

<b>Identificação dos Anos com Eventos Extremos (Quantil 15 - 85)</b>		
<b>Clusters</b>	<b>Muito Seco</b>	<b>Muito Chuvoso</b>
<b>0</b>	1982; 1990 1993; 1998 2001	1985; 1988 1989; 1995 2004
<b>1</b>	1990; 1993 1996; 2001 2003	1985; 1989 1992; 2000 2005
<b>2</b>	1982; 1990 1993; 1996 2003	1985; 1989 1992; 2004 2005
<b>3</b>	1987; 1993 1996; 1998 2003	1985; 1989 1992; 1999 2000
<b>4</b>	1984; 1990 1993; 1996 2007	1985; 1989 1992; 1999 2009

Verifica-se que o ano de 1993 é considerado como um ano “muito seco” para todo o estado. Observa-se também que foi o ano de menores volumes de precipitação de acumulado

em quase todos os *clusters* (0, 1, 2 e 3, com 160mm, 323mm, 512mm, 900mm, respectivamente), exceto para o *cluster* 4 com o acumulado de 757mm.

Resultados semelhantes foram encontrados por Barbosa (2000), descrevendo a grande seca de 1993 como a mais rigorosa para o estado, atingindo fortemente o setor agropecuário com perdas totais de safras e morte de inúmeros animais. Andrade *et al.* (2012) falam sobre a seca do ano 92/93 como desastrosa para o município de Itaetê, causando a morte de animais e perda total para os principais produtos da região, como, mamona, milho e feijão. Ao tratar das grandes perdas sobre a safra de feijão, Barbosa (2000) destaca a cidade de Irecê, conhecida como “cidade do feijão”, pela sua grande produtividade, que decretou estado de emergência em 1993 pelas inúmeras perdas.

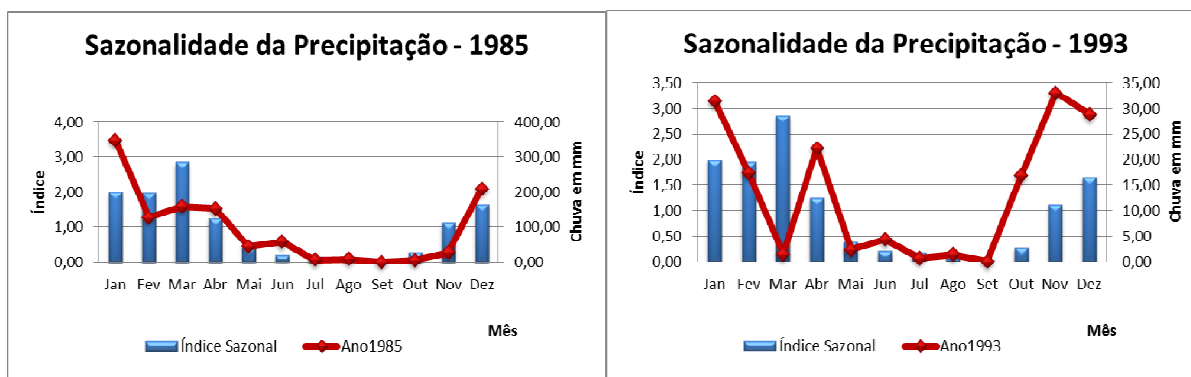
A seca do ano de 1993 foi registrada em outros estados da região Nordeste. Segundo Xavier *et al.* (1999), em estudos com série pluviométricas históricas do estado de Ceará, a seca de 1993 teve a característica marcante de ser mais rigorosa do que a de 1982/83, considerada até então a mais desastrosa seca ocorrida neste estado e em grande parte do Nordeste. Semelhantemente, Monteiro *et al.* (2012) identificaram os anos mais secos e mais chuvosos, aplicando a técnica dos quantis, para o município de Mossoró no Rio Grande do Norte.

Com relação aos anos identificados como “muito chuvosos” (Tabela 12), verifica-se que ocorre uma alta dispersão entre a ocorrência desse evento extremo entre as três décadas estudadas. No entanto, os anos que mais apresentaram volumes pluviométricos para serem considerados como “muito chuvosos” em quase todos os *clusters* (1, 2, 3 e 4) foram 1985, 1989 e 1992, sendo que para o *cluster* 0, apenas o ano de 1992 não foi considerado como “muito chuvoso” em relação aos demais.

O ano de 1985, o mais chuvoso em todos os *clusters* (0, 1, 2, 3 e 4), apresentou volumes pluviométricos anuais de 1148mm, 1146mm, 1337mm, 1860mm e 1530mm, respectivamente. Observa-se o grande impacto causado principalmente para o *cluster* 0 (região com os mais baixos volumes pluviais), que apresentou neste ano um acumulado de 1148mm, o que corresponde a um desvio positivo em relação ao limiar estabelecido para a classe “muito chuvoso” (633mm) de 515mm. Isso significa um percentual superior a 81%.

Outra análise realizada foi a sazonalidade da precipitação. Nesta etapa do trabalho, são destacados os anos que apresentaram discrepância com relação ao comportamento sazonal

na precipitação da região de referência. Desta forma, ressalta-se como subsídio de análise o ano mais seco e o mais chuvoso de cada *cluster*. Os demais resultados contendo a sazonalidade com destaque, para os demais anos de todos os *clusters*, encontram-se na seção Apêndice III, IV, V, VI e VII.



**Gráfico 5:** Sazonalidade da precipitação dos anos de 1985 e 1993 – *Cluster 0*.

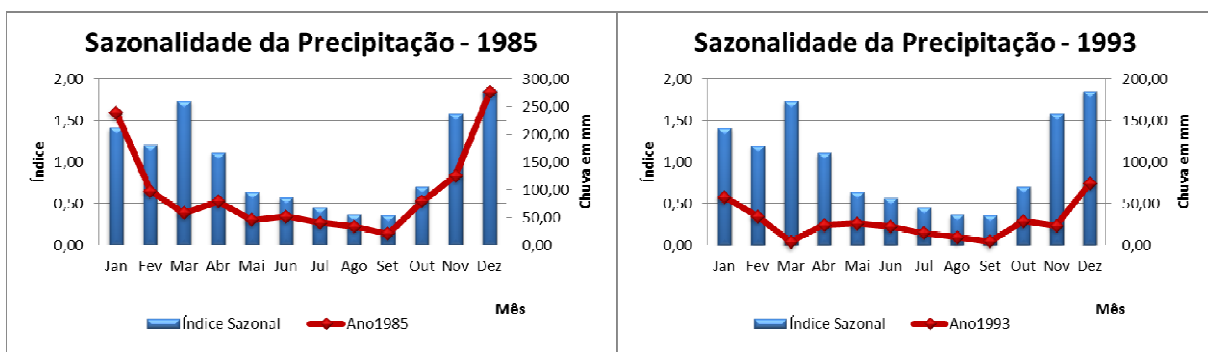
O Gráfico 5 mostra a relação da sazonalidade da precipitação mensal para a série histórica do *cluster 0* e a distribuição pluviométrica mensal referente aos anos de 1985 e 1993. Os índices sazonais encontrados acompanham o cenário pluviométrico de distribuição temporal do *cluster 0*, com estações secas e chuvosas bem definidas e o mês de março como o mais chuvoso.

Observa-se que a distribuição pluviométrica ocorrida no ano de 1985, em relação à sazonalidade das chuvas para toda série histórica, apresentou um volume de aproximadamente 400mm para o mês de janeiro, o que representa em torno de 400% acima da média climatológica para o mês. Esta é a principal causa do desalinhamento da curva de precipitação deste ano com relação à sazonalidade.

Com relação ao ano de 1993, este apresentou os menores volumes de chuva da série histórica, não acompanhando a sazonalidade. Foram registrados volumes abaixo da média histórica em todos os meses do ano, principalmente para os meses da estação chuvosa, novembro a abril, comprometendo assim toda safra agrícola. O valor máximo da precipitação mensal acumulada para este ano foi de 33,15 mm no mês de novembro, quando o esperado era 42,40mm.

O Gráfico 6 ilustra a distribuição pluviométrica do ano mais chuvoso (1985) e o mais seco (1993) para o *cluster* 1. Observa-se que a curva da precipitação do ano de 1985 acompanha o movimento da sazonalidade mensal da série. E o mês de máxima precipitação deste ano foi condizente com o mês de máxima precipitação da distribuição sazonal, visto também por Molion *et al.* (2002).

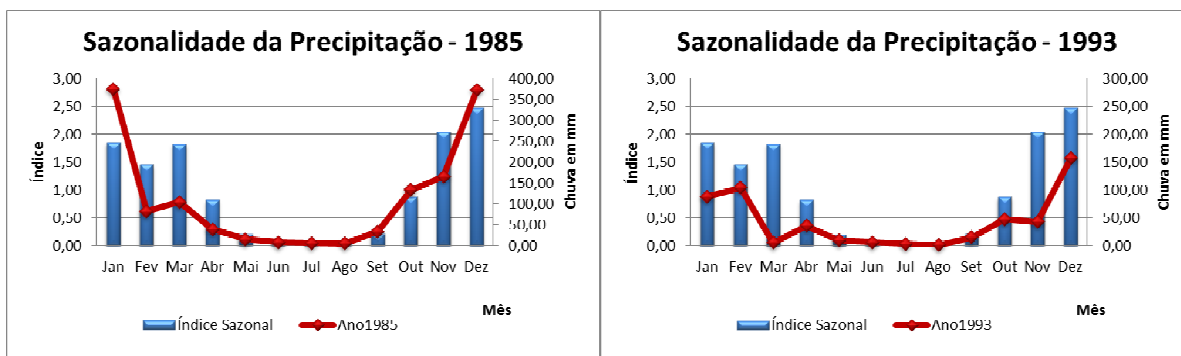
Nota-se no Gráfico 6, referente ao ano de 1993 deste mesmo cluster, que a curva da precipitação, apesar de acompanhar o movimento da sazonalidade apresenta queda em meses do período chuvoso para região (novembro e março), principalmente para o mês de março, registrando um acumulado pluviométrico de 3,8mm, quando a média esperada é de 99,4mm.



**Gráfico 6:** Sazonalidade da precipitação dos anos de 1985 e 1993 – *Cluster* 1.

O Gráfico 7 relaciona os anos de 1985 (mais chuvoso) e 1993 (mais seco) com a sazonalidade para o *cluster* 2. A curva da precipitação deste ano acompanha o movimento da sazonalidade do período, no entanto verifica-se que o volume pluvial alcançou valores máximos de aproximadamente 400mm em dois meses seguidos (dezembro e janeiro) da estação chuvosa.

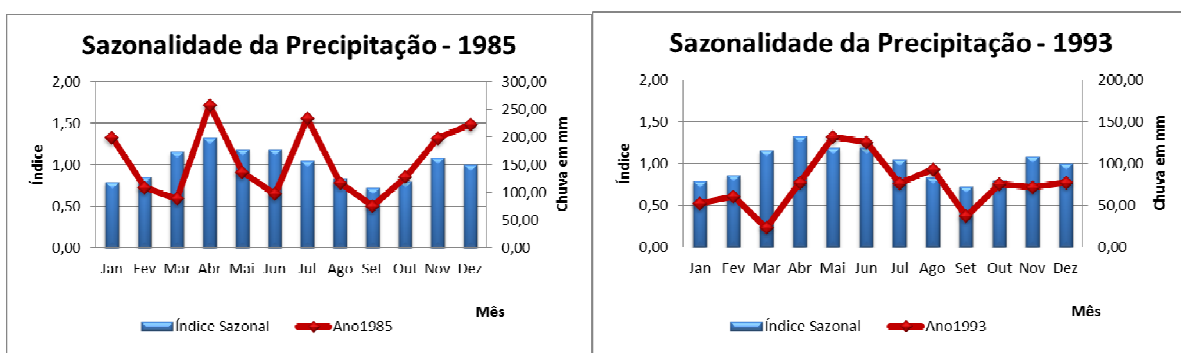
Para o ano de 1993 (mais seco) foram registrados volumes pluviais máximos de 150mm no mês de dezembro, concordando com o mês mais chuvoso da estação chuvosa desse *cluster*. No entanto a estação chuvosa deste ano foi comprometida com queda pluviométrica principalmente nos meses de novembro e março, destaque para o mês de março com volumes totais de 5mm (Gráfico 7).



**Gráfico 7:** Sazonalidade da precipitação dos anos de 1985 e 1993 – *Cluster 2*.

A análise do Gráfico 8 refere-se à sazonalidade do *cluster 3* e os anos identificados como mais chuvoso e mais seco. Observa-se uma variabilidade entre a curva da precipitação do ano de 1985 e a distribuição sazonal da série, principalmente nos meses de julho e dezembro. Verificou-se um aumento no volume pluviométrico nestes meses quando o esperado seria uma diminuição da precipitação. Além disso, o volume pluviométrico para este ano em todos os meses foi superior ao registrado na sazonalidade da série histórica deste *cluster*.

Observa-se um cenário oposto para os registros pluviométricos do ano de 1993 (mais seco) com o volume mensal acumulado abaixo do registrado na sazonalidade da série. Com o agravante da redução da precipitação no mês de março, pertencente à estação chuvosa.



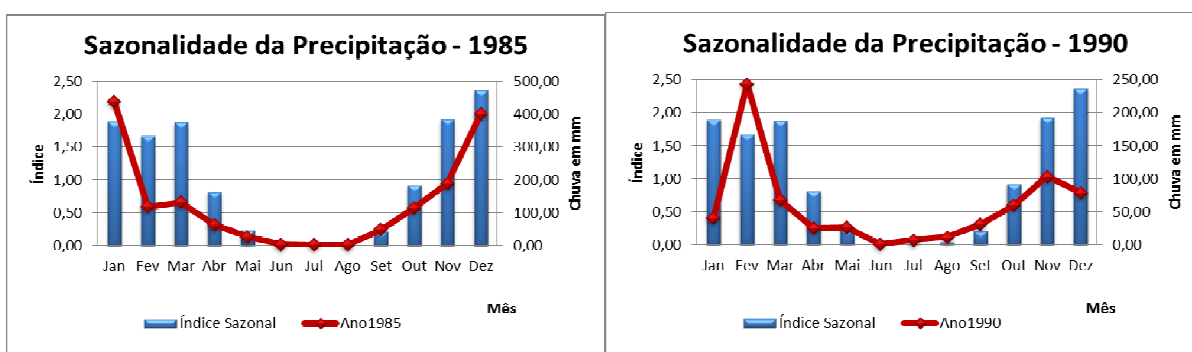
**Gráfico 8:** Sazonalidade da precipitação dos anos de 1985 e 1993 – *Cluster 3*.

O Gráfico 9 representa a relação entre a sazonalidade e a distribuição pluvial para os anos de destaque como o mais chuvoso (1985) e o mais seco (1990) para o *cluster 4*. Verifica-se que ocorre uma variabilidade em 1990 com relação à sazonalidade registrada para o período e uma semelhança no comportamento da curva para o ano de 1985.



Os registros de variabilidade da precipitação do ano de 1990 com a sazonalidade para o período é visto em quase todos os meses. Os baixos volumes pluviométricos registrados, principalmente para os meses da estação chuvosa proporcionaram a este ano a classificação do mais seco de toda a série.

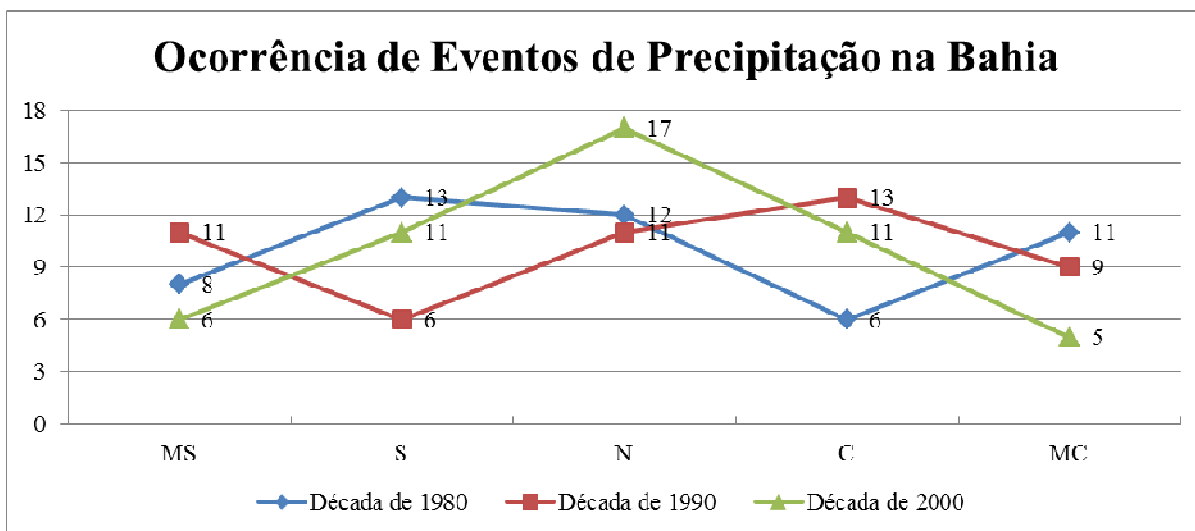
Apesar da semelhança com o comportamento da distribuição sazonal, o ano de 1985 (mais chuvoso) registrou o maior volume pluviométrico mensal de todo o estado da Bahia, alcançando o volume de 437,40mm no mês de janeiro. Um percentual de quase 100% em relação ao índice sazonal para este *cluster*.



**Gráfico 9:** Sazonalidade da precipitação dos anos de 1985 e 1990 – *Cluster 4*.

Com o objetivo identificar a frequência de ocorrência de eventos extremos “muito secos” e “muito chuvosos” no estado, visando identificar um possível aumento ou redução destes, dividiu-se a série em três décadas, a saber: 1981 a 1990; 1991 a 2000, 2001 a 2010.

O Gráfico 10 mostra a distribuição da ocorrência dos eventos classificados como muito seco, seco, normal, chuvoso e muito chuvoso para as três décadas, no estado. Os eventos referentes às categorias (muito seco, seco, normal, chuvoso e muito chuvoso) totalizam 50 eventos por década (10 anos vezes 5 *clusters*).



**Gráfico 10:** Distribuição de Eventos de Chuva por década, na Bahia. Tem-se para MS = Muito Seco; S = Seco; N = Normal; C = Chuvoso; MC = Muito Chuvoso.

No Gráfico 10 observa-se que na década de 1980 houve a maior ocorrência de eventos extremos “muito chuvosos” em relação às demais, com 22%, dos eventos para esta década, com destaque para os anos de 1985 e 1989 presentes em todos os clusters. Verifica-se também, que apesar disso, o número de ocorrências de anos “secos” foi o maior comparado com as outras décadas, com 26% dos eventos dessa década. E o segundo maior de anos classificados como “normais”, com 24%.

A década de 1990 revela um cenário oposto quando comparado com a década de 1980. Tem-se o período de maior concentração de eventos “muito secos”, apresentando 22% dos eventos para essa década e os eventos “chuvosos” com 26%, como aqueles de maior ocorrência nessa década. A quantidade de anos classificados como “normais” foi a mais baixa comparada com as outras décadas, atingindo o percentual de 22% dos eventos dessa década.

Ao analisar a década de 2000, nota-se que esta é a de maior ocorrência de eventos “normais”, com 34% de ocorrência em relação aos demais eventos para esta década, e a maior em relação aos anos classificados como “normais” em relação às demais décadas. O número de eventos “secos” e “chuvosos” é semelhante em relação à mesma década, isto é, em torno de com 22%. Foi verificado o menor número de eventos extremos “muito chuvosos” e “muito secos” com relação às demais décadas.

Ao comparar as três décadas em estudo é possível inferir que a quantidade de eventos extremos “muito secos” é maior na década de 1990, a categoria de eventos “muito chuvosos” é maior na década de 1980 e, na década de 2000, foi registrada a superioridade para os eventos “normais”.

Partindo-se do estudo com os anos classificados como “muito secos” e “muito chuvosos”, prevaleceu a década de 1990 com mais anos com eventos extremos “muito secos”. Os anos classificados como “muito chuvosos” foram 1985, 1988, 1989, 1992, 1995, 1999, 2000, 2004, 2005 e 2009. Destes, a década de 2000 foi a que apresentou o maior número de anos com estes eventos.

Também para região Nordeste, para os estados do Rio Grande do Norte e Paraíba, Santos e Brito (2007) utilizaram dados pluviométricos diários do período de 1935 a 2000, dados de IDVN de 1980 a 2000 e dados de anomalias de TSM para o período de 1950 a 2000, e encontraram resultados semelhantes ao correlacionar os índices de eventos extremos com as anomalias da temperatura de superfície do mar (TSM), o que revelou uma tendência de aumento da precipitação total anual.

Os registros do Relatório Especial de Eventos Extremos do IPCC também divulgaram um aumento de eventos extremos positivos de precipitação em algumas áreas da região Nordeste do Brasil, ocorridos no período de 1961 a 2000. O mesmo relatório traz projeções para o período de 2071 a 2100 com os mesmos resultados (IPCC, 2012).

As informações sobre a dinâmica das chuvas para o período entre 1981 a 2010 no estado da Bahia e a classificação dos anos com ocorrência de eventos extremos (muito secos e muito chuvosos) foram comparadas com os registros dos sistemas oceânicos (El Niño e La Niña), atuante para o período em análise (Tabela 13).

Ao analisar a Tabela 13, observa-se que o sistema oceânico mais atuante para o período em análise foi o El Niño, variando entre fraco (com cinco ocorrências), moderado (com duas ocorrências) e forte (com três ocorrências), de acordo com NOAA (2012). Com menos ocorrências, o fenômeno La Niña também variou em intensidade de fraco a forte. Tal observação é de grande importância, devido à forte influência dos fenômenos de El Niño nos eventos de secas severas na região Nordeste do Brasil (SANTOS *et al.*, 2012).

**Tabela 13:** Relação dos anos de ocorrência e intensidade dos fenômenos de El Niño e La Niña.

<b>Ano</b>	<b>Episódio</b>	<b>Intensidade</b>
1979 - 1980	El Niño	Fraco
1982 - 1983	El Niño	Forte
1983 - 1984	La Niña	Fraco
1984 - 1985	La Niña	Fraco
1986 - 1988	El Niño	Moderado
1988 - 1989	La Niña	Forte
1990 - 1993	El Niño	Forte
1994 - 1995	El Niño	Fraco
1995 - 1996	La Niña	Moderado
1997 - 1998	El Niño	Forte
1998 - 2001	La Niña	Forte
2002 - 2003	El Niño	Moderado
2004 - 2005	El Niño	Fraco
2006 - 2007	El Niño	Fraco
2009 - 2010	El Niño	Fraco

Os anos classificados com ocorrências de eventos “muito secos” para todo estado da Bahia (1982; 1984; 1987; 1990; 1993; 1996; 1998 e 2003) estão em quase sua totalidade associados à ocorrência do fenômeno El Niño, como por exemplo, os episódios nos anos de 1982, 1987, 1990, 1993, 1998 e 2003. A década de 90, registrada como de maior ocorrência dos eventos secos, foi também a década com maior ocorrência de fenômeno El Niño, com intensidade forte, para o período analisado.

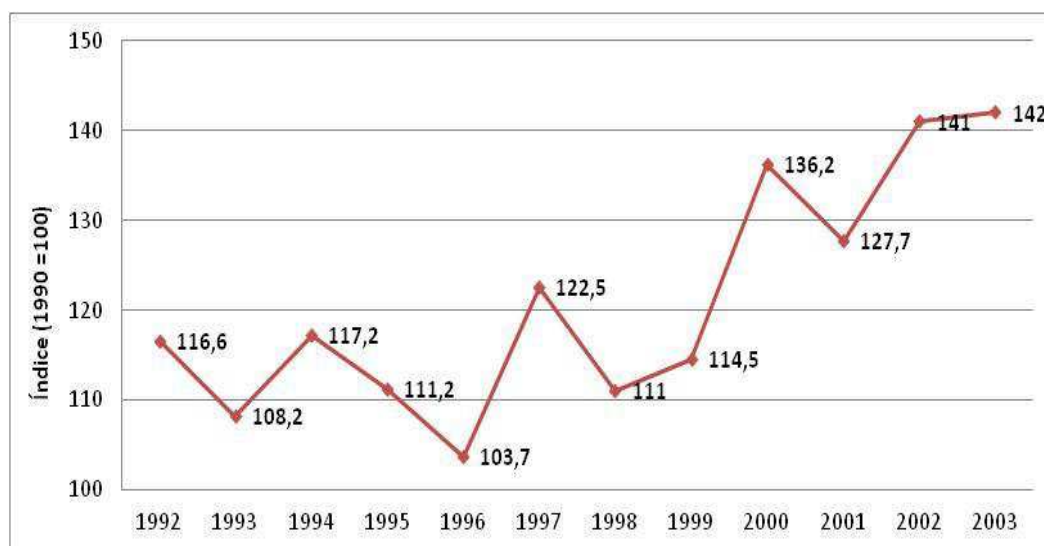
Resultados semelhantes da associação do El Niño com eventos de seca no Nordeste também foram encontrados por Xavier (2003) para o Ceará e Monteiro (2012) para o Rio Grande do Norte.

Diniz *et al.* (2011) analisaram uma série pluviométrica de 10 anos (1996 a 2005) do município de Feira de Santana na Bahia, verificando correlação entre os anos de seca da região e os anos de El Niño. Os autores atribuíram a queda na safra de feijão, milho e mandioca, devido à grande seca que houve em 1997/98, ao fenômeno El Niño de intensidade forte ocorrido nesse período.

