

UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA
DEPARTAMENTO DE CIÊNCIAS EXATAS
MESTRADO EM MODELAGEM EM CIÊNCIAS DA TERRA E DO
AMBIENTE

CHARLES NOVAES DE SANTANA

**O uso de ferramentas fractais e redes
complexas no estudo da variabilidade
pluviométrica do Nordeste do Brasil**

Dissertação apresentada como requisito parcial
para a obtenção do grau de
Mestre em Ciências Ambientais

Prof. Dr. José Garcia Vivas Miranda
Orientador

Prof^a . Dr^a . Joselisa Maria Chaves
Co-orientador

Feira de Santana, janeiro de 2008

CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Santana, Charles Novaes de

O uso de ferramentas fractais e redes complexas no estudo da variabilidade pluviométrica do Nordeste do Brasil / Charles Novaes de Santana. – Feira de Santana: PPGM da UEFS, 2008.

84 f.: il.

Dissertação (mestrado) – Universidade Estadual de Feira de Santana. Mestrado em Modelagem em Ciências da Terra e do Ambiente, Feira de Santana, BR-BA, 2008. Orientador: José Garcia Vivas Miranda; Co-orientador: Joselisa Maria Chaves.

1. Variabilidade pluviométrica. 2. Redes complexas. 3. Fractais. I. Miranda, José Garcia Vivas. II. Chaves, Joselisa Maria. III. O uso de ferramentas fractais e redes complexas no estudo da variabilidade pluviométrica do Nordeste do Brasil.

UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA

Prof. José Carlos Barreto de Santana: Reitor

Profa. Eneida de Moraes Marcílio Cerqueira: Pró-Reitora de Pesquisa e Pós-Graduação

Profa. Joselisa Maria Chaves: Diretor do Departamento de Ciências Exatas

Profa. Marjorie Csekö Nolasco: Coordenadora do Colegiado do PPGM

“É impossível explicar honestamente as belezas contidas nas leis da natureza, de uma forma que as pessoas possam senti-las, sem que elas tenham uma boa compreensão da Matemática.”

— RICHARD FEYNMANN

AGRADECIMENTOS

Se apoiar em ombros de gigantes é uma regra pra quem faz ciência. Portanto, agradeço a todos que disponibilizaram bibliografias sobre climatologia e sobre as ferramentas da física estudadas, bem como à ADENE, à ANA e ao NCAR pela disponibilização dos dados utilizados.

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo financiamento da bolsa de Mestrado, e à Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB) pelo financiamento do projeto de Mestrado, ainda que problemas políticos tenham atrasado a efetivação da ajuda desta Fundação.

Agradeço ao *MAIS UEFS*, que deu provas vivas de que o mundo pode ser mudado, e que a academia é um dos lugares onde isso pode começar a acontecer.

Não posso deixar de agradecer a todos do PPGM, meus colegas estudantes (Elane, Danilo, João, Geovana, Ozineide, Clea, David), aos incríveis funcionários do Programa (Edson, Chawelly, Juciane) que fizeram mais que o possível para dar conta das infinitas burocracias de um Mestrado tão novo como o nosso, aos professores (Barros, Uchoa, Malena, Nildon), e em especial a Joselisa, por ter me *adotado* nos primeiros momentos nesta nova casa, a Washington, por ter me feito acreditar que as *possibilidades* (de aprovação, de bolsa, de intercâmbio, ...) se tornariam realidade, e a Marjorie, que se mostrou incansável na luta por um Programa de qualidade, principalmente no ano de 2007.

Agradeço a todos que ajudaram a construir este trabalho, com críticas e sugestões, desde os primeiros Congressos dos quais participamos. Em especial, a meus grandes amigos do INPE, Marina e Rogério, aos amigos do Grupo de Meteorologia da Universitat de les Illes Balears, Romualdo e Victor Homar, aos Professores Clemente Tanajura, Rosane Chaves e Humberto Barbosa, por palavras de incentivo e esclarecimento a este humilde *forasteiro* na Meteorologia; e aos membros da banca de qualificação, Thierry Lobão e Emanuel Reis, pelas enriquecedoras contribuições.

Agradeço aos meus amigos de UFBA. Aos amigos do curso de Ciência da Computação, uma *galera-do-bem* que marcou minha vida. Aos meus alunos da UFBA, que suportaram com paciência e dedicação as dificuldades de ter um professor-mestrando. E a todos os membros do Grupo FESC e *adjacências* (GRH, LFNA, CONES), que tornam reuniões de orientação momentos incrivelmente divertidos e transformadores usando todos os recursos da famosa *Multidisciplinaridade*.

Por fim, mas nunca menos importante, agradeço a minha família. Aos inúmeros primos, primas, sobrinhos, sobrinhas, irmãos, irmã, amigos-irmãos, que participaram de cada etapa da realização deste sonho. Em especial a Garcia: o amigo mais orientador que eu já tive; Neno: meu pai, e maior orgulho que tenho na vida; minha mãe: minha maior incentivadora e torcedora oficial, a quem todo o amor que eu

puder retribuir será infinitamente menor ao que ela merece; Alex: aquele que mais sofreu, certamente, ao ter que suportar todos os dias os principais efeitos-colaterais que um mestrado pode trazer na vida de uma pessoa; e a minha *noivinha-quase-marida*, Isis, que a cada dia me faz confundir os sonhos individuais com os coletivos.

SUMÁRIO

LISTA DE FIGURAS	8
LISTA DE TABELAS	12
LISTA DE ABREVIATURAS E SIGLAS	13
RESUMO	14
ABSTRACT	16
1 INTRODUÇÃO	17
2 FUNDAMENTAÇÃO TEÓRICA	19
2.1 Características climáticas do Nordeste Brasileiro (NEB)	19
2.2 Correlações estatísticas	21
2.3 Redes complexas	24
2.3.1 Teoria dos Grafos	24
2.3.2 Teoria de Redes Complexas	29
2.4 Geometria Fractal	33
2.4.1 Dimensão Fractal	34
2.5 Trabalhos anteriores sobre Sistemas Complexos aplicados à análise da pluviometria	38
2.5.1 Correlação temporal da chuva no Nordeste	38
2.5.2 Correlação espacial da chuva no Nordeste	39
2.6 Trabalhos sobre modelagem em Redes Complexas	39
2.6.1 Correlação espaço-temporal de terremotos	39
2.6.2 Correlação espaço-temporal entre neurônios	40
3 MATERIAIS E MÉTODOS	41
3.1 Origem dos dados	41
3.2 Tratamento dos dados	42
3.3 Estratégias de construção das redes	44
3.3.1 Limitação espacial	44
3.3.2 Correlação linear	44
3.4 Análise das Redes	46
3.4.1 Análise visual	46
3.4.2 Análise topológica (estatística)	46
3.4.3 Análise de características Fractais	46
3.5 Resumo metodológico	47

4	LIMITAÇÃO ESPACIAL - LE	49
4.1	Análise visual	49
4.2	Análise topológica (estatística)	53
5	CORRELAÇÃO LINEAR ESTATÍSTICA - CLE	57
5.1	Análise das características Fractais	58
5.2	Identificação da correlação crítica	64
5.2.1	Geração das redes complexas	67
5.2.2	Cálculo dos índices caracterizadores das redes	67
5.2.3	Identificação e análise das redes críticas	72
5.3	Análise Topológica das redes críticas.	77
6	CONCLUSÕES	80
	REFERÊNCIAS	82

LISTA DE FIGURAS

Figura 2.1: Exemplo de uma rede complexa	24
Figura 2.2: Ilustração representando o problema das pontes de Königsberg. . .	25
Figura 2.3: Grafo G_1 , representado pela matriz de adjacências M_1	25
Figura 2.4: Matriz de adjacências M_1 , representativa do grafo G_1	25
Figura 2.5: Exemplo de um grafo de ordem 2 e tamanho 1	26
Figura 2.6: O grau do vértice B (k_B) é igual a 4 e o grau médio ($\langle k \rangle$) do grafo é igual a 2.8	26
Figura 2.7: Configuração de um grafo conexo.	27
Figura 2.8: Configuração de um grafo desconexo, composto de 2 componentes. . .	27
Figura 2.9: Exemplo de busca de um caminho mínimo em um grafo	27
Figura 2.10: Tomando-se todos os caminhos mínimos existentes neste grafo (G), aquele de maior tamanho, e que representa o diâmetro do grafo, é o caminho mínimo entre os vértices A e C , que vale 3. . .	28
Figura 2.11: Configuração de um grafo completo	28
Figura 2.12: Clique formado pelo subgrafo BDE	28
Figura 2.13: Rede aleatória	30
Figura 2.14: Distribuição de Graus dos vértices de uma rede seguindo uma distribuição de Poisson	30
Figura 2.15: Exemplo de uma rede de mundo pequeno.	31
Figura 2.16: Os vértices vizinhos do vértice D formam um clique. Por isso, o CA de D vale 1	32
Figura 2.17: O vértice B do grafo possui 4 vértices adjacentes. Para que estes vizinhos formassem um clique, seriam necessárias 6 arestas conectando-os. Como apenas dois deles formam um clique (aresta entre os vértices D e E), o CA de B vale $\frac{1}{6}$	32
Figura 2.18: Exemplo de uma rede livre de escala. FONTE(ALBERT; JEONG; BARABÁSI, 1999)	33
Figura 2.19: Distribuição de graus em lei de potência. FONTE(ALBERT; JEONG; BARABÁSI, 1999)	33
Figura 2.20: Exemplo de similaridade	34
Figura 2.21: Exemplo de escala para medida de comprimento	35
Figura 2.22: Exemplo de escala para medida de área	35
Figura 2.23: Quatro gerações da curva de Koch. No limite de infinitas gerações torna-se um fractal FONTE: (MIRANDA, 1997)	37
Figura 3.1: Mapa da Bahia contendo as 514 estações da ADENE	41
Figura 3.2: Mapa do NE contendo as 1400 estações do NCAR	42

Figura 3.3:	Exemplo de um arquivo do tipo <i>FPP</i>	43
Figura 3.4:	Histograma com as Estações da ADENE que contêm dados de chuva para cada dia do intervalo de 1979 a 2000.	43
Figura 3.5:	Histograma com as Estações do NCAR que contêm dados de chuva para cada dia do intervalo de 1904 a 1983	43
Figura 3.6:	Estudo das características fractais da rede, baseado na quantidade de conexões existentes a diferentes distâncias (raios)	47
Figura 3.7:	Fluxograma que mostra as etapas de construção e análise das redes complexas de estações pluviométricas	48
Figura 4.1:	Rede gerada a partir da análise de dados de chuva da ADENE para períodos de inverno.	50
Figura 4.2:	Rede gerada a partir da análise de dados de chuva da ADENE para períodos de primavera.	50
Figura 4.3:	Rede gerada a partir da análise de dados de chuva da ADENE para períodos de verão.	51
Figura 4.4:	Rede gerada a partir da análise de dados de chuva da ADENE para períodos de outono.	51
Figura 4.5:	Divisão da Região Sul do Nordeste em áreas homogêneas. As regiões A, B, E, F, G são similares, apresentando ciclo anual bem definido, com período chuvoso em dezembro e março e alta variabilidade pluviométrica ao longo do ano - FONTE: (CHAVES, 1999).	52
Figura 4.6:	Distribuição sazonal do número de vértices das redes de estações pluviométricas.	53
Figura 4.7:	Distribuição sazonal do Grau médio das redes de estações pluviométricas.	54
Figura 4.8:	Distribuição sazonal do Caminho Mínimo Médio das redes de estações pluviométricas.	55
Figura 4.9:	Distribuição sazonal do Diâmetro das redes de estações pluviométricas.	55
Figura 4.10:	Relação entre Caminho mínimo médio e diâmetro.	56
Figura 4.11:	Distribuição sazonal do Coeficiente de Aglomeração Médio das redes de estações pluviométricas.	56
Figura 5.1:	Distribuição dos tamanhos das conexões (em Km) de redes construídas pelo método <i>CLE</i> , com $r_{limite} = 0,7$, para diferentes meses do ano de 1979.	57
Figura 5.2:	Cálculo da distribuição de arestas de uma rede completa com 514 vértices, para diferentes distâncias - expoente $b = 1,98942$ com $erro = 0,01003$	59
Figura 5.3:	Exemplo de uma rede com vértices distribuídos em forma de Reticulado, completa e regular, com 25 vértices.	59
Figura 5.4:	Cálculo da distribuição de arestas de um Reticulado com 529 vértices, para diferentes distâncias - expoente $b = 1,885$ com $erro = 0,05925$	60
Figura 5.5:	Cálculo da distribuição de arestas da rede de dados da Bahia para janeiro de 79, 80, 81, 82 e 83	61

Figura 5.6: Cálculo da distribuição de arestas da rede de dados do Norte do Nordeste para janeiro de 79, 80, 81, 82 e 83	62
Figura 5.7: Valores dos expoentes das distribuições de arestas das redes de dados da Bahia e do Norte do Nordeste para janeiro de 79, 80, 81, 82 e 83	62
Figura 5.8: Cálculo da distribuição de arestas da rede relativa a dados de todo o Nordeste para janeiro de 79, 80, 81, 82 e 83	63
Figura 5.9: Sobreposição dos valores dos expoentes (com respectivos valores de <i>erro</i>) das distribuições de arestas das redes relativas a dados da Bahia e do Nordeste inteiro para janeiro de 79, 80, 81, 82 e 83	64
Figura 5.10: Variação do número de arestas com o crescimento do r_{limite}	65
Figura 5.11: Exemplo de rede completa, com 3 vértices e Caminho Mínimo Médio igual a 1	65
Figura 5.12: Exemplo de rede com 3 vértices e Caminho Mínimo Médio igual a 1,33	66
Figura 5.13: Exemplo de rede conexa, com 5 vértices e Caminho Mínimo Médio igual a 1,7	66
Figura 5.14: Exemplo de rede desconexa, com 5 vértices e Caminho Mínimo Médio igual a 1	66
Figura 5.15: Redes de estações pluviométricas da região Nordeste, geradas para o mês de janeiro de 1979 com r_{limite} igual a 0,55, 0,65, 0,75, 0,85.	67
Figura 5.16: Comportamento do $\langle l \rangle$, do $\langle K \rangle$ e do $\langle C \rangle$ para redes de estações pluviométricas da região Nordeste, para os meses de Janeiro a Abril de 1979, geradas com diferentes valores de r	68
Figura 5.17: Comportamento do $\langle l \rangle$, do $\langle K \rangle$ e do $\langle C \rangle$ para redes de estações pluviométricas da região Nordeste, para os meses de Maio a Agosto de 1979, geradas com diferentes valores de r	69
Figura 5.18: Comportamento do $\langle l \rangle$, do $\langle K \rangle$ e do $\langle C \rangle$ para redes de estações pluviométricas da região Nordeste, para os meses de Setembro a Dezembro de 1979, geradas com diferentes valores de r	70
Figura 5.19: Valores do $\langle l \rangle$ para o mês de Abril, nos anos de 1979 a 1982, relacionado com diferentes valores de r	72
Figura 5.20: Valores do $\langle l \rangle$ para o mês de Abril, nos anos de 1979 a 1982, relacionado com diferentes valores de r	73
Figura 5.21: Valores do $\langle l \rangle$ para o período de Novembro, Dezembro e Janeiro, definido pelas médias, relacionado com diferentes valores de r	74
Figura 5.22: Valores do $\langle l \rangle$ para o período de Fevereiro, Março e Abril, definido pelas médias, relacionado com diferentes valores de r	74
Figura 5.23: Valores do $\langle l \rangle$ para o período de Maio, Junho e Julho, definido pelas médias, relacionado com diferentes valores de r	75
Figura 5.24: Valores do $\langle l \rangle$ para o período de Agosto, Setembro e Outubro, definido pelas médias, relacionado com diferentes valores de r	75
Figura 5.25: Relação entre $\langle l \rangle$ e D dos meses de Janeiro e Agosto dos anos de 1979 a 1982	77
Figura 5.26: Relação entre $\langle C \rangle$ e l dos meses de Janeiro e Agosto dos anos de 1979 a 1982	78

Figura 5.27: Relação entre $\langle C \rangle$ e K dos meses de Janeiro e Agosto dos anos de 1979 a 1982 79

LISTA DE TABELAS

Tabela 5.1: Expoentes das distribuições de arestas das redes relativas a dados do Sul (Bahia), do Norte e de toda a região Nordeste	60
Tabela 5.2: Valores do $r_{critico}$ para cada mês do ano.	76

LISTA DE ABREVIATURAS E SIGLAS

ANA	Agência Nacional de Águas
ADENE	Agência Nacional para o Desenvolvimento do Nordeste
NCAR	National Center for Atmospheric Research
NEB	Nordeste Brasileiro
VCAN	Vórtice Ciclônico de Altos Níveis
ZCAS	Zona de Convergência do Atlântico Sul
ZCIT	Zona de Convergência Intertropical

RESUMO

Climaticamente, a região Nordeste do Brasil é marcada pela predominância de clima semi-árido, caracterizado por períodos de secas severas intercalados por períodos de chuvas intensas. Eventos climáticos como a chuva resultam da interação de vários fenômenos físicos que, em sua maioria, pode ser descrita individualmente pelas leis básicas da mecânica e termodinâmica de forma satisfatória. Por esse motivo, um imenso progresso tem sido observado, nos últimos anos, com relação à previsão de tempo e clima utilizando algoritmos mais precisos em recursos computacionais de larga escala. Estes algoritmos levam em consideração as variáveis que são relevantes para a circulação atmosférica e oceânica além de uma grande quantidade de dados físicos obtidos de um conjunto denso de estações distribuídas ao redor do mundo. Com objetivo de prover a interpretação dos dados destes algoritmos, é necessário proceder com análises sofisticadas dos dados armazenados e simulados, como correlações estatísticas temporais e espaciais, propriedades de escalas, propriedades topológicas da distribuição espacial de eventos, etc. Os resultados falam sobre a relevância estatística dos dados, efeitos locais e globais, padrões típicos e outros recursos relacionados ao fenômeno. Neste trabalho, nós exploramos o uso da Teoria de Redes Complexas para a análise e interpretação de eventos de chuva, baseando-nos somente em registros de dados de um conjunto de estações pluviométricas da região Nordeste do Brasil. Este método é inspirado em uma proposta para caracterizar sequências de eventos sísmicos, eventos em que, assim como no fenômeno das chuvas, a grande quantidade de variáveis físicas envolvidas motiva a análise usando métodos da Teoria de Sistemas Complexos. Os nós das redes geradas são as estações meteorológicas onde há dados de chuva no período analisado, enquanto as arestas são criadas de acordo com critérios de correlação temporal e espacial entre eventos de chuva ocorridos em diferentes estações pluviométricas. Calculamos os índices mais comuns de caracterização de redes complexas, tais como: diâmetro, caminho mínimo médio, coeficiente de aglomeração médio. As redes conectam estações a diferentes distâncias, e a fim de estudar a causalidade não-local desse fenômeno foram calculados índices fractais de caracterização. Os valores de diâmetro e de caminho mínimo médio são menores para os meses de inverno e primavera, típicos de chuva mais localizada no litoral; enquanto que para os meses de verão e outono, típicos de chuva mais distribuída em toda a região, os valores são maiores. A dimensão fractal calculada para dados do Sul do Nordeste (Bahia) é semelhante à calculada para dados do Norte do Nordeste (demais estados da Região), mas ambas são diferentes das dimensões fractais de redes completas e regulares hipotéticas, o que demonstra que a distribuição das estações pluviométricas não é homogênea. Estes resultados sugerem o estudo mais aprofundado deste método de análise de dados pluviométricos, que,

a princípio, consegue identificar padrões pluviométricos através da modelagem em Sistemas Complexos.

Palavras-chave: Variabilidade pluviométrica, redes complexas, fractais.

**The use of fractal tools and complex networks to study rainfall
variability in Northeast Brazil**

ABSTRACT

Brazil Northeast's climate is usually described as semi-arid, characterized by hard dry seasons intermingled by hard rainfall seasons. Some areas of Northeast present annual pluviometric measure about 400 mm in mean, while in others the annual pluviometric measure is about 2000 mm. Rain events result from interplay of several physical phenomena, most of which can be individually described on the basic laws of mechanics and thermodynamics in a rather adequate way. Because of this, a huge progress has been achieved in recent years in relation to weather forecast with the use of very precise algorithms in large scale computing resources. They take into account the variables that are relevant for the atmospheric and ocean circulation and input of large amount of physical data obtained from a dense set of stations scattered around the world. In order to improve the interpretation of the accurate data resulting from the description of atmospheric phenomena and rain events, it is necessary to proceed with sophisticated analyses of recorded and simulated data, as spatial and temporal statistical correlations, scale properties, topological properties of spatial event distribution, ad so on. They indicate the extent of statistical relevance of the data, local and global effects, typical patterns, and other topological features related to the phenomena. In this work, we explore the usefulness of complex network framework for the analysis and understanding of rain events, based solely on recorded data from a set of stations in Northeast Brazil. The method is inspired on a proposal to characterize actual sequences of earthquake events where, like precipitation phenomena, the available data stems from complex systems with a very large number of physical variables. The potential network nodes are the meteorological stations where the rain events have been recorded, while the network edges are placed according to rules that take into account temporal and spatial correlation criteria between events occurring at different stations, for a time span as large as one month. We evaluate usual network properties based on diameter, node degrees, clustering coefficient, minimal inter-node distance along network edges. This allows for a characterization of networks based on seasonality and on spatial span of the region where the stations are distributed. The obtained results are discussed, taking into account the known precipitation patterns of the investigated region.