Referente aos anos classificados como “muito chuvosos” (1985, 1988, 1989, 1992, 1995, 1999, 2000, 2004, 2005 e 2009), nota-se que os mesmos coincidem em quase sua totalidade com os anos de ocorrência de La Niña (1985; 1988, 1989, 1995, 1999, 2000 e 2001).

Tais resultados são coerentes com a atuação do fenômeno para região Nordeste, que favorece a intensificação das chuvas na região.

Os registros da Superintendência de Estudos e Informações da Bahia, divulgados em relatórios “Bahia em Números” apresenta um gráfico de produtividade do estado e atribui aos anos com queda na produtividade agrícola aos anos (1990, 1993, 1996, 1998, 2001 e 2003) acometidos por adversidades climáticas. Ao comparar o cenário apresentado no Gráfico 11 (SEI, 2010), observa-se que o termo adversidade climática usado pelo órgão do estado da Bahia se refere diretamente aos eventos extremos de seca, sendo o principal fator climático para redução da produtividade da região.



**Gráfico 11:** Produtividade agrícola do estado da Bahia. Fonte: SEI, 2010.

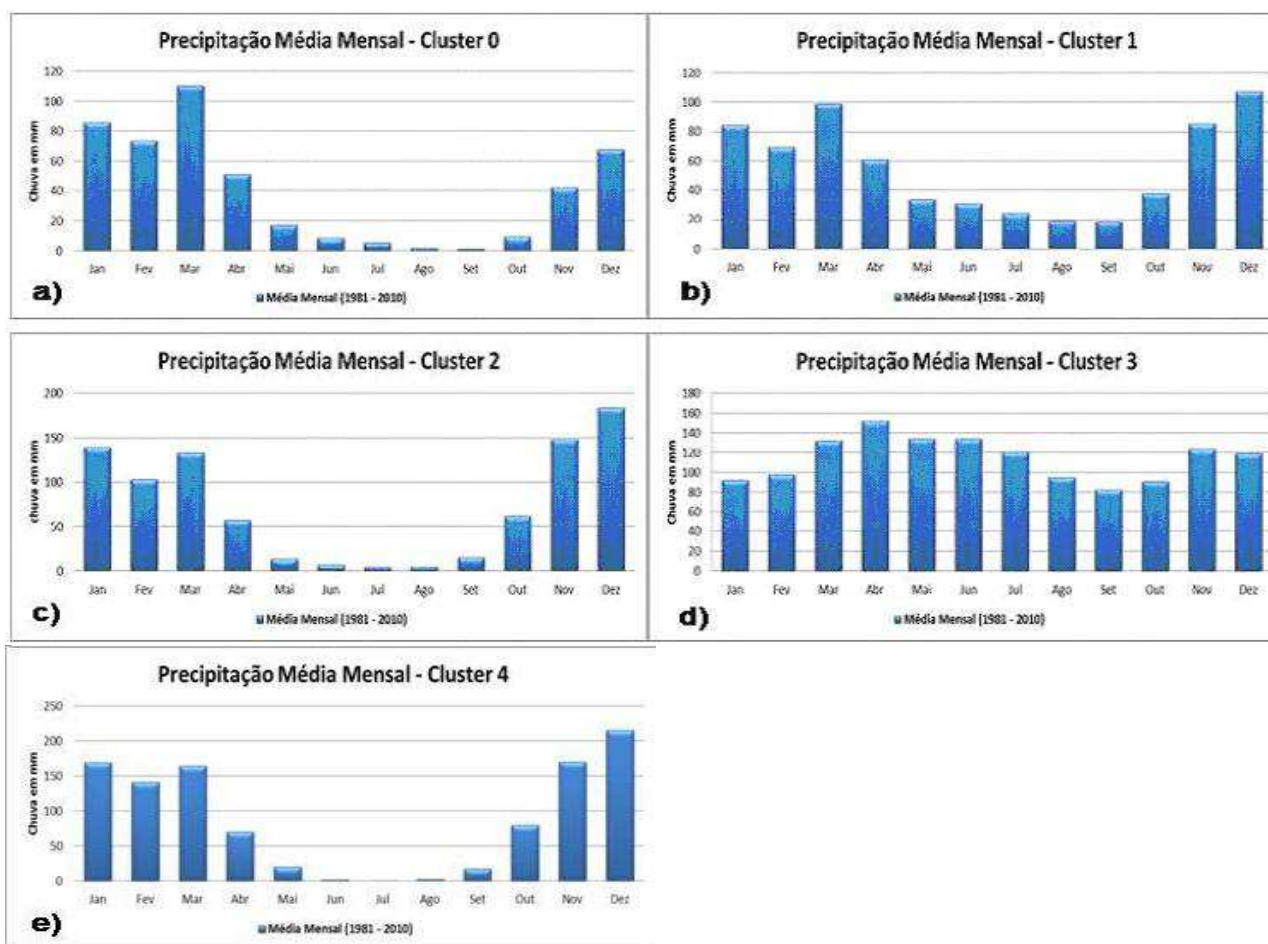
Verifica-se no Gráfico 11 e nos relatos da SEI (2010) que os impactos causados na agricultura pelos eventos de seca, apresentam queda de -7,3% em 1993, -5,1% em 1995, -6,7% em 1996, -9,4 em 1998 e -6,2% em 2001, destacando as perdas mais expressivas nas safras do feijão, mandioca, cana-de-açúcar e cacau, que estão entre os dez principais produtos no *ranking* de produção agrícola baiano.

O ano de 2003, também classificado como “seco”, apresentou queda na produtividade dos principais produtos agrícolas. No entanto, constatou-se um leve incremento na produção, propiciado pelos resultados positivos das culturas de algodão, milho, mamona,

batata inglesa e soja, devido à expansão nas áreas plantadas e nos avanços dos sistemas de irrigação na região oeste da Bahia (SEI, 2010).

#### 4.4.2 Análise Mensal

Os resultados da identificação de eventos extremos referentes à escala mensal correspondem à aplicação da técnica dos quantis, de ordem quantílica  $Q(0,85)$ , sobre os dados de precipitação pluviométrica dos meses que compõem a estação chuvosa de cada região pluviometricamente homogênea. Verifica-se no Gráfico 12 (a), (b), (c), (d) que os *clusters* apresentam épocas chuvosas diferentes. A diferença é observada nos aspectos quais e quantos meses compõem a época chuvosa de cada *cluster*, e a diversificação da distribuição das chuvas no estado da Bahia.



**Gráfico 12:** (a), (b), (c), (d), (e) – Distribuição da precipitação pluviométrica acumulada mensal para o período de 1981 – 2010.

Os resultados a respeito da distribuição pluviométrica de cada *cluster*, de acordo com a Figura 18, em consonância com a literatura especializada sobre os regimes de chuva da região (KOUSHY, 1979; CAVALCANTI *et al.*, 2009; TANAJURA *et al.*, 2010), fundamentaram a seleção dos meses pertencentes à estação chuvosa de cada *cluster*.

Tem-se como estação chuvosa para o *cluster 0* os meses de dezembro a abril, totalizando 150 meses na série de trinta anos; *cluster 1*, de outubro/novembro a abril, com 210 meses; *cluster 2*, de outubro a abril, com 210 meses; *cluster 3*, de março a julho e novembro, dezembro, com 210 meses; e para o *cluster 4*, de outubro a abril, com 210 meses.

Com base nos valores da Tabela 14, observa-se a diferença entre os valores de limiares de eventos extremos, tanto para classificação de “muito seco” quanto para “muito chuvoso”. Verifica-se que o menor limite entre os *clusters* para ser classificado como um evento extremo “muito seco” é menor ou igual a 18,71mm (*cluster 0*, com média de 77,43mm). Já o maior limite entre os *clusters* para ser considerado como “muito seco”, encontra-se no *cluster 3*, cujo limiar é menor ou igual a 79,25mm.

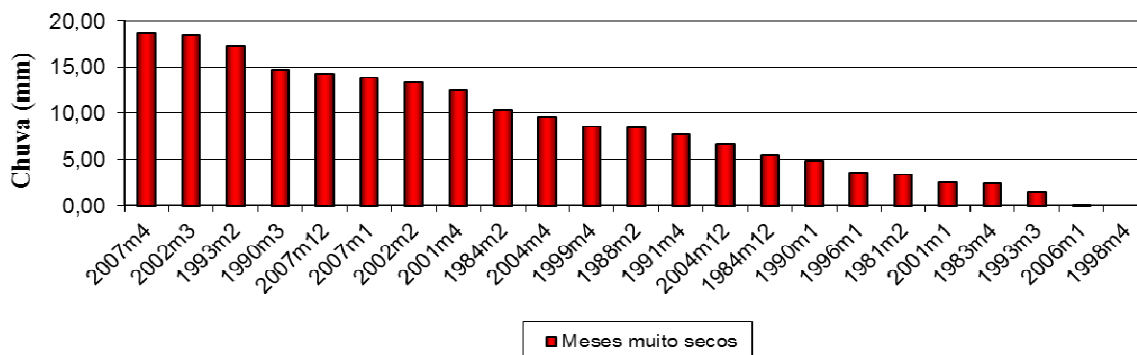
Para o cenário de eventos extremos “muito chuvosos”, constata-se que o menor limiar de precipitação entre os *clusters* é 133,03mm, pertencente ao *cluster 1* e o maior é 238,23mm, pertencente ao *cluster 4*. Mesmo sendo o *cluster 3* o que apresenta o maior volume pluviométrico em todo estado, no aspecto de volume pluviométrico de período chuvoso o *cluster 4* se destaca com o maior volume, com média de 144,86mm comparada a 131,07mm do *cluster 3*.

**Tabela 14:** Definição de limiares de chuva para o estado da Bahia.

Cluster	Limiares de Eventos Extremos	
	Muito Seco Q(0,15)	Muito Chuvoso Q(0,85)
0	$P \leq 18,70$	$P \geq 145,60$
1	$P \leq 21,80$	$P \geq 133,00$
2	$P \leq 37,80$	$P \geq 199,60$
3	$P \leq 79,00$	$P \geq 184,00$
4	$P \leq 57,60$	$P \geq 238,20$

Os resultados da identificação dos meses com ocorrência de eventos extremos “muito secos” e “muito chuvosos” em seus respectivos *clusters* estão apresentados nos Gráficos 13 e

14 para o *cluster* 0, Gráficos 15 e 16 para o *cluster* 1, 17 e 18 para o *cluster* 2, 19 e 20 para o *cluster* 3, 21 e 22 para o *cluster* 4.



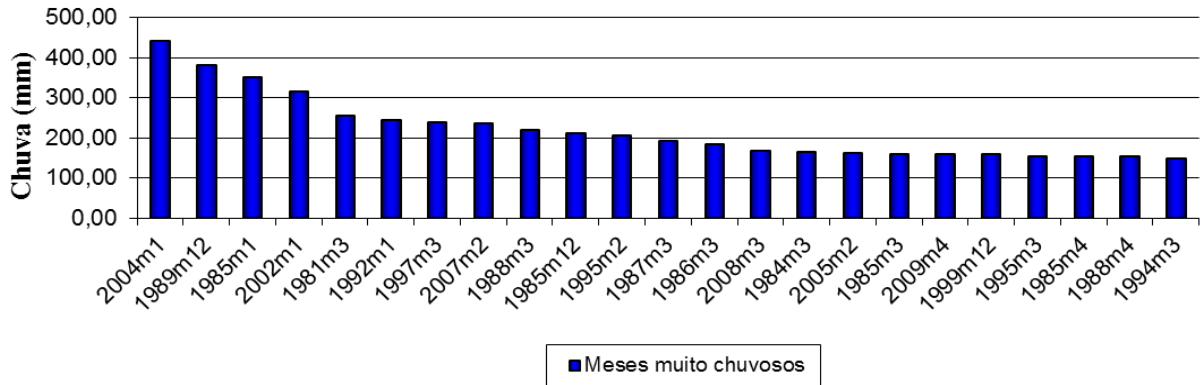
**Gráfico 13:** Acumulado de chuva mensal para os meses classificados como muito secos no *cluster* 0.

Para a estação chuvosa do *cluster* 0 foram identificados os meses com seus respectivos anos que foram classificados como “muito secos”. Observa-se que a maior incidência de eventos secos ocorre no mês de abril (com 31%), considerado como mês de transição da estação chuvosa para seca. A segunda maior ocorrência (22%) concentra-se no mês de janeiro, o segundo mês da estação chuvosa, o que pode causar danos severos à agricultura da região (Gráfico 13).

Nota-se que o mês de abril do ano de 2007 possui o maior volume pluvial dessa abordagem, e o mês de abril do ano de 1998 é o mês mais seco de toda série de 30 anos de dados pluviométricos. Segundo o Boletim Climanálise do CPTEC (2012), no mês de abril de 1998 houve escassez de chuva em todo Nordeste, atingindo principalmente o semiárido. As regiões mais afetadas desde o início da estação chuvosa foram o nordeste da Bahia, o sertão de Alagoas, Pernambuco, Paraíba e o centro-sul do Maranhão. Na Bacia do Rio São Francisco, as precipitações também ficaram abaixo da média em todos os postos hidrométricos.

O Gráfico 14 mostra os meses com ocorrência de eventos extremos “muito chuvosos” para o *cluster* 0. A máxima precipitação acumulada mensal nesse período alcança 440,70mm, registrada em janeiro de 2004. Tem-se o mês março como o mês mais chuvoso desse período (com 44% das ocorrências de eventos muito chuvosos). Resultados similares foram encontrados por Barbosa (2000) ao analisar os eventos de seca na Bahia.

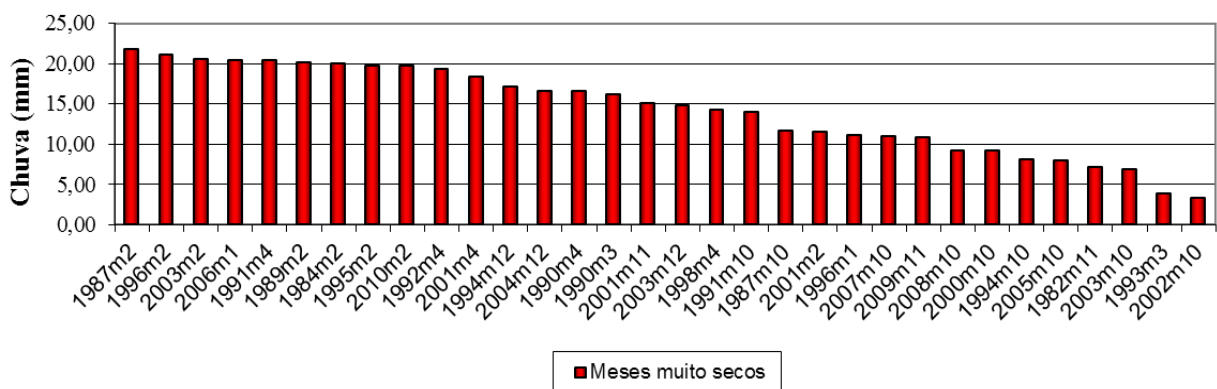




**Gráfico 14:** Acumulado de chuva mensal para os meses classificados como muito chuvosos - *cluster 0*.

Destaca-se o mês de janeiro de 2004 como o mês mais chuvoso da série de 1981 a 2010. De acordo com o CPTEC (2012), as chuvas intensas que ocorreram em toda a Região foram devidas ao deslocamento de frentes frias para o norte, à atuação da Zona de Convergência Intertropical (ZCIT) e à presença de Vórtices Ciclônicos em Altos Níveis (VCAN) sobre o Oceano Atlântico. Em algumas localidades, os máximos de precipitação excederam em mais que 300 mm a média climatológica do mês. Na cidade de Barra, na Bahia, choveu 280 mm (a média do mês é 108 mm).

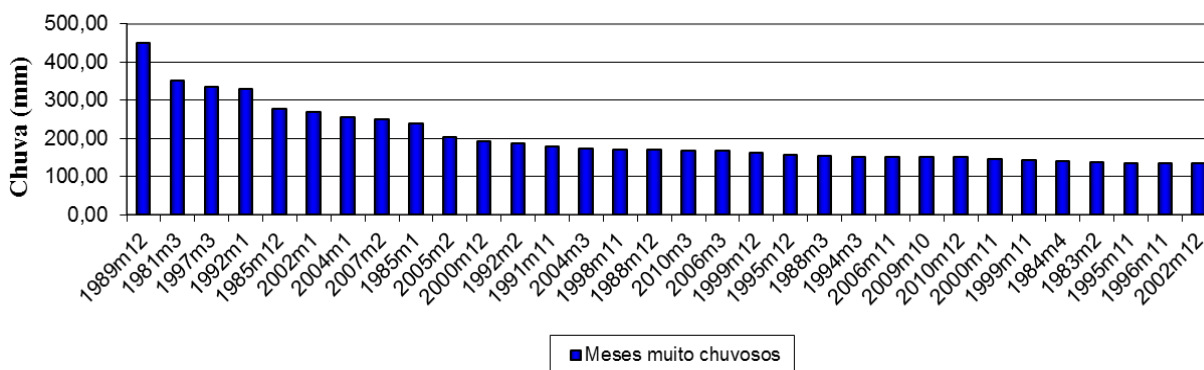
O *cluster 1*, representante de grande parte da região do semiárido nordestino, também apresenta baixos volumes de chuva. Os meses pertencentes à quadra chuvosa que foram classificados como “muito secos” apresentam volumes mensais abaixo de 80mm, variando até 3,30mm (Gráfico 15).



**Gráfico 15:** Acumulado de chuva mensal para os meses classificados como muito secos - *cluster 1*.

Identificado como o mais seco dessa série para o *cluster 1*, o mês de outubro de 2010 foi um mês de estiagem em grande parte da região Nordeste. A presença de vórtices e cavados em altos níveis foi desfavorável à ocorrência de chuvas, em particular sobre o estado da Bahia, onde os desvios negativos excederam os 50 mm (CPTEC, 2012).

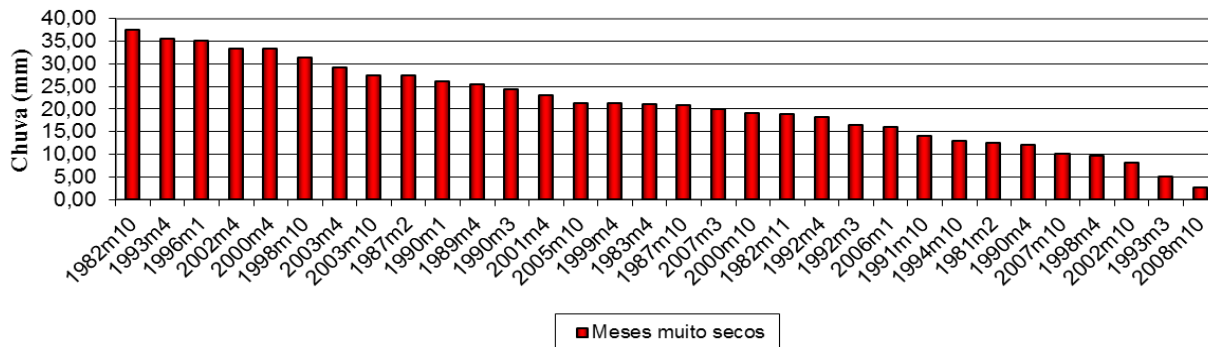
Para este mesmo cluster, o Gráfico 16 revela que os meses classificados como “muito chuvosos” variaram entre o máximo de 447,50mm em dezembro de 1989 a 133,50mm em dezembro de 2002. Observa-se a maior ocorrência de eventos “muito chuvosos” no mês de dezembro, com 28% de frequência na série. Tal constatação também foi vista por Barbosa (2000) que detectou o mês de dezembro como o mais chuvoso para essa região.



**Gráfico 16:** Acumulado de chuva mensal para os meses classificados como muito chuvosos - *cluster 1*.

No mês de março de 1997, foram observados desvios positivos de até 500 mm acima da normal climatológica, sendo registrados totais mensais de 600 mm. Esse grande volume de chuva concentrado nesse mês foi causado pela entrada de dois sistemas frontais no sul da Bahia, que se acoplaram à ZCAS, causando chuvas no interior e no sul da região (CPTEC, 2012).

A análise do Gráfico 17, referente ao *cluster 2*, mostra um aumento no volume pluviométrico para os meses considerados como “muito secos” em comparação com os clusters 0 e 1 (pertencentes à região semiárida). Apresenta uma variação pluviométrica de 37,50mm (outubro de 1982) a 2,60mm (outubro de 2008). A maior concentração de meses “muito secos” ocorre nos meses de outubro com 34% de frequência e abril também com 34%.

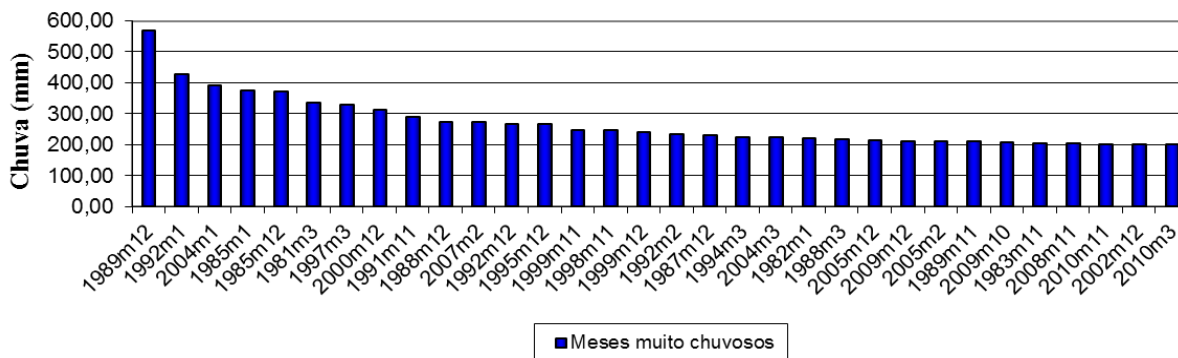


**Gráfico 17:** Acumulado de chuva mensal para os meses classificados como muito secos - *cluster 2*.

De acordo com o Boletim Climanalise, este período costuma ser de estiagem na maior parte da Região. Por esta razão, as chuvas são climatologicamente reduzidas e ocorrem baixos valores de umidade relativa do ar. No dia 5 de outubro de 2006, a umidade do ar atingiu 10% em Bom Jesus da Lapa-BA. A atuação de cavados e vórtices ciclônicos em altos níveis também contribuiu para a diminuição das chuvas nessa região (CPTEC, 2012).

Ao analisar o Gráfico 18, sobre os meses classificados como “muito chuvosos” para o *cluster 2*, nota-se que esse período alcança totais pluviais de até 562,30mm (dezembro de 1989) variando para 199,70mm (março de 2008). O mês com maiores ocorrências de eventos extremos positivos é o mês de dezembro, com 34% de frequência.

Verifica-se ao comparar os dados apresentados nos Gráficos 14 e 15, referentes aos eventos “muito secos” e “muito chuvosos”, que para os meses da estação chuvosa deste *cluster*, há uma diferença pluviométrica de 559,70mm. Foram registrados 562,30mm no mês mais chuvoso e 2,60mm no mês mais seco.



**Gráfico 18:** Acumulado de chuva mensal para os meses classificados como muito chuvosos - *cluster 2*.

De acordo com os registros divulgados pelo CPTEC (2012), as chuvas intensas que ocorreram em toda a região foram devidas ao deslocamento de frentes frias para o norte e a forte atuação da Zona de Convergência Intertropical (ZCIT). Na cidade de Barra, na Bahia, choveu 280 mm, a média do mês é de 108 mm.

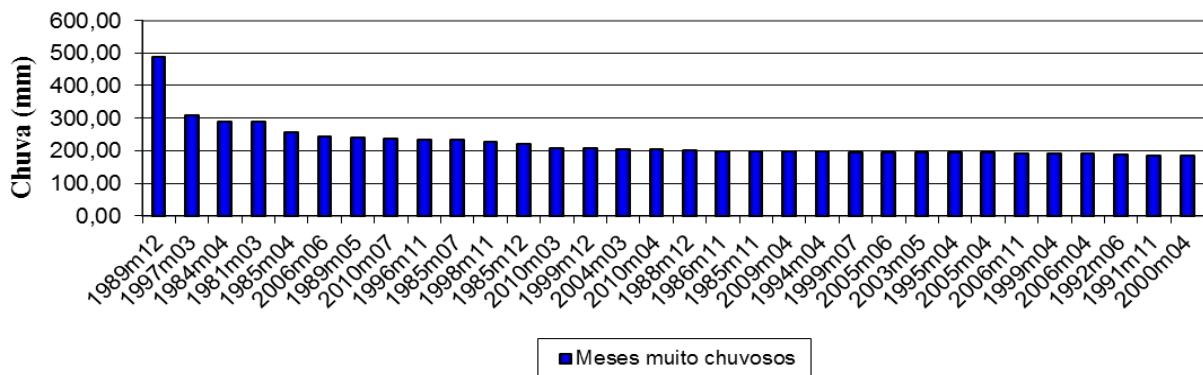
Os baixos índices pluviométricos registrados na quadra chuvosa da região do litoral da Bahia (*cluster 3*) variam entre 79,30mm e 23mm. Analisando o Gráfico 19, constata-se que há baixa variabilidade entre a distribuição pluvial na quadra chuvosa, característica do regime pluvial local, uniformidade e regularidade de chuvas, de acordo com Barbosa (2000). O evento registrado como mais seco ocorreu em março de 1993 (Gráfico 19).



**Gráfico 19:** Acumulado de chuva mensal para os meses classificados como muito secos - *cluster 3*.

Relatos do CPTEC (2012) atribuem a escassez de chuva no mês de dezembro de 2009 à persistência de cavados e vórtices ciclônicos em altos níveis, também contribuindo para a ocorrência de chuvas abaixo da média na maior parte da Bahia. Cinco episódios de Vórtices Ciclônicos em Altos Níveis (VCAN) ocorreram no decorrer do mês de dezembro. O VCAN reduziu a nebulosidade sobre grande parte do Nordeste e atuou em conjunto com o segundo episódio de ZCAS, associado, por sua vez, ao sistema frontal à superfície.

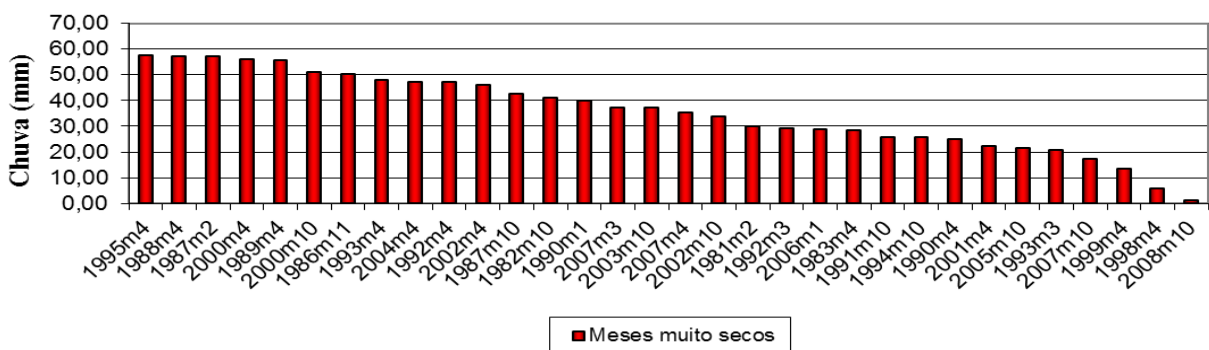
Observa-se no Gráfico 20 os meses classificados como “muito chuvosos” do cluster 3, com um destaque para o mês de dezembro de 1989, que apresentou volumes de chuva de 489mm, destoando dos demais meses de toda a série. O ano de 1989 foi registrado como o ano mais chuvoso do *cluster 3*.



**Gráfico 20:** Acumulado de chuva mensal para os meses classificados como muito chuvosos - *cluster 3*.

Durante o mês de março de 1997 (segundo maior registro de evento muito chuvoso), as chuvas intensas foram causadas pela entrada de dois sistemas frontais no sul da Bahia, que se acoplaram à ZCAS, causando chuvas no interior e no sul da região. Nos dias 09, 11 e 12 deste mesmo mês, ocorreram também distúrbios de leste, que atingiram o litoral do estado da Bahia (CPTEC, 2012).

O Gráfico 21 mostra uma variação nas chuvas dos meses classificados como “muito secos” do *cluster 4*. O volume total variou entre 57,40mm a 1,40mm. Os meses com maior frequência nessa análise são outubro e abril, os quais são meses de transição entre as estações secas e chuvosas, refletindo um atraso na estação chuvosa ou um término precoce do “período das águas”.



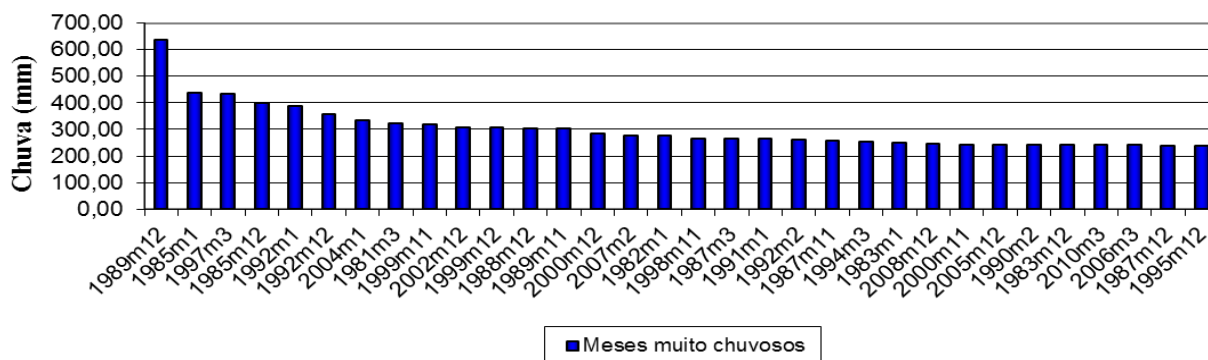
**Gráfico 21:** Acumulado de chuva mensal para os meses classificados como muito secos - *cluster 4*.

No mês considerado como mais seco da série estudada, outubro de 2008, o Centro de Previsão de Tempo e Estudos Climáticos afirmou que este período costuma ser de estiagem na maior parte da Região. Por esta razão, as chuvas são climatologicamente reduzidas e ocorrem

baixos valores de umidade relativa do ar. A atuação de cavados e vórtices ciclônicos em altos níveis também contribuiu para a diminuição das chuvas nessa região.

A configuração de Vórtices Ciclônicos em Altos Níveis (VCAN) foi notada em cinco episódios no decorrer do mês de outubro. Os episódios observados a partir do dia 15 de outubro de 2008 contribuíram para a ocorrência de anomalias negativas de precipitação no leste da Região Centro-Oeste atingindo o oeste da Bahia (CPTEC, 2012).

Dados do Gráfico 22 revelam que o mês mais chuvoso da estação chuvosa do *cluster* 4 é o mês de dezembro, com 37,50% de frequência. Destaca-se o evento ocorrido no mês de dezembro de 1989 com precipitação de 637,70mm, sobressaindo-se com relação aos volumes registrados nos demais meses desse período.



**Gráfico 22:** Acumulado de chuva mensal para os meses classificados como muito chuvosos - *cluster* 4.

Foi feita uma análise do mês de março de 1997 pelo CPTEC (2012), relatando que durante este mês, as chuvas intensas destacaram-se sobre o estado da Bahia, desvios positivos de até 500mm acima da normal climatológica (no centro deste estado), com totais mensais de 600mm. Estas precipitações foram causadas pela entrada de dois sistemas frontais no sul da Bahia, que se acoplaram à ZCAS, causando chuvas no interior e no sul da região.

Os resultados das análises dos gráficos referentes aos meses classificados como “muito secos” e “muito chuvosos” revelaram alta variabilidade das chuvas no período chuvoso de todos os *clusters*, comprometendo o comportamento pluvial desse período. As informações extraídas sobre a ocorrência de eventos extremos de chuva durante a estação chuvosa é de grande relevância principalmente para o setor agrícola, altamente dependente dos períodos

chuvosos, logo uma modificação na distribuição das chuvas ou uma redução no volume desta, é o suficiente para desorganizar toda a atividade econômica.

A partir da identificação dos meses com ocorrência de eventos extremos “muito secos” (negativos) e “muito chuvosos” (positivos), foi possível descobrir a sequência das ocorrências e com que frequência aconteceram os episódios de extremos de chuva na série histórica de 30 anos de dados pluviométricos do estado da Bahia.

Para isto, foi utilizado o algoritmo *Generalized Sequential Patterns* (GSP), em que o parâmetro “suporte” foi ajustado para 3%. Como resultado, foram detectados padrões nas ocorrências dos eventos extremos para cada *cluster*. Os padrões relacionam os eventos extremos e o mês em que ocorreram. Observa-se que os padrões encontrados diferem nos aspectos de ordem da sequência de eventos, quantidade de ocorrências, e meses de suas ocorrências, em cada região pluviometricamente homogênea.

Para os padrões de sequência de tamanho um, constatou-se que os meses que mais se destacaram com ocorrência de eventos negativos (muito secos) e positivos (muito chuvosos), coincidem com os meses classificados como os mais chuvosos e os mais secos da aplicação da técnica do quantil.

A Tabela 15 mostra que, de forma similar aos resultados da classificação de eventos extremos por quantil, os meses mais secos foram aqueles de transição entre as estações chuvosa e seca, enquanto o mês com maior ocorrência de evento positivo foi o mais chuvoso, também classificado pelo uso do quantil e referenciado pela literatura especializada nos regimes pluviais do estado da Bahia.

**Tabela 15:** Identificação dos meses com mais ocorrências de eventos extremos para uma sequência de tamanho um.

Clusters	Eventos Extremos Negativos		Eventos Extremos Positivos	
	Mês	Frequência (%)	Mês	Frequência (%)
0	Abril	23,33	Março	33,33
1	Outubro	30	Dez e Mar	23,33
2	Abr e Out	36,66	Dezembro	36,66
3	Dezembro	36,66	Abril	30
4	Abril	46,66	Dezembro	40

A Tabela 16 ilustra os padrões identificados pelo algoritmo GSP com sequências de tamanhos dois e três, e aqueles mais frequentes em toda série histórica de 1981 a 2010 para cada *cluster*. As sequências de tamanho três são encontradas nos *clusters* 2, 3 e 4 e a sequência de tamanho dois em todos os *clusters*.

Ao analisar os resultados com o uso do GSP, é confirmada a alta variabilidade de eventos negativos (secos) e positivos (chuvosos) existente no período chuvoso de todos os *clusters*. Verifica-se (Tabela 16) que os eventos positivos e negativos podem acontecer simultaneamente. Por exemplo, a sequência “EN m2, EP m3 > 3”, com frequência 3, significa que um evento negativo ocorreu no mês de fevereiro, seguido de um evento positivo no mês de março.

Por outro lado, também foram descobertos padrões do mesmo tipo evento em meses sequenciais ou não sequenciais. Por exemplo, a sequência “EN m2, EN m3 > 2” tem frequência 2, em que eventos negativos ocorreram nos meses de fevereiro e março. Já a sequência “EN m4, EN m12 > 2” tem frequência 2, em que eventos negativos ocorreram nos meses de abril e dezembro.

Observa-se que as maiores ocorrências de padrões encontrados para o *cluster* 1 foram um evento negativo em fevereiro, seguido de um evento positivo em dezembro (por três vezes); um evento negativo em outubro seguido de outro negativo no mês de dezembro; e um evento positivo em novembro seguido de outro positivo em dezembro, em que os dois últimos também ocorreram por três vezes.



**Tabela 16:** Sequências de eventos extremos que mais ocorreram na série temporal de 1981 a 2010.

Cluster 0	Cluster 1	Cluster 2	Cluster 3	Cluster 4
{{(EN m2)}{(EP m3)}> (3)}	{{(EN m2)}{(EP m12)}> (3)}	{{(EN m4)}{(EN m10)}{(EP m12)}> (2)}	{{(EN m3)}{(EN m11)}{(EN m12)}> (3)}	{{(EN m4)}{(EP m11)}{(EP m12)}> (3)}
{{(EN m2)}{(EN m3)}> (2)}	{{(EN m10)}{(EN m12)}>(3)}	{{(EN m4)}{(EP m11)}{(EP m12)}> (2)}	{{(EN m3)}{(EP m11)}{(EN m12)}> (2)}	{{(EP m1)}{(EN m4)}{(EP m12)}> (2)}
{{(EN m4)}{(EN m12)}>(2)}	{{(EP m11)}{(EP m12)}>(3)}	{{(EN m4)}{(EP m12)}> (5)}	{{(EN m3)}{(EN m7)}{(EN m12)}> (2)}	{{(EN m10)}{(EP m11)}{(EP m12)}> (2)}
{{(EN m1)}{(EN m4)}> (2)}	{{(EP m3)}{(EP m12)}> (2)}	{{(EN m10)}{(EP m12)}> (4)}	{{(EN m3)}{(EN m4)}{(EN m7)}> (2)}	{{(EN m4)}{(EN m10)}{(EP m12)}> (2)}
	{{(EP m3)}{(EN m12)}> (2)}	{{(EN m4)}{(EN m10)}> (4)}	{{(EP m4)}{(EN m11)}{(EN m12)}> (2)}	{{(EP m2)}{(EN m3)}{(EN m4)}> (2)}
	{{(EP m2)}{(EN m10)}> (2)}	{{(EN m4)}{(EP m11)}> (4)}	{{(EP m4)}{(EP m7)}{(EP m12)}> (2)}	{{(EN m4)}{(EP m12)}> (8)}
	{{(EN m2)}{(EN m10)}> (2)}	{{(EN m10)}{(EP m11)}> (3)}	{{(EP m4)}{(EN m7)}{(EP m11)}> (2)}	{{(EN m10)}{(EP m12)}> (5)}
	{{(EN m2)}{(EP m11)}> (2)}	{{(EN m3)}{(EN m4)}> (3)}	{{(EN m3)}{(EN m12)}> (5)}	{{(EN m4)}{(EP m11)}> (4)}
	{{(EN m10)}{(EP m11)}>(2)}	{{(EP m1)}{(EP m12)}> (2)}	{{(EN m11)}{(EN m12)}> (4)}	{{(EP m11)}{(EP m12)}> (4)}
	{{(EN m4)}{(EP m11)}> (2)}	{{(EP m11)}{(EP m12)}> (2)}	{{(EN m3)}{(EN m11)}> (3)}	{{(EP m1)}{(EP m12)}> (3)}
	{{(EN m1)}{(EP m11)}> (2)}	{{(EP m2)}{(EN m10)}> (2)}	{{(EN m3)}{(EN m7)}> (3)}	{{(EP m1)}{(EN m4)}> (3)}
		{{(EP m2)}{(EP m12)}> (2)}	{{(EP m4)}{(EN m12)}> (3)}	{{(EN m4)}{(EN m10)}> (3)}
		{{(EP m2)}{(EN m3)}> (2)}	{{(EP m4)}{(EP m7)}> (3)}	{{(EP m2)}{(EN m4)}> (3)}
			{{(EP m4)}{(EP m11)}> (3)}	{{(EN m3)}{(EN m4)}> (3)}
			{{(EP m4)}{(EP m12)}> (3)}	
			{{(EN m7)}{(EP m11)}> (3)}	
			{{(EP m3)}{(EN m6)}> (2)}	
			{{(EN m3)}{(EP m11)}> (2)}	
			{{(EP m4)}{(EN m11)}> (2)}	
			{{(EP m4)}{(EN m7)}> (2)}	
			{{(EP m4)}{(EN m5)}> (2)}	
			{{(EP m7)}{(EP m12)}> (2)}	
			{{(EP m11)}{(EN m12)}> (2)}	
			{{(EN m7)}{(EN m12)}> (2)}	
			{{(EN m4)}{(EN m7)}> (2)}	
			{{(EN m4)}{(EN m5)}> (2)}	

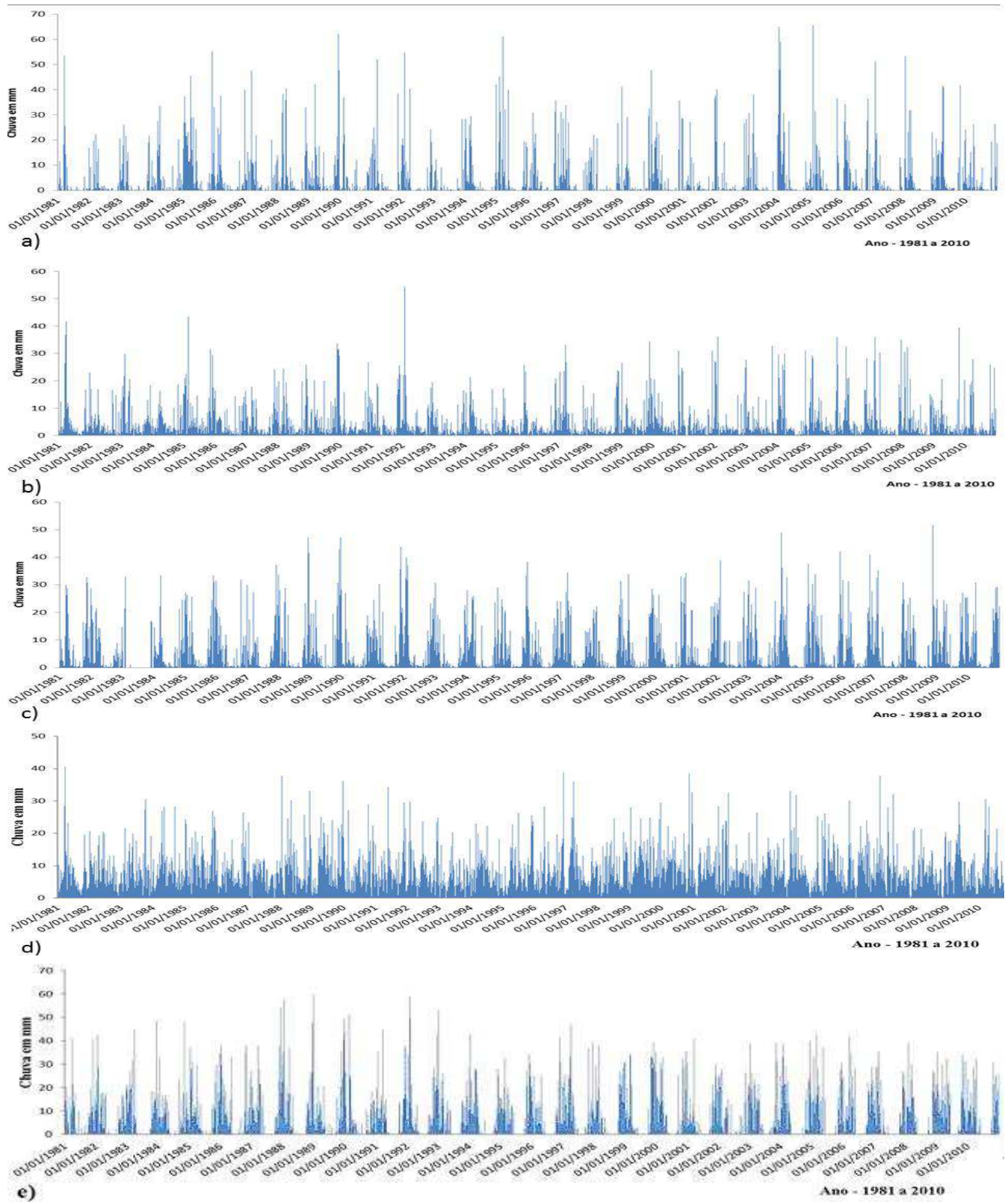
Ainda com relação à Tabela 16, o *cluster 2* apresentou padrões semelhantes aos do *cluster 4* para algumas sequências de tamanho 2. O padrão de maior ocorrência para ambos os *clusters* foi um evento negativo em abril e um positivo em dezembro. Mesmo para uma sequência de tamanho três, houve semelhança, pois ocorreu um evento negativo em abril, seguido de um positivo em novembro e outro positivo em dezembro.

O *cluster 3* foi o que mais apresentou sequência de tamanho três, com destaque para a sequência de um evento negativo no mês de março, seguido de outro negativo em novembro e em dezembro, com frequência igual a 3. As maiores ocorrências de sequências com tamanho dois, para este cluster, foi um evento extremo negativo no mês de março seguido de outro negativo no mês de dezembro.

Com base nos resultados dos padrões identificados pelo algoritmo GSP, para todos os *clusters*, constata-se que as maiores ocorrências de sequência de tamanho dois são para os eventos negativos ocorridos nos meses de transição das estações secas e chuvosas.

#### **4.4.3 Análise Diária**

A Figura 14: (a), (b), (c), (d) e (e), ilustra a distribuição do volume da precipitação diária referente ao período de 1981 a 2010, para cada região pluviometricamente homogênea. Pode-se visualizar, nos gráficos, picos de volumes extremos de precipitação e uma alta variabilidade no comportamento pluvial ao longo de toda série histórica. Tais eventos serão abordados com mais detalhamento nesta seção.



**Figura 14:** (a), (b), (c), (d) e (e) - Distribuição pluviométrica diária para os *clusters* 0, 1, 2, 3 e 4, respectivamente.

A Tabela 17 mostra a classificação da intensidade acumulada em 24 horas relacionada às ordens quantílicas, considerando para um evento extremo negativo o total acumulado de chuva abaixo de 2mm/dia, classificado pelo quantil  $Q(0,05)$ , e um evento extremo positivo quando a precipitação for maior que o volume pluviométrico definido pela ordem quantílica  $Q(0,95)$  indicada para cada *cluster*.

**Tabela 17:** Definição de limiares de eventos extremos para um acumulado em 24horas/dia.

Clusters	Limiares de Eventos Extremos	
	Muito Seco (mm)	Muito Chuvoso (mm)
<b>0</b>	P < 2,00	P > 31,60
<b>1</b>	P < 2,00	P > 21,0
<b>2</b>	P < 2,00	P > 24,50
<b>3</b>	P < 2,00	P > 16,00
<b>4</b>	P < 2,00	P > 27,00

Ao definir os limiares de intensidade de chuva em 24 horas, foram identificados os dias com ocorrência de eventos extremos. Para um total de 10.957 dias (em trinta anos), o *cluster* 0 apresentou 9.719 dias “muito secos” (precipitação abaixo de 2mm), com 89% dos dias classificados como eventos extremos negativos. Neste *cluster*, foram identificados 62 dias com eventos “muito chuvosos”, menos de 1% .

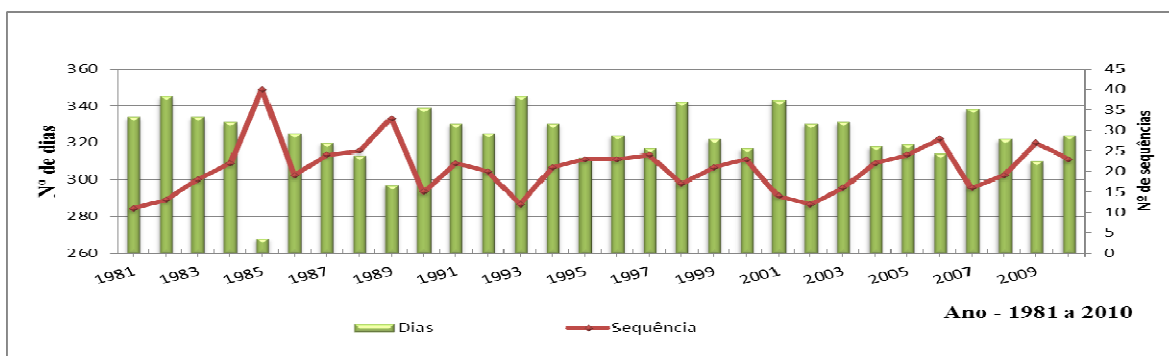
No *cluster* 1, foram identificados 8.575 dias com eventos extremos negativos, 78,26% de toda série histórica e 122 dias correspondendo aos dias classificados como eventos positivos, compondo 1% da série. Verifica-se que no *cluster* 2 foram identificados 8.377 dias “muito secos” (76,45%) e 130 dias “muito chuvosos” correspondendo a 1,20%. Para o *cluster* 3 foram detectados 5.138 dias “muito secos” (47%) e 277 dias “muito chuvosos”, representando 2,50% da série. Por fim, para o *cluster* 4, foram classificados 7.790 dias com eventos negativos (71%) e 160 dias “muito chuvosos”, correspondendo a 1,50%.

Resultados de alta concentração, em todos os *clusters*, dos dias classificados como “muito secos” também foram encontrados nos trabalhos de Souza *et al.* (2012), ao estudar a ocorrência de eventos extremos na cidade de Recife adotando como dia “muito seco” valores de precipitação diária inferior a 2mm. Os autores constataram para série de 47 anos no período de 1961 a 2008, que 62% correspondem aos dias secos.

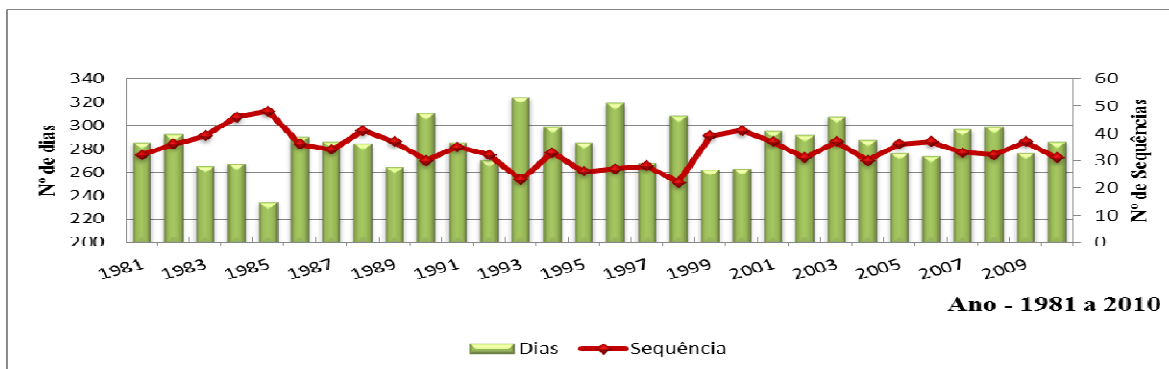
Após quantificar as ocorrências de eventos extremos negativos (muito secos), observa-se uma similaridade com o comportamento pluvial do estado da Bahia relatada por Barbosa (2000) e pela Superintendência de Estudos e Informações da Bahia (2010), onde o volume pluviométrico diminui das extremidades para o interior do território baiano, verificando uma maior ocorrência de eventos negativos no interior do estado.