Keywords: rainfall variability, complex networks, fractals.

1 INTRODUÇÃO

O desenvolvimento da Computação nas últimas décadas permitiu (e ainda permite) o avanço de pesquisas em diversas outras áreas da ciência. O alto processamento, alta capacidade de armazenamento e transmissão de dados, a portabilidade dos equipamentos entre outros avanços, vêm mudando o modo de *fazer ciência*.

Entre as aplicações da computação às diversas áreas da ciência, as atividades de modelagem têm relativo destaque. Um modelo pode ser definido como uma forma para reproduzir, imitar ou simular algum objeto, fenômeno ou situação. No dia-a-dia, lidamos com modelos o tempo todo: nas artes (pintura, escultura, ...), na moda, nas engenharias, na matemática.

Modelar matematicamente significa criar uma imitação matemática para *coisas do mundo*. A modelagem matemática é uma etapa importante para representação computacional do *mundo real*. Para interpretar matematicamente o *mundo real* é preciso formular as descrições abstratas do fenômeno a ser estudado. Depois estas abstrações precisam ser discretizadas, em forma de algoritmo, para enfim serem implementadas computacionalmente. É o que nos diz o *Paradigma dos 4 Universos*. Como “mundo real“ entendemos a natureza que nos cerca, as atividades cotidianas (economia, comércio, ...), etc.

De forma similar, diversos conceitos originalmente desenvolvidos dentro da Física Estatística têm sido estendidos e redirecionados para aplicações nas mais diversas áreas de conhecimento, tais como a biologia, a economia e as geociências, conferindo-lhe um caráter fortemente multidisciplinar.

Dentro deste contexto, e impulsionada pela propagação dos métodos computacionais, nos anos 90 houve o surgimento da área de pesquisa denominada de Sistemas Complexos, alicerçada na Teoria do Caos e na área dos Fenômenos Não-Lineares. No contexto dos sistemas complexos, atualmente a Teoria de Redes Complexas tem se mostrado uma ferramenta adequada para análise de diversos sistemas de grande escala onde seja possível relacionar seus elementos segundo algum critério.

Durante a graduação, também sob a orientação de Professor José Garcia Vivas Miranda, iniciamos o desenvolvimento de uma ferramenta de análise da variabilidade pluviométrica do Estado da Bahia segundo critérios de correlação espacial e temporal, através da Teoria de Redes Complexas, resultando em uma Monografia de Conclusão do Curso de Bacharelado em Ciência da Computação da Universidade Federal da Bahia(SANTANA, 2005). Utilizamos métodos de análise topológica e estatística de redes complexas como ferramentas de descrição da variabilidade pluviométrica da região.

O presente trabalho tem como objetivo expandir os estudos realizados anteriormente (SANTANA, 2005), ampliando a base de dados utilizada, fundamentando

e melhorando as ferramentas físicas abordadas, na direção de identificar aplicações práticas deste novo método de análise.

Com uma base de dados maior e mais robusta, a interpretação das redes foi beneficiada, permitindo-nos observar padrões temporais na variação dos índices caracterizadores de Redes Complexas que não foram percebidos nos trabalhos anteriores. E além de aprimorarmos o método desenvolvido anteriormente, baseado na Teoria de Redes Complexas, acrescentamos a Geometria Fractal com ferramenta de análise dos dados, o que permitiu analisar características de causalidade não-local na distribuição pluviométrica na Região Nordeste.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo serão apresentados os principais fundamentos teóricos utilizados no desenvolvimento deste trabalho.

2.1 Características climáticas do Nordeste Brasileiro (NEB)

A região Nordeste está localizada na faixa tropical, aproximadamente entre 1° e 18° de latitude Sul e entre 35° e 47° de longitude Oeste, abrangendo os estados do Maranhão, Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe e Bahia. Essa região é marcada pela adversidade de um clima com secas e enchentes ocasionais, logo a precipitação e sua variabilidade constituem elemento fundamental do seu clima, sobretudo devido à influência que exercem nas atividades humanas.

Vários autores classificam o clima do NEB como semi-árido, apesar de existirem faixas litorâneas onde a precipitação chega a 2000 mm anuais em média. Devido a sua localização numa faixa tropical, a extremo leste da América do Sul, o NEB está sujeito à influência de vários sistemas atmosféricos, cuja interação confere características climáticas muito peculiares a esta região (BARBOSA, 1999).

Os totais de precipitação são, na maior parte do NEB, inferiores a 800 mm, caindo para menos de 500 mm em alguns setores do interior. Em contraste, a parte oriental da região, abrangida pelas isoietas superiores a 1000 mm constitui a zona mais úmida do Nordeste.

Além da variabilidade espacial, o clima da região se caracteriza também por grandes flutuações na pluviosidade anual. Nos diz Sugahara (1982) que em uma parte considerável da região, a variabilidade relativa é maior que 30 por cento. Disto resulta que as precipitações para cada ano podem se afastar grandemente dos valores normais, sendo este fenômeno mais expressivo em algumas áreas do interior, como as do Rio Grande do Norte e Pernambuco.

É característica de clima semi-árido uma certa indefinição no padrão pluviométrico. Certas regiões semi-áridas são mais úmidas, outras mais secas, e outras, ainda, são um meio termo entre as duas. Uma das razões atribuídas ao semi-árido no Nordeste é a existência do anticiclone semipermanente do Atlântico Sul, que torna a região Nordeste uma área na qual diversos sistemas de circulação se dissipam ao se aproximarem da divergência anticiclônica que domina esta área (SUGAHARA, 1982).

Chaves (1999) cita as circulações de Walker e Hadley como dois dos principais sistemas atmosféricos com atuação no NEB. A circulação de Walker possui ramo ascendente sobre a Bacia Amazônica e está associada ao alto índice de chuvas nessa

região, e a circulação de Hadley possui ramo ascendente associado à Zona de Convergência Intertropical (ZCIT). Ambas possuem ramo descendente sobre o NEB e proximidades do Atlântico subtropical sul. Citando Aragão(1975), Chaves afirma que não existe déficit de umidade sobre a região semi-árida do Nordeste, e que a ausência de um mecanismo dinâmico que provoque movimentos ascendentes seria a principal razão da existência do semi-árido Nordestino.(ARAGÃO, 1975)

Além disso, a influência da orografia costeira também é considerada um fator importante a reduzir a precipitação desta região, pois o relevo da costa nordestina, embora apresentando elevação moderada, poderia reduzir sensivelmente o transporte de umidade do Atlântico para o interior do NORdeste, sobretudo quando se sabe que, em média, nesta região, os ventos de baixos níveis sopram de leste para oeste (SUGAHARA, 1982).

Gomes Filho (1979) mostrou que os efeitos locais da topografia e as diferenças de albedo tendem a manter baixa a pluviosidade da região, confirmando hipótese levantada por Charney(1975)(FILHO, 1979)(CHARNEY, 1975).

Outro aspecto marcante do clima nordestino é a grande concentração de chuvas em poucos meses. Isto torna o regime de chuvas altamente vulnerável às secas (SUGAHARA, 1982). Markham (1972) afirma que sob regimes menos sazonais, a ausência de chuvas num período poderia ser compensada em período posterior. Entretanto, no Nordeste, devido à extrema sazonalidade da precipitação, se as chuvas não vêm em uma estação esperada, não há possibilidade real de recuperação durante o resto do ano(MARKHAM, 1972).

Sugahara (1982) nos diz que a distribuição espacial heterogênea dos máximos pluviométricos mensais pode ser decorrente de diversos sistemas de circulação que incidem na região. Nos setores Norte e Central da região, a pluviometria mensal atinge um máximo durante março e abril. Diversos autores como Markham (1972) admitem que este máximo está intimamente vinculado ao deslocamento extremo meridional da Zona de Convergência Intertropical (ZCIT), que ocorre geralmente neste período.

Segundo Kousky (1979), a chuva na região Sul do Nordeste pode estar associada às penetrações de sistemas frontais que, quando penetram nas baixas latitudes do Brasil, interagem com a umidade tropical da Bacia Amazônica ou do litoral oriental e produzem chuvas torrenciais na região(KOUSKY, 1979).

Chaves (1999) também afirma que o período de máxima precipitação nessa região do Nordeste é entre os meses de novembro a janeiro, com pico em dezembro. Ela associa essa máxima precipitação à penetração de Sistemas Frontais, que fornecem mecanismos de indução à convecção tropical e estão, algumas vezes, relacionados à Zona de Convergência do Atlântico Sul (ZCAS), aos Vórtices Ciclônicos de Altos Níveis (VCAN) e à Zona de Convergência Intertropical (ZCIT).

Os VCAN penetram no interior do NEB na maioria das vezes próximo a Salvador, perdendo intensidade à medida que se deslocam para Oeste, produzindo céu claro na região central e sul do NEB e chuvas na região Norte. A maior parte do Sul do Nordeste, segundo Chaves(1999), apresenta baixo índice pluviométrico e alta variabilidade pluviométrica espacial e temporal.

Sobre a vasta área costeira, desde o extremo sul do Rio Grande do Norte até o limite sul de Sergipe, Yamazaki e Rao (1977) afirma que o máximo pluviométrico ocorre entre maio e julho. Yamazaki e Rao (1977) associam este máximo aos sistemas de nuvens que eles detectaram sobre o Atlântico Sul tropical, que se movem de leste

para oeste com uma velocidade média aproximada de 2° a 3° por dia. Este máximo pode, também, ser relacionado às chuvas frontais e às chuvas ligadas com a circulação local.(YAMAZAKI; RAO, 1977).

Chaves (1999) afirma que na faixa litorânea do Sul do NEB os totais anuais de precipitação ultrapassam os 1000 mm, decrescendo fortemente em direção ao interior, onde o total anual da precipitação fica entre 400 mm e 800 mm, valores que caracterizam regiões semi-áridas. A partir da região da Chapada Diamantina, com direção à região Amazônica, a precipitação anual volta a crescer, alcançando valores acima de 1000 mm, onde a chuva está relacionada à convecção Amazônica.

Sobre a precipitação da região Norte do Nordeste, Barbosa(1999) afirma que o máximo mensal é atingido durante os meses de março e abril e as áreas litorâneas, de aproximadamente 5° S a 18° S, apresentam chuvas máximas entre maio e junho. Segundo ele, a existência de diferentes regimes de precipitação para regiões distintas do NEB indica que mais de um mecanismo é responsável pelas chuvas da região.

Além disso, pode-se notar na região regimes de chuva anuais com dois máximos. Isto ocorre no setor setentrional da Bahia (dezembro e março) e na área costeira desse estado (dezembro e maio). Strang (1972) associa esses máximos à existência de mais de um fator dinâmico que atua em épocas diferentes.(STRANG, 1972).

A modelagem de dados de precipitação permite observar estas e outras características climáticas da Região Nordeste. Neste trabalho, usamos como suporte à modelagem a correlação estatística de séries temporais e dois métodos da Teoria de Sistemas Complexos: Fractais e Redes Complexas. A seguir, segue breve explanação sobre os conceitos de correlação estatística utilizados.

2.2 Correlações estatísticas

Com muita frequência, na prática, verifica-se que existe uma relação entre duas ou mais variáveis, representando objetos de estudo. Por exemplo: os pesos dos adultos do sexo masculino dependem, em certo grau, de suas alturas; o comprimento e a área de círculos dependem de seus raios; e a pressão de uma determinada massa de gás depende de sua temperatura e seu volume. Deseja-se, frequentemente, expressar essas relações sob forma matemática, por meio do estabelecimento de uma equação que ligue as variáveis.

Para auxiliar a determinação de uma equação que relacione as variáveis, um primeiro passo consiste em colecionar dados que indiquem os valores correspondentes das variáveis consideradas. Por exemplo, sendo X e Y duas variáveis que representem peso e altura de um homem adulto respectivamente, então, uma amostra de N indivíduos apresentaria os pesos X_1, X_2, \dots, X_n e as alturas Y_1, Y_2, \dots, Y_n .

Uma segunda etapa consiste em local os pontos $(X_1, Y_1), (2, Y_2) \dots, (X_n, Y_n)$ em um sistema de coordenadas cartesianas. O conjunto de pontos resultante é denominado *diagrama de dispersão*. No diagrama de dispersão é possível, frequentemente, visualizar uma curva regular que se aproxime dos dados. Essa curva é denominada de *ajustamento*. Quando o ajustamento se assemelha a uma linha reta, diz-se que há uma relação linear entre as variáveis, caso contrário, diz-se que há uma relação não-linear. (SPIEGEL, 1977)

Para evitar o critério individual na construção de retas, parábolas ou outras curvas de ajustamento que se adaptem ao conjunto de dados, é necessário instituir uma definição da melhor curva de ajustamento. Para conseguir uma definição possível,

considera-se os dados representados pelos pontos $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$. Para um valor dado de X , por exemplo X_1 , haverá uma diferença entre Y_1 e o valor correspondente determinado na curva C . Representa-se esta diferença por D_1 , que é muitas vezes designado como desvio, erro ou resíduo e pode ser positivo, negativo ou nulo. De modo semelhante, em correspondência com os valores de X_2, \dots, X_n , obtém-se os desvios D_2, \dots, D_n . (SPIEGEL, 1977)

Uma medida de qualidade do ajustamento da curva C aos dados apresentados é proporcionada pela quantidade $D_1^2 + D_2^2 + \dots + D_n^2$. Se ela é pequena, o ajustamento é bom; se é grande, o ajustamento é mau. Portanto, de todas as curvas que se ajustam a um conjunto de pontos, a que tem a propriedade de apresentar o menor valor de $D_1^2 + D_2^2 + \dots + D_n^2$ é chamada *melhor curva de ajustamento*. Diz-se que uma curva que apresenta esta propriedade, ajusta os dados no sentido dos mínimos quadrados e é denominada curva de mínimos quadrados. Em consequência, uma reta que apresenta essa propriedade é denominada *reta de mínimo quadrado*. (SPIEGEL, 1977)

A reta de mínimos quadrados que se ajusta ao conjunto de pontos X_1, X_2, \dots, X_n e Y_1, Y_2, \dots, Y_n tem a equação:

$$Y = aX + b \quad (2.1)$$

A medida que resume o grau de relacionamento entre duas variáveis é chamada de coeficiente de correlação. Quanto maior o módulo da correlação, mais forte é o relacionamento entre as duas variáveis envolvidas. Se as variáveis se relacionam segundo uma equação linear, esta medida é chamada de *coeficiente de correlação linear* (FONSECA; MARTINS, 1996). Valores positivos de correlação linear entre duas variáveis indicam que quando uma variável cresce a outra decresce, enquanto que valores negativos de correlação linear indicam que quando uma delas cresce a outra também cresce.

Existem alguns métodos para calcular o Coeficiente de correlação linear, cada um deles identifica características específicas das variáveis estudadas. São exemplos de coeficientes de correlação linear.

- Pearson

O coeficiente de Pearson é o mais utilizado coeficiente de correlação linear entre duas variáveis. Este coeficiente varia entre -1 (correlação negativa) e $+1$ (correlação positiva), sendo o 0 o valor que indica ausência de correlação. Este coeficiente é indicado pela letra r e segue a seguinte fórmula:

$$r = \frac{\sum XY - \sum X \sum Y}{\sqrt{n \sum X^2 - (\sum X)^2} \sqrt{n \sum Y^2 - (\sum Y)^2}} \quad (2.2)$$

- Spearman

Em estatística, o coeficiente de Spearman é uma medida de correlação não-paramétrica, ou seja, ela avalia o quão bem uma função pode descrever o relacionamento entre duas variáveis, sem fazer nenhuma suposição sobre a distribuição de frequência das variáveis. Esta medida de correlação não exige que o relacionamento entre duas variáveis seja linear, nem requer que as variáveis sejam medidas no mesmo intervalo de tempo. É um método que usa somente

a ordenação das variáveis - o que chama-se de posto - e não faz quaisquer suposições ou restrições quanto à distribuição de valores.

Nos casos em que a disposição dos dados no plano cartesiano não forma uma nuvem bem comportada, com alguns pontos bem distantes dos demais, ou em que parece existir uma relação crescente ou decrescente em formato de curva, o coeficiente de correlação por postos de Spearman é mais apropriado a ser aplicado.

Em princípio, a correlação de Spearman equivale à correlação de Pearson aplicada à ordenação dos valores da variável. Para o cálculo desta correlação entre duas variáveis X e Y , o primeiro passo é ordenar em ordem crescente os valores das variáveis. A posição que x_i ocupa na sequência ordenada é chamada *posto*(x_i); As diferenças entre dois postos, d_i , são calculadas pela equação

$$d_i = (\text{posto}(x_i) - \text{posto}(y_i)) \quad (2.3)$$

e o coeficiente de correlação R é determinado usando a equação:

$$R = 1 - \frac{6 \sum d_i^2}{n^3 - n} \quad (2.4)$$

Os valores da correlação de Spearman variam entre -1 (completo desacordo entre os conjuntos) e $+1$ (perfeita concordância entre os dois conjuntos).

- Kendall

O coeficiente de Kendall é uma medida de correlação que, assim como o coeficiente de Spearman, se baseia em conjuntos ordenados de dados. Em dois conjuntos ordenados de dados X e Y , chama-se de elemento concordante a todos os X_j maiores que X_i para $j > i$. E chama-se de elemento discordante a todos os Y_j menores que Y_i para $j > i$. Também varia entre -1 (autocorrelação negativa) e $+1$ (perfeita correlação).

O coeficiente de Kendall é descrito pela seguinte equação:

$$\tau = \frac{2k}{n^2 - n} \quad (2.5)$$

onde k é a diferença entre o número de concordantes e o de discordantes.

Utilizamos as informações fornecidas pela correlação linear estatística de séries históricas de dados de precipitação acumulada como princípio para a modelagem destes dados em Redes Complexas. A seguir, apresentamos alguns conceitos que embasaram a escolha por este método de modelagem.

2.3 Redes complexas

Na última década, diversos estudos foram publicados sobre a utilização de redes na modelagem de sistemas complexos. A princípio, qualquer sistema que possa ser dividido em elementos relacionados segundo alguma regra pode ser modelado em redes. A linha de pesquisa que trata do estudo das redes segundo os conceitos da Física Estatística chama-se *Redes Complexas* (Figura 2.1).

Esta teoria tem sua origem na Teoria dos Grafos, e por isso apresentaremos a seguir alguns conceitos de Grafos importantes para a compreensão do nosso trabalho.

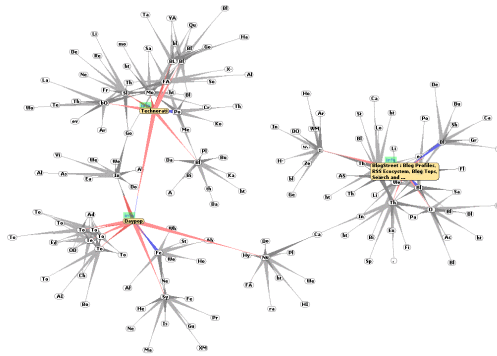


Figura 2.1: Exemplo de uma rede complexa

2.3.1 Teoria dos Grafos

Muitas situações podem ser convenientemente descritas através de diagramas constituídos de um conjunto de pontos, juntamente com linhas que ligam alguns pares destes pontos. A abstração matemática de situações onde pode-se representar elementos e suas relações através de diagramas como este dá lugar ao conceito de *grafo* (LUCCHESI et al., 1979).

Os primeiros fundamentos da *Teoria dos Grafos* surgiram em meados do século XIX, a partir de um problema algorítmico, quando Eüller se dispôs a responder o que hoje é conhecido como o *Problema da ponte de Königsberg*.

A cidade de Königsberg, capital da antiga Prússia, foi construída à volta do Rio Pregel. No meio do rio havia uma ilha chamada Kniephof. Para unir todas as partes da cidade foram construídas 7 pontes, como mostra a Figura 2.2. Os habitantes daquela cidade gostavam de passear pelas pontes e sempre buscavam uma forma de atravessar todas as pontes apenas uma vez em um passeio.

Pediram então ajuda ao grande matemático suíço Eüller. Ao utilizar um modelo em grafos para generalizar este problema, Eüller provou que não havia tal caminho. Ele verificou que existe o desejado trajeto quando, e somente quando, em cada região concorrer um número par de pontes. A solução deste problema é considerada o primeiro teorema em *Teoria dos Grafos* (SZWARCFITER, 1984).

Este foi apenas o primeiro de muitos trabalhos que ajudaram a fundamentar a *Teoria dos Grafos*.