Desta forma, encontra-se a maioria dos eventos negativos registrados nos *clusters* do centro do estado (*cluster 0* com 89% e *cluster 1* com 78%), seguido do *cluster 2*, com 76%, depois o *cluster 4*, na extremidade oeste do estado, com 71% e por fim o *cluster 3*, litoral baiano, com 47%.

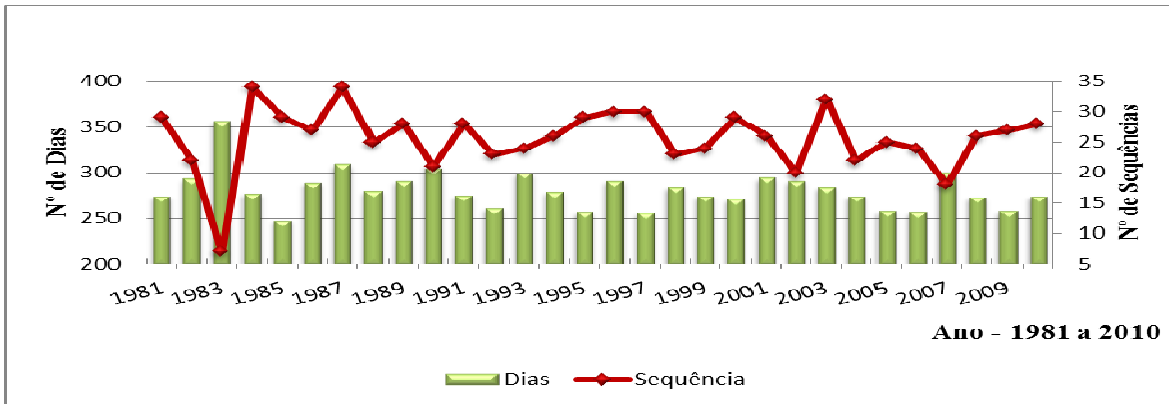
Os Gráficos 23, 24, 25, 26 e 27 indicam a quantidade de dias secos (< 2mm) ocorridos em cada ano da série histórica de 1981 a 2010, para todos os *clusters* do estado da Bahia. Verifica-se também nesses gráficos, o número de sequências dos registros de volumes pluviais menores que 2mm em cada ano.



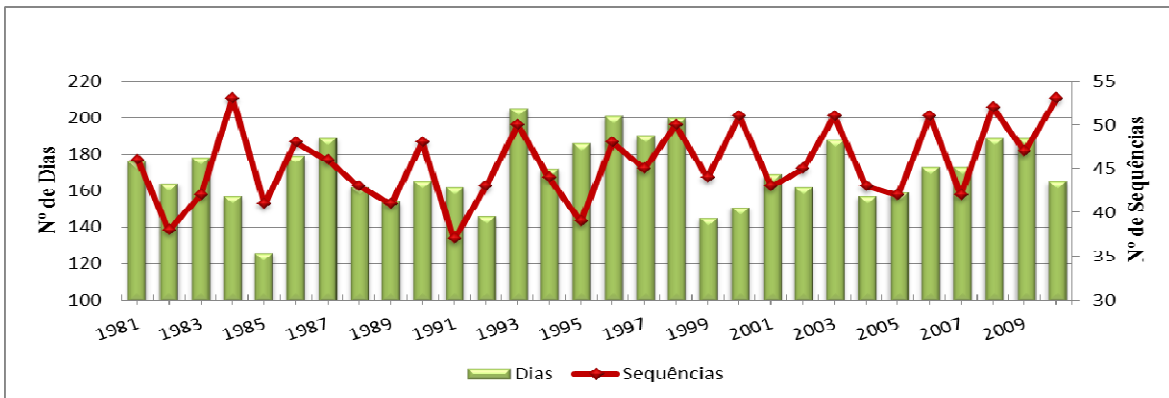
**Gráfico 23:** Distribuição do total de dias secos e número de sequências com dias secos consecutivos - 1981 a 2010 – *Cluster 0*.



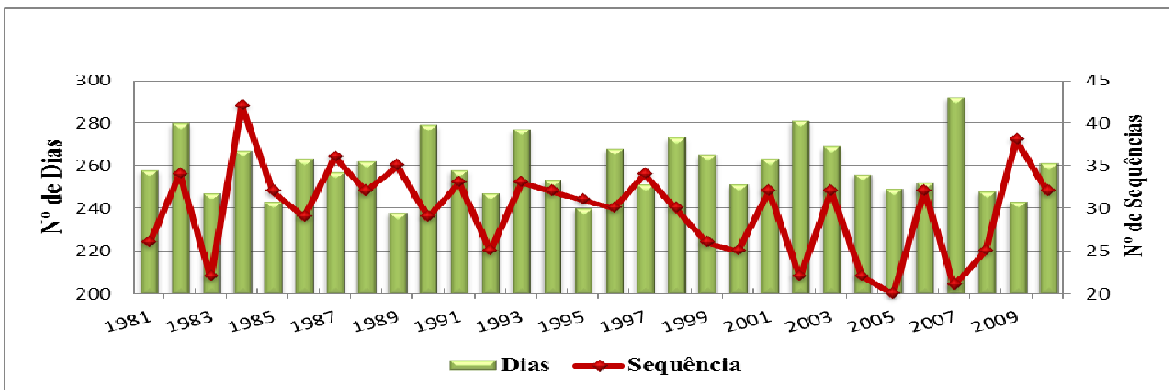
**Gráfico 24:** Distribuição do total de dias secos e número de sequências com dias secos consecutivos - 1981 a 2010 – *Cluster 1*.



**Gráfico 25:** Distribuição do total de dias secos e número de sequências com dias secos consecutivos - 1981 a 2010 – Cluster 2.



**Gráfico 26:** Distribuição do total de dias secos e número de sequências com dias secos consecutivos - 1981 a 2010 – Cluster 3.



**Gráfico 27:** Distribuição do total de dias secos e número de sequências com dias secos consecutivos - 1981 a 2010 – Cluster 4.

Verifica-se que os resultados da quantificação dos dias secos, em cada ano, coincidem com os resultados da classificação anual por meio da técnica dos quantis. Os anos considerados como de maior ocorrência de dias secos ( $< 2\text{mm}$ ) são similares aos anos identificados como “muito secos”, e os anos com menor ocorrência de dias secos foram semelhantes aos anos classificados como “muito chuvosos”. Resultados similares foram encontrados nos trabalhos de Barbosa (2000), sobre os registros das secas no estado.

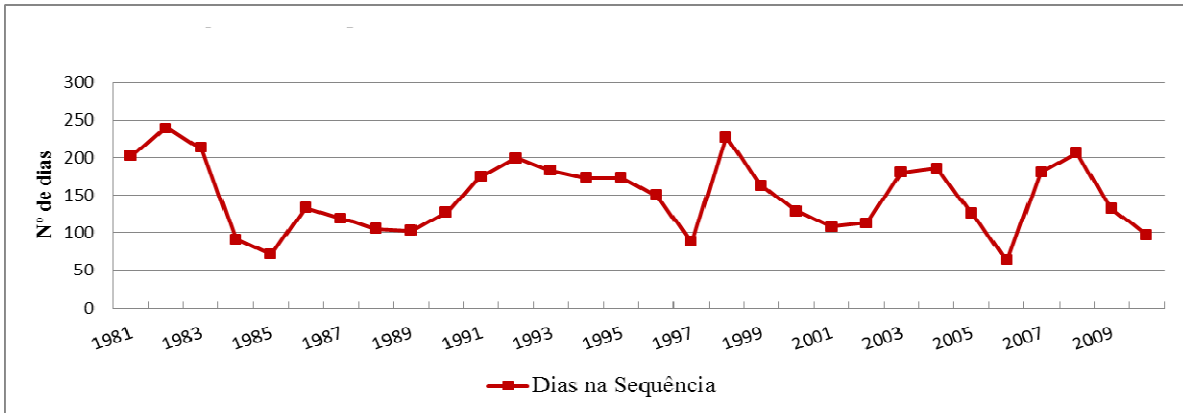
Não é possível afirmar que quanto maior o número de dias secos, maior é o número de sequências e vice-versa. No entanto, em alguns casos tende a um aumento no número de sequências quando o número de dias secos for menor.

Ainda com base nos gráficos acima, pode-se inferir que exceto para o *cluster 3*, todos os outros *clusters* apresentaram número de dias secos acima de 200, aproximadamente 55% dos anos foram registrados volumes pluviométricos abaixo de 2mm.

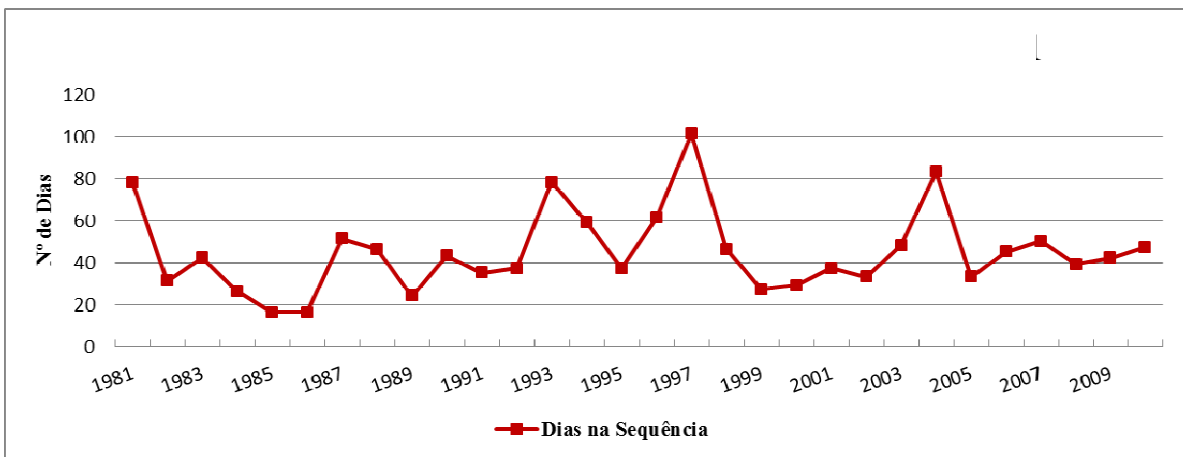
Os Gráficos 28, 29, 30, 31 e 32 mostram o número máximo de dias secos (precipitação menor que 2mm) consecutivos, registrado em cada ano da série histórica de cada zona pluviométrica da Bahia.

Os resultados obtidos com a quantificação do número máximo de dias secos consecutivos permitem fazer inferências sobre períodos de excessivas secas em cada região do estado. Resultados similares foram descritos por Dufek (2008), ao utilizar essa metodologia para identificar tendências positivas ou negativas de eventos extremos de precipitação para América do Sul.

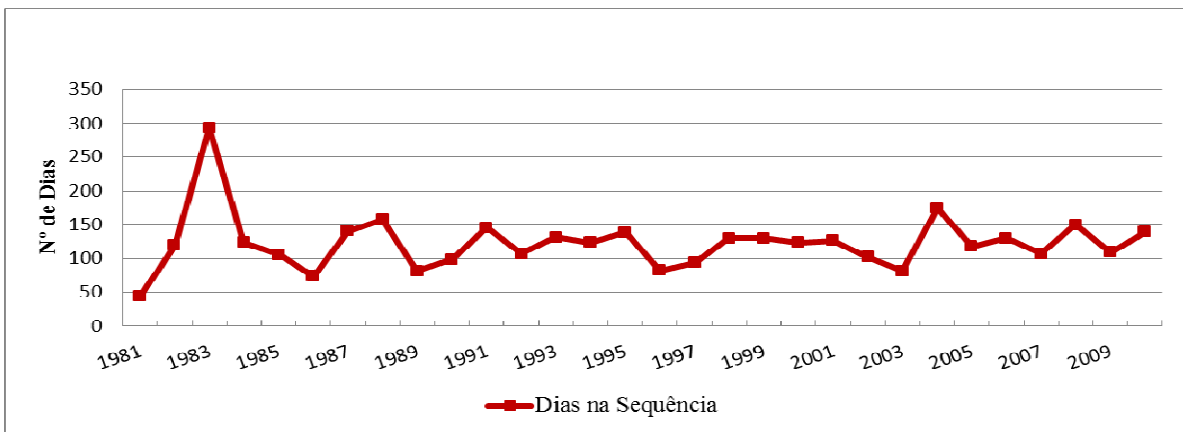
Verifica-se uma variação nos registros de dias secos consecutivos em cada *cluster*. Destacam-se os *clusters 0* e *2* com as maiores ocorrências, alcançando 239 dias secos consecutivos no *cluster 0* e 293 dias no *cluster 2*. Constata-se que o *cluster 3* apresentou os menores registros de dias consecutivos secos, haja vista ser o *cluster* de maior volume pluviométrico e uma regularidade de chuvas ao longo do ano (MOLION *et al.*, 2002).



**Gráfico 28:** Distribuição do número máximo de dias secos consecutivos em cada ano (1981 – 2010) para o cluster 0.

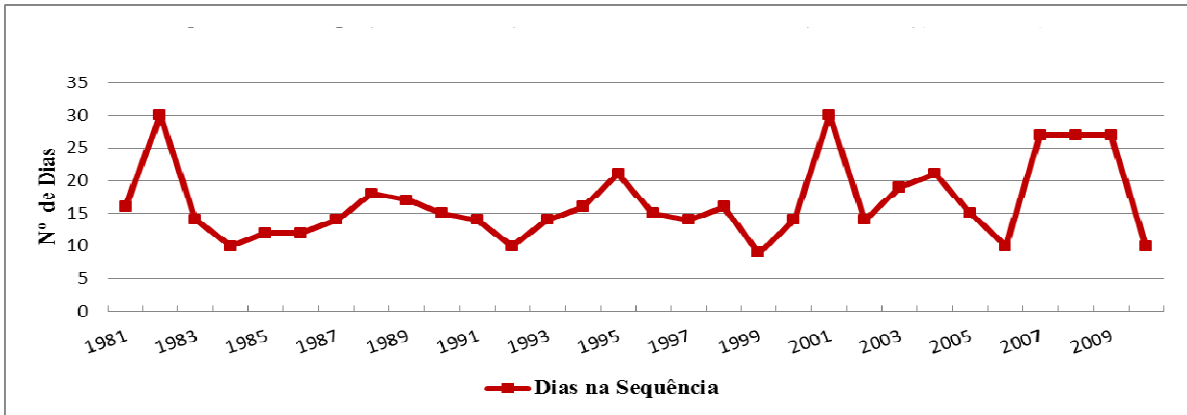


**Gráfico 29:** Distribuição do número máximo de dias secos consecutivos em cada ano (1981 – 2010) para o cluster 1.

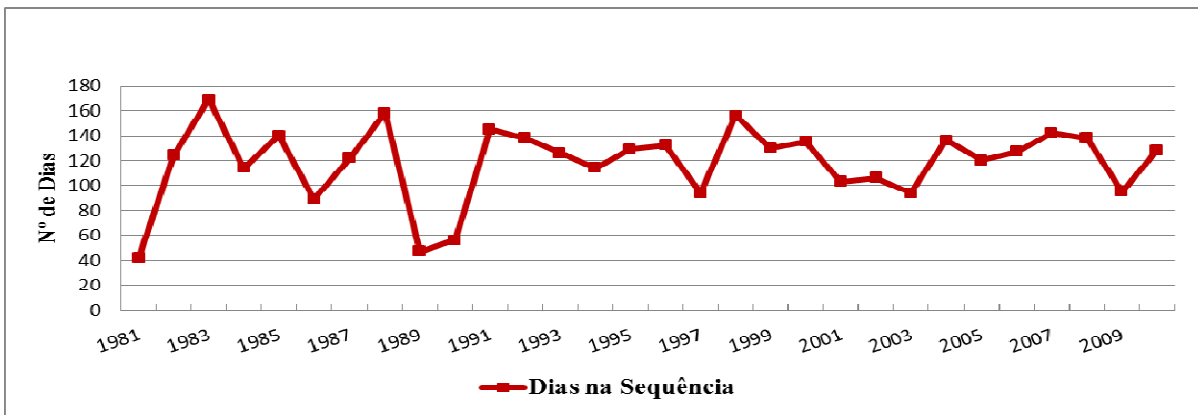


**Gráfico 30:** Distribuição do número máximo de dias secos consecutivos em cada ano (1981 – 2010) para o cluster 2.





**Gráfico 31:** Distribuição do número máximo de dias secos consecutivos em cada ano (1981 – 2010) para o *cluster 3*.



**Gráfico 32:** Distribuição do número máximo de dias secos consecutivos em cada ano (1981 – 2010) para o *cluster 4*.

Em todos os *clusters* observa-se uma variabilidade na ocorrência do número de dias secos consecutivos ao longo da série temporal estudada. Os anos que se destacaram com os maiores registros de dias secos consecutivos, podendo considerar como períodos de secas excessivas, foram os mesmos classificados pela técnica do Quantil para as demais escalas trabalhadas. Da mesma forma, para as análises dos registros de anos com menos dias secos consecutivos, estes foram classificados como os anos “muito chuvosos”.

## 5. CONCLUSÃO

Os resultados alcançados com o desenvolvimento deste trabalho permitiram a validação da hipótese proposta. Portanto, é possível encontrar padrões de ocorrência de eventos extremos de precipitação por meio de técnicas de mineração de dados, com o propósito de identificar o comportamento e a frequência das ocorrências desses eventos, em séries históricas de dados climáticos.

Além disso, o objetivo proposto foi atingido. Foram identificadas cinco zonas pluviometricamente homogêneas no estado da Bahia, com os dados de precipitação correspondente ao período de 1981 a 2010. Para tal, foram adotadas técnicas de agrupamento por meio do algoritmo *k-means*, que delimitou com coerência as regiões pluviométricas do estado, mostrando-se altamente eficiente para esta aplicação.

Avaliando os valores acumulados anuais, mensais e diários da precipitação, para todo período, observou-se uma alta variabilidade na distribuição pluviométrica em todos os *clusters*. A maior variação foi registrada nos *clusters* 0 e 1, pertencentes à região semiárida do estado, que apresenta como característica climática alta variabilidade pluviométrica. A menor variação foi registrada no *cluster* 3, região pertencente à faixa litorânea, que apresenta como característica a regularidade e altos volumes pluviais durante todo ano. Foi observado também, que em todos os *clusters* as estações secas e chuvosas são bem definidas.

Foram identificados por meio da técnica dos quantis, os limiares de classificação de eventos extremos para as escalas anual, mensal e diária, nas cinco zonas homogêneas. Os limiares para determinar os extremos foram diferentes para cada *cluster*. Os anos com registros de anomalias extremas de precipitação, para as três escalas, apresentaram similaridade de ocorrência em todos os *clusters*.

Foi observada uma predominância da classificação de eventos extremos “muito secos” na década de 90. A quantificação de eventos “muito chuvosos” foi registrada com predominância na década de 80, porém a década que apresentou mais anos com ocorrência desses eventos foi década de 2000.

Os anos classificados como “muito secos” estiveram em quase sua totalidade associados com os anos registrados com ocorrência do fenômeno El Niño. Os anos com registro de El Niño, de intensidade forte, refletiu na ocorrência de eventos “muito secos” em

todos os *clusters*. Por outro lado, a ocorrência do fenômeno La Niña está associada com os anos de eventos extremos “muito chuvosos”, com destaque para os anos de 1985 e 1989, sob uma atuação do fenômeno com intensidade forte, sendo considerados os anos mais chuvosos em todo estado.

Avaliando a classificação dos eventos extremos na escala mensal, foi verificado que a maior ocorrência de eventos extremos “muito secos” aconteceu nos meses de transição de estações, e os meses de maior ocorrência de eventos “muito chuvosos” foram aqueles considerados como meses de máximas precipitações no período chuvoso, também registrado em literaturas especializadas. Não obstante, foram registradas ocorrências de anomalias negativas de precipitação para outros meses, ao longo do período chuvoso de cada cluster, sendo altamente prejudicial para o setor agrícola da região, levando a perdas parciais e totais de safras.

Ainda na escala mensal, foram encontrados padrões sequenciais da ocorrência dos eventos extremos, nos meses pertencentes à estação chuvosa de cada cluster. Para uma sequência de tamanho um, foram encontradas as maiores ocorrências de eventos extremos negativos nos meses de transição das estações, e para os eventos de extremos positivos as maiores ocorrências foram nos meses com máximo volume pluvial do período. Os maiores tamanhos de sequência registrados foram dois e três, revelando a frequência com que ocorre determinada sequência de eventos.

Numa análise em escala diária, pode-se afirmar que predomina a ocorrência de eventos extremos “muito secos” em todo estado, variando de 47% a 89% dos dias da série de 1981 a 2010. Os maiores registros de dias secos consecutivos para cada ano mostraram similaridade com os anos classificados como “muito secos”, quando comparado com as escalas anuais e mensais.

## REFERÊNCIAS BIBLIOGRÁFICAS

ADEJUWON, J.O.; ODEKUNLE, T. O. Variability and severity of the “Little Dry Season” in south western Nigeria; **Journal of Climate** n19, p. 1 – 8, 2006

AGRAWAL R., IMIELINSKI T., SWAMI A. **Mining Association Rules Between Sets of Items in Large Databases**. Proc. of the 1993 ACM SIGMOD Conf. on Management of Data. 1993.

AGRAWAL, R.; SRIKANT R. Mining Sequential Patterns. Proc. 11th Int’l Conf. Data Eng., March 1995;

AGRAWAL, R.; SRIKANT, R. **Mining sequential patterns: Generalizations and performance improvements**. In P. M. G. Apers, M. Bouzeghoub, and G. Gardarin, editors, Proc. 5th Int. Conf. Extending Database Technology, EDBT, volume 1057, pages 3-17. Springer-Verlag, 1996. 25-29 p.

ALENCAR, R.; BARBOSA, J.; SOARES, J. Desempenho dos métodos dos quantis e SPI para o monitoramento da precipitação no Alto Sertão da Paraíba. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 14., 2006, Fortaleza. **Anais...** Rio de Janeiro: Sociedade Brasileira de Meteorologia, 2006.

ALTAMIRANO, R. J. A. **Climatologia dos eventos chuvosos e secos severos, extremos e muito extremos usando o índice de precipitação normalizada (SPI) para as regiões Centro-oeste, Sudeste e Sul do Brasil**. Dissertação de Mestrado em Meteorologia. Instituto de Pesquisas Espaciais, São José dos Campos, SP. 2010. 186p.

AMO, S. Técnicas de Mineração de Dados. In: Congresso da Sociedade Brasileira de Computação, 24, 2004, Salvador. **Anais do XXIV Congresso da Sociedade Brasileira de Computação**, 2004.

ANA - Agência Nacional das Águas. Disponível em: <http://hidroweb.ana.gov.br>. Acesso em: agosto de 2011.

ANDRADE, H. O.; FERNANDES, R. C. Contribuição ao Estudo da Variabilidade Pluviométrica no Município de Itaetê – BA. **Revista Geonorte**, Edição Especial 2, V.2, N.5, p.1066 – 1075 , 2012.

ANDRÉ, R. G. B.; MARQUES, V. S.; PINHEIRO, F. M. A.; FERRAUDO, A. S. Identificação de Regiões Pluviometricamente Homogêneas no Estado do Rio de Janeiro, Utilizando Valores Mensais. **Revista Brasileira de Meteorologia**, v. 23, nº 4, 501-509, 2008.

ARAUJO, H. A.; RODRIGUES, R. S. **Regiões características do Estado da Bahia para previsão de tempo e clima**, p.1-13, out. 2000. Disponível em: <www.inga.ba.gov.br>. Acesso em: 26 de junho de 2011

AOUAD, M.S. **Tentativa de classificação climática para o Estado da Bahia**. Rio de Janeiro: IBGE, 1983. 80 p.

AZEVEDO, P.V.; SILVA, B.B.; RODRIGUES, M.F.G.. Previsão Estatística das Chuvas de Outono no Estado do Ceará. **Revista Brasileira de Meteorologia**, v. 13, no. 1, 19-30, 1998.

AZEVEDO, P. V. de; SILVA, F. D. S. Risco climático para o cultivo do algodoeiro na região Nordeste do Brasil. **Revista Brasileira de Meteorologia**, v.22, p.408-416, 2007.

BARBOSA, D.V.N. **Os Impactos da Seca de 1993 no Semiárido Baiano: Caso de Irecê**. Salvador: SEI, 2000. 98p.

BARBOSA, J. P. M. **Avaliação de técnicas empíricas e estatísticas de identificação de extremos de precipitação para o litoral paulista e entorno**. Dissertação de Mestrado em Engenharia Agrícola. Universidade Estadual de Campinas. Campinas, SP, 2008. 110p.

BERLATO, M.A.; FONTANA D.C. **El Niño e La Niña: Impactos no clima, na vegetação e na agricultura do Rio Grande do Sul**. Aplicações de previsões climáticas na agricultura. Porto Alegre: Ed. da UFRGS, 2003, 110 p.

BERRY, M. J. A.; LINOFF, G. Data mining techniques: for marketing, sales and customer support. USA: **Wiley Computer Publishing**, 1997.

BLAIN, G. C. **Avaliação e adaptação do Índice de Severidade de Seca de Palmer (PDSI) e do Índice Padronizado de Precipitação (SPI) às condições climáticas do Estado de São Paulo**. 122 p. Dissertação (Mestrado em Agricultura Tropical e Subtropical)-Instituto Agrônomo de Campinas, Campinas. SP, 2005.