Chama-se grafo $G(V, E)$ a um conjunto finito não-vazio $V(G)$ de elementos chamados vértices, um conjunto $E(G)$ de elementos chamados arestas e uma função de incidência $F(G)$ que associa a cada aresta α de G um par não ordenado de vértices

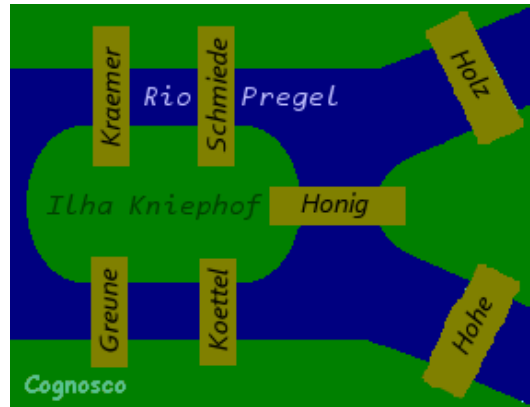


Figura 2.2: Ilustração representando o problema das pontes de Königsberg.

(não necessariamente distintos) de G , chamados de *extremos de α* (LUCCHESI et al., 1979) (SZWARCFITER, 1984).

Qualquer sistema que envolva relações entre elementos (arestas entre vértices), a princípio, pode ser modelado com sucesso utilizando a *Teoria dos Grafos*.

Grafos podem ser representados por diagramas, onde cada vértice é representado por um ponto e cada aresta por uma linha ligando os pontos que representam seus extremos. Muitos termos utilizados na *Teoria dos Grafos* (LUCCHESI et al., 1979) (SZWARCFITER, 1984) advêm da representação em diagramas.

Um *grafo* $G(V, E)$ com n vértices, pode ser representado por uma matriz M , quadrada ($n \times n$), onde cada elemento $M(i, j)$ da matriz representa a existência ou não de uma ligação (aresta) entre os vértices i e j do grafo. Se existe aresta entre os vértices i e j , então $M(i, j) = 1$, caso contrário, $M(i, j) = 0$. Chama-se M de *Matriz de adjacências* (Figura 2.4).

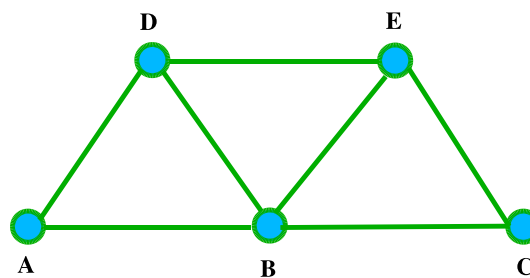


Figura 2.3: Grafo G_1 , representado pela matriz de adjacências M_1

	A	B	C	D	E
A	0	1	0	1	0
B	1	0	1	1	1
C	0	1	0	0	1
D	1	1	0	0	1
E	0	1	1	1	0

Figura 2.4: Matriz de adjacências M_1 , representativa do grafo G_1

Caminho Mínimo Médio, *Diâmetro*, *Grau Médio*, *Distribuição de Graus* são termos definidos pela *Teoria dos Grafos* e utilizados na sua análise topológica. Através da análise de um *Grafo* usado como modelo para um sistema é possível inferir características deste sistema.

- *Ordem e tamanho de um grafo*

O número de vértices em $V(G)$ é denominado *ordem do grafo* G (n) e o número total de arestas em $E(G)$ é chamado de *tamanho do grafo* G (M). Sejam v e w dois vértices de um grafo G , e seja vw uma aresta relacionando estes dois vértices, conforme exemplo na Figura 2.5. Neste caso, a *ordem do grafo* G é igual a 2 (dois) e o tamanho do grafo G é igual a 1 (um).



Figura 2.5: Exemplo de um grafo de ordem 2 e tamanho 1

- *Grau de um vértice e grau médio de um grafo*

Dois vértices são ditos adjacentes se são os extremos de uma aresta α de um grafo. O número de vértices adjacentes a um vértice v de um grafo G é denominado *grau do vértice* v , e representado por (k_v) . O valor médio entre os valores do *grau* de cada vértice de um grafo G é chamado de *Grau médio de um grafo* G , e representado por $(\langle k \rangle)$. Na Figura 2.6, o vértice em destaque tem *grau* (k_i) igual a 4 (quatro) e o grafo tem *grau médio* $(\langle k \rangle)$ igual a 2.8.

Denomina-se *grafo regular* a todo grafo cujos vértices possuem todos o mesmo grau. Se estes vértices possuem todos grau 3, por exemplo, denominamos o grafo como *3-regular*; se possuem grau 4, *4-regular* e assim por diante.

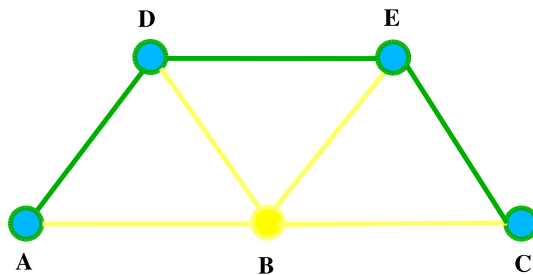


Figura 2.6: O grau do vértice B (k_B) é igual a 4 e o grau médio $(\langle k \rangle)$ do grafo é igual a 2.8

- *Conectividade de grafo, caminho mínimo médio e diâmetro de um grafo*

Chama-se caminho C em um grafo G a toda seqüência finita e não-vazia $(v_0, \alpha_1, v_1, \alpha_2, v_2, \dots, \alpha_n, v_n)$ cujos termos são alternadamente vértices v_i , distintos dois a dois, e arestas α_j , e tal que toda aresta α_i tem como extremos os vértices v_{i-1} e v_i , para $1 \leq i \leq n$. Dizemos que C é um caminho de v_0 a v_n (LUCCHESI et al., 1979) (SZWARCFITER, 1984).

O comprimento de um caminho C é o inteiro n , que representa o número de arestas de C . No caso de $n = 0$, considera-se o caminho como *degenerado*.

Se em um grafo existe sempre um caminho (independente do comprimento) que leva de qualquer vértice a qualquer outro vértice, dizemos que esse grafo é *conexo*, como mostra a Figura 2.7. Por outro lado, se entre algum par de vértices de um grafo não houver nenhum caminho, dizemos que esse grafo é *desconexo*, como mostra a Figura 2.8. Em um grafo *desconexo*, se existe um

caminho de um vértice v a um conjunto de vértices C_1 , dizemos que o vértice v e os vértices do conjunto C_1 pertencem à mesma *componente do grafo*.

Chama-se de *caminho mínimo* entre os vértices i e j ($L(i, j)$), ou *distância* entre i e j ($d(i, j)$), ao menor caminho existente do vértice i ao vértice j sem repetir vértices (LUCCHESI et al., 1979) (SZWARCFITER, 1984). No grafo da Figura 2.9, partindo do vértice A ao vértice D , o menor caminho existente neste grafo (G) tem tamanho 2.

Tomando-se a média dos caminhos mínimos entre um vértice i e todos os outros vértices de um grafo G , tem-se o caminho mínimo médio do vértice i (L_i). Tomando-se a média entre os L_i para todos os vértices i de um grafo G , tem-se o caminho mínimo médio do *grafo* ($\langle l \rangle$) (CALDEIRA, 2005) (LUCCHESI et al., 1979).

Chama-se *diâmetro* de G à maior das distâncias entre vértices de G , ou seja, o maior *caminho mínimo* entre vértices de G . Representamos o diâmetro de um grafo G por D (LUCCHESI et al., 1979) (Figura 2.10).

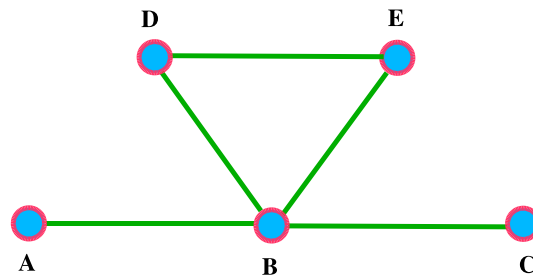


Figura 2.7: Configuração de um grafo conexo.

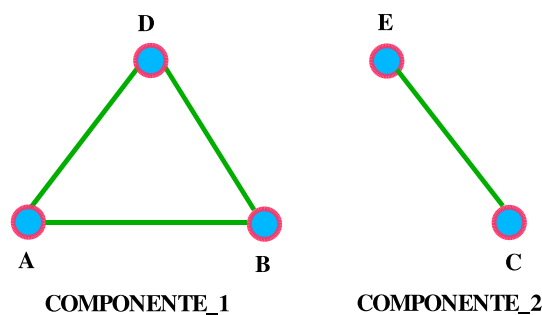


Figura 2.8: Configuração de um grafo desconexo, composto de 2 componentes.

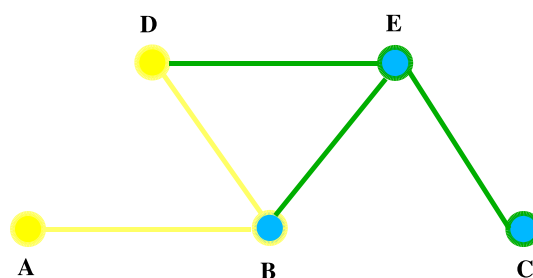


Figura 2.9: Exemplo de busca de um caminho mínimo em um grafo

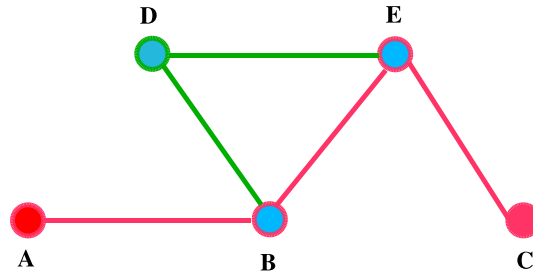


Figura 2.10: Tomando-se todos os caminhos mínimos existentes neste grafo (G), aquele de maior tamanho, e que representa o diâmetro do grafo, é o caminho mínimo entre os vértices A e C , que vale 3.

- *Completude de um grafo*

Chama-se *Grafo completo* a todo grafo que possui uma aresta associada a cada par de seus vértices, como mostra o exemplo da Figura 2.11.

À diferença entre o número de arestas de um grafo G e o número de arestas necessárias para que G seja completo dá-se o nome de *número de pares desconectados de G* (NC).

Um *subgrafo* $G_2(V_2, E_2)$ de um grafo $G_1(V_1, E_1)$ é um grafo tal que $V_2 \subset V_1$ e $E_2 \subset E_1$, ou seja G_2 é um subconjunto de G_1 .

Denomina-se *clique* em um grafo G a todo subgrafo de G que seja completo, conforme exemplo da Figura 2.12.

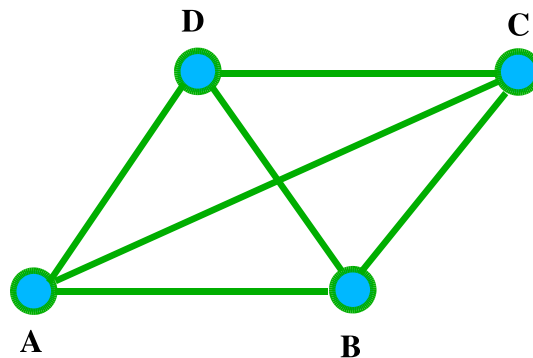


Figura 2.11: Configuração de um grafo completo

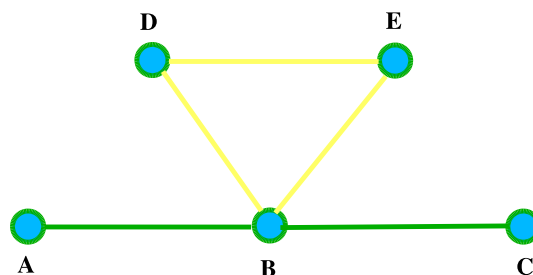


Figura 2.12: Clique formado pelo subgrafo BDE

2.3.2 Teoria de Redes Complexas

Com o desenvolvimento das ferramentas computacionais e a visível ruptura dos limites interdisciplinares, a *Teoria dos Grafos* passou a ser cada vez mais utilizada na modelagem de sistemas. À medida em que a complexidade dos sistemas modelados crescia aumentava o número de vértices e arestas a serem representados nos modelos. Isso fez com que se tornasse usual utilizar ferramentas da Física Estatística (invariância de escala) no estudo dos modelos.

Quando os Grafos passaram a ser analisados utilizando-se conceitos da Física Estatística, surgiu a *Teoria de Redes Complexas*, que hoje é largamente utilizada na análise de sistemas complexos (sistemas biológicos, sistemas sociais, sistemas ambientais, sistemas linguísticos, sistemas geofísicos, ...) (BARABÁSI; ALBÉRT, 2002) (CALDEIRA, 2005).

As redes complexas permitem analisar mais que propriedades individuais de vértices e arestas em pequenos grafos (dezenas, no máximo centenas de vértices). Permitem observar propriedades estatísticas de larga escala em grafos com grande quantidade de vértices (milhões, às vezes bilhões de vértices). (NEWMAN, 2003)

Se em um grafo com 100 vértices é interessante saber qual *vértice* não pode ser removido sem alterar a conectividade da estrutura, em uma rede complexa com milhões de vértices é mais adequado questionar qual *percentagem de vértices* não pode ser removida sem alterar a conectividade da estrutura. (NEWMAN, 2003)

Dessa forma, a teoria de redes complexas permite:

- Primeiro, encontrar e destacar propriedades estatísticas (como caminho mínimo médio, distribuição de graus e coeficiente de aglomeração média) que caracterizam a estrutura e o comportamento de sistemas em rede, oferecendo formas adequadas de medir essas propriedades;
- Segundo, criar modelos de redes que ajudem a entender o significado dessas propriedades;
- Terceiro, prever o comportamento do sistema modelado em redes, baseado no comportamento das propriedades estatísticas.

Como dissemos acima, as redes complexas podem ser classificadas de acordo com as suas propriedades estatísticas, e esta classificação ajuda a entender o comportamento do fenômeno modelado. Abaixo, descreveremos brevemente as principais classificações de redes complexas.

- *Redes Aleatórias*

Redes aleatórias são aquelas onde a composição das arestas é feita de forma aleatória. Para um número fixo de vértices n , existe igual probabilidade p que se forme cada uma das $\frac{n(n-1)}{2}$ arestas possíveis (Figura 2.13). Uma forma de identificar esse comportamento é analisar a distribuição dos graus de cada vértice da rede (NEWMAN, 2003). Se a distribuição de graus dos vértices da rede segue uma *distribuição de Poisson*¹ (Figura 2.14) diz-se que essa rede é aleatória (BARABÁSI; ALBÉRT, 2002).

¹A distribuição de Poisson tem um pico proeminente, indicando que a maioria dos vértices possui o mesmo grau, próximo de um valor médio. Dos dois lados do pico a distribuição diminui exponencialmente, tornando desvios significativos da média extremamente raros

As redes aleatórias foram estudadas por diversos matemáticos, principalmente por Paul Erdős e Alfréd Rényi, e servem para modelar diversos sistemas complexos que não apresentam critérios definidos de relacionamentos entre seus elementos (conexão entre seus vértices), por exemplo alguns fenômenos sociais.

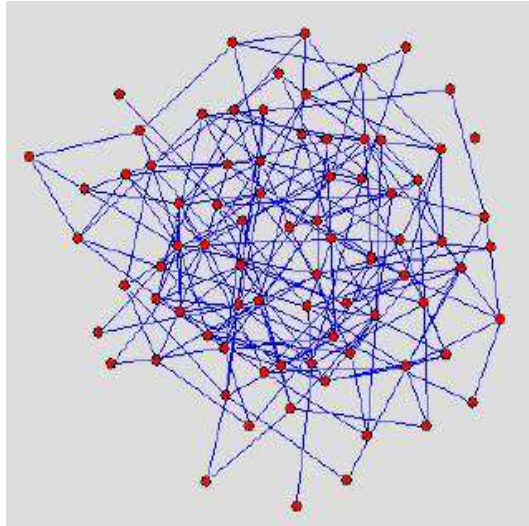


Figura 2.13: Rede aleatória

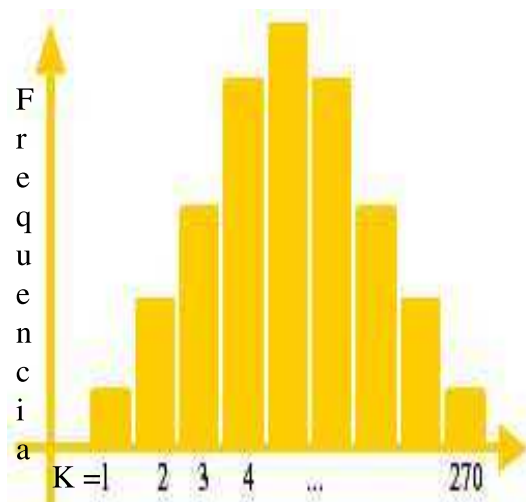


Figura 2.14: Distribuição de Graus dos vértices de uma rede seguindo uma distribuição de Poisson

- *Redes de mundo pequeno*

Redes onde a distância média entre seus vértices é muito pequena, comparada à quantidade de vértices, são classificadas como *redes de mundo pequeno*. Fenômenos como o *seis graus de separação*², proposto em 1929 pelo escritor

²Teoria segundo a qual são necessárias apenas 6 pessoas para contactar qualquer pessoa do mundo com qualquer outra pessoa do mundo.

Karinthy³ e comprovado em experimento de Stanley Milgram em 1967⁴, são a origem dessa classificação para o fenômeno estudado pelo Doutor em Sociologia Duncan Watts, que relacionou fenômenos sociais às descobertas feitas por Barabási e Albért sobre Redes Complexas (CALDEIRA, 2005) (NEWMAN, 2003) (BARABÁSI; ALBÉRT, 2002).

Redes de mundo pequeno podem apresentar distribuição de Poisson na distribuição de graus de seus vértices, demonstrando características aleatórias. Porém, sua principal característica é que apresentam um pequeno valor de caminho mínimo médio e um alto grau de completude, como mostra o exemplo da Figura 2.15 (NEWMAN, 2003).

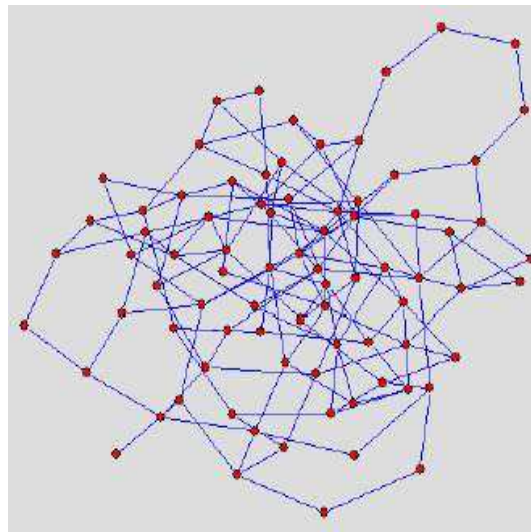


Figura 2.15: Exemplo de uma rede de mundo pequeno.

A fim de quantificar o grau completude de uma rede, utiliza-se o *Coefficiente de Aglomeração*. O Coeficiente de Aglomeração de um vértice i ($CA(i)$) representa o quanto os vértices adjacentes a i formam um clique, e é definido como a razão entre o número de arestas existentes entre os vértices adjacentes a i e o número total de arestas possíveis entre os vértices adjacentes a i (Equação 2.6) (BARABÁSI; ALBÉRT, 2002) (Figura 2.16 e Figura 2.17).

$$CA(i) = \frac{2E_i}{k_i(k_i - 1)} \quad (2.6)$$

Onde:

E_i é o número de arestas existentes entre os vértices adjacentes ao vértice i ; e $\frac{k_i(k_i-1)}{2}$ é o número total de arestas possíveis entre os vértices vizinhos de i , e k_i é o grau do vértice i .

³Num livro ele criou uma rede de relacionamentos em que participavam o próprio Kalinthy, o rei Gustavo da Hungria, e Henri Ford, entre outras pessoas espalhadas pelo mundo.

⁴Milgram escolheu duas pessoas, aleatórias, como destinatárias de duas cartas. Enviou estas cartas para pessoas desconhecidas, explicando o experimento, a fim de que estas pessoas enviassem a carta para conhecidos seus que considerasse pudessem conhecer os destinatários. De um total de 160 cartas enviadas por Milgram, 46 chegaram aos destinatários. Calculando a média de pessoas que participaram dessa corrente, Milgram chegou ao número 5,5.(MILGRAM, 1967)

É trivial perceber que $CA(i)$ se aproxima de 1 à medida que o subgrafo (ou subrede) formado pelos vértices adjacentes a i se aproxima de um clique.

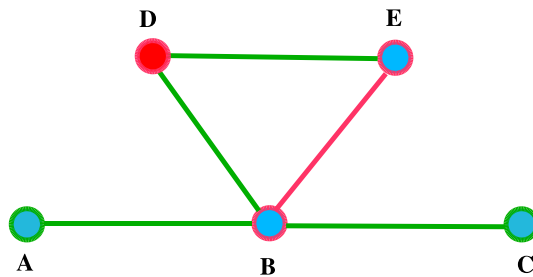


Figura 2.16: Os vértices vizinhos do vértice D formam um clique. Por isso, o CA de D vale 1

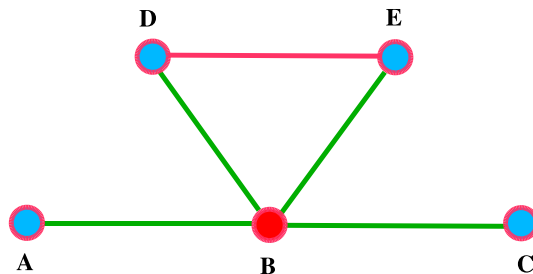


Figura 2.17: O vértice B do grafo possui 4 vértices adjacentes. Para que estes vizinhos formassem um clique, seriam necessárias 6 arestas conectando-os. Como apenas dois deles formam um clique (aresta entre os vértices D e E), o CA de B vale $\frac{1}{6}$.

Tomando-se a média entre os $CA(i)$ de cada vértice i da rede, tem-se o *Coefficiente de aglomeração médio* da rede ($\langle C \rangle$), que representa o grau de conectividade, ou de completude, de uma rede.

Portanto, redes de mundo pequeno apresentam elevado valor de $\langle C \rangle$.

- *Redes Livres de escala*

Redes de modelagem de diversos sistemas (naturais, sociais, tecnológicos, ambientais) apresentam como característica a existência de um número relativamente pequeno de vértices com grau muito alto e um número grande de vértices com graus muito baixo (CALDEIRA, 2005).

Esses vértices altamente conectados são chamados de conectores, e essas redes são chamadas de *redes livres de escala* (Figura 2.18), pois não apresentam um valor médio no número de conexões (k) de seus vértices. Não apresentam uma escala característica. A distribuição de graus de seus vértices segue uma lei de potência (Figura 2.19).

Nesse tipo de rede, aparentemente alguns vértices têm preferência durante a criação de arestas (conexões). Esse comportamento de conexão preferencial é encontrado na Web, onde existem mais links para determinadas páginas que para outras, e nas correntes migratórias, onde poucas cidades recebem imigrantes de muitos lugares (ALBERT; JEONG; BARABÁSI, 1999) (CASTRO, 2007).

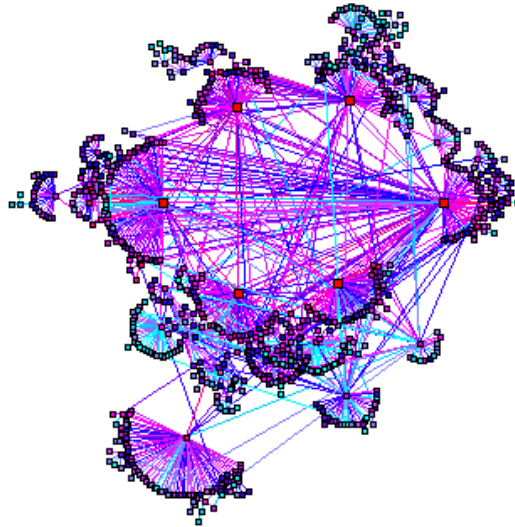


Figura 2.18: Exemplo de uma rede livre de escala. FONTE(ALBERT; JEONG; BARABÁSI, 1999)

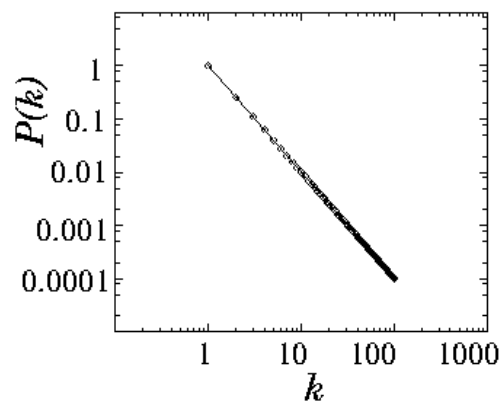


Figura 2.19: Distribuição de graus em lei de potência. FONTE(ALBERT; JEONG; BARABÁSI, 1999)

2.4 Geometria Fractal

Neste trabalho, a distribuição do tamanho das arestas das redes de correlação exibem um padrão de autosemelhança típico dos sistemas com características fractais. A seguir, apresentamos algumas características de sistemas desse tipo que balizaram essas análises nesta dissertação.

Fractal (uma palavra cunhada por Benoit Mandelbrot em 1975) e seus conceitos têm se tornado, nos últimos 30 anos, uma ferramenta primordial nas ciências naturais: Física, Biologia, Química, Geologia, Meteorologia e Ciências dos Materiais.

Segundo Mandelbrot, "*Um fractal é uma figura feita de partes similares ao todo de alguma forma.*". Mas existem diversas definições sobre o termo de Fractal. A idéia básica por trás desse conceito é a de auto-similaridade, conforme ilustrado na Figura 2.20.

Este conceito implica na propriedade de invariância de escala, ou seja, que mudando a escala de medida de um objeto, determinadas características permanecem

constantes, que os objetos têm características similares quando comparados com eles mesmos a diferentes escalas: *a parte é semelhante ao todo* (GLEICK, 1987).

Mas a maior contribuição desse conceito é uma definição de dimensão, em que objetos heterogêneos podem ser convenientemente medidos. O conhecimento desenvolvido a partir da formulação da geometria fractal permitiu que os modelos dos fenômenos naturais pudessem estar mais topologicamente ajustados às formas naturais.

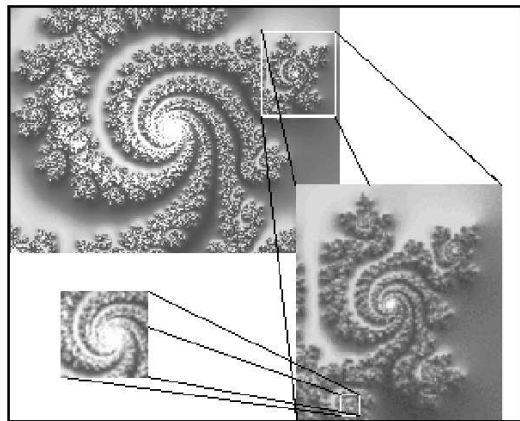


Figura 2.20: Exemplo de similaridade

2.4.1 Dimensão Fractal

Existem diversas maneiras de se introduzir o conceito de fractal. Eles se baseiam na idéia de medida. Por exemplo, uma folha de papel pode ser considerada bidimensional; uma esfera maciça é tridimensional; mas qual seria a dimensão de uma bolinha de papel amassado?

Para responder a esta questão é preciso entender alguns conceitos de dimensão indispensáveis ao conceito de fractal.

- *Dimensão Topológica - D_t* : é definida de forma recursiva. Um ponto, por definição, tem dimensão topológica zero. A dimensão topológica de um corpo qualquer é igual à dimensão do corpo necessário para dividi-lo em duas partes desconexas mais um. No caso de uma linha, para dividi-la em duas partes desconexas é preciso retirar um ponto (de dimensão zero), portanto, sua dimensão topológica é $D_t = 0 + 1 = 1$; no caso de um plano, para dividi-lo em duas partes desconexas é preciso retirar-lhe uma linha, portanto sua dimensão topológica $D_t = 1 + 1 = 2$; no caso de uma esfera, é necessário cortar-lhe com um plano, portanto $D_t = 2 + 1 = 3$. (MIRANDA, 1997)
- *Dimensão de Imersão - D_i* : também chamada de *Dimensão Euclideana*, é a menor dimensão em que o objeto pode estar imerso. Por exemplo, um arame reto tem $D_t = 1$ e $D_i = 1$, caso o dobramos em forma de um aro, ele terá $D_t = 1$ e $D_i = 2$, e se o dobrarmos em forma de um óculos, terá $D_t = 1$ e $D_i = 3$. (MIRANDA, 1997)
- *Dimensão Fractal - D* : ou dimensão de medida, ou de Hausdorff Besicovitch, ou de similaridade. Foi introduzida na década de 20 e está relacionada com

o conceito de medida. Por exemplo, se queremos medir uma reta, escolhemos uma escala ε e seu comprimento L será definido por

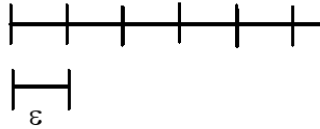


Figura 2.21: Exemplo de escala para medição de comprimento

$$L = N(\varepsilon)\varepsilon \quad (2.7)$$

Onde $N(\varepsilon)$ é o número de vezes que a escala ε é aplicada, como no caso da Figura 2.21, onde o comprimento seria aproximadamente 5ε .

No caso de uma área A (ver 2.22), teremos:

$$A(\varepsilon) = N(\varepsilon)\varepsilon^2 \quad (2.8)$$

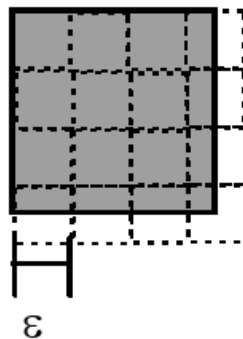


Figura 2.22: Exemplo de escala para medição de área

E analogamente para um volume V :

$$V(\varepsilon) = N(\varepsilon)\varepsilon^3 \quad (2.9)$$

Podemos perceber nas Figuras 2.21 e 2.22, que ao reduzirmos a escala ε a medida (comprimento, volume, etc) se torna mais precisa. Portanto, podemos definir de forma genérica a medida M :

$$M = \lim_{\varepsilon \rightarrow 0} N(\varepsilon)\varepsilon^{Dm} \quad (2.10)$$

Onde Dm é a dimensão do instrumento de medição, ou seja, se estamos medindo com retas $Dm = 1$, se usamos planos $Dm = 2$.

Avaliando o $N(\varepsilon)$, temos que, no caso de uma reta, à medida que ε diminui o número de intervalos $N(\varepsilon)$ necessários para cobrir toda a figura aumenta com a razão inversa, ou seja:

$$N(\varepsilon) \sim \varepsilon^{-1} \quad (2.11)$$

Para um plano temos que

$$N(\varepsilon) \sim \varepsilon^{-2} \quad (2.12)$$

Dessa forma temos que:

$$N(\varepsilon) = C\varepsilon^{-D} \quad (2.13)$$

Onde C é uma constante de proporcionalidade.

Logo, ao usarmos ε como escala relativa $\varepsilon = \varepsilon/L$ (onde L é o comprimento total)

$$M = \lim_{\varepsilon \rightarrow 0} C\varepsilon^{-D}\varepsilon^{Dm} \quad (2.14)$$

$$M = \lim_{\varepsilon \rightarrow 0} C\varepsilon^{(Dm-D)} \quad (2.15)$$

Ou seja, a medida só será definida para $Dm = D$, pois para $Dm > D$, no limite de $\varepsilon \rightarrow 0$ a medida será zero, e para $Dm < D$ será infinito. O que é razoável pois ao tentarmos medir uma reta com planos obteremos área zero e o inverso (um plano com retas) obteremos comprimento infinito. (MIRANDA, 1997)

Ao tentar medir a Figura 2.20, calculando seu contorno com segmentos de tamanho ε ($Dm = 1 = Dt$), à medida que ε diminui, mais detalhes são encontrados; desta forma seu comprimento será infinito. Logo de acordo com o que foi dito, sua dimensão D deve ser maior que 1. Tentando medi-la com placas de lado ε ($Dm = 2 = Di$), obtemos zero, por se tratar de um objeto de topologia unidimensional.

E assim, concluímos que a dimensão de contorno da Figura 2.20 deve ser algo entre 1 e 2.

Na verdade, essa é uma idéia considerada como uma das definições do termo fractal:

Ou seja, em D está a idéia de preenchimento do espaço, assim, à medida que D aumenta, o grau de preenchimento do espaço, ou desordem ou irregularidade, aumenta, podendo assumir valores não inteiros, diferentes da idéia comum sobre dimensão. (MIRANDA, 1997)

Agora é possível responder à pergunta sobre a bolinha de papel. Sabendo que:

$$N(\varepsilon) = C\varepsilon^{-D} \quad (2.16)$$

Através da linearização temos:

$$\log[N(\varepsilon)] \sim -D \log \varepsilon \quad (2.17)$$

e usando a massa no lugar de $N(\varepsilon)$ e o raio das bolinhas no lugar de ε , podemos construir várias bolinhas com massa e raio conhecidos. Sua dimensão será a inclinação da reta massa *versus* raio (o valor encontrado é aproximadamente 2,4).(MIRANDA, 1997)

Até então a definição de fractal estava relacionada com a capacidade de se obter a mesma figura (ao menos estatisticamente) ao se efetuar uma mudança de escala. Há fractais onde uma transformação de escala, com o mesmo fator para todas as direções, leva a sistemas similares, aos quais chamamos de fractais auto-similares, como podemos notar em estruturas do tipo da curva de Koch (Figura 2.23(MIRANDA, 1997)

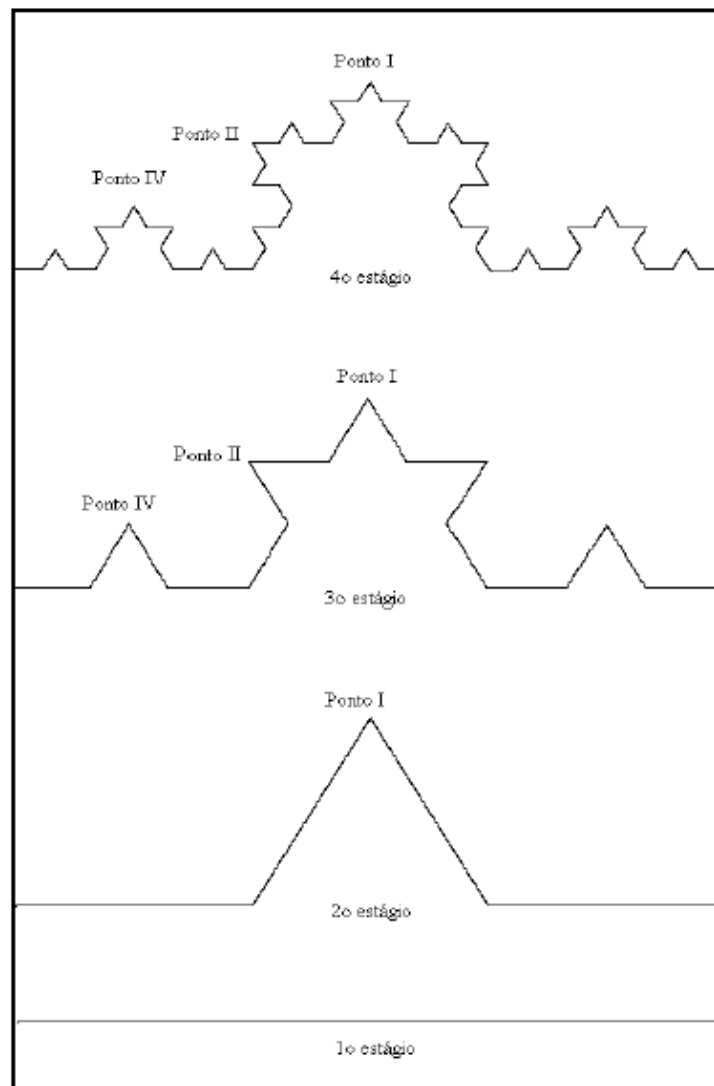


Figura 2.23: Quatro gerações da curva de Koch. No limite de infinitas gerações torna-se um fractal FONTE: (MIRANDA, 1997)

Na Figura 2.23 temos a ilustração dos estágios de formação da curva de Koch (MIRANDA, 1997). No estágio inicial, tem-se um segmento de reta; divide-se o segmento de reta em três segmentos de igual comprimento; desenha-se um triângulo equilátero (fazendo um ângulo de π radianos, ou 60°) em que o segmento central servirá de base; apagando este segmento base logo em seguida. Tem-se então, no segundo estágio, 4 segmentos de reta. Procedendo da mesma forma para cada um desses 4 segmentos formam-se 16 novos segmentos menores no terceiro estágio. A curva de Koch é o limite para o qual tende essa construção, repetindo as operações referidas, sucessivamente, para cada segmento. Logo, para uma variação de $1/3$ no comprimento dos segmentos de reta, o número de segmentos aumenta com um fator de 4 vezes, ou seja $N(1) = 1, N(1/3) = 4, N(1/9) = 16, \dots$. Utilizando a relação 2.17, obtemos que:

$$D = -\frac{\log[N(\varepsilon)]}{\log(\varepsilon)} = \frac{\log(16)}{\log(9)} \sim 1,2618. \quad (2.18)$$

2.5 Trabalhos anteriores sobre Sistemas Complexos aplicados à análise da pluviometria

Existem inúmeros trabalhos que utilizam Sistemas Complexos para analisar dados climáticos, mas dois em especial tratam exatamente da mesma área de estudo que nos propomos. Esta seção descreve estes trabalhos e os principais resultados que serviram como motivação para nosso trabalho.

2.5.1 Correlação temporal da chuva no Nordeste

Como muitos outros eventos climáticos e geofísicos, as chuvas mostram algum grau de invariância de escala temporal e espacial estatística. Sendo assim, importantes informações podem ser obtidas de dados históricos de precipitação, baseado em métodos matemáticos e conceitos da geometria fractal (MIRANDA et al., 2004).

O método R/S é um método matemático comumente usado para análise de seqüências temporais, e tem sido largamente usado para análise de dados climáticos. Através desse método é determinado o expoente de Hurst (H), que está relacionado com a dimensão fractal do perfil associado à série temporal de dados pluviométricos (FEDER, 1988).

De acordo com Feder(1988): Um sistema *persistente* é aquele onde um incremento *positivo* no passado aumenta a possibilidade de um incremento *positivo* no futuro, e considera-se assim todo sistema com $H > 1/2$; um sistema *antipersistente* é aquele onde um incremento *positivo* no passado aumenta a possibilidade de um incremento *negativo* no futuro, e considera-se assim todo sistema com $H < 1/2$; todo sistema com $H = 1/2$ é considerado um sistema sem memória (*movimento browniano*).

Segundo Miranda(1997), utilizando dados de séries diárias de precipitação do Nordeste Brasileiro computados pelo *NCAR* (National Center for Atmospheric Research, Colorado, EUA), foi calculado o H para o acumulado anual de chuvas no período de 1903 a 1982 (MIRANDA, 1997) (MIRANDA et al., 2004).

Os resultados desses cálculos mostraram que, do ponto de vista da correlação temporal, a chuva, de maneira geral, se mostra persistente ($H > 1/2$) para escalas de até 11 anos. Mas para escalas acima de 11 anos foi encontrado um H médio de 0,377,

indicando a existência de um ciclo de 11 anos na correlação do acumulado anual de chuva (MIRANDA, 1997) (MIRANDA et al., 2004). Associou-se esta periodicidade de 11 anos ao ciclo de oscilação decadal do pacífico ⁵

Segundo Miranda *et al* (2004), dessa forma é possível concluir que, não obstante a grande variabilidade de regimes pluviométricos da região, os padrões obtidos, com relação a persistência, sofrem pouca variação e são praticamente independentes do regime anual de precipitação, da marcha estacional, da região de influência dos sistemas de circulação, do relevo, em suma, independentes da estrutura climática do local (MIRANDA et al., 2004).

2.5.2 Correlação espacial da chuva no Nordeste

A análise da variação da semivariância com a escala, chamada semivariograma, permite caracterizar a variação de uma determinada característica de um sistema a diferentes escalas.

Utilizando os mesmos dados de séries diárias de precipitação usados para caracterizar a correlação temporal da chuva no Nordeste, Miranda (2004) realizou um estudo para determinar a correlação espacial da chuva na região, através da análise de semivariograma.

Analisando o total mensal de pluviometria de estações pluviométricas separadas por diferentes distâncias obteve-se a semivariância para cada escala diferente de separação das estações (MIRANDA et al., 2004).

Relacionado os valores da semivariância com suas respectivas escalas, descobriu-se que existe uma correlação, seguindo uma lei de potência, entre estações pluviométricas distantes de até 150 Km entre si. A partir dessa distância verificou-se ser impossível definir se há ou não uma correlação entre a variação de chuva de duas estações (MIRANDA et al., 2004).

2.6 Trabalhos sobre modelagem em Redes Complexas

A seguir, descrevemos alguns trabalhos que utilizaram a Teoria de Redes Complexas para modelagem de série temporal de dados espacializados, uma motivação para a nossa proposta metodológica.

2.6.1 Correlação espaço-temporal de terremotos

A análise de séries temporais de eventos naturais é uma das diversas aplicações de redes complexas a fenômenos naturais. Abe (2004) analisou seqüências temporais de dados de eventos sísmicos na Califórnia e no Japão, modelando-os em redes complexas (ABE; SUZUKI, 2004).

Considerou-se vértice da rede toda localidade onde houve evento sísmico no período analisado. Para cada dia do período, criava-se cliques entre as localidades onde houve evento sísmico naquele dia. Repetia-se o processo para dias subsequentes e ao fim do período tinha-se um conjunto de relacionamentos extremamente complexo (ABE; SUZUKI, 2004).

⁵Tendência de a temperatura da superfície do Oceano Pacífico - o maior do planeta e que influencia no clima do mundo inteiro - mudar de fase a cada 20 ou 30 anos, que está associada, assim como outras grandes variações climáticas do planeta, à influência dos ciclos solares de 11 anos (sunspot cycles).