BLAIN, G. C; BRUNINI, O. Análise da escala temporal de monitoramento das secas agrícolas e meteorológicas no Estado de São Paulo. **Revista Brasileira de Meteorologia**, v. 22, n. 2, p. 255-261, 2007a.

BLAIN, G. C; BRUNINI, O. Análise comparativa dos Índices de Seca de Palmer, Palmer adaptado e Índice Padronizado de Precipitação no Estado de São Paulo. **Revista Brasileira de Meteorologia**, v. 22, n. 1, p. 105-111, 2007b.

BOSCHI, R. S.; OLIVEIRA, S. R. de M.; ASSAD, E. D. Técnicas de Mineração de Dados para Análise Pluvial Decenal do Rio Grande do Sul. **Revista Engenharia Agrícola**. v. 31, n.6, p. 1189 – 1201, 2011.

BRAGA, C.C; MELO, M.L.D; MELO, E.C.S. Análise de Agrupamento Aplicada a Distribuição da Precipitação no Estado da Bahia.10º Congresso Brasileiro de Meteorologia. Brasília-DF. **Anais...** Sociedade Brasileira de Meteorologia. p.1857-62, 1998.

BRASIL. Ministério da Integração Nacional. Atualiza a relação dos municípios pertencentes à região Semiárida do Fundo Constitucional de Financiamento do Nordeste (FNE). Portaria nº 89, de 16 de março de 2005. **Diário Oficial da União**, Brasília, 17 de março de 2005a. Disponível em:

<<http://www.mi.gov.br/desenvolvimentoregional/publicacoes/delimitacao.asp>> Acesso em: 05 de dezembro de 2011.

BRASIL. Ministério da Integração Nacional. Secretaria de Políticas de Desenvolvimento Regional. **Nova Delimitação do Semiárido do Brasil**. Brasília, 2005b. 35p. Disponível em: <<http://www.mi.gov.br/desenvolvimentoregional/publicacoes/delimitacao.asp>> Acesso em: 05 de dezembro de 2011.

BRASIL. Ministério do Meio Ambiente. **Vulnerabilidade Ambiental. Desastres naturais ou fenômenos induzidos?** Organização do texto: Rosely Ferreira dos Santos. Brasília, 2007. 192p.

BRUNINI, O.; BLAIN, G.; CIARELLI, A.; DOS SANTOS, L.; SARPA R.; LOREGIAN, E. Avaliação do Índice de Severidade de Seca de Palmer para quantificação da seca agrícola no Estado de São Paulo. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 12., 2002, Foz de Iguaçu. **Anais...** Rio de Janeiro: SBMet, 2002a. 1 CD-ROM.

BRUNINI, O.; BLAIN, G.; CIARELLI, A.; DOS SANTOS, L.; SARPA R.; LOREGIAN, E. Monitoramento das condições de seca no Estado de São Paulo utilizando o Índice Padronizado de Precipitação – SPI. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 12., 2002, Foz de Iguaçu. **Anais...** Rio de Janeiro: SBMet, 2002b. 1 CD-ROM.

BUCENE, L. C.; RODRIGUES, L. H. A; MEEIRA, C.A.A. **Mineração de dados climáticos para previsão de geada e deficiência hídrica para as culturas do café e da cana-de-açúcar para o Estado de São Paulo**. Campinas, 2002. Disponível em: <http://www.infoteca.cnptia.embrapa.br/handle/doc/8675>. Acesso em: 26 de junho de 2011.

CAMERON, C., MORRINGTON-DAVIES, G., VELDE, V. **Gerenciando Extremos Climáticos e Desastres na América Latina e no Caribe: Lições do relatório SREX**. Rede de Conhecimento de Clima e Desenvolvimento, CDKN, 2012. Disponível em: [www.cdkn.org/srex](http://www.cdkn.org/srex). Acesso em novembro de 2012.

CARDOSO, A. O.; DIAS, P. L. S. Identificação de Trimestres Extremos no Regime Pluviométrico no Sul e Sudeste do Brasil e Relação com Anomalias com a TSM. **Revista Brasileira de Meteorologia**, v.19, nº1, 149-162p. 2004.

CARVALHO, O de. O impacto social da seca no Nordeste. In: Conferência Nacional Latino-americana da Desertificação. **Anais...** Fortaleza, 1994, 23p.

CARVALHO, L.M.V; JONES C; LIEBMANN, B. Extreme precipitation events in southeast south America and large-scale convective patterns in the south atlantic convergence zone. **Journal of Climate**, nº15. 2377-3394p, 2002.

CARVALHO, L., JONES, C., LIEBMAN B. The South Atlantic Convergence Zone: intensity, form, persistence, and relationship with intraseasonal to interannual Activity and Extreme Rainfall. **Journal of Climate**, v.17, p. 88-108. 2004.

CAVALCANTI, I. F. A.; FERREIRA, N. J.; DIAS, M. A. F.; JUSTI, M. G. A. **Terra e Clima no Brasil**. São Paulo. Editora Oficina de Textos, 464p. 2009.

CENTRO DE ESTATÍSTICA E INFORMAÇÕES DA BAHIA (CEI). **Risco de Seca na Bahia**. Salvador, 1991. 110 p.

CENTRO DE PLANEJAMENTO DA BAHIA (CEPLAB). **Atlas Climatológico do Estado da Bahia**. Salvador, 1976. 237 p.

CHAPMAN, P.; CLINTON, J.; KERBER, R.; KHABAZA, T.; REINARTZ, T.; SHEARER, C.; WIRTH, R. **CRISP-DM 1.0: step-by-step data mining guide**. Illinois: SPSS, 2000. 78p.

CPTEC. Centro de Previsão de Tempo e Estudos Climáticos. **Boletim Climanálise**. Disponível em: <http://climanalise.cptec.inpe.br/~rclimanl/boletim/> Acesso em out. de 2012.

DAS G.; LIN K., MANNILA, H.; RENGANATHAN, G.; SMYTH, P. Rule discovery from time series. In: Proceedings of the International Conference on Knowledge Discovery and Data Mining (KDD-98), **Anais...4.**, 1998. pp 16-22.

DHANYA, C. T.; KUMAR, D. N. Data mining for evolution of association rules for droughts and floods in India using climate inputs. **Journal of Geophysical Research**, v. 114, p. 1-15, 2009.

DINIZ, M. de C. S. **Variabilidade Climática e Ocorrência de Cheias na Zona Semiárida da Bacia Hidrográfica do São Francisco**. Dissertação (Mestrado) – Centro Estadual de Meteorologia do Estado da Bahia, Salvador, 2008.

DINIZ, A. F.; SANTOS, R. L. Repercussões das secas de 1997 e 1998 no semiárido baiano: o caso de Feira de Santana (Bahia-Brasil). **Revista Geográfica de América Central**, Costa Rica, p. 1 - 17, 29 jul. 2011.

DUFEK, A. S. **Índices de Extremos Climáticos de Temperatura e Chuva na América do Sul: Clima Presente e Validação do Modelo RegCM3**. Dissertação (Mestrado em Meteorologia) Universidade de São Paulo, 2008. 120p.

EM-DAT – Emergency Events Database. The OFDA/CRED International Disaster Database. Santa Maria. Disponível em: <<http://www.em-dat.net/>>. Acesso em: novembro de 2012.

ESPÍRITO SANTO, C.; SATYAMURTY, P. Eventos extremos de precipitação na Região Sudeste do Brasil e redondezas no período de 1997-2001. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 12., 2002, Foz de Iguaçu. **Anais...** Rios de Janeiro: SBMet, 2002. págs. 397-402.

ESRI – **Gis and Mapping Software**- Arc-Map. versão 9.3 (2009). Disponível em: <http://www.esri.com/>. Acesso em: dezembro de 2011.

ESTER, M., KRIEGEL, H. P., SANDER, J., XU, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In: International Conference on Knowledge Discovery and Data Mining (KDD-96), 2, 1996, Portland. **Anais...** AAAI Press. pp. 226–231.

FAO – Organização das Nações Unidas para Alimentação e Agricultura. Disponível em: <https://www.fao.org.br/destaques2.asp> Acesso em: outubro de 2012.

FARIAS, R. F. de L.; ALVES, K. M. A. da S.; NÓBREGA. R. S. Climatologia de Ocorrência de Eventos Extremos de Precipitação na Mesorregião do Sertão Pernambucano. **Revista Geonorte**, edição Especial 2, v.1, nº5, 930-941p. 2012.

FAYYAD, U.; PIATETSKI-SHAPIRO, G.; SMYTH, P. From Data Mining to Knowledge Discovery: an overview. In: **Advances In Knowledge Discovery & Data Mining**. Menlo Park: American Association for Artificial Intelligence, p. 1-34, 1996.

FECHINE, J. A. L.; GALVINCIO, J. D. Agrupamento da Precipitação Mensal da Bacia Hidrográfica do Rio Brígida – PE, Através da Multivariada. **Revista Brasileira de Geografia Física**. vol.01, nº1. p. 39-46, 2008.

FERREIRA, G. A.; MELLO, N.G. S. Principais Sistemas Atmosféricos Atuantes sobre a Região Nordeste do Brasil e a Influência dos Oceanos Pacífico e Atlântico no Clima da Região. In: **Revista Brasileira de Climatologia**, vol.1 nº1. p. 15-28, 2005.

FURTADO, D. A. **Mineração de Padrões Sequenciais Múltiplos**. Dissertação (Mestrado em Ciência da Computação). Universidade Federal de Uberlândia, Minas Gerais. 2005. 110 p.

GAN, M.A; KOUSKY, V.E. **Um Estudo Observacional sobre as Baixas Frias da Alta Troposfera nas Latitudes Subtropicais do Atlântico Sul e Leste do Brasil**. INPE, São José dos Campos, SP. 1982. 25 p.



GAN, M. A.; KOUSKY, V. E.; ROPELEWSKI, C. F. The South America Monsoon Circulation and its relationship to rainfall over West-Central Brazil. **Journal of Climate**, v. 17, p. 47-66, 2004.

GÓIS, G. **Caracterização da seca e seus efeitos na produção da cultura do milho para as diferentes Regiões do Estado de Minas Gerais**. Dissertação (Mestrado em Meteorologia Agrícola). Universidade Federal de Viçosa, Minas Gerais. 2005. 141 p.

GONÇALVES, N.M.S., **Impactos pluviiais e desorganização do espaço urbano em Salvador**. In: MONTEIRO, C.A.F.; MENDONÇA, F. (Org) *Clima Urbano*. São Paulo: 192p, 2003.

GRIMM, A; FERRAZ, S; GOMES, J. **Precipitation anomalies in Souther Brazil associated with El Niño and La Niña events**. **Journal of Climate**. v. 11, p. 2863-2880, 1998.

GRIMM, A. M.; TEDESCHI, R. G. ENSO and Extreme Rainfall Events in South America. **Journal of Climate**. v. 22, p. 1589 - 1609, 2009.

HALL, M. A.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. **The WEKA Data Mining Software: An Update**. SIGKDD Explorations, v. 11, n. 1. 2009.

HAN, J.; KAMBER, M. **Data Mining: Concepts and Techniques**. San Francisco: Morgan Kaufmann Publishers, 770p. 2011.

HAYES, J. M.; SVOBODA, M.D.; WILHITE, D. A.; VANYARKHO, O.V. Monitoring the 1996 drought using the Standardized Precipitation Index. Bulletin of the **American Meteorological Society**, v. 80, n. 3, p. 429-438, 1999.

HOLANDA, C.V.M, OLIVEIRA, E. Programa para Homogeneização de Dados – PROHD. In: Simpósio de Hidrologia, 3, 1979, Brasília. **Anais...** Porto Alegre: Associação Brasileira de Recursos Hídricos, p. 810-845, 1979.

IBGE- Instituto Brasileiro de Geografia e Estatística. Clima da Bahia. Disponível em: [www.ibge.gov.br](http://www.ibge.gov.br). Acessado em julho de 2012.

IPCC – Intergovernmental Panel on Climate Change. **Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation**. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. FIELD, C.B., BARROS, V., STOCKER, T.F., QIN, D., DOKKEN, D. J., EBI, K.L., MASTRANDEA, M.D., MACH, K.J., PLATTNER, G.-K., ALLEN, S.K., TIGNOR, M., MIDGLEY, P.M. Cambridge University Press, Cambridge, UK, and New York, NY, USA, 582 pp. Disponível

em: [http://ipcc-wg2.gov/SREX/images/uploads/SREX-All\\_FINAL.pdf](http://ipcc-wg2.gov/SREX/images/uploads/SREX-All_FINAL.pdf). Acesso em: novembro, 2012.

KODAMA, Y. M. Large – scale common features of subtropical precipitation zone (the Baiu frontal zone, the SPCZ, and the SACZ). Part I: Characteristics of subtropical frontal zones. **Journal Meteorological Society Japan**, v. 70, n. 4, p. 813-836, 1992.

KOUSKY, V.E. **Frontal influences on northeast Brazil**. São Paulo: INPE, 1979. 16 p.

KOUSKY, V. E. e MOURA, A.D. **Previsão de precipitação no Nordeste do Brasil: o aspecto dinâmico**. São Paulo: INPE, 1981. 13 p.

KUMAR, D. N.; DHANYA, M. C. T. Data Mining and its Applications for Modelling Rainfall Extremes. **Journal of Hydraulic Engineering**, .v. 15, p.25 – 51, 2009.

KUMAR, K.; SOUSA, F. A study of droughts at Campina Grande base don the Standardized Precipitation Index. In: CONGRESSO BRASILEIRO DE METEOROLOGIA. Foz de Iguaçu. **Proceedings...** Rio de Janeiro: SBMet, 2002.

LIMA, R. A. F. de A.; BRITO, J. I. B.; SOARES, J. F. P. Desempenho dos métodos dos quantis e SPI para o monitoramento da precipitação no Alto Sertão da Paraíba. In: CONGRESSO BRASILEIRO DE METEOROLOGIA. Santa Catarina **Anais...** Rio de Janeiro: SBMet, 2006.

LIU, J. N. K.; LI, B. N. L.; DILLON, T. S. An improved naïve Bayesian classifier technique coupled with a novel input solution method. **IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews**, v. 31, n. 2, May, p. 249-256, 2001.

LOPES, P.M.O., SILVA, B.B., RODRIGUES, M.F.G. Identificação de grupos pluviometricamente homogêneos no Estado do Rio Grande do Norte. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 9,Campos do Jordão-SP. **Anais...** Sociedade Brasileira de Meteorologia. p.272-274,1996.

MARCUZZO, F. F. N; ANDRADE, L. R.; MELO, D. C. R. Métodos de Interpolação Matemática no Mapeamento de Chuvas no Estado do Mato Grosso. **Revista Brasileira de Geografia Física**. n.4, 793-804p. 2011.

MARENGO, J. A.; OBREGÓN G. **Relatório Nº 2**. Caracterização do Clima no Século XX no Brasil: Tendências de Chuvas e Temperaturas Médias e Extremas. Centro de Previsão de Tempo e Estudos Climáticos (CPTEC/INPE). São Paulo, 2007.

MARTINS, P.; SILVA, D. S. **Um estudo sobre a mineração de padrões sequenciais**. Disponível em: <http://homepages.dcc.ufmg.br>. Acesso em: 18 de junho de 2011.

- MAY, W. Variability and extremes of daily rainfall during the Indian summer monsoon in the period 1901-1989. **Global and Planetary Change**, v. 44. p. 83-105, 2004.
- MEIRA, C. A. A. Modelos de alerta para controle da ferrugem do cafeeiro em lavouras com alta carga pendente. **Revista Pesquisa Agropecuária Brasileira**. v. 44, n. 3, p.233 – 242, março, 2009.
- MELO, G. R. de, et al. **Análise de Índice de Seca para o Nordeste do Brasil**. Abril, 2010<Disponível em: <http://www.ime.unicamp.br/sinape>> Acessado em: setembro de 2011.
- MOLINEROS, J.; DE WOLF, E.; FRANCL, L.; MADDEN, L.; LIPPS, P. Modeling epidemics of fusarium head blight: trials and tribulations. **Phytopathology**, v.95, p.71, 2005.
- MOLION, L.C.B. Secas, eterno retorno. **Ciência Hoje**, Rio de Janeiro, v. 3, n. 18, p. 26-32, maio/jun. 1985.
- MOLION, L. C. B.; BERNARDO, S. O. Uma Revisão da Dinâmica das Chuvas no Nordeste Brasileiro. **Revista Brasileira de Meteorologia**, v. 17, n. 1, p. 1-10, 2002.
- MONTEIRO, C.A.de F. **Clima e excepcionalismo: conjecturas sobre o desempenho da atmosfera como fenômeno geográfico**. Florianópolis: Ed. da UFSC, 241 p. 1991.
- MONTEIRO, J. B.; ROCHA, A. B.; ZANELLA, M. E. Técnica dos Quantis para Caracterização de Anos Secos e Chuvosos (1980 – 2009): Baixo Curso do Apodi – Mossoró/RN. **Revista do Departamento de Geografia – USP**, v.23. p. 232 – 249. 2012.
- MORAIS B.; CARVALHO, R.; LIMA, G. Previsão de seca usando redes neurais artificiais. In: SIMPÓSIO DE RECURSOS HÍDRICOS DO NORDESTE, 9., Salvador, Bahia. **Anais...**, 2008.
- MORAES, V. B.; JESUS JÚNIOR, W.C; MORAES, W. B.; CECÍLIO, R.A. Potenciais impactos das mudanças climáticas globais sobre a agricultura. **Revista Trópica- Ciências Agrárias e Biológicas**, v.5, nº2, p. 3-14, 2011.
- MUZA, M. N. **Variabilidade intrasazonal e interanual dos eventos extremos de precipitação e seca no sul e sudeste do Brasil durante o verão austral**. Dissertação (Mestrado em Meteorologia) – Universidade de São Paulo, São Paulo, 2005. 120p.
- MUZA, M. N.; CARVALHO, L. M. V. Diagnóstico de extremos climáticos no Centro- Leste da América do Sul. In: CONGRESSO BRASILEIRO DE METEOROLOGIA. Florianópolis. **Anais...** Rio de Janeiro: SBMet, 2006. 1 CD-ROM
- MUZA, M. N.; CARVALHO, L. M. V.; JONES, C.; LIEBMANN, B. Intraseasonal and Interannual Variability of Extreme Dry and Wet Events over Southeastern South America and

the Subtropical Atlantic during Austral Summer. **Journal of Climate** v.22. p.1682 – 1699, 2009.

NEVES, H. M. A; CRUZ, A. R. A Sazonalidade da Procura Turística na Ilha de Porto Santo. **Estudos Regionais**, nº 17, p.25 – 44, 2007.

NG, R. T.; HAN, J. Clarans: A method for clustering objects for spatial data mining. **Knowledge and Data Engineering, IEEE Transactions on**, v. 14, nº 5, 2002.

NOAA- National Oceanic and Atmospheric Administration. **Dados El Niño**. Disponível em: <http://www.elnino.noaa.gov>. Acesso em dezembro de 2012.

NOBRE, C. A. , MOLION., L.C.B.. **The Climatology of Droughts and Drought Prediction**,. In: Impacts of Climatic Variations on Agriculture, v.2 : Assesments in semi-arid regions, M. P. Parry, T.R. Carter e N. T. Konijn (eds.), D. Reidel Pub. Co., 764p, 1988.

NONATO, R. T. **Aplicação de Mineração de Dados na Identificação de Áreas Cultivadas com Cana – de- açúcar em Imagens de Sensoriamento Remoto no Estado de São Paulo**. Dissertação (Mestrado em Engenharia Agrícola) – Universidade Estadual de Campinas, São Paulo, 2010. 147p.

PALMER, W.C. Meteorological drought. **US Weather Bureau Res.** Nº 45. Washington, 58p, 1965.

PAUL, P.A.; MUNKVOLD, G.P. A model-based approach to preplanting risk assessment for gray leaf spot of maize. **Phytopathology**, v.94, p.1350-1357, 2004.

PATROCÍNIO, S. F. **Previsão de secas para o Nordeste do Brasil**. Fortaleza. Funceme, 2008.

PINKAYAN, S. **Conditional probabilities of occurrence of Wet and Dry Years Over a Large Continental Area**. Colorado: State University, Boulder – Co, 1966. (Hidrology papers, n.12).

RADAMBRASIL - Mapa Geológico e Mapa de Potencial de Recursos Hídricos da Folha SC 24/25 - Aracaju/Recife - 1983.

REPELLI, C. A; SILVEIRA, N. F; ALVES, J. M; NOBRE, C. A Índice de anomalia de precipitação para o Estado do Ceará. In: CONGRESSO BRASILEIRO DE AGROMETEOROLOGIA, 7., 1991, Viçosa, MG. **Resumos...** Viçosa: Sociedade Brasileira de Agrometeorologia/ Universidade Federal de Viçosa.

REZENDE, S. O.; PUGLIESI, J. B., MELANDA, E. A., PAULA, M. F. Mineração de Dados. In: **Sistemas Inteligentes. Fundamentos e aplicações**. Organização: Solange Oliveira Rezende. Barueri, SP: Manole, 2005.

RIOS, F. C. dos. **Mineração de Dados Aplicada à Definição de Zonas de Manejo por Produtividade no Âmbito da Agricultura de Precisão**. 2011, 92p. Trabalho de Graduação (Graduação em Engenharia da Computação). Universidade do Estado de Santa Catarina – UDESC, Joinville, 2011.

ROCHA, A. B. da. **Análise Multitemporal da dinâmica do uso e ocupação do baixo curso do rio Apodi-Mossoró-RN (1989-2009)**, 2011. 113p. Dissertação (Mestrado) - Programa de Pós-Graduação em Geografia, Universidade Estadual do Ceará, Ceará, 2011.

SAITO, S. M.; SAUSEN, T. M. Análise dos Prejuízos Socioeconômicos e Ambientais Causados pelas Inundações no Rio Grande do Sul. 13º Congresso Brasileiro de Geologia de Engenharia e Ambiental, 2011. São Paulo, **Anais...** São Paulo: Associação Brasileira de Geologia de Engenharia Ambiental.

SANTOS, C. A. C., BRITO, J. I. B. Análise dos índices de extremos para o semiárido do Brasil e suas relações com TSM e IVDN. **Revista Brasileira de Meteorologia**, v. 22, n. 3, p. 303-312, 2007

SANTOS, D. N; SILVA, V. P. R.; SOUSA, F. A. S.; SILVA, R. A. Estudo de Alguns Cenários Climáticos para o Nordeste Brasileiro. **Revista Brasileira de Engenharia Agrícola Ambiental**. v.14, nº5, p.492-500, 2010.

SANTOS, C. A. C. dos; MANZI, A. O.; Eventos extremos de precipitação no estado do Ceará e suas relações com a temperatura dos oceanos tropicais. **Revista Brasileira de Meteorologia**, v.26, p.157-165, 2011.

SANTOS, E.; MATOS, H.; ALVARENGA, J.; SALES, M.C.L. A Seca no Nordeste no ano de 2012: Relato sobre a Estiagem na Região e o Exemplo de Prática de Convivência com o Semiárido no Distrito de Iguaçú/ Canindé – CE. **Revista Geonorte**, Edição Especial 2, V.1, N.5, p.819 – 830, 2012.

SAREWITZ, D. et al. **WORKSHOP: Extreme Events Developing a Research Agenda for the 21st Century**. Bolder, 2000. Disponível em <http://www.esig.ucar.edu/extremes>. Acesso em novembro de 2011.

SECRETARIA DA AGRICULTURA, IRRIGAÇÃO E REFORMA AGRÁRIA (SEAGRI-BA). **Produção Agrícola da Bahia**. Disponível em: [http://www.seagri.ba.gov.br/investir\\_oportunidade.asp](http://www.seagri.ba.gov.br/investir_oportunidade.asp). Acesso em: novembro de 2012.

SILVA, L. A. **A influência do vórtice ciclônico de altos níveis (VCAN) sobre a precipitação do nordeste do Brasil e as características associadas**. Dissertação de Mestrado em Meteorologia, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 131 pg., 2005.

SILVA, R. A.; PAPAJORGJI, P. J.; GUIMARAES, A. M.; ROCHA, J. C. F. DAGER: Uma Ferramenta para Mineração de Dados Georreferenciados. **Revista de Engenharia e Tecnologia**. v.4, n.2, p. 113-118, 2012.

SILVA, R. M. A. Entre o Combate a Seca e a Convivência com o Semiárido: políticas publicas e transição paradigmática. **Revista Econômica do Nordeste**. v.38, n.3, 466-486p. 2007.

SILVA, S.T., BRAGA, C.C., ARAGÃO, M.R.S. Influência do El Nino - Oscilação Sul na distribuição espacial e temporal da precipitação no Estado da Paraíba. In: CONGRESSO BRASILEIRO DE METEOROLOGIA, 9,Campos do Jordão-SP. **Anais...** Sociedade Brasileira de Meteorologia p. 236-239,1996.

SILVA, V. de P. R. On climate variability in Northeast of Brazil. **Journal of Arid Environments**, v.58, n.4, p.575-596, 2004.

SILVA, V. de P. R. da; AZEVEDO, P. V. de; SILVA, B. B. da. Surface energy and evapotranspiration of a mango orchard grown in a semiarid environment. **Agronomy Journal**, v.1, p.1391- 1396, 2007.

SILVA, V. de P. R. da; PEREIRA, E. R. R.;ALMEIDA, R. S. R. Estudo da Variabilidade Anual e Intra-Anual da Precipitação na Região Nordeste do Brasil. **Revista Brasileira de Meteorologia**, v.27, n.2, p.163-172, 2012

SOUZA, J. L., de AMORIM, R.F.C., CARVALHO, S.M.R., PEREIRA, J.O., Curi, P.R.C. Agrupamento de estações pluviométricas do Estado de Alagoas, utilizando-se análise multivariada. **Rev. Brasileira de Meteorologia**, v. 7, n.2, p.603-12, 1992.