Analisando-se a distribuição de graus, verificou-se que as redes construídas apresentavam características de redes *livre de escala*. Isso significa que algumas poucas localidades tiveram muitos eventos sísmicos, enquanto outras muitas localidades tiveram poucos eventos sísmicos no período.

Essa idéia de que poucos locais têm muito mais eventos sísmicos que outros muitos locais já era conhecida pelos geólogos. O que esse trabalho trouxe de mais importante foi a validação da metodologia de sequências temporais de dados utilizando redes complexas. Essa característica foi muito importante para o desenvolvimento de nosso trabalho (ABE; SUZUKI, 2004).

2.6.2 Correlação espaço-temporal entre neurônios

A definição da estratégia de construção das redes é, talvez, o ponto crucial de qualquer trabalho da área. Chialvo *et al* utilizaram da correlação linear de Pearson para definir a estratégia de construção de uma rede de neurônios (EGUILUZ et al., 2005).

Induzindo a atividade cerebral através de movimentos com os dedos e som de músicas, e monitorando esses sinais de atividade cerebral através de uma ressonância magnética funcional, foi obtida a sequência temporal de ativação cerebral em diferentes regiões do cérebro (EGUILUZ et al., 2005).

Considerou-se cada região do cérebro como vértice de uma rede. Foi feito o cálculo da correlação linear entre a sequência de dados de cada par de regiões do cérebro. Conectaram-se os vértices relativos a duas regiões a e b do cérebro se a correlação linear de Pearson entre os dados das regiões a e b era superior a um determinado patamar, definido pela equipe (EGUILUZ et al., 2005).

Dessa forma foram formadas redes de correlação de regiões do cérebro, baseada na série temporal de ativação dessas regiões. A utilização do método de correlação linear foi também ponto importante no desenvolvimento de nosso trabalho.

3 MATERIAIS E MÉTODOS

Este capítulo visa descrever a metodologia utilizada para a construção e análise das redes representativas da pluviometria do semi-árido nordestino, detalhando o formato dos dados analisados, as estratégias escolhidas para a construção das redes e as ferramentas computacionais utilizadas para construção e análise dessas redes.

3.1 Origem dos dados

Nosso estudo é baseado em séries históricas de precipitações pluviométricas de estações pluviométricas distribuídas em toda a região Nordeste do Brasil. As informações foram obtidas com a Agência Nacional para o Desenvolvimento do Nordeste (ADENE) e com o *National Center for Atmospheric Research* (NCAR) dos Estados Unidos. Da ADENE, utilizamos dados de 514 estações pluviométricas distribuídas no Estado da Bahia, com dados de chuva datados desde 1979 até 2000. Estas estações podem ser visualizadas na Figura 3.1. Do NCAR, utilizamos dados de 1400 estações pluviométricas distribuídas em 7 estados da região Nordeste, com dados de chuva datados desde 1904 até 1983, conforme mostra a Figura 3.2.



Figura 3.1: Mapa da Bahia contendo as 514 estações da ADENE

Nem todas as estações possuem dados correspondentes ao mesmo período. Algumas, por exemplo, possuem dados entre 1911 e 1940, enquanto outras possuem dados entre 1950 e 2000. Essas diferenças fizeram necessário identificar o período com maior número de estações com dados simultâneos para ser analisado para cada

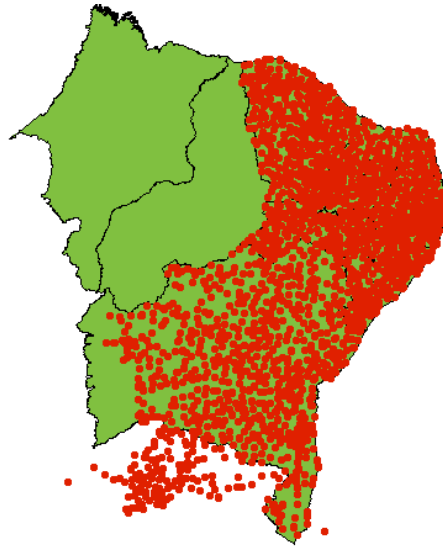


Figura 3.2: Mapa do NE contendo as 1400 estações do NCAR

fonte de dados.

3.2 Tratamento dos dados

Com o intuito de simplificar a análise dos dados de chuva, armazenamos os dados de todas as estações de cada fonte de dados em arquivos de extensão *.FPP*. A Figura 3.3 mostra uma parte de um desses arquivos. O cabeçalho arquivo contém o número total de dias com dados de chuva para todas as estações (N). A seguir, linha a linha, vêm os dados de todas as estações dispostos da seguinte forma: código da estação, cidade de localização da estação¹, a localização geográfica da estação (latitude e longitude) e a seqüência de precipitação diária da estação para os N dias (caso a estação não tenha dados em um dia i , considera-se que a precipitação daquela estação no dia i é igual a -2.0).

A fim de utilizar dados consistentes na construção das redes, selecionamos o período com maior quantidade de dados de pluviometria, chamado de *período ótimo*, através da análise de histogramas que mostram a quantidade de estações com dados pluviométricos em cada intervalo do período total, conforme mostram as Figuras 3.4 e 3.5. Para os dados da ADENE, o período ótimo corresponde ao intervalo entre os dias 0 e 3000, relativo aos anos entre 1979 e 1986; para os dados do NCAR, o período ótimo corresponde ao intervalo entre os dias 23000 e 24500, relativo aos anos entre 1966 e 1969; mas quando consideramos dados simultâneos das duas fontes de dados, o período ótimo é entre os anos 1979 e 1983.

¹Para simplificar a representação da cidade, optamos por utilizar o código da estação novamente, no lugar do nome da cidade por extenso.

```

34224
839028
839028
0.089394 0.677778
0 -2.000000
1 -2.000000
2 -2.000000
3 -2.000000
4 -2.000000
5 -2.000000
6 -2.000000
7 -2.000000
8 -2.000000
9 -2.000000
10 -2.000000
11 -2.000000
12 -2.000000
13 -2.000000
14 -2.000000
15 -2.000000
16 -2.000000
17 -2.000000
18 -2.000000
19 -2.000000
20 -2.000000
21 -2.000000
22 -2.000000
23 -2.000000
24 -2.000000
25 -2.000000
26 -2.000000
27 -2.000000
28 -2.000000
29 -2.000000
30 -2.000000
31 -2.000000
  
```

Figura 3.3: Exemplo de um arquivo do tipo *FPP*

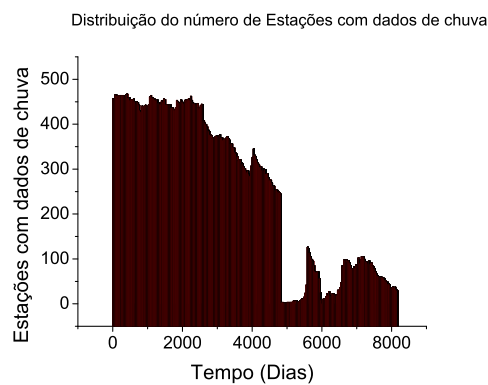


Figura 3.4: Histograma com as Estações da ADENE que contêm dados de chuva para cada dia do intervalo de 1979 a 2000.

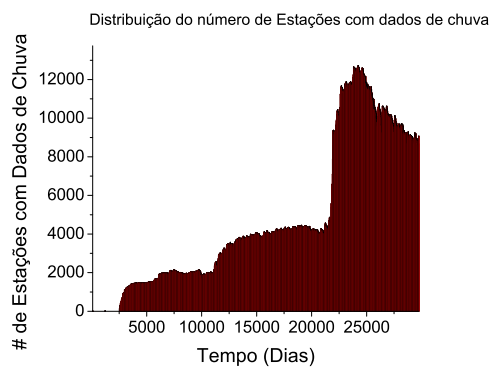


Figura 3.5: Histograma com as Estações do NCAR que contêm dados de chuva para cada dia do intervalo de 1904 a 1983

3.3 Estratégias de construção das redes

Para construir as redes representativas dos dados de pluviometria do semi-árido, utilizamos duas estratégias:

- Uma baseada na limitação espacial (LE) do raio de correlação espacial da chuva no Nordeste (MIRANDA, 1997) e no trabalho sobre análise de seqüência temporal de eventos sísmicos segundo redes complexas (ABE; SUZUKI, 2004);
- Outra baseada na correlação linear estatística (CLE) entre as seqüências de dados de pluviometria de duas estações pluviométricas, seguindo estratégia utilizada por Eguíluz (2005) na análise da seqüência temporal de atividade cerebral mediante redes complexas (EGUILUZ et al., 2005).

Para as duas estratégias, a forma escolhida para a representação matemática das redes foi a matriz de adjacências. E a representação computacional dessa matriz foi o padrão seguido pelo programa *Pajek*, que é um software para visualização de grafos (BATAGELI; MRVAR, 1996).

A seguir apresentamos detalhes de cada uma das estratégias utilizadas:

3.3.1 Limitação espacial

Considerando que a chuva não é um evento isolado no tempo, na construção das redes não analisamos os dados de precipitação de um dia X isoladamente, mas os dados de um período que chamamos de *janela temporal*.

Definimos um raio temporal T que representa a quantidade de dias antes e depois de X que formarão essa *janela temporal*. Assim, para construirmos a rede representativa da precipitação de um dia X , analisamos o intervalo entre o dia $(X - T)$ e o dia $(X + T)$. Variamos o valor do raio temporal T entre 1 e 10 dias, a fim de estudarmos as variações nas diferentes redes construídas.

Consideramos que toda estação onde houve chuva durante a *janela temporal* analisada é um vértice de nossa rede. Relacionamos duas estações onde houve chuva num mesmo período segundo a distância ²entre elas, que deve ser no máximo de uma distância $D = 150$ km (MIRANDA et al., 2004). Na rede, representamos essa relação entre duas estações criando arestas conectando os vértices que as representam. O procedimento é repetido para dias subseqüentes, isto é, variando X , mantendo-se T constante.

3.3.2 Correlação linear

Utilizando dados de chuva de diferentes meses foram calculados, para cada par de estações pluviométricas, os coeficientes de correlação de Kendall, Spearman e Pearson.

Como vimos na seção 2.2, cada um destes coeficientes de correlação apresenta sensibilidade diferente com relação aos dados estudados. Por isso, as redes geradas por cada um deles apresentaram características diferentes, o que motivou a escolha do mais apropriado para este trabalho.

Para a maioria dos pares de estações, o coeficiente de Kendall apresentou valores menores ou iguais a 0,1. Isso significa que se este coeficiente fosse usado com critério de construção das redes, estas redes teriam poucas conexões.

²Raio característico da correlação espacial das chuvas no semi-árido, segundo MIRANDA(2004).

Para o coeficiente de Spearman, a maioria dos pares de estações apresentou valores maiores ou iguais a 0,9. Isso significa que as redes construídas usando este coeficiente como critério de correlação, teriam muitas conexões.

O coeficiente de Pearson apresentou valores mais variados, representando melhor a diversidade de pluviometria da região Nordeste durante o ano: meses de chuva mais distribuídas no espaço apresentaram coeficiente de Pearson maiores para a maioria dos pares de estações, enquanto os meses de chuva menos distribuídas no espaço apresentaram coeficiente de Pearson menores, de forma que escolhemos este método para trabalhar.

Conforme já vimos na seção 2.2, o coeficiente de Pearson calculado é representado pela equação:

$$r(x, y) = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{\sum x_i^2 - \sum x_i^2 \sum y_i^2 - \sum y_i^2}} \quad (3.1)$$

Onde:

- n é o número de dias analisados.
- x_i representa a quantidade de chuva registrada na estação x no dia i ;
- y_i representa a quantidade de chuva registrada na estação y no dia i .

Foram consideradas vértices da rede, as estações pluviométricas onde houve chuva no período analisado. Foram criadas arestas para cada dois vértices x e y , para os quais o valor absoluto da correlação linear entre as sequências de dados de chuva das estações representadas por estes vértices, $r(x, y)$, fosse maior ou igual a um r_{limite} . No trabalho de Eguiluz(2005) foi utilizado $r_{limite} = 0,7$, mas nós preferimos definir o valor deste parâmetro baseado na evolução da rede para diferentes limites.

À medida que o valor do r_{limite} usado para construir as redes é maior, menos arestas são consideradas, porque o critério para criação de arestas se torna mais exigente, conforme mostra a Fig 5.10. A variação no número de arestas causa variação nos valores de todos os outros índices.

Para decidir sobre o valor do r_{limite} a ser utilizado, verificamos a variação do *Caminho Mínimo Médio*, do *Grau Médio* e do *Coefficiente de Aglomeração Médio*, porque estes índices permitem identificar a conectividade da rede.

A conectividade das redes traz informações sobre a variação da chuva no período analisado, já que as arestas representam a existência de alta correlação na variação da quantidade de chuva das estações pluviométricas. Redes muito conectadas representam períodos em que a variação da quantidade de chuva foi parecida na maioria das estações, ou seja, representam períodos de chuva mais igualmente distribuída na região. O valor do r_{limite} , portanto, deveria favorecer a representação mais precisa da distribuição pluviométrica do período analisado. Isso faz do r_{limite} um valor crítico, que chamaremos, a partir de agora, de $r_{critico}$, a partir do qual o comportamento da rede muda bastante.

3.4 Análise das Redes

As redes geradas pelos dois métodos passaram por análise visual, com a ajuda do Software Pajek, e por análise de sua topologia, mediante cálculo dos índices caracterizadores de grafos e redes complexas.

3.4.1 Análise visual

Para a análise visual das redes utilizou-se o software Pajek, que permite, entre outras coisas, desenhar os vértices das redes segundo a localização geográfica (latitude e longitude) das estações pluviométricas que representam. Esse recurso foi de extrema importância para a compreensão do sentido geográfico das redes geradas.

3.4.2 Análise topológica (estatística)

A estrutura topológica da rede é analisada mediante a determinação de índices de caracterização: número de vértices, caminho mínimo médio, diâmetro, grau médio, coeficiente de aglomeração (BARABÁSI; ALBÉRT, 2002). Esses índices foram calculados através do *NetAllbin*, um programa computacional produzido para suprir as necessidades de diversos trabalhos de pesquisa do grupo FESC³ envolvendo Redes Complexas.

O *NetAllbin* foi feito com a linguagem de programação *C++*. Possui uma classe chamada *Grafo*, que armazena a matriz de adjacências que representa a rede analisada e se utiliza de algoritmos clássicos para cálculo de caminho mínimo, diâmetro e outros índices caracterizadores de grafos.

O arquivo de saída do *NetAllbin* fornece o coeficiente de aglomeração ($CA(i)$), o caminho mínimo médio ($L(i)$) e o grau ($k(i)$) de cada vértice i da rede analisada.

Além disso, o *NetAllbin* também retorna para a saída padrão o número de arestas (E), o diâmetro (D), o coeficiente de aglomeração médio ($\langle C \rangle$), o caminho mínimo médio ($\langle l \rangle$), o grau médio ($\langle K \rangle$) e o número de pares desconectados (NC) de cada rede analisada.

3.4.3 Análise de características Fractais

Para uma rede qualquer, contabilizamos o número de relações (arestas) com tamanho menor ou igual a um dado raio r . Definimos um valor inicial $r = 150$ e o variamos de forma exponencial (aumentando seu valor em 10%) até o tamanho da maior aresta da rede. A contabilidade do número de arestas e para cada escala r diferente, fornece a *Distribuição cumulativa dos tamanhos das arestas*, que nesse caso segue uma Lei de Potência. Distribuições que seguem Lei de Potência são aquelas nas quais uma variável é proporcional a uma potência da outra, nesse caso o número de arestas acumuladas (e) é proporcional a uma potência da distância r , na forma $e = kr^d$, onde a constante k e o expoente d são desconhecidos. O valor desse expoente está associado ao valor da dimensão fractal do fenômeno analisado, e seu estudo para várias redes geradas em períodos e em territórios diferentes permite identificar características como a invariância de escala nas redes.

³Grupo de Física Estatística e Sistemas Complexos da UFBA

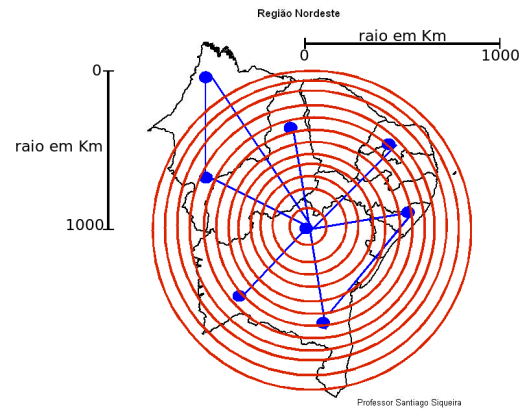


Figura 3.6: Estudo das características fractais da rede, baseado na quantidade de conexões existentes a diferentes distâncias (raios)

3.5 Resumo metodológico

Esta seção se propõe a resumir as etapas da construção e análise das redes complexas de estações pluviométricas, representados no fluxograma da Figura 3.7 e itemizados a seguir:

- PASSO 1: Conversão dos dados da ADENE e do NCAR para o formato *FPP*
- PASSO 2: Definição dos $r_{critico}$ para construção das redes complexas pelo método *CLE*
- PASSO 3: Construção das redes complexas pelos dois métodos: *CLE* e *LE*
- PASSO 4: Análise das características fractais das redes geradas pelo método *CLE*
- PASSO 5: Análise dos índices caracterizadores de Redes complexas das redes geradas pelos dois métodos
- RESULTADO: Associação dos resultados obtidos com características da variabilidade pluviométrica do **NEB**

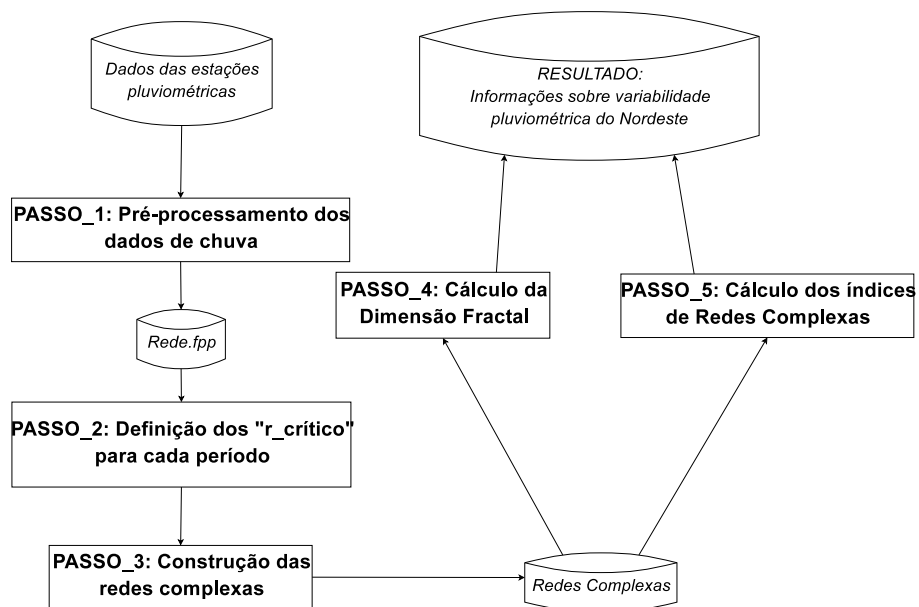


Figura 3.7: Fluxograma que mostra as etapas de construção e análise das redes complexas de estações pluviométricas

4 LIMITAÇÃO ESPACIAL - LE

Neste capítulo, apresentamos os resultados para as redes construídas pelo método *LE*, demonstrando características visuais e topológicas, pelos índices estatísticos caracterizadores de redes complexas.

4.1 Análise visual

Pela análise visual, através do software Pajek verificamos que as arestas conectam vértices não muito distantes, representando localmente a dinâmica da chuva, conforme é observado nas redes de exemplo das figuras 4.1 a 4.4.

- As redes relativas aos dados de verão e de outono (ver figuras 4.3 e 4.4) possuem vértices mais distribuídos ao longo do mapa da Bahia do que as redes relativas aos dados de inverno e primavera. O que representa o fato de a chuva durante o verão e o outono ser mais distribuída territorialmente na Bahia, enquanto no inverno e na primavera a chuva é mais concentrada em algumas localidades, principalmente no litoral .
- Além disso, os diferentes comportamentos das redes de estações pluviométricas para diferentes períodos do ano refletem os comportamentos climáticos das diferentes regiões da Bahia, descritas por Chaves(1999).

Segundo Chaves (1999), as regiões Oeste, central, sudoeste da Bahia apresentam alta similaridade climática e alta variabilidade pluviométrica anual, apresentando ciclo anual bem definido e máximo pluviométrico em dezembro e em março (CHAVES, 1999), conforme mostra a Figura 4.5.

Nas redes de estações pluviométricas geradas é possível identificar esse mesmo comportamento. Poucas arestas na região oeste, central e sudoeste, nas redes relativas aos dados de inverno e primavera, como mostram as Figuras 4.1 e 4.2; e muitas arestas nessas mesmas regiões nas redes relativas aos dados de verão e outono, como mostram as Figuras: 4.3 e 4.4.