SOUSA JÚNIOR, I. F. **A influência da urbanização no clima da cidade de Campina Grande, PB**. Campina Grande: DCA, 2006. 94p. Dissertação Mestrado.

SOUZA, W. M; AZEVEDO, P. V.; ARAÚJO, L. E. Classificação da Precipitação Diária e Impactos Decorrentes dos Desastres Associados às Chuvas na Cidade do Recife –PE. **Revista Brasileira de Geografia Física**, v.02, p. 250 – 268, 2012.

SOUZA, Z. N. de; CERRI, D. G. P; COLET, M. J.; RODRIGUES, L. H. A.; MAGALHÃES, P. S. G.; MANDONI, R. J. A. Análise dos atributos do solo e da produtividade da cultura da cana-de-açúcar com o uso da geoestatística e árvore de decisão. **Revista Ciência Rural**, v. 40, n.4,p.840 – 847, 2010.

SRIKANT, R.; AGRAWAL, R. **Mining Sequential Patterns Generalizations and Performance Improvements**. In Proceedings of the Fifth Int'l Conference on Extending Database Technology (EDBT). Avignon, France, 1996. Disponível em: <http://www.almaden.ibm.com/>. Acesso em: Setembro de 2011.

SUPERINTENDÊNCIA DE ESTUDOS ECONÔMICOS E SOCIAIS DA BAHIA (SEI). **Bahia em números**. v.10, Salvador, 2010. 110 p.

SUPERINTENDÊNCIA DE ESTUDOS ECONÔMICOS E SOCIAIS DA BAHIA (SEI). **Uso Atual das Terras. Bacias do Extremo Sul e do Rio Jequitinhonha**. Salvador, 2008. 104p

SUPERINTENDÊNCIA DE ESTUDOS ECONÔMICOS E SOCIAIS DA BAHIA. (SEI). **Sistemas de Dados Estatísticos**. 2000. Disponível em: <http://www.sei.ba.gov.br/side/consulta>. Acesso em: 20 de maio de 2011.

SUPERINTENDÊNCIA DO DESENVOLVIMENTO DO NORDESTE – SUDENE. **As Secas do Nordeste (uma abordagem histórica de causas e efeitos)**. Recife, 1992. 81p

STEINBACH, M., P.-N.; TAN, V.; KUMAR, S.; KLOOSTER, C. POTTER. **Discovery of climate indices using clustering**, ACM SIGKDD Conferência em KDD, Washington, 2003.

TANAJURA, C. A. S.; GENZ, F.; ARAÚJO, H. A. Mudanças Climáticas e Recursos Hídricos na Bahia: Validação da Simulação do Clima Presente do HADRM3P e Comparação com os Cenários A2 e B2 para 2070 – 2100. **Revista Brasileira de Meteorologia**, v.25, n.3, 345 - 358, 2010.

TEDESCHI, R. **Impacto de episódios El Niño e La Niña sobre a frequência de eventos extremos de precipitação e vazão na América do Sul**. Dissertação (Mestrado em Engenharia de Recursos Hídricos e Ambiental, Setor de Tecnologia) Universidade Federal de Paraná, 2008. 160 p.

TEIXEIRA, M. **Atividade de ondas sinópticas relacionada a episódios de chuvas intensas na região sul do Brasil**. Dissertação (Mestrado em Meteorologia) – Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2005.

THORNTHWAITE, C.W. & MATHER, J.R. The water balance. **Climatology**, Centerton, v.8, n.1, p.1 – 40, 1955.

TIEPPO, R. C.; NUNES, C. C. P.; DALLACORT, R.; FIETZ, C. R.; COMUNELLO, E.; CREMON, C. Análise de interpoladores na geração de mapas de precipitação para o Estado de Mato Grosso. **Anais 3º Simpósio de Geotecnologias no Pantanal**, p. 95 -105, 2010.

VENDRUSCUO, L. G.; OLIVEIRA, S. R. de M.; ESQUERDO, J. C. D. M.; ANTUNES, J. F. G. Análise de padrões sequenciais em série histórica do Rio Paraguai. In: 2º Simpósio de Geotecnologias no Pantanal, 2009, Corumbá. **Anais...**, 2009, p. 323– 332.

WITTEN, I.H.; FRANK, E.; HALL, M.A. **Data mining: practical machine learning tools and techniques**. 3rd ed. San Francisco: Morgan Kaufmann, 2011. 629p.

XAVIER, T; XAVIER, A. **Classificação de anos secos e chuvosos na Região Nordeste do Brasil e sua distribuição espacial**, 1984. Disponível em: <<http://www.agro.unitau.br/seer/index.php/ambigua/article/viewPDFInterstitial/300/368>> Acesso em setembro de 2011.

XAVIER, T. M. B. S.; XAVIER, A. F. S. Caracterização de Períodos Secos ou Excessivamente Chuvosos no Estado Ceará Através da técnica dos Quantis: 1964-1998 **Revista Brasileira de Meteorologia**, Brasil, v. 14, n. 2, p. 63-78, 1999.

XAVIER, T. de Ma. B.S. **Tempo de Chuva – Estudos Climáticos e de Previsão para o Ceará e Nordeste Setentrional**, ABC Editora, Fortaleza-Ceará, 478 p. 2001.

XAVIER, T. M. B. S.; SILVA, J. F.; REBELLO, E. R. G. **A Técnica dos Quantis e suas aplicações em Meteorologia, Climatologia e Hidrologia, com ênfase para as regiões brasileiras**. Thesaurus Editora de Brasília Ltda. Brasília, 2002, 141 p.

XAVIER, T. Ma. B. S. ; XAVIER, A. F. S.; DIAS, P. L. S. ; DIAS, M. A. F. S. . Interrelações dos Eventos ENOS (ENSO), da ZCIT (ITCZ) no Atlântico e da Chuva nas Bacias Hidrográficas do Ceará. **Revista Brasileira de Recursos Hídricos**, ABRH-Porto Alegre-RS, v. 8, n. 2, p. 111-126, 2003.

XAVIER, T. M. B.; XAVIER, A. F. S, ALVES, J. M. B. **Quantis e Eventos Extremos. Aplicações em Ciências da Terra e Ambientais**. 2007. 278p.



## APÊNDICE I – Tabelas de classificação de eventos extremos em escala anual

a) *Cluster 0*

b) *Cluster 1*

<b>Classificação</b>	<b>Limites de Precipitação (mm)</b>	<b>Classificação</b>	<b>Limites de Precipitação (mm)</b>
<b>Muito Seco</b>	$P \leq 273$	<b>Muito Seco</b>	$P \leq 478$
<b>Seco</b>	$273 \geq P \leq 397$	<b>Seco</b>	$478 \geq P \leq 595$
<b>Normal</b>	$397 \geq P \leq 534$	<b>Normal</b>	$595 \geq P \leq 751$
<b>Chuvoso</b>	$534 \geq P \leq 633$	<b>Chuvoso</b>	$751 \geq P \leq 834$
<b>Muito Chuvoso</b>	$P \geq 633$	<b>Muito Chuvoso</b>	$P \geq 834$

c) *Cluster 2*

d) *Cluster 3*

<b>Classificação</b>	<b>Limites de Precipitação (mm)</b>	<b>Classificação</b>	<b>Limites de Precipitação (mm)</b>
<b>Muito Seco</b>	$P \leq 646$	<b>Muito Seco</b>	$P \leq 1202$
<b>Seco</b>	$646 \geq P \leq 794$	<b>Seco</b>	$1202 \geq P \leq 1291$
<b>Normal</b>	$794 \geq P \leq 940$	<b>Normal</b>	$1291 \geq P \leq 1415$
<b>Chuvoso</b>	$940 \geq P \leq 1075$	<b>Chuvoso</b>	$1415 \geq P \leq 1542$
<b>Muito Chuvoso</b>	$P \geq 1075$	<b>Muito Chuvoso</b>	$P \geq 1542$

d) *Cluster 4*

<b>Classificação</b>	<b>Limites de Precipitação (mm)</b>
<b>Muito Seco</b>	$P \leq 820$
<b>Seco</b>	$820 \geq P \leq 958$
<b>Normal</b>	$958 \geq P \leq 1127$
<b>Chuvoso</b>	$1127 \geq P \leq 1250$
<b>Muito Chuvoso</b>	$P \geq 1250$

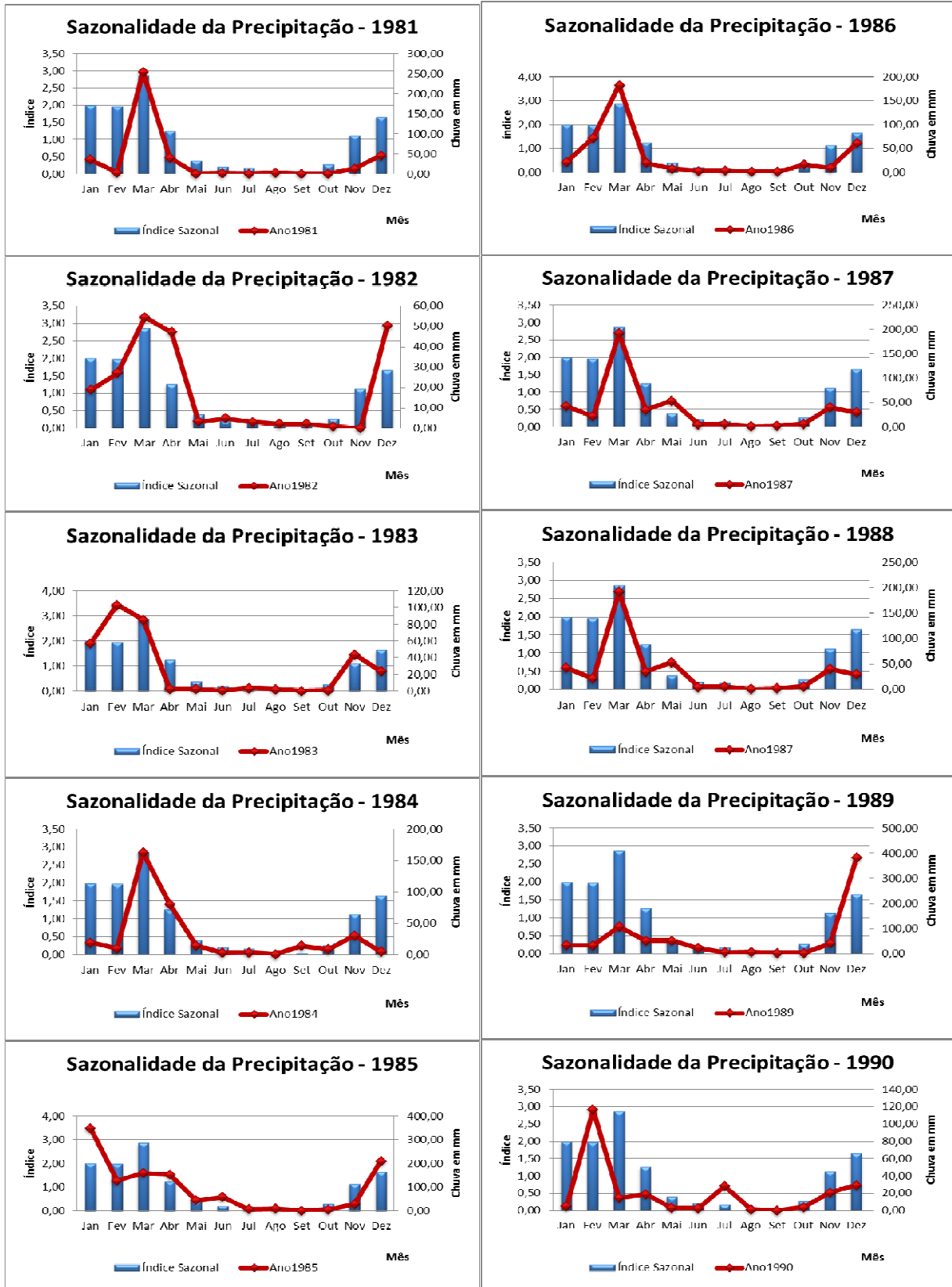
**APÊNDICE II – Tabela 17. Relação dos municípios e estações meteorológicas utilizadas.**

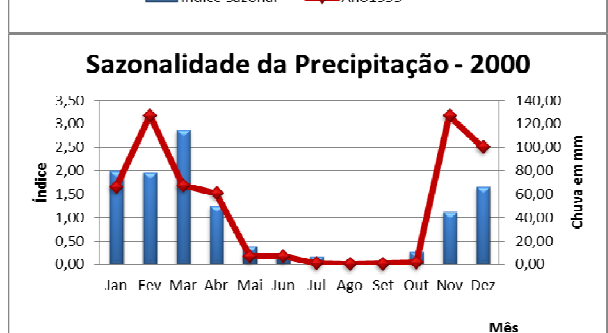
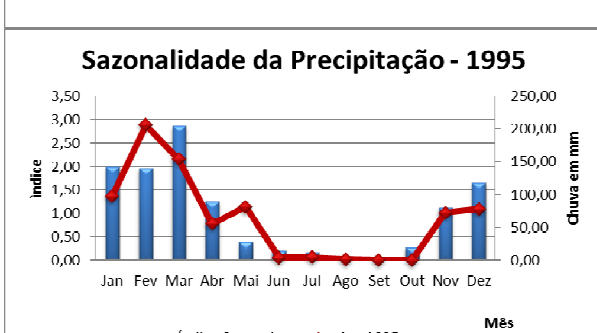
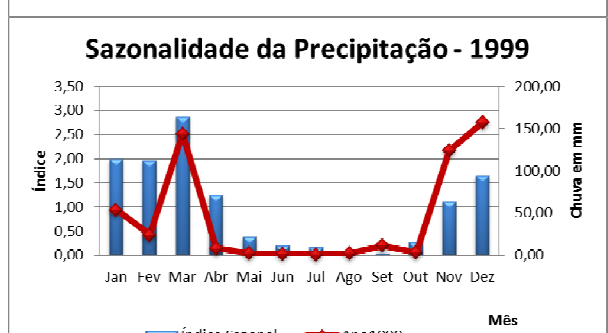
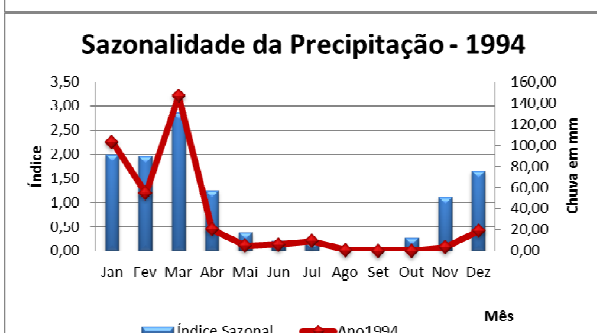
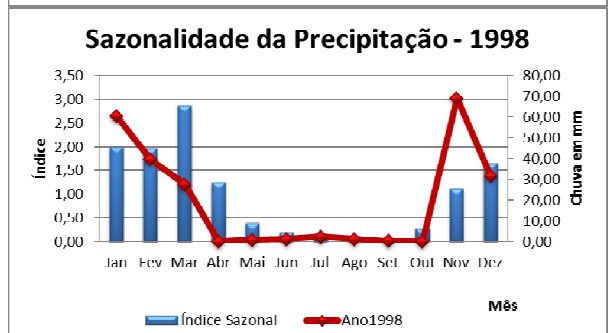
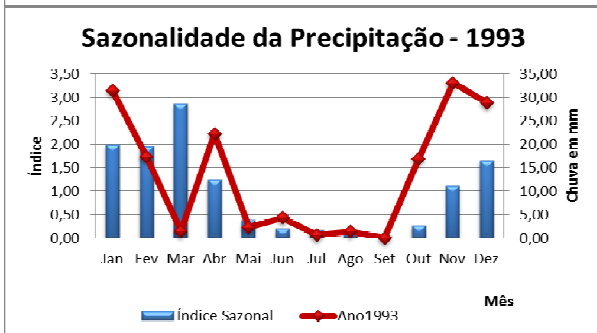
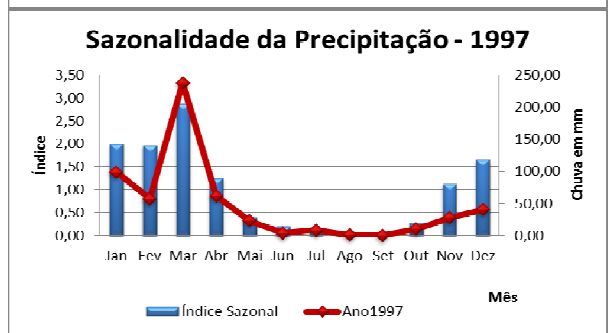
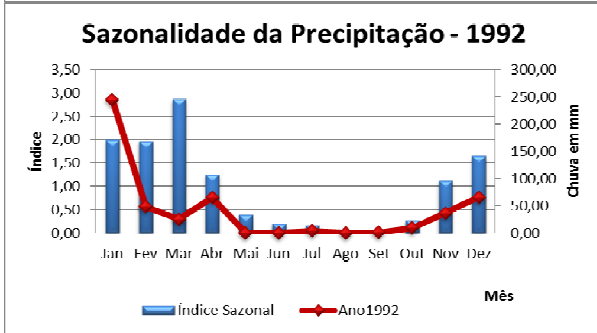
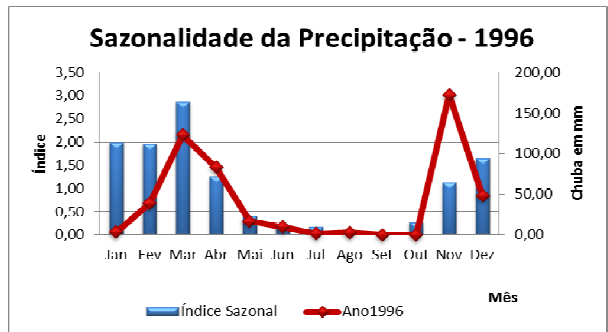
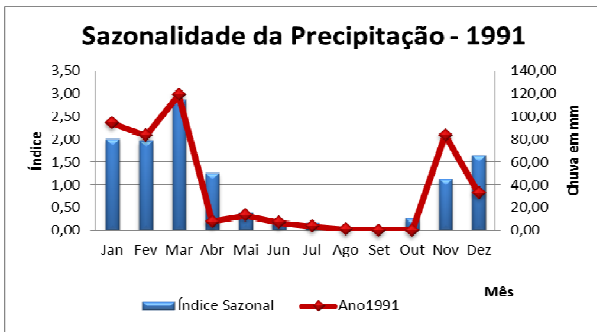
Município	Cod_Est	Lat	Long	Alt	Cluster	Chuva
<b>Abaré</b>	839031	-8,623	-39,239	300	0	510
<b>Curaçá</b>	939017	-9,123	-39,936	365	0	464
<b>Juazeiro</b>	940018	-9,560	-40,647	377	0	484
<b>Juazeiro</b>	940024	-9,406	-40,503	370	0	439
<b>Juazeiro</b>	940025	-9,491	-40,204	378	0	473
<b>Juazeiro</b>	940028	-9,677	-40,604	397	0	462
<b>Queimadas</b>	1039019	-10,974	-39,633	279	1	523
<b>Itaguaçu da Bahia</b>	1042012	-10,986	-42,342	414	1	577
<b>Cipó</b>	1138014	-11,098	-38,513	131	1	483
<b>Miguel Calmon</b>	1140000	-11,433	-40,604	562	1	659
<b>Piritiba</b>	1140010	-11,553	-40,600	1	1	673
<b>Ibipeba</b>	1142017	-11,329	-42,305	490	1	659
<b>Ibipeba</b>	1142020	-11,356	-42,274	489	1	640
<b>Rafael Jambeiro</b>	1239007	-12,585	-39,518	159	1	610
<b>Itaeté</b>	1240012	-12,990	-40,963	299	1	656
<b>Iaçu</b>	1240016	-12,762	-40,211	237	1	524
<b>Itaeté</b>	1241001	-12,935	-41,066	317	1	665
<b>Lençóis</b>	1241017	-12,498	-41,296	355	1	869
<b>Oliveira dos Brejinhos</b>	1242016	-12,252	-42,763	440	1	578
<b>Jequié</b>	1340003	-13,882	-40,140	199	1	514
<b>Santa Cruz da Vitória</b>	1439044	-14,959	-39,808	243	1	922
<b>Planalto</b>	1440009	-14,744	-40,492	989	1	966
<b>Tanhaçu</b>	1441000	-14,099	-41,291	341	1	570
<b>Itaju do Colônia</b>	1539008	-15,139	-39,724	182	1	763
<b>Itaju do Colônia</b>	1539016	-15,154	-39,769	180	1	788
<b>Vitória da Conquista</b>	1540003	-15,341	-40,933	520	1	661
<b>Jequié</b>	15	-13,870	-40,070	305	1	556
<b>Morpará</b>	1143002	-11,563	-43,275	418	2	788
<b>Andaraí</b>	1241008	-12,802	-41,328	330	2	994
<b>Brotas de Macaúbas</b>	1242015	-12,004	-42,628	837	2	707
<b>Ibotirama</b>	1243011	-12,181	-43,223	420	2	711
<b>Cândido Sales</b>	1541001	-15,505	-41,229	676	2	733
<b>Medeiros Neto</b>	1740005	-17,372	-40,226	1	2	970
<b>Barra</b>	31	-11,080	-43,170	401	2	703
<b>Mansidão</b>	1143010	-11,341	-43,828	411	2	806
<b>Cotegipe</b>	1144005	-11,614	-44,158	421	2	818
<b>Sítio do Mato</b>	1243000	-12,867	-43,381	413	2	774

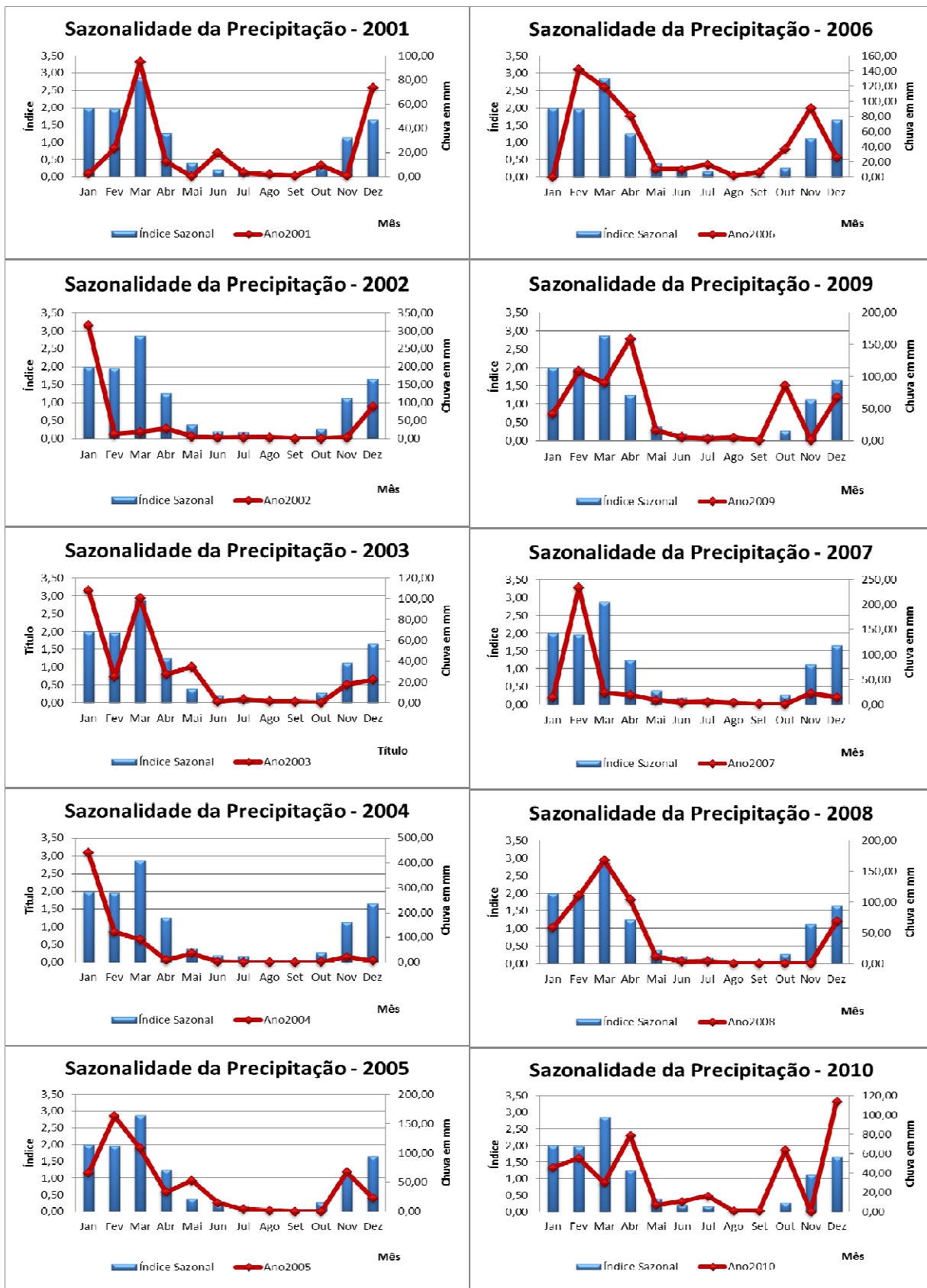
<b>Santana</b>	1343008	-13,293	-43,909	442	2	769
<b>Santa Maria da Vitória</b>	1344002	-13,278	-44,559	520	2	970
<b>Jaborandi</b>	1344013	-13,711	-44,633	610	2	1027
<b>Correntina</b>	1344014	-13,336	-44,652	566	2	906
<b>Coribe</b>	1344015	-13,567	-44,306	447	2	802
<b>Correntina</b>	1344016	-13,451	-44,566	479	2	966
<b>Santa Maria da Vitória</b>	1344017	-13,401	-44,198	437	2	900
<b>Correntina</b>	1345000	-13,734	-45,403	679	2	871
<b>Carinhanha</b>	1443002	-14,304	-43,768	440	2	761
<b>Cocos</b>	1444017	-14,268	-44,522	500	2	922
<b>Correntina</b>	41	-13,330	-44,620	549	2	882
<b>Jandaira</b>	1137027	-11,542	-37,566	15	3	1617
<b>Inhambupe</b>	1138002	-11,783	-38,348	158	3	916
<b>Esplanada</b>	1237000	-12,058	-37,756	34	3	1675
<b>Camaçari</b>	1238000	-12,570	-38,057	34	3	1639
<b>Araças</b>	1238010	-12,210	-38,201	80	3	1290
<b>Alagoinhas</b>	1238042	-12,251	-38,515	130	3	1157
<b>Camaçari</b>	1238046	-12,603	-38,138	13	3	1579
<b>Teodoro Sampaio</b>	1238051	-12,300	-38,644	116	3	1098
<b>Wenceslau Guimarães</b>	1339000	-13,601	-39,492	171	3	1477
<b>Ituberá</b>	1339012	-13,811	-39,169	114	3	1759
<b>Aratuípe</b>	1339020	-13,567	-39,002	16	3	1953
<b>Wenceslau Guimarães</b>	1339027	-13,685	-39,479	135	3	1132
<b>Mutuípe</b>	1339030	-13,225	-39,504	212	3	956
<b>Nazaré</b>	1339034	-13,031	-39,032	35	3	1640
<b>Valença</b>	1339038	-13,367	-39,077	59	3	2033
<b>Ibicaraí</b>	1439001	-14,810	-39,471	151	3	1365
<b>Floresta Azul</b>	1439002	-14,860	-39,658	210	3	996
<b>Gongogi</b>	1439006	-14,327	-39,469	88	3	996
<b>Ipiaú</b>	1439014	-14,171	-39,690	142	3	932
<b>Itajuípe</b>	1439023	-14,678	-39,389	107	3	1488
<b>São José da Vitória</b>	1539002	-15,086	-39,342	153	3	1391
<b>Itapebi</b>	1539006	-15,961	-39,526	80	3	1133
<b>Mascote</b>	1539010	-15,564	-39,302	41	3	1409
<b>Camacan</b>	1539014	-15,603	-39,517	87	3	1107
<b>Santa Cruz Cabralia</b>	1639000	-16,264	-39,581	180	3	1103
<b>Prado</b>	1739001	-17,350	-39,220	6	3	1402
<b>Itamaraju</b>	1739005	-17,045	-39,545	80	3	1228
<b>Nova Viçosa</b>	1739006	-17,808	-39,663	59	3	1287
<b>Alcobaça</b>	1739010	-17,516	-39,228	10	3	1515
<b>Vereda</b>	1740006	-17,188	-39,982	1	3	1126
<b>Itanhém</b>	1740008	-17,163	-40,327	1	3	1107