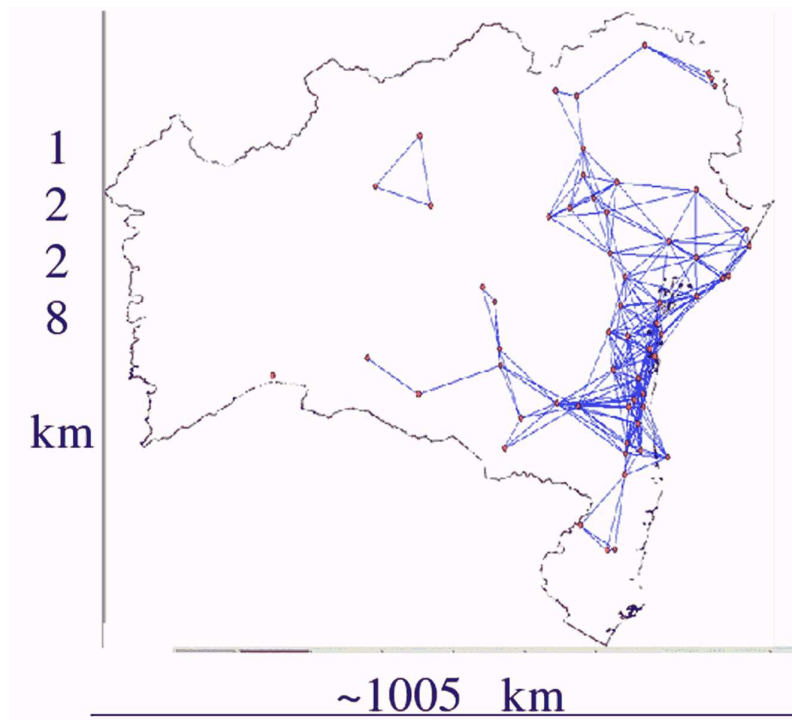


Figura 4.1: Rede gerada a partir da análise de dados de chuva da ADENE para períodos de inverno.

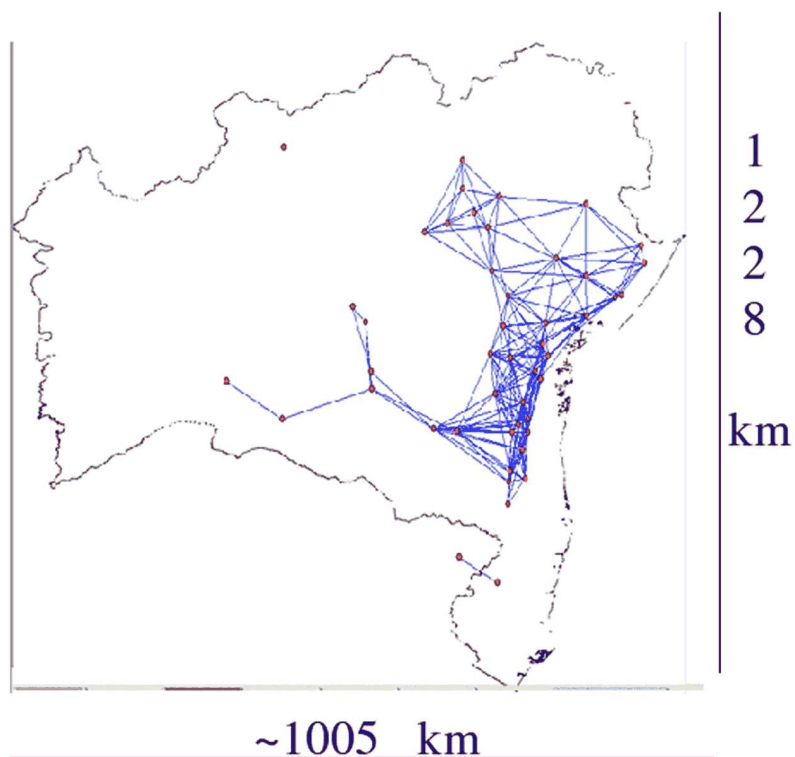


Figura 4.2: Rede gerada a partir da análise de dados de chuva da ADENE para períodos de primavera.

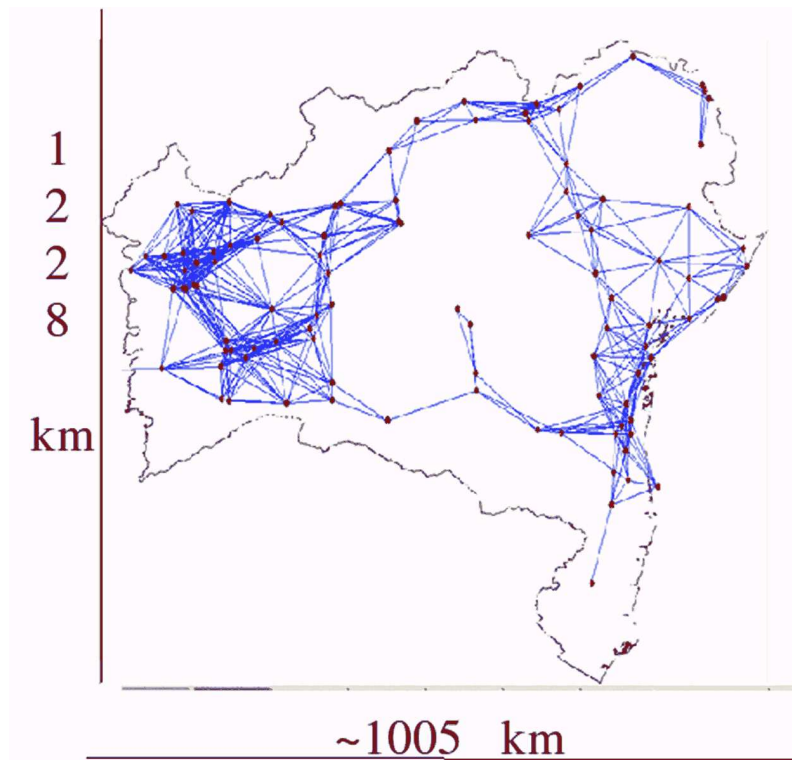


Figura 4.3: Rede gerada a partir da análise de dados de chuva da ADENE para períodos de verão.

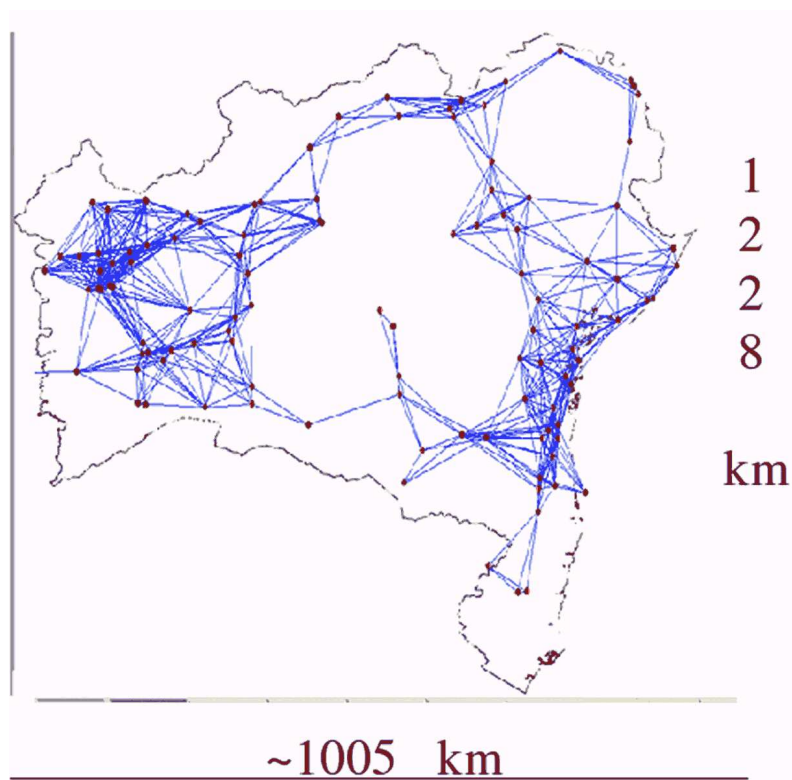


Figura 4.4: Rede gerada a partir da análise de dados de chuva da ADENE para períodos de outono.

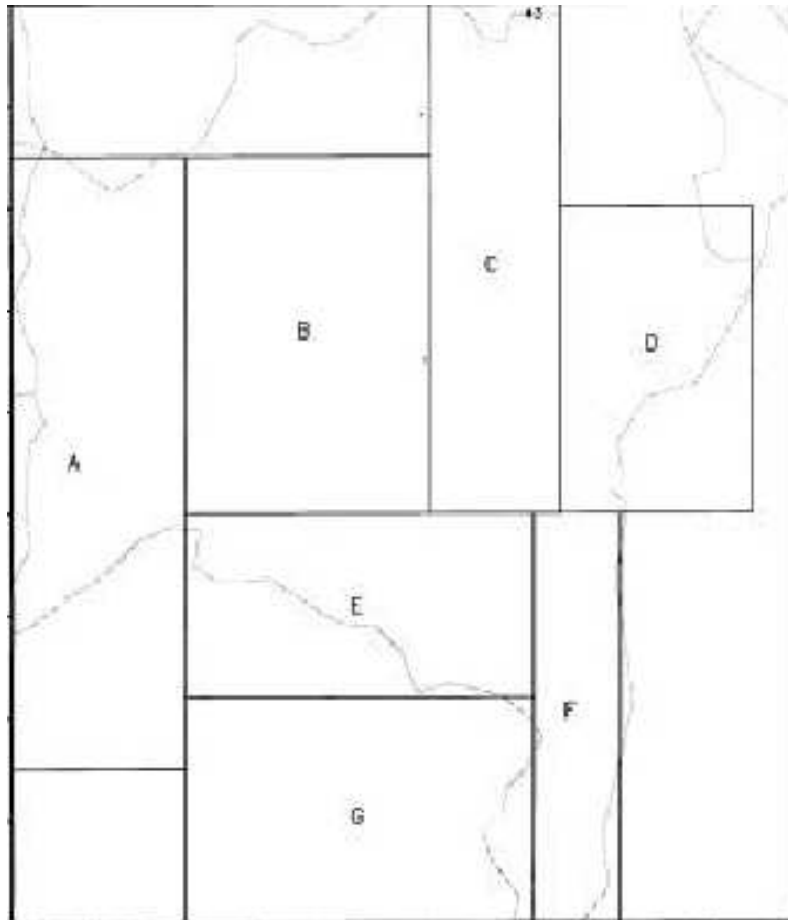


Figura 4.5: Divisão da Região Sul do Nordeste em áreas homogêneas. As regiões A, B, E, F, G são similares, apresentando ciclo anual bem definido, com período chuvoso em dezembro e março e alta variabilidade pluviométrica ao longo do ano - FONTE: (CHAVES, 1999).

4.2 Análise topológica (estatística)

Analisamos estatisticamente os índices caracterizadores de grafos, utilizando o programa NetAllbin e obtivemos os seguintes resultados para a variação do dia X no intervalo de 2 anos (1994 e 1995), utilizando dados da ADENE.

- Conforme observado na análise visual, as redes com grande número de vértices demonstram ser mais comuns durante o Outono e o Verão do que no Inverno e na Primavera, o que significa que o número de estações onde houve chuva no Outono e no Verão foi maior que o número de estações onde houve chuva no Inverno e na Primavera, conforme vemos na Figura 4.6.

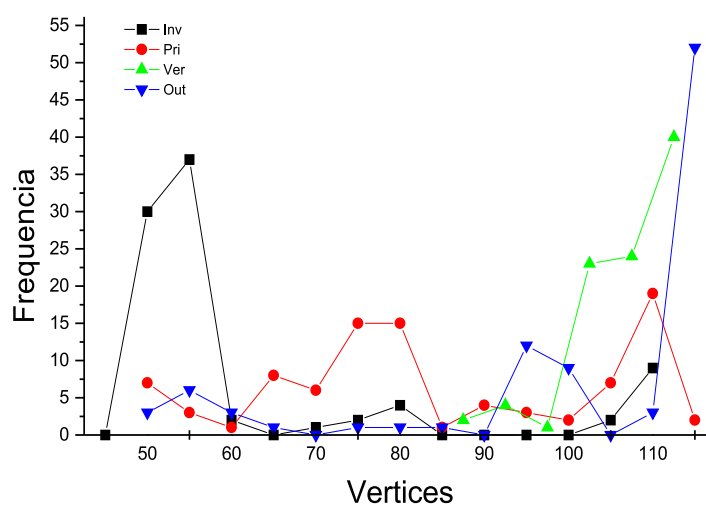


Figura 4.6: Distribuição sazonal do número de vértices das redes de estações pluviométricas.

- O grau médio das redes relativas aos dados de Inverno demonstra ser menor que das redes relativas aos dados de outras estações do ano. O grau médio de uma rede de estações pluviométricas significa o número médio de relações (arestas) de cada estação pluviométrica que compõe essa rede, ilustrada na Figura 4.7.
- O caminho mínimo de um vértice de uma rede estações pluviométricas significa qual a distância (número de arestas) mínima a ser percorrida para chegar desse vértice a qualquer outro vértice da rede. Fazendo a média dos caminhos mínimos de todos os vértices da rede é obtido o caminho mínimo médio da rede. Redes com valores altos de caminho mínimo médio demonstram ser mais comuns durante o Outono do que nas outras estações do ano, como mostrado na Figura 4.8.
- O maior caminho mínimo médio de uma rede é chamado de diâmetro. As redes relativas aos dados de Outono têm maior diâmetro que as redes relativas aos dados de outras estações. Entretanto, o diâmetro mais comum das redes das chuvas do Outono é próximo ao diâmetro mais comum das redes das chuvas

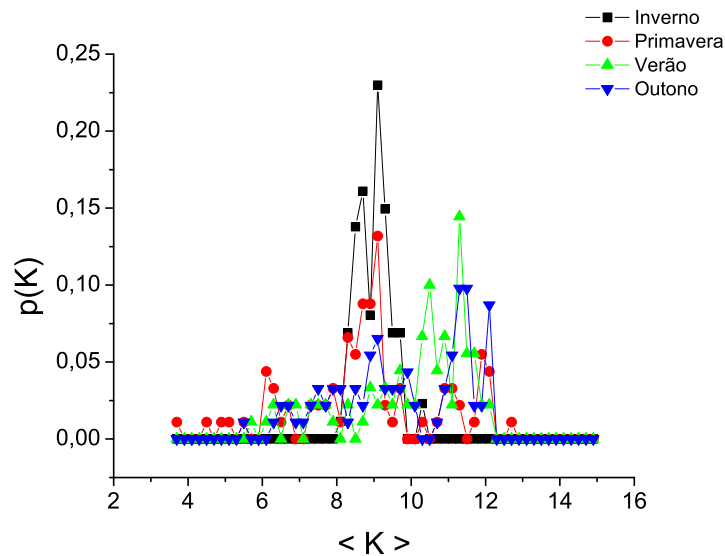


Figura 4.7: Distribuição sazonal do Grau médio das redes de estações pluviométricas.

das outras estações do ano. O diâmetro a distância máxima a ser percorrida para chegar de qualquer vértice a qualquer outro vértice da rede, conforme a Figura 4.9.

- Verificamos que é possível classificar de qual estação do ano os dados de chuva tratam se analisarmos especificamente a relação entre o *Caminho Mínimo Médio* e o *Diâmetro*. As redes relativas aos dados de verão e outono têm *caminho mínimo médio* e *diâmetro* maiores que as redes relativas aos dados de inverno e primavera, como mostra a Figura 4.10.
- O Coeficiente de aglomeração médio de uma rede mede o quanto uma rede é completa. O coeficiente de aglomeração médio demonstra ser maior nas redes de chuva do Outono do que nas redes de chuvas das outras estações do ano, como mostrado pela Figura 4.11.

Todos esses resultados estão de acordo com o que diz a literatura sobre a distribuição espacial da chuva no Nordeste: a chuva no verão e outono age de forma mais abrangente, enquanto a chuvas no inverno e primavera têm comportamento mais local.

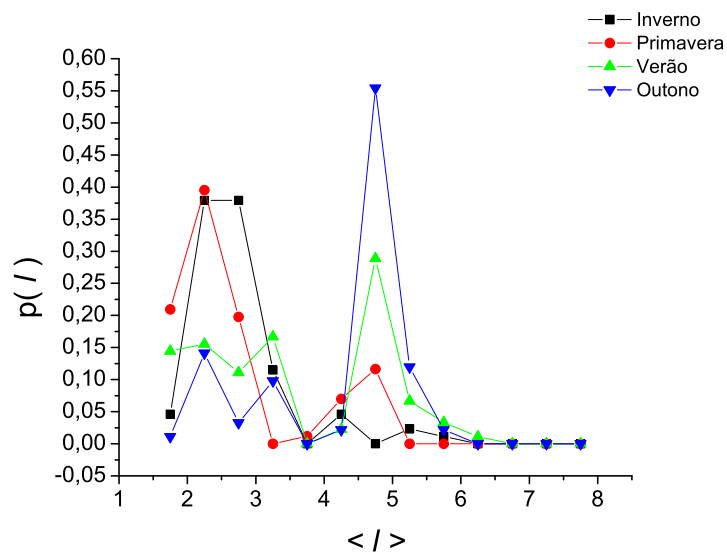


Figura 4.8: Distribuição sazonal do Caminho Mínimo Médio das redes de estações pluviométricas.

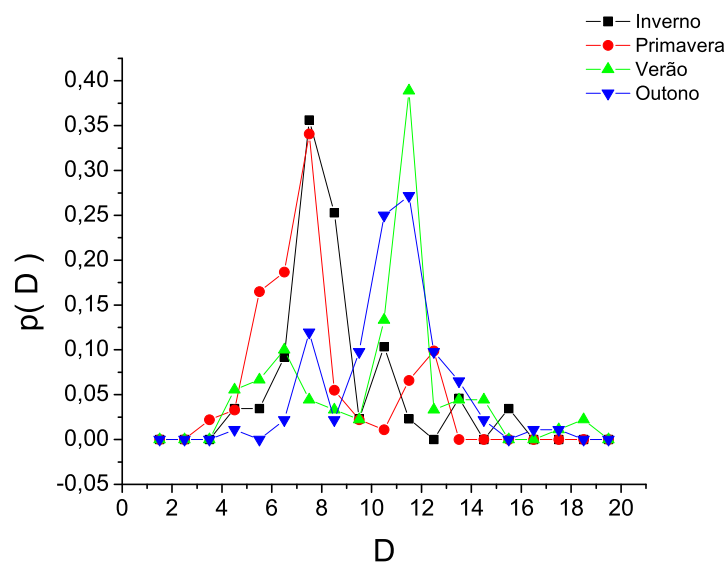


Figura 4.9: Distribuição sazonal do Diâmetro das redes de estações pluviométricas.

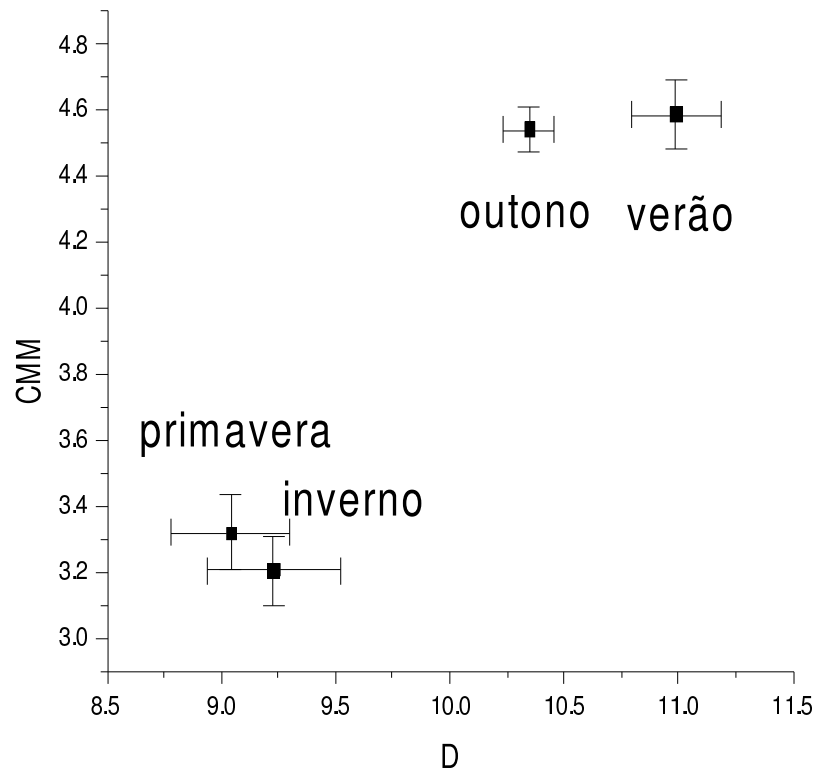


Figura 4.10: Relação entre Caminho mínimo médio e diâmetro.

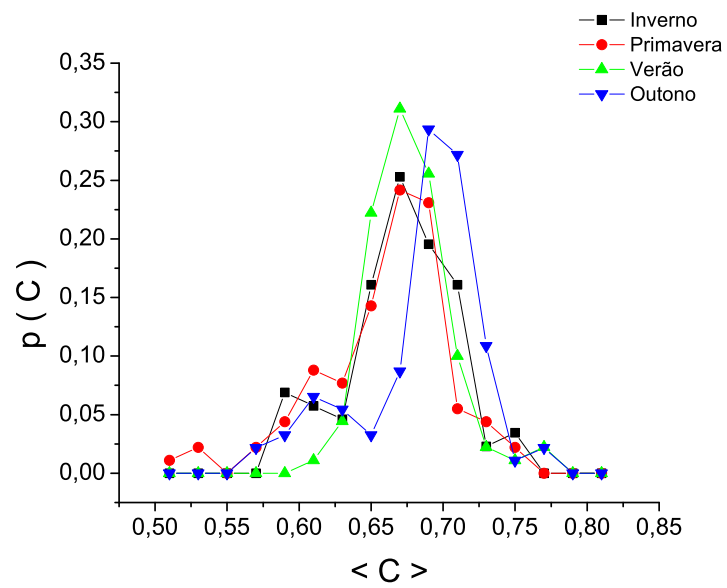


Figura 4.11: Distribuição sazonal do Coeficiente de Aglomeração Médio das redes de estações pluviométricas.

5 CORRELAÇÃO LINEAR ESTATÍSTICA - CLE

Neste capítulo, apresentamos os resultados para as redes construídas pelo método *CLE*. Este método difere-se do método *LE* pelo critério utilizado para a criação de arestas. Enquanto aquele impõe um limite espacial de $150Km$ para as conexões, este apenas considera a correlação estatística do acumulado pluviométrico em diferentes regiões, liberando, em tese, para conexões de diferentes tamanhos, inclusive maiores que $150Km$.

A fim de verificar a relação entre os tamanhos das conexões dos dois métodos, calculou-se a distribuição do tamanho das conexões de redes geradas pelo método *CLE*, considerando critério de correlação $r_{limite} = 0,7$, sugerido por Eguluz(2005), para os meses do ano de 1979. A Figura 5.1 mostra essa distribuição para os meses de janeiro, abril, agosto e outubro, típicos de verão, outono, inverno e primavera respectivamente.

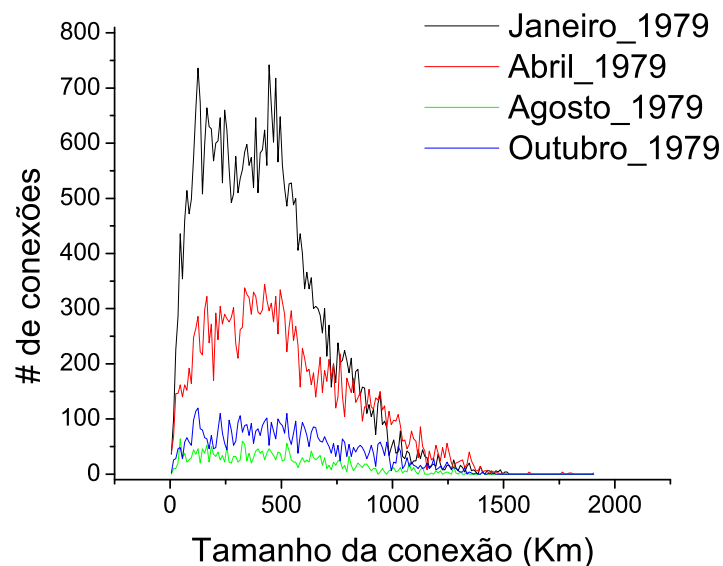


Figura 5.1: Distribuição dos tamanhos das conexões (em Km) de redes construídas pelo método *CLE*, com $r_{limite} = 0,7$, para diferentes meses do ano de 1979.

Percebe-se que todas as redes analisadas exibem conexões muito maiores que $150Km$. A maioria dessas conexões tem entre $200Km$ e $600Km$, mas algumas conexões relacionam vértices distantes até $1000Km$ entre si. Isso indica que duas

regiões a essa distância uma da outra têm sua variação pluviométrica correlacionada positivamente, demonstrando uma causalidade não-local da chuva na região Nordeste.

Além disso, o histograma demonstra que os meses de janeiro e abril, tipicamente meses com chuva mais distribuída na região Nordeste (CHAVES, 1999), apresentam maior quantidade de conexões mais distantes que os meses de agosto e outubro, típicos de inverno e primavera, quando a chuva não é tão distribuída.

Consideramos os diferentes tamanhos das conexões como diferentes escalas das redes de estações pluviométricas. Logo a Geometria Fractal e seus conceitos de invariância de escala tornam-se ferramenta de análise bastante indicada ao estudo dessas redes.

A seguir detalharemos os passos, bem como os principais resultados extraídos da observação das redes de chuva sob o olhar da Geometria Fracta.

5.1 Análise das características Fractais

O primeiro passo para a análise das características fractais das redes de estações pluviométricas foi estabelecer um parâmetro para comparação. Criamos duas redes completas hipotéticas.

Ambas as redes estão associadas aos dados da ADENE, que abrange 514 estações pluviométricas distribuídas no estado da Bahia. Uma delas é construída com os mesmos dados da ADENE, portanto a distribuição espacial dos seus vértices é a mesma das redes construídas com o método *CLE*. A outra rede possui os vértices distribuídos na forma de um reticulado, ou seja, todos os vértices estão à mesma distância um do outro.

Consideramos o raio inicial $raio = 10Km$ e quantificamos o número de conexões existentes a no máximo essa distância. Variamos o $raio$ em 10% de seu valor até atingirmos a distância máxima entre duas estações pluviométricas da rede e quantificamos o número de conexões para cada $raio$ diferente. Ao final, obtivemos duas distribuições do número de conexões que seguem lei de potência conforme mostram as Figuras 5.2 e 5.4.

- Rede Completa - expoente $b = 1,98942$ com $erro = 0,01003$
- Rede Reticulado - expoente $b = 1,885$ com $erro = 0,05925$

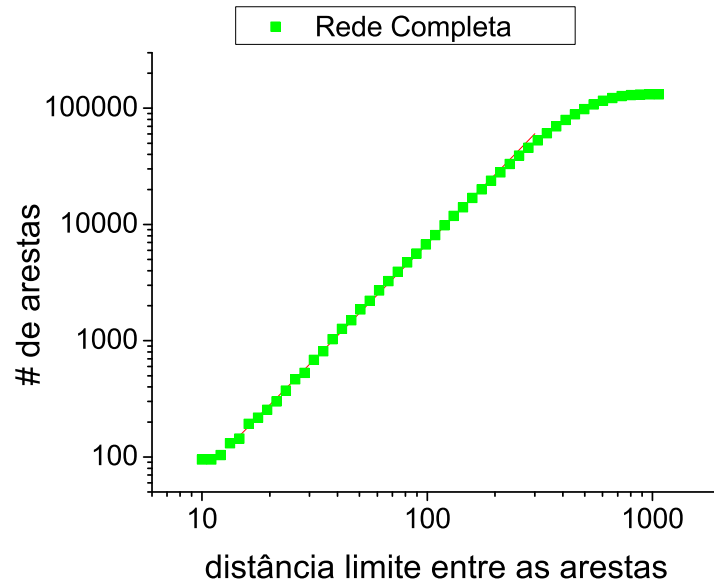


Figura 5.2: Cálculo da distribuição de arestas de uma rede completa com 514 vértices, para diferentes distâncias - expoente $b = 1,98942$ com $erro = 0,01003$

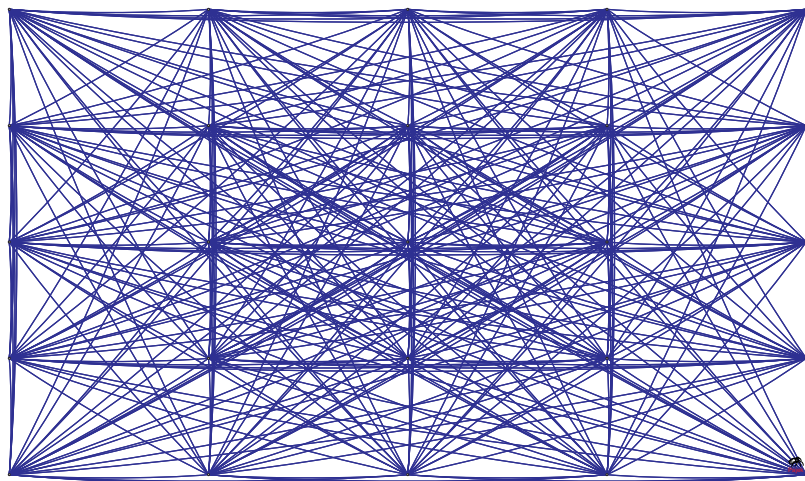


Figura 5.3: Exemplo de uma rede com vértices distribuídos em forma de Reticulado, completa e regular, com 25 vértices.

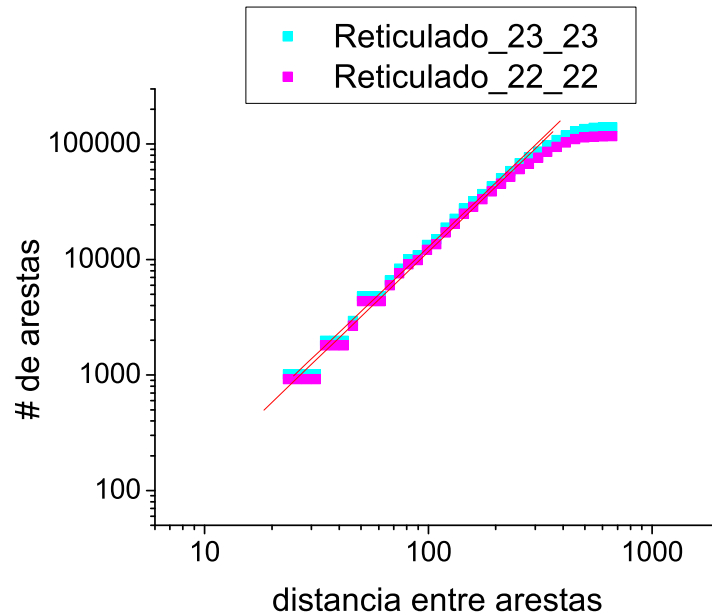


Figura 5.4: Cálculo da distribuição de arestas de um Reticulado com 529 vértices, para diferentes distâncias - expoente $b = 1,885$ com $erro = 0,05925$

Tomamos como parâmetros estes valores e analisamos da mesma forma as redes obtidas com os dados da ADENE e do NCAR. Usamos os dados da ADENE para analisar as redes do Estado da Bahia e os dados do NCAR para analisar as redes do restante do Nordeste (Norte do Nordeste) e de toda a região (Norte e Sul). Dessa forma, pudemos analisar o fenômeno da variação pluviométrica utilizando escalas iguais (Bahia e Norte do Nordeste, que têm aproximadamente a mesma dimensão - $1000Km \times 1000Km$) e escalas diferentes (Bahia e Nordeste inteiro, que tem dimensão $1000Km \times 2000Km$, maior que a da Bahia).

Como apenas entre os anos de 1979 e 1983 existem dados simultâneos da ADENE e do NCAR, utilizamos dados desse período para analisar as características fractais. Construímos redes mensais para todo o período, totalizando 60 redes.

Analisando as redes da Bahia e do Norte do Nordeste, verificamos a existência de lei de potência com expoentes diferentes dos encontrados com as redes completa e regular. Analisamos redes referentes ao mês de janeiro de 79 a 82. O expoente b da rede da Bahia variou entre 1,54 (no ano de 1979) e 1,75 (no ano de 1981) e o expoente da rede do Norte do Nordeste variou entre 1,55 (no ano de 1979) e 1,85 (no ano de 1980). Esses dados são mostrados na Tabela 5.1.

Local/Data	1979	1980	1981	1982
Norte do NE	1,55	1,85	1,7	1,72
Sul do NE	1,54	1,68	1,75	1,69
Nordeste	1,57	1,42	1,65	1,55

Tabela 5.1: Expoentes das distribuições de arestas das redes relativas a dados do Sul (Bahia), do Norte e de toda a região Nordeste

A rede regular exibe um comportamento uniforme, em que todos os vértices têm a mesma *importância* na rede, e na rede completa todos os vértices estão sujeitos à mesma *lei*, à mesma regra de conexão. E as conexões entre estações pluviométricas nas redes de estações pluviométricas se devem à variação da chuva. A diferença entre os valores do expoente dessas redes e os expoentes das redes completa e regular indicam que as conexões entre cada uma das redes acontecem de forma diferente. Ou seja, indicam que a variação da chuva não acontece de forma uniforme na Bahia e nem no Norte do Nordeste ou ainda que as estações pluviométricas não estão uniformemente distribuídas no espaço, privilegiando talvez áreas com maior população ou com mais riqueza.

Além disso, quando consideramos o erro referente ao cálculo do expoente das redes da ADENE e do NCAR, verificamos que existe sobreposição de valores, com exceção do ano de 1980. Apesar de se tratar de regiões diferentes, elas exibem o mesmo comportamento de distribuição de conexões.

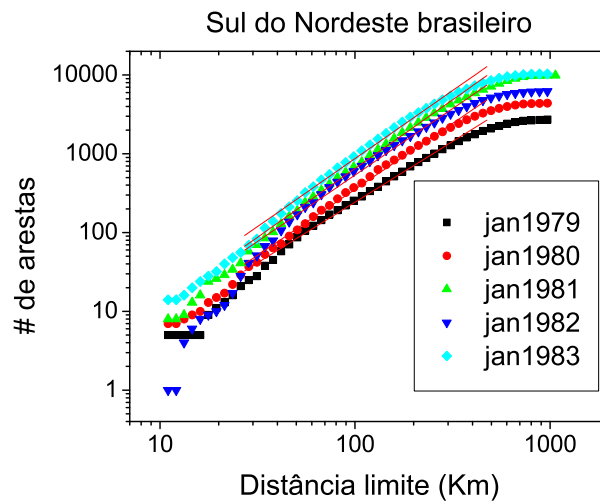


Figura 5.5: Cálculo da distribuição de arestas da rede de dados da Bahia para janeiro de 79, 80, 81, 82 e 83

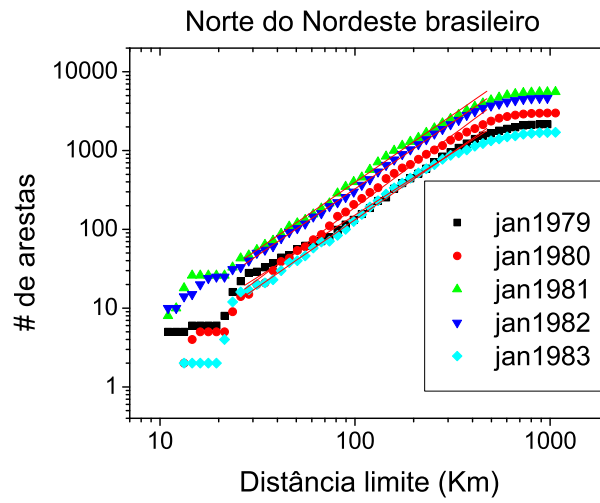


Figura 5.6: Cálculo da distribuição de arestas da rede de dados do Norte do Nordeste para janeiro de 79, 80, 81, 82 e 83

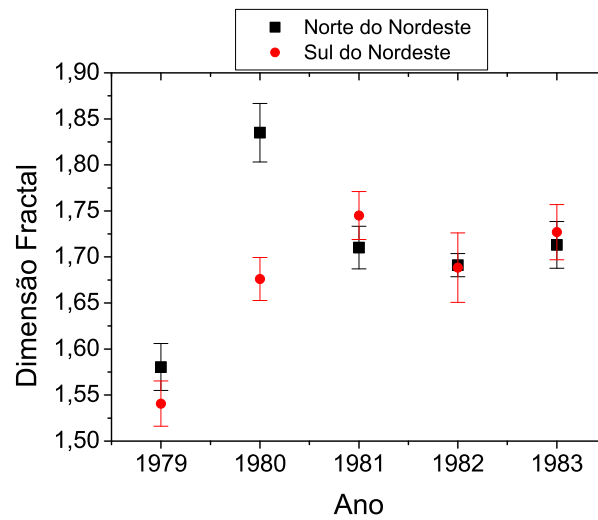


Figura 5.7: Valores dos expoentes das distribuições de arestas das redes de dados da Bahia e do Norte do Nordeste para janeiro de 79, 80, 81, 82 e 83

Analisando as redes da Bahia e as de todo o Nordeste, também verificamos a existência de lei de potência com expoentes diferentes dos encontrados para as redes completa e regular. Também analisamos redes referentes ao mês de janeiro de 79 a 83. O expoente b da rede da Bahia variou entre 1,54 (no ano de 1979) e 1,75 (no ano de 1981) e o expoente da rede do Nordeste inteiro variou entre 1,42 (no ano de 1980) e 1,65 (no ano de 1981), conforme mostrado na Tabela 5.1.

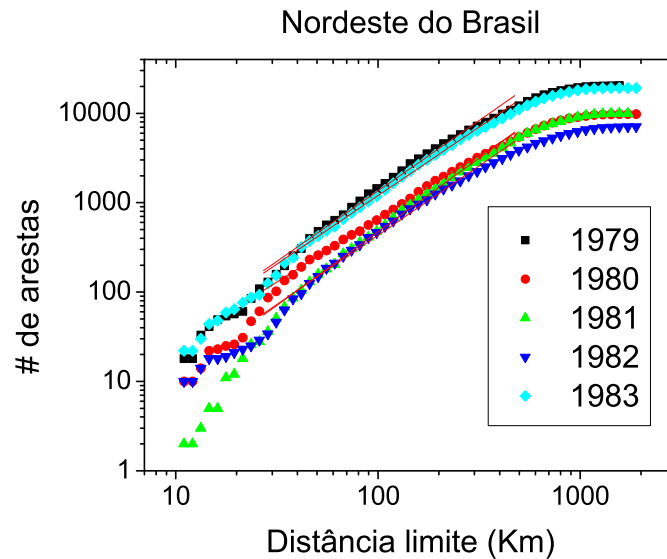


Figura 5.8: Cálculo da distribuição de arestas da rede relativa a dados de todo o Nordeste para janeiro de 79, 80, 81, 82 e 83

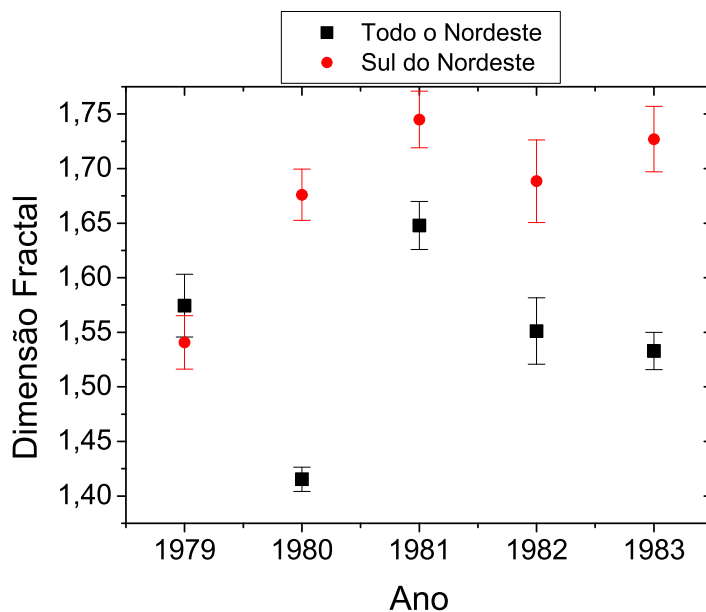


Figura 5.9: Sobreposição dos valores dos expoentes (com respectivos valores de *erro*) das distribuições de arestas das redes relativas a dados da Bahia e do Nordeste inteiro para janeiro de 79, 80, 81, 82 e 83

Também nessa comparação, a diferença entre os valores do expoente dessas redes e os expoentes das redes completa e regular indicam que as conexões entre cada uma das redes acontecem de forma diferente, e que a variação da chuva em todo o Nordeste também não acontece de forma uniforme ou ainda que as estações em toda a região Nordeste não estão uniformemente distribuídas no espaço.

A diferença entre esta comparação e a comparação anterior é que, quando consideramos o erro referentes ao cálculo do expoente das redes da Bahia e do Nordeste inteiro, não existe sobreposição de valores, com exceção do ano de 1979. O fato de estarmos lidando com escalas diferentes (o Nordeste inteiro é muito maior que o estado da Bahia) pode ser uma razão para esta diferença.

5.2 Identificação da correlação crítica

Conforme dissemos na seção 3.3.2, e como demonstra a figura 5.10 à medida que o valor do r_{limite} usado para construir as redes é maior, menos arestas são consideradas, porque o critério para criação de arestas se torna mais exigente, e a variação no número de arestas causa variação nos valores de todos os outros índices.

Quando a diminuição do número de arestas de uma rede implica em crescimento do valor do *Caminho Mínimo Médio*, pode significar que deixaram de ser criados atalhos entre vértices, conforme mostram as figuras 5.11 e 5.12.