<b>Eunapólis</b>	9	-16,290	-39,580	152	3	932
<b>Angical</b>	1144014	-11,984	-44,711	447	4	891
<b>Formosa do Rio Preto</b>	1145001	-11,048	-45,202	489	4	963
<b>Formosa do Rio Preto</b>	1145004	-10,993	-45,527	451	4	931
<b>Barreiras</b>	1145013	-11,896	-45,612	713	4	1117
<b>Barreiras</b>	1145014	-11,853	-45,122	458	4	927
<b>Barreiras</b>	1244011	-11,154	-45,009	444	4	933
<b>São Desidério</b>	1244019	-12,389	-44,932	502	4	1106,
<b>Barreiras</b>	1245004	-12,135	-45,104	490	4	855
<b>São Desidério</b>	1245005	-12,411	-45,120	502	4	993
<b>São Desidério</b>	1245007	-12,431	-45,086	512	4	1064
<b>Correntina</b>	1346006	-13,752	-46,140	947	4	1166
<b>Correntina</b>	1346007	-13,330	-46,062	824	4	1135

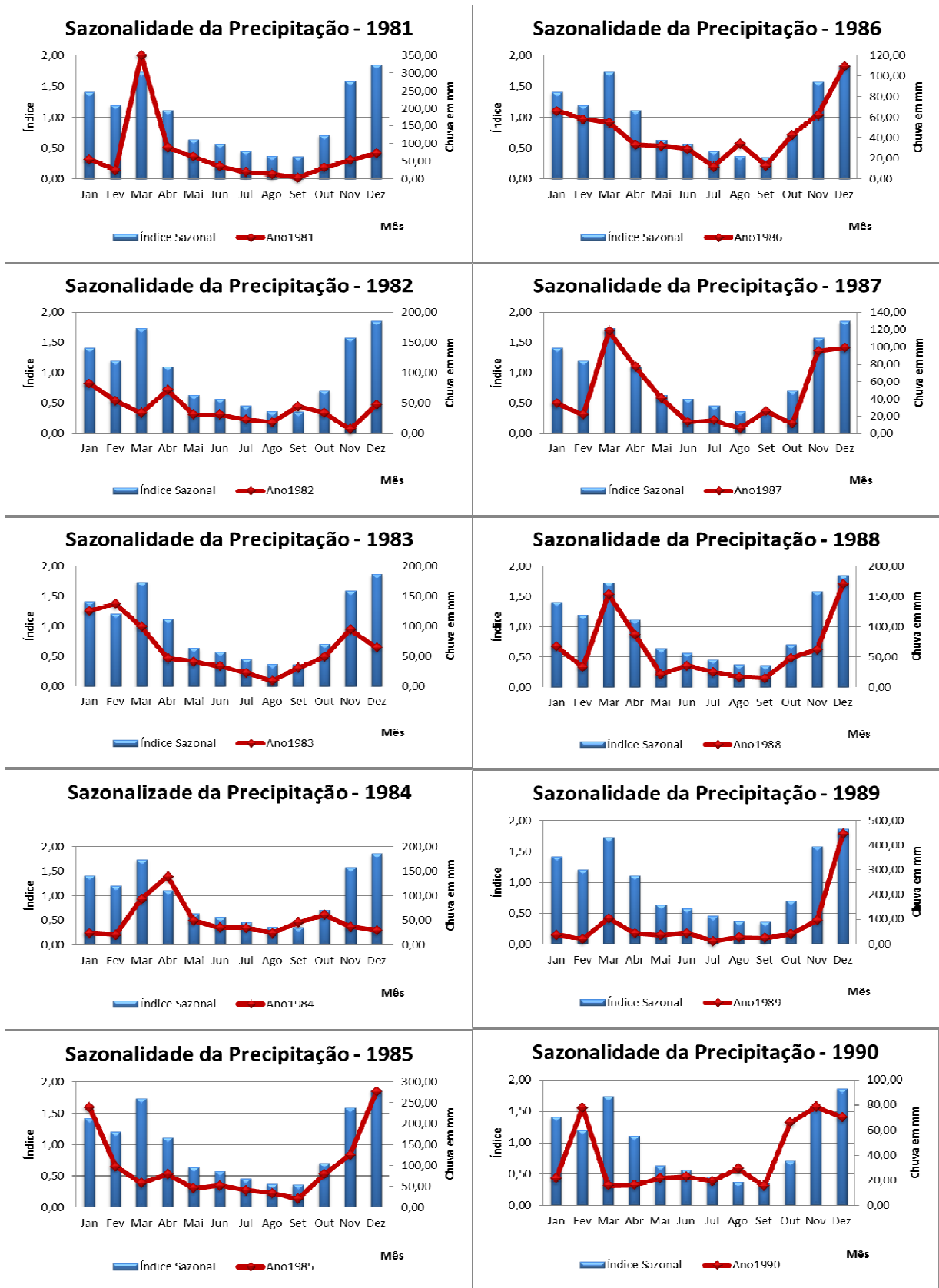
### APÊNDICE III – Sazonalidade Cluster 0



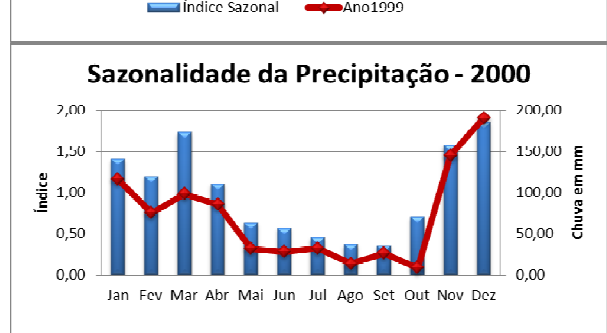
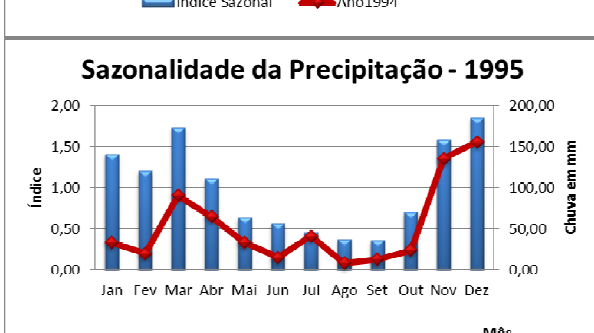
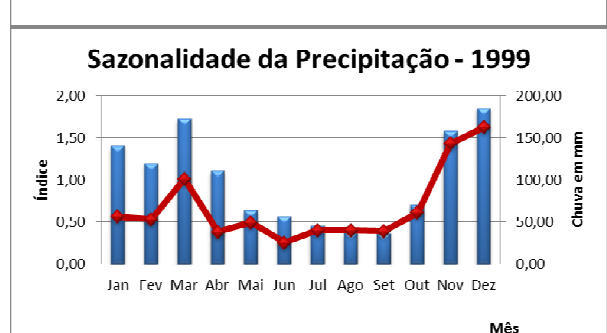
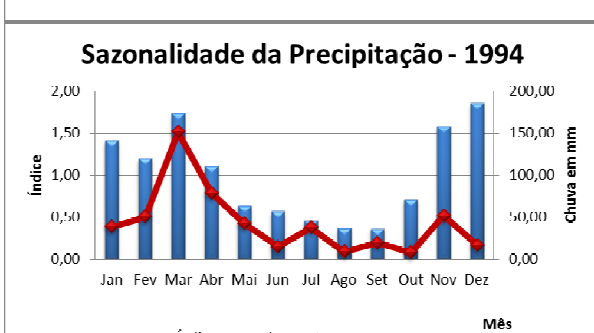
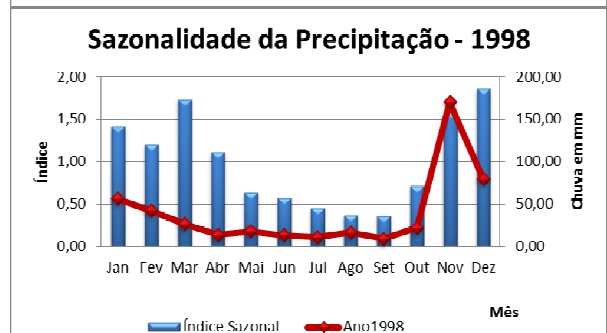
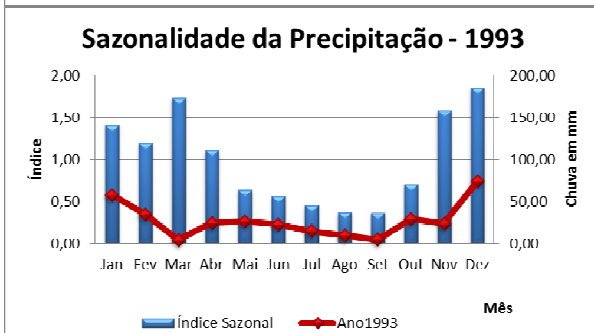
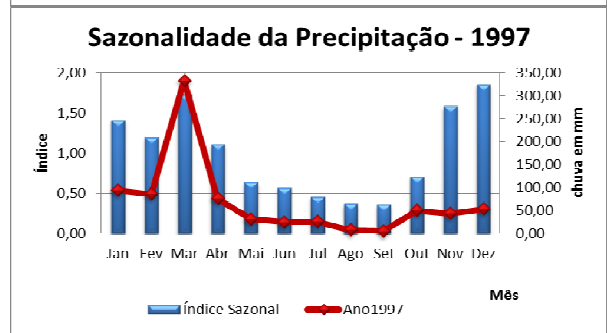
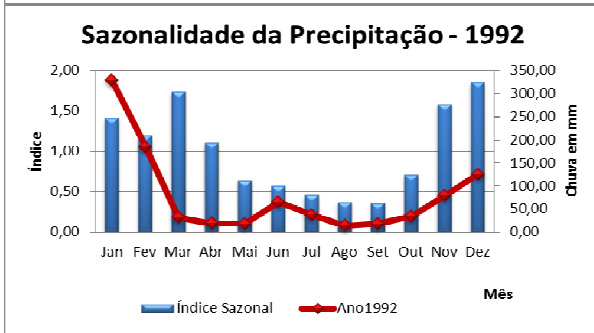
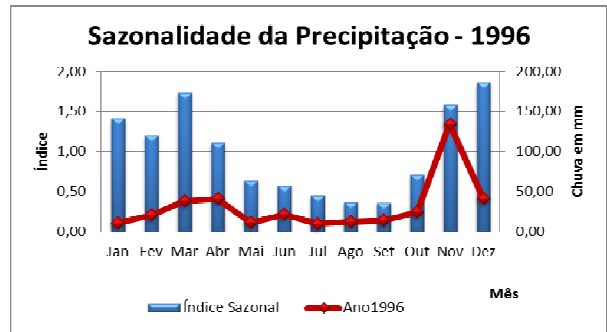
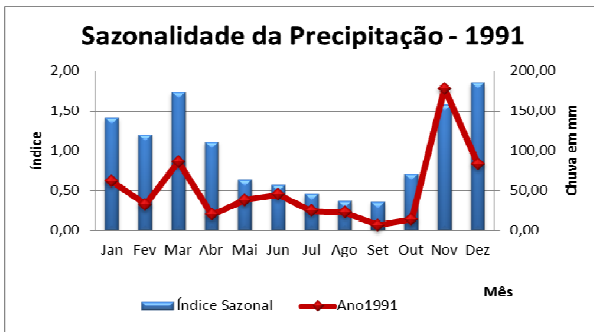


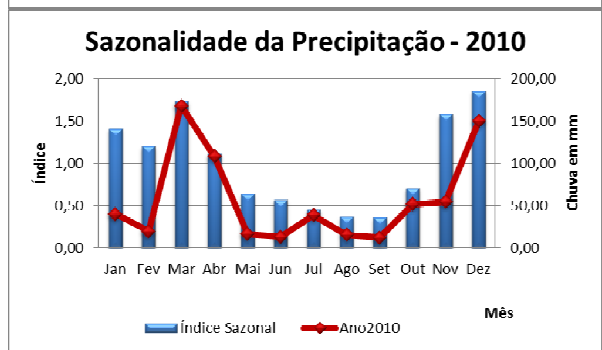
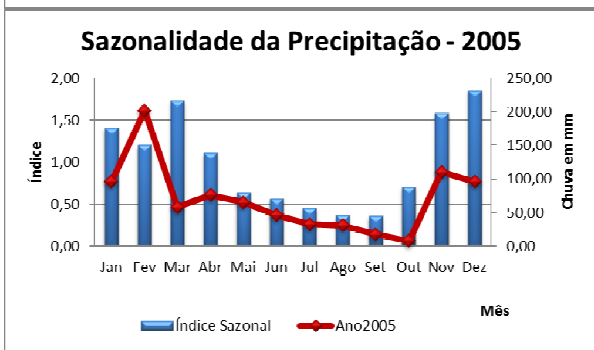
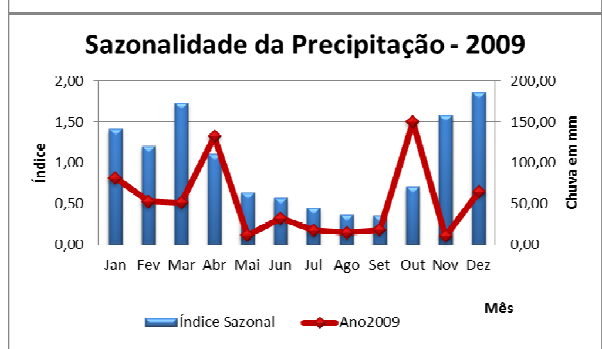
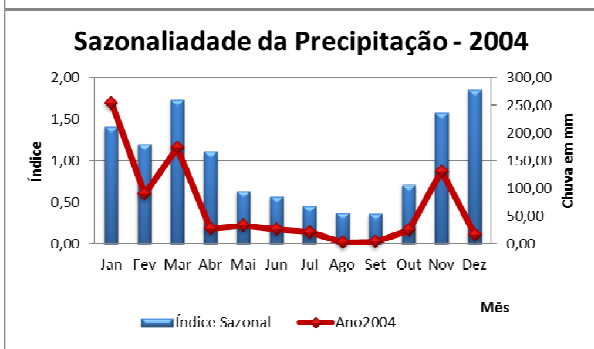
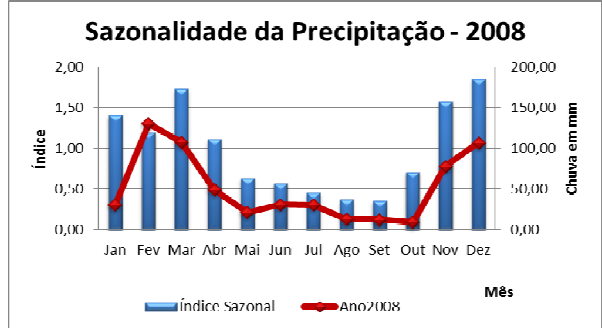
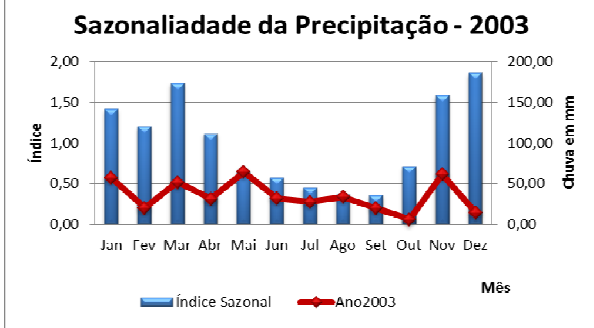
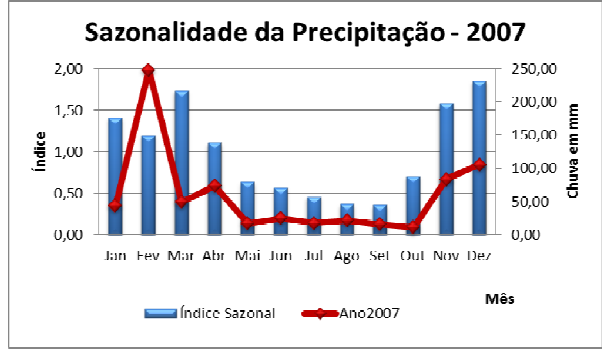
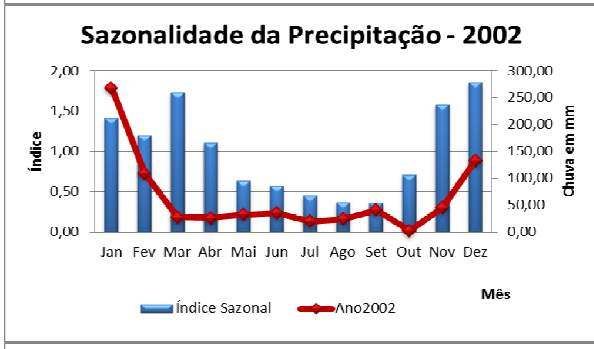
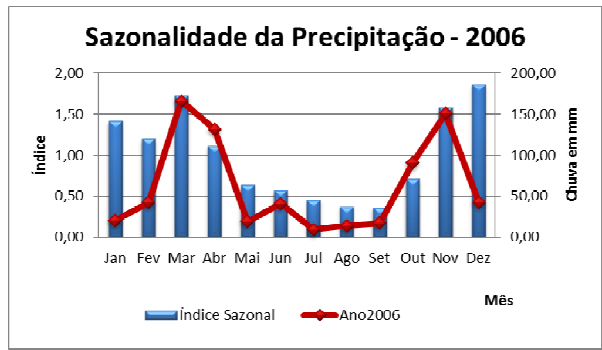
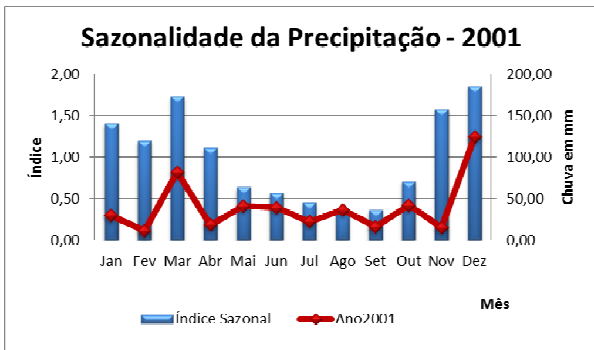


## APÊNDICE IV – Sazonalidade *Cluster 1*

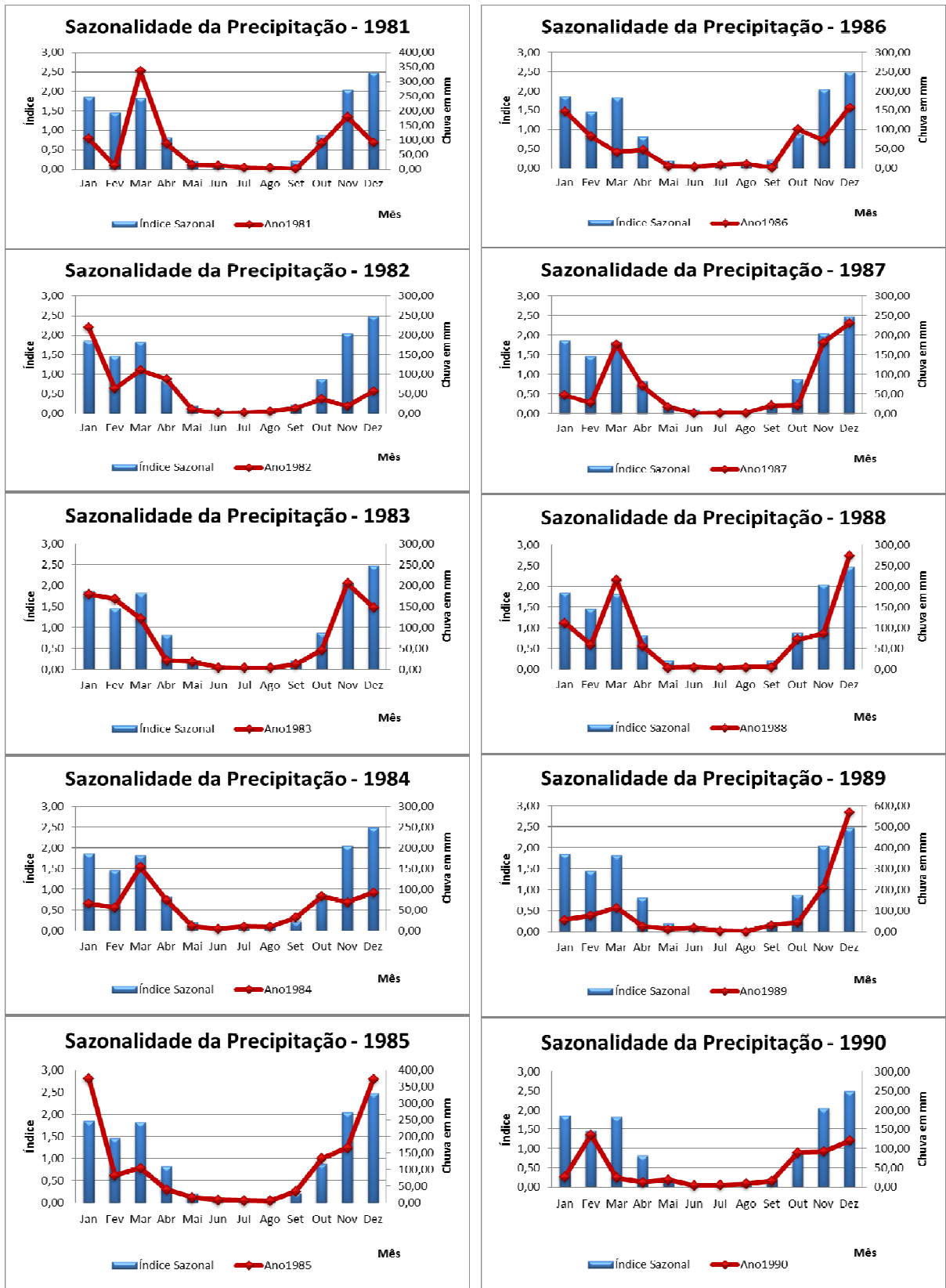


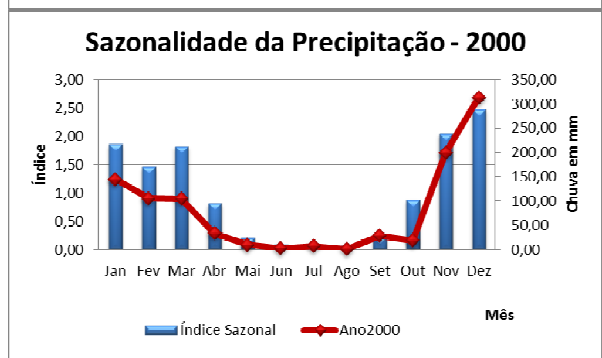
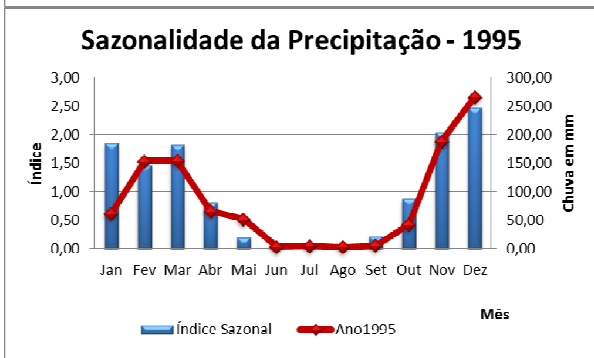
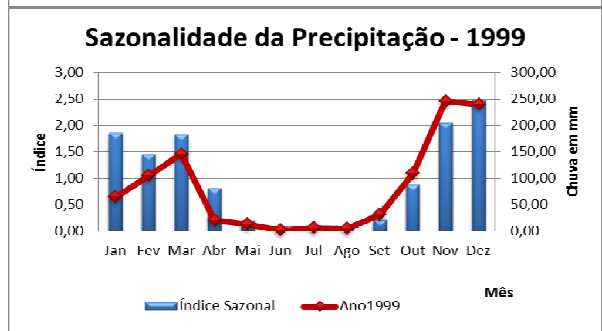
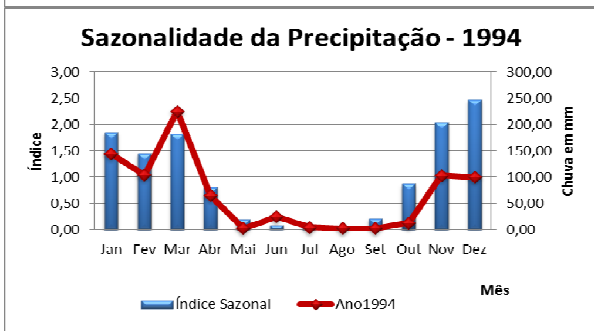
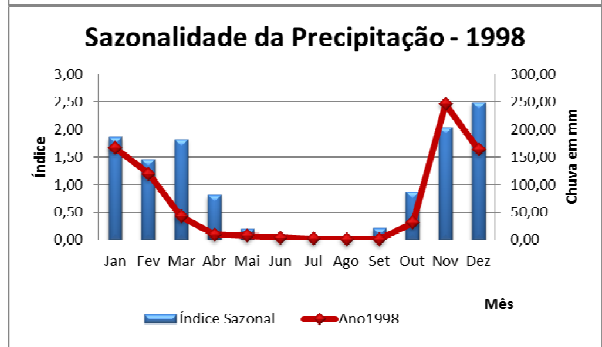
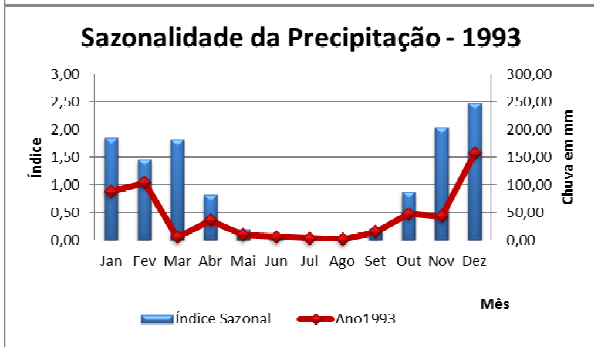
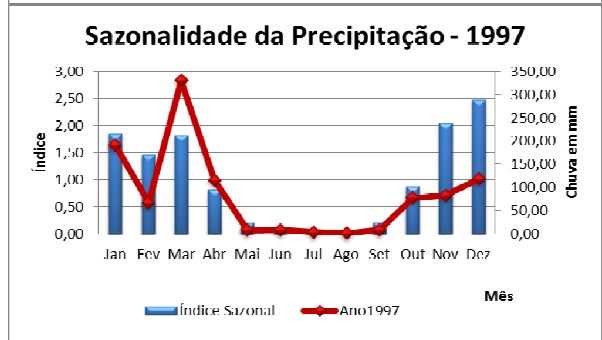
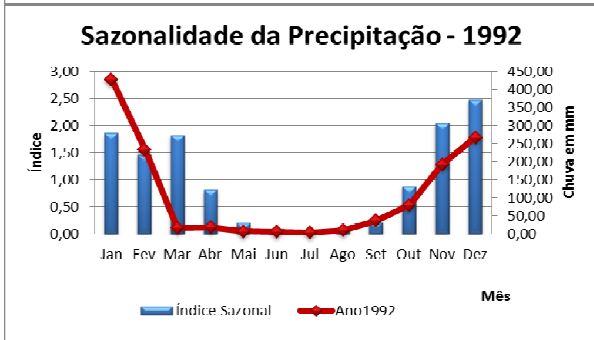
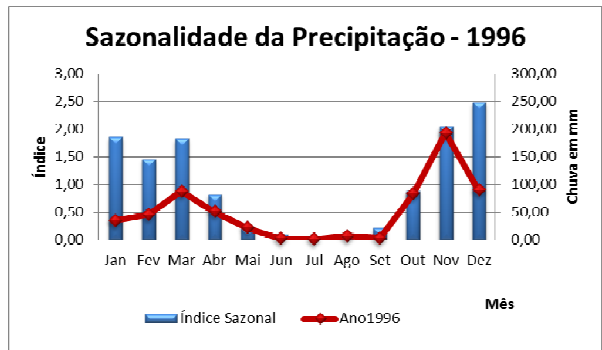
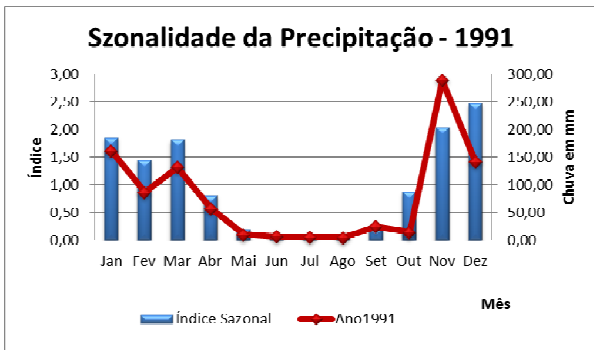


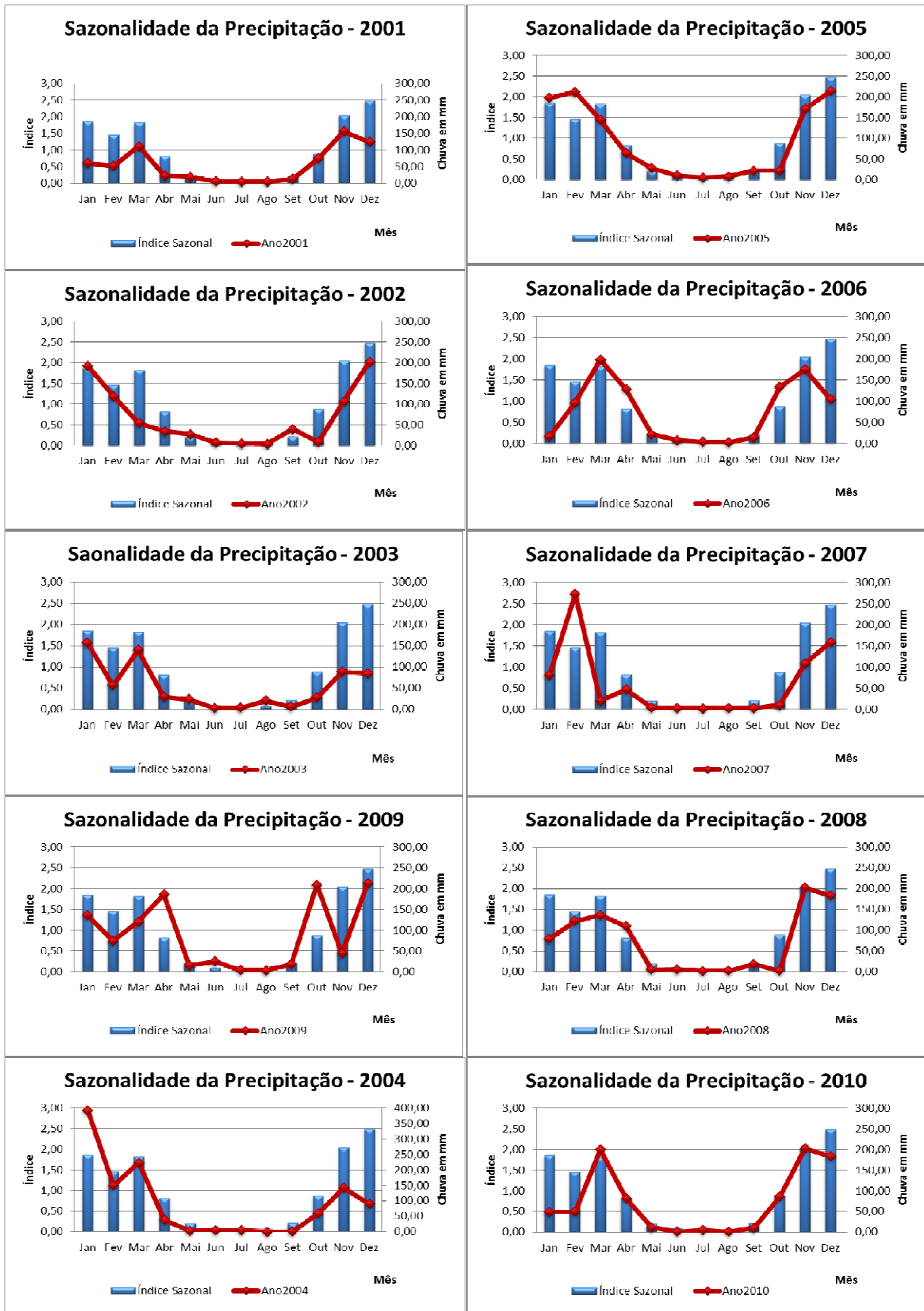




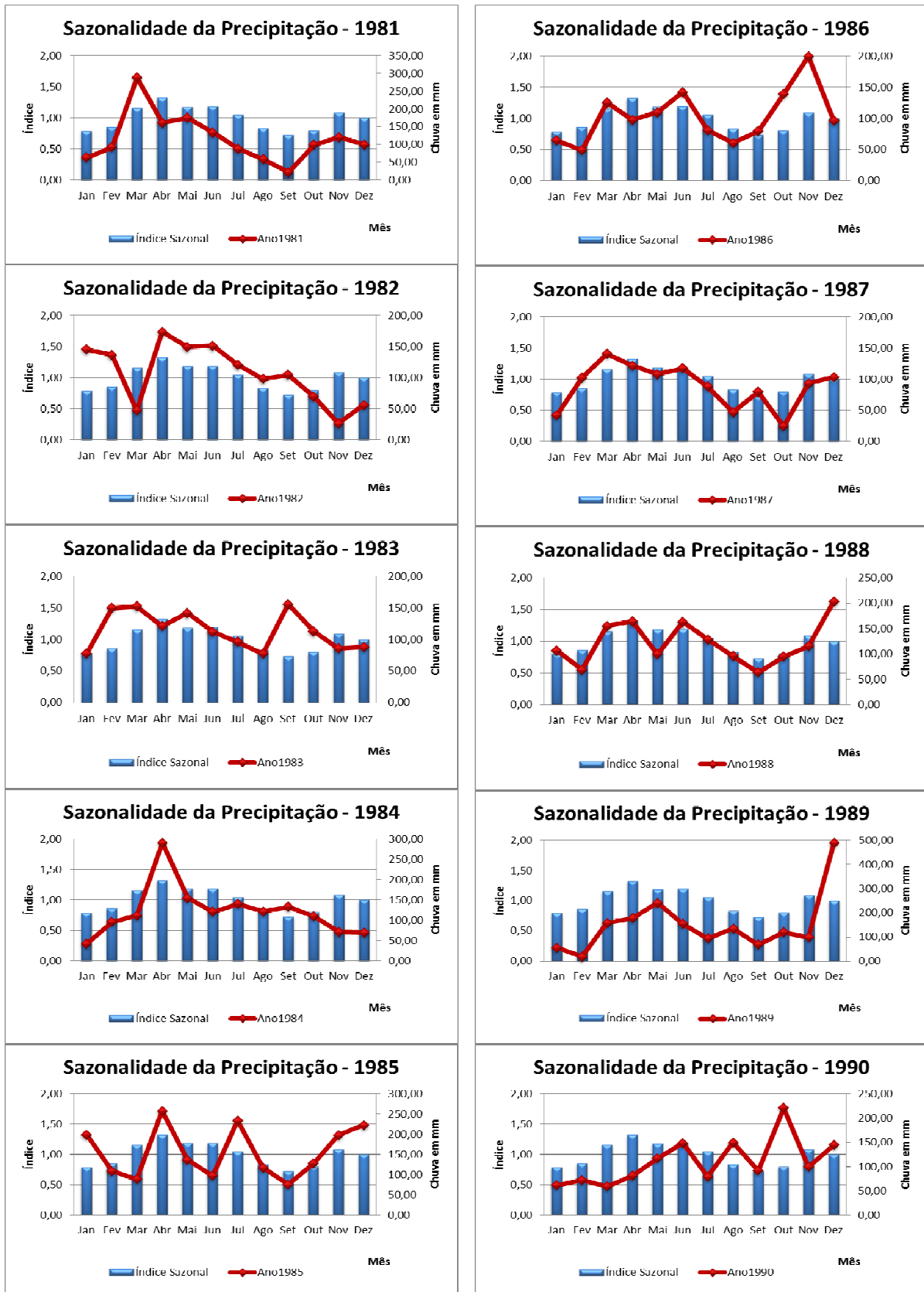
## APÊNDICE V – Sazonalidade *Cluster 2*

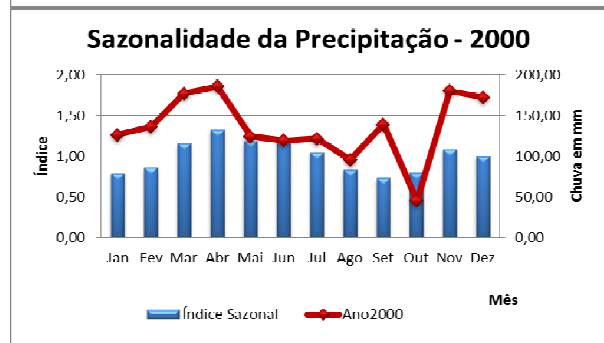
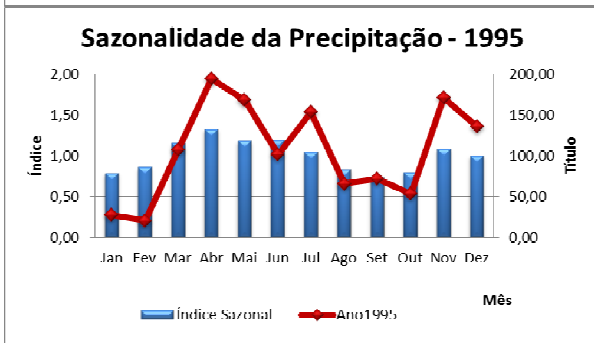
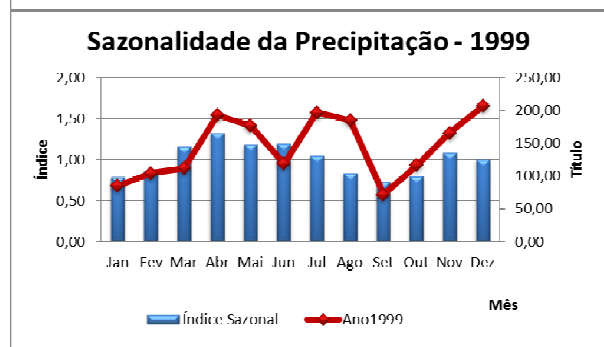
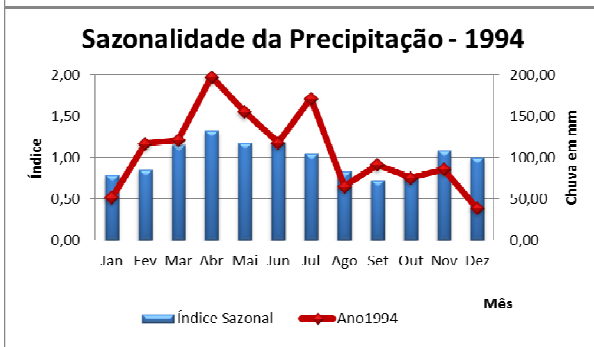
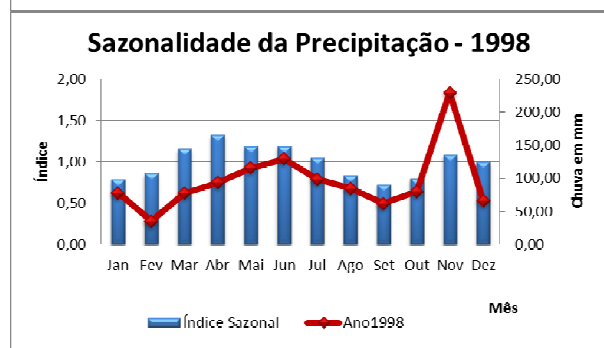
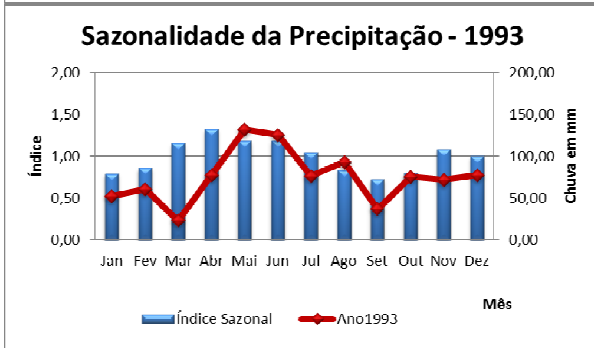
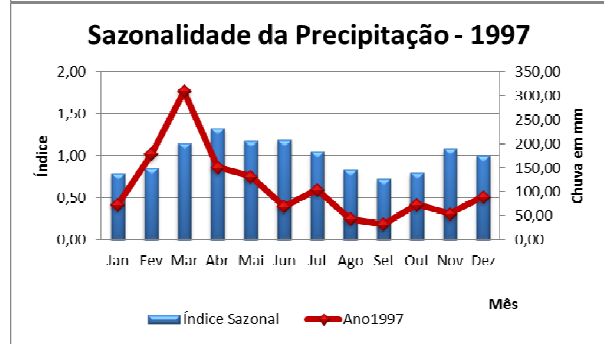
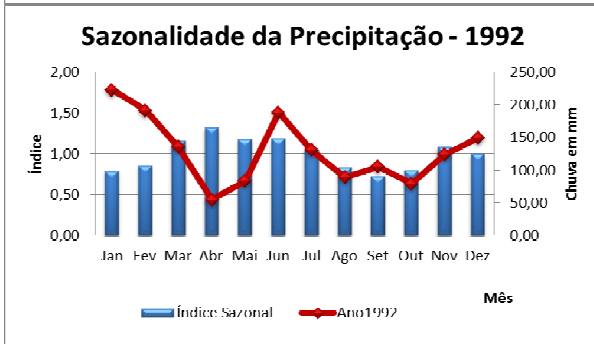
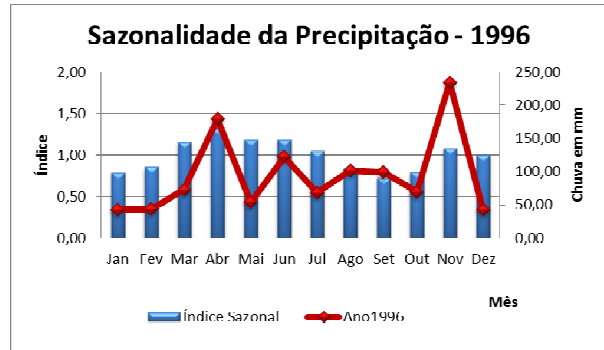
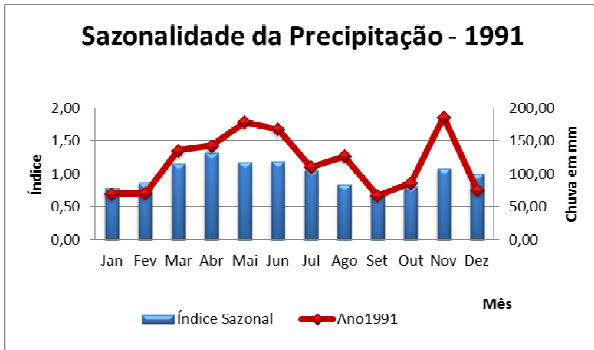


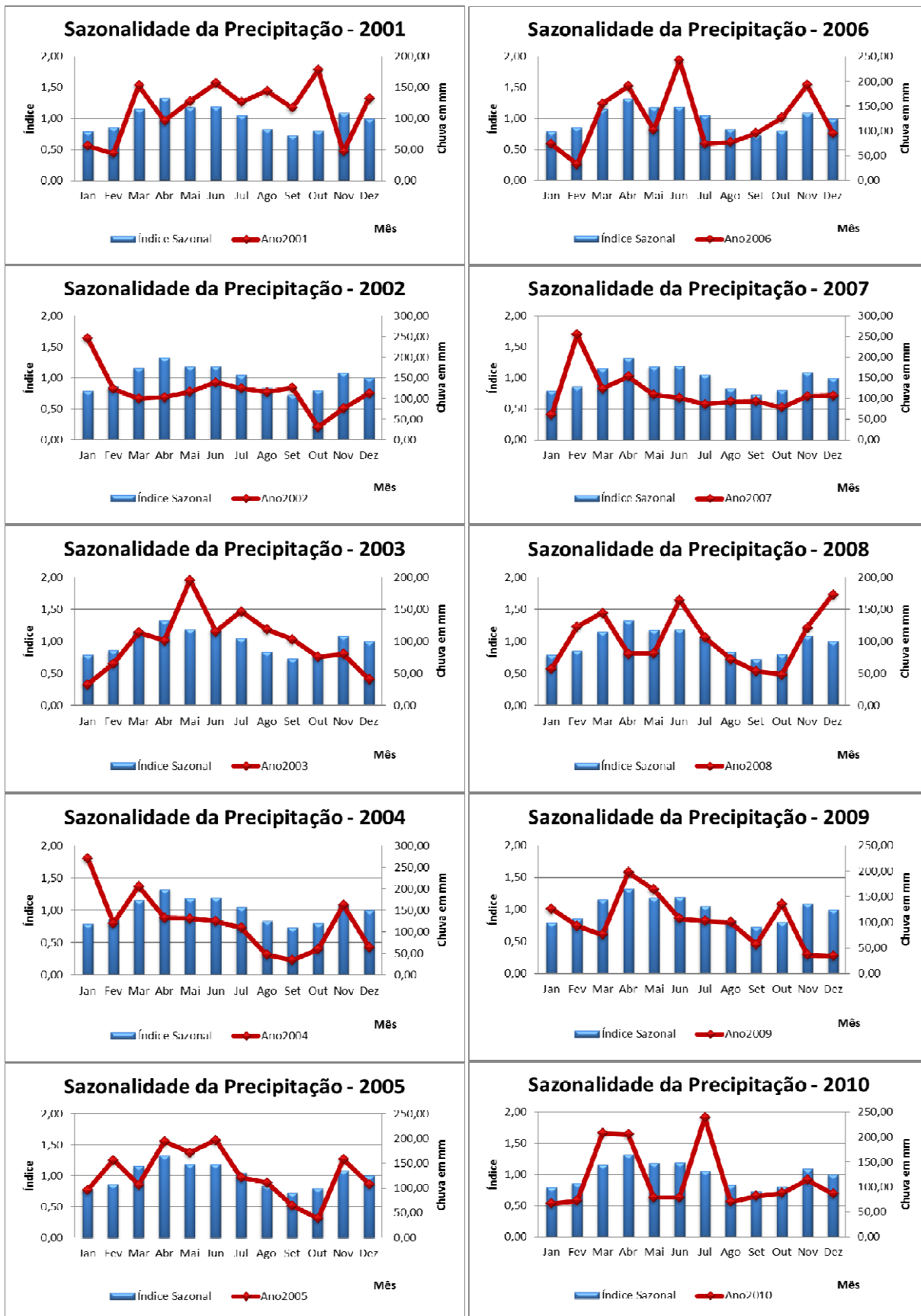




## APÊNDICE VI- Sazonalidade *Cluster 3*









## APÊNDICE VII – Sazonalidade *Cluster 4*