Quando a diminuição no número de arestas não implica em aumento do valor do *Caminho Mínimo Médio* pode significar que a perda de aresta fez com que a rede se dividisse, como mostram as figuras 5.13 e 5.14. A esse efeito de fragmentação da rede, chamamos *criticalidade*.

Entretanto, para analisar a criticalidade, é preciso observar outros índices de

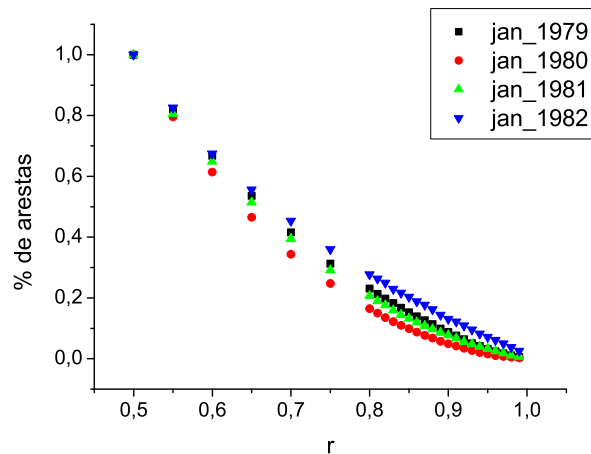


Figura 5.10: Variação do número de arestas com o crescimento do r_{limite}

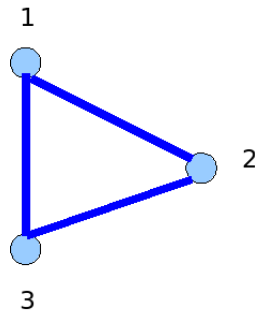


Figura 5.11: Exemplo de rede completa, com 3 vértices e Caminho Mínimo Médio igual a 1

caracterização das redes que trazem mais informações sobre quantidade e qualidade de conexões, tais como *Grau Médio* e *Coefficiente de Aglomeração Médio*.

O estudo desses índices pode apontar a existência de um r_{limite} crítico a ser considerado na construção das redes. Redes construídas utilizando esse critério de correlação podem trazer mais informações sobre o fenômeno modelado (variabilidade pluviométrica) em sua estrutura topológica. A escolha desse r , ao qual chamaremos de $r_{critico}$ se baseou no estudo da topologia das redes, seguindo as etapas descritas a seguir.

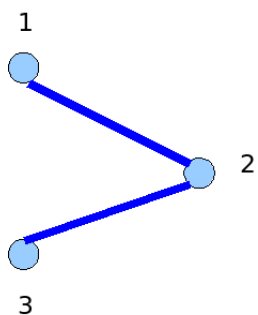


Figura 5.12: Exemplo de rede com 3 vértices e Caminho Mínimo Médio igual a 1,33

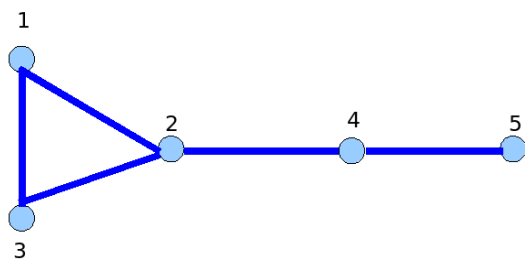


Figura 5.13: Exemplo de rede conexa, com 5 vértices e Caminho Mínimo Médio igual a 1,7

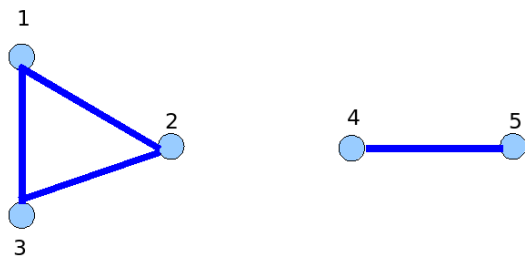


Figura 5.14: Exemplo de rede desconexa, com 5 vértices e Caminho Mínimo Médio igual a 1

5.2.1 Geração das redes complexas

Foram geradas redes relativas a cada mês dos anos de 1979 a 1982, para diferentes valores de r - no nosso caso usamos os seguintes valores de r : 0,5; 0,55; 0,6; 0,65; 0,7; 0,75; 0,8; 0,81; 0,82; 0,83; 0,84; 0,85; 0,86; 0,87; 0,88; 0,89; 0,90; 0,91; 0,92; 0,93; 0,94; 0,95; 0,96; 0,97; 0,98; 0,99.

Os valores de r são todos maiores que 0,5 para garantir que redes representassem correlações significativas. E variamos esse valor de 0,05 até atingir 1,0, a fim de identificar o valor mais representativo da correlação. Para valores maiores que 0,8 as redes exibiam comportamentos críticos, em que o crescimento do Caminho Mínimo Médio mudava de direção (deixava de crescer para decrescer), o que representa que as arestas perdidas não formavam apenas atalhos, mas garantiam a conectividade da rede, ou seja, que a rede havia se dividido. A partir desse valor de correlação, variamos o r de 0,01 até atingir 1,0, para estudar esse comportamento crítico com mais detalhes.

A Figura 5.15 exemplifica a variação da estrutura das redes de acordo com mudanças no valor de r para valores de r iguais a 0,55, 0,65, 0,75, 0,85.

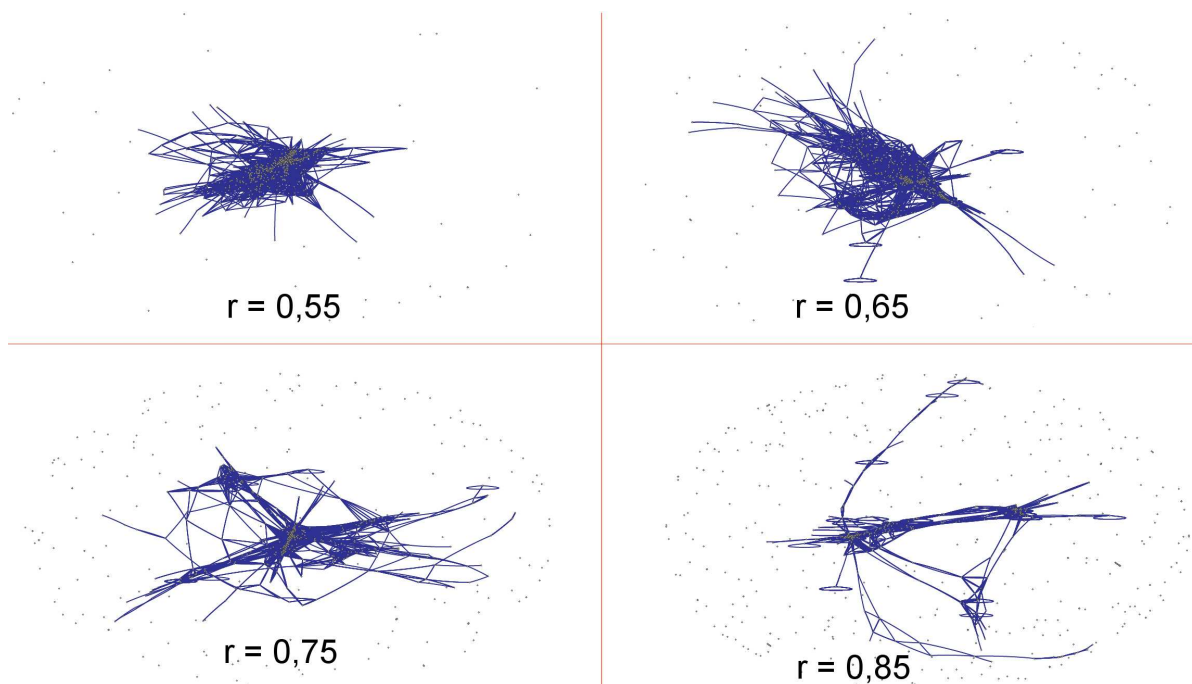


Figura 5.15: Redes de estações pluviométricas da região Nordeste, geradas para o mês de janeiro de 1979 com r_{limite} igual a 0,55, 0,65, 0,75, 0,85.

5.2.2 Cálculo dos índices caracterizadores das redes

Para estas redes geradas, calculamos valores dos índices de caracterização de redes complexas: Caminho Mínimo Médio $\langle l \rangle$; Coeficiente de Aglomeração Médio $\langle C \rangle$; Diâmetro D ; Grau Médio $\langle K \rangle$; Número de Pares Desconectados; Número de Arestas.

Identificamos, para cada uma das redes geradas, quais valores de r estavam associados à mudança de comportamento do *Caminho Mínimo Médio* e comparamos com a estrutura topológica das redes, demonstrada pelos valores de $\langle K \rangle$ e de $\langle C \rangle$, como mostram os gráficos das figuras 5.16 a 5.18.

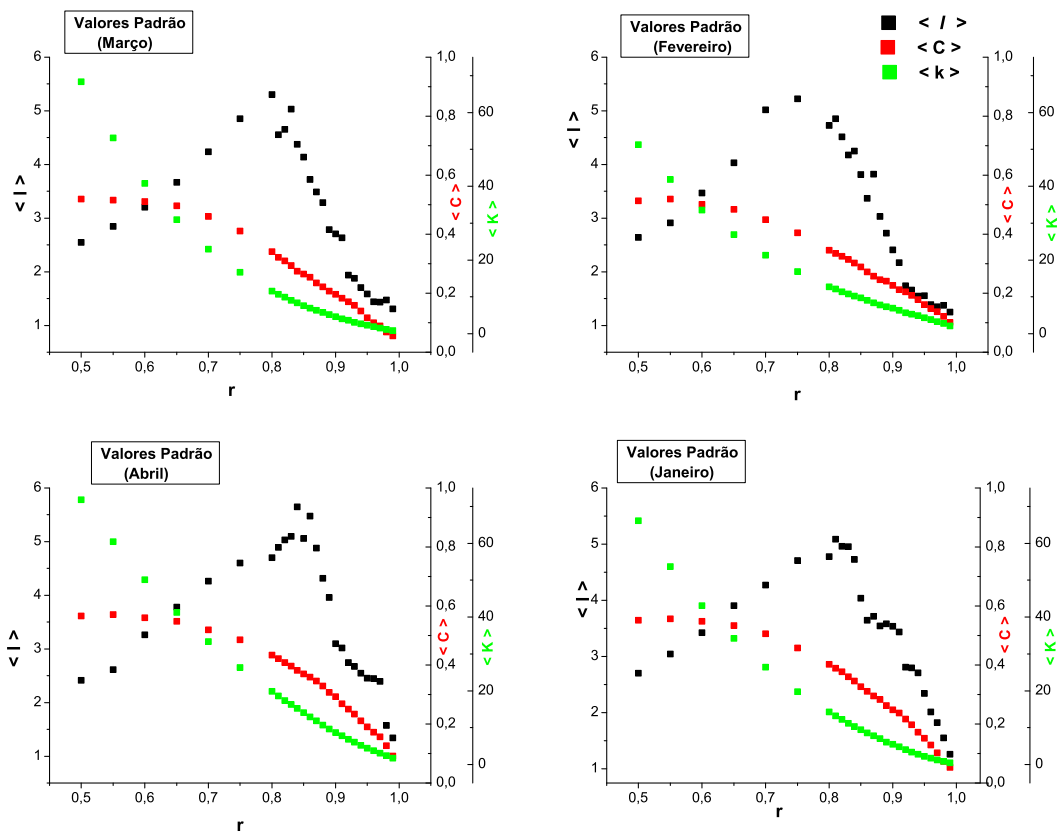


Figura 5.16: Comportamento do $\langle l \rangle$, do $\langle K \rangle$ e do $\langle C \rangle$ para redes de estações pluviométricas da região Nordeste, para os meses de Janeiro a Abril de 1979, geradas com diferentes valores de r

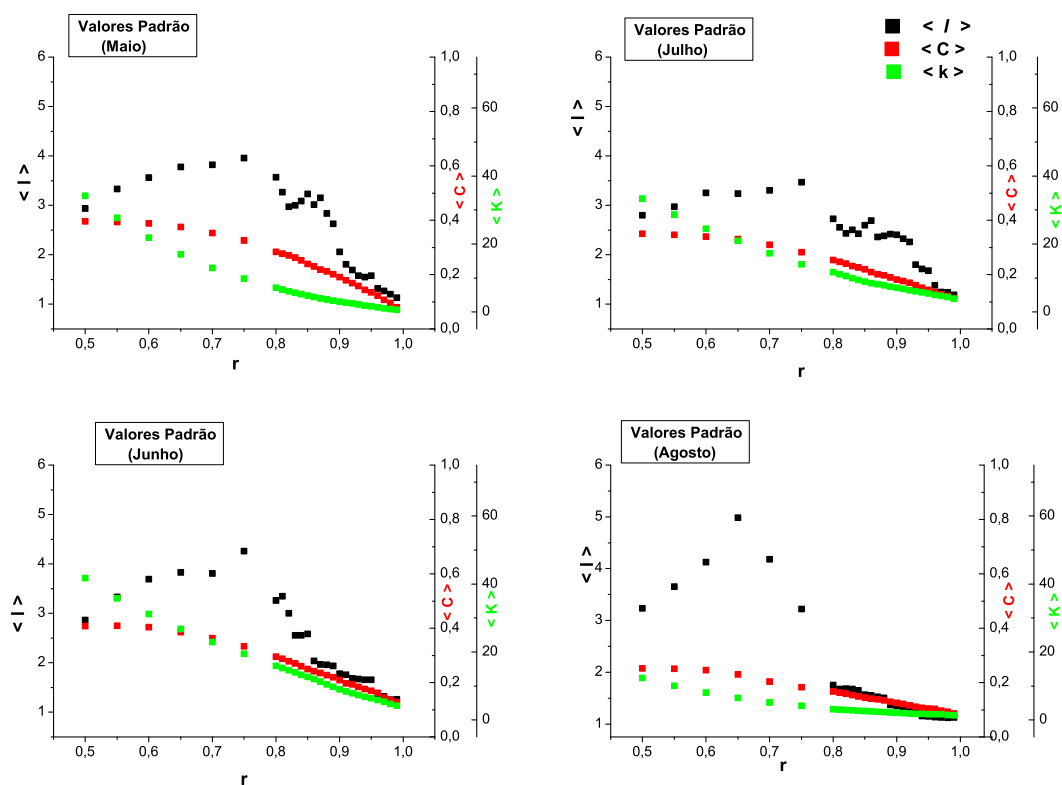


Figura 5.17: Comportamento do $\langle l \rangle$, do $\langle K \rangle$ e do $\langle C \rangle$ para redes de estações pluviométricas da região Nordeste, para os meses de Maio a Agosto de 1979, geradas com diferentes valores de r

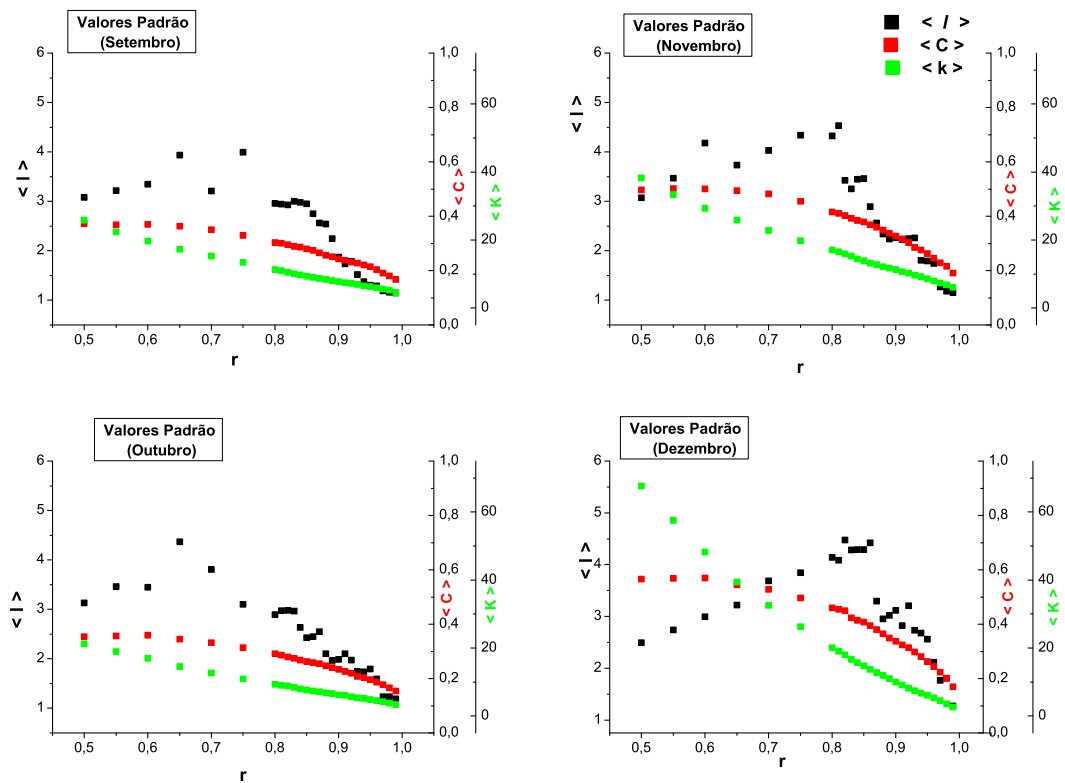


Figura 5.18: Comportamento do $\langle l \rangle$, do $\langle K \rangle$ e do $\langle C \rangle$ para redes de estações pluviométricas da região Nordeste, para os meses de Setembro a Dezembro de 1979, geradas com diferentes valores de r

Entre os meses de maio a outubro percebe-se que os valores de $\langle K \rangle$ e $\langle C \rangle$ são muito menores que os verificados em outros meses do ano. Por exemplo: Para o mês de janeiro, a rede construída com r de 0,7 tem $\langle K \rangle$ igual a 25, enquanto para o mês de agosto, a rede construída com mesmo valor de r tem $\langle K \rangle$ igual a 5. Mas, para essas mesmas redes, o $\langle C \rangle$ vale 0,52 em janeiro e 0,22 em agosto.

Percebe-se também que tanto $\langle K \rangle$ e $\langle C \rangle$ diminuem à medida que o critério de correlação cresce. Isso está associado ao fato de que quanto mais exigente é o critério de correlação, menos arestas serão consideradas. Como a retirada de arestas está relacionada ao crescimento do $\langle l \rangle$, pode-se sugerir duas hipóteses sobre a *perda* de arestas:

- **Hipótese 1** - A rede complexa, quando o r é igual a 0,5, é conexa e formada por um grande aglomerado de arestas. As arestas *perdidas* formavam atalhos, ligando pequenas subredes dentro de uma mesma rede conexa. Nos pontos em que o crescimento do $\langle l \rangle$ muda de direção (o $\langle l \rangle$ decresce com o aumento do critério de correlação, em vez de crescer) ocorre uma *quebra* nas redes, e são perdidas arestas que não formavam apenas atalhos entre subredes de uma mesma rede conexa, mas que representavam as arestas-conectividade da rede. Cada subrede pode estar associada a um conjunto de localidades em que a variabilidade pluviométrica é semelhante, o que em climatologia se chama *regionalização*.
- **Hipótese 2** - A rede complexa é desconexa e formada por várias pequenas subredes. As arestas “perdidas” formavam atalhos dentro dessas subredes e a mudança de direção do crescimento do $\langle l \rangle$ está associada à quebra dessas subredes. Ou seja, mesmo com critérios de correlação menores, as redes já representam uma divisão de regiões com variabilidade pluviométrica semelhante.

A diferença entre essas duas hipóteses é que a segunda sugere que as redes conseguem representar diferentes regiões com variabilidade pluviométrica semelhante, mesmo com critérios de correlação menores (redes formadas por várias subredes desconexas) enquanto a primeira hipótese sugere que a identificação de regiões com variabilidade pluviométrica semelhante só acontece a partir de um certo *rlimite*, onde passa a ocorrer a quebra da rede.

Uma forma de validar essas hipóteses é avaliar a variação do $\langle K \rangle$ da rede. Redes desconexas, como descritas na hipótese 2, tendem a ter valores de $\langle K \rangle$ menores, mesmo para valores de $\langle C \rangle$ maiores, enquanto que as descritas na hipótese 1 tendem a ter valores de $\langle K \rangle$ maiores. Isso porque o valor do $\langle K \rangle$ é um valor absoluto, que está associado diretamente ao *número* de conexões da rede, enquanto o $\langle C \rangle$ é um valor relativo, variando entre 0 e 1, e que fala sobre a qualidade das conexões. Se uma rede possui várias subredes desconexas entre si, estas sub-redes podem ter valores altos de $\langle C \rangle$, desde que suas conexões internas sejam coesas, mas dificilmente terão $\langle K \rangle$ alto porque o número de conexões é limitado pelo número de vértices presentes em cada sub-rede.

Um problema é como identificar *altos valores* ou *baixos valores* de $\langle K \rangle$. Para determinadas redes, 70 pode ser um valor baixo de $\langle K \rangle$, enquanto que para outras pode ser um valor alto. De forma semelhante, para redes com menos vértices, 0,6 é um valor baixo de $\langle C \rangle$, enquanto que para redes com milhares de vértices esse seria um valor alto.

Analizamos os valores de $\langle K \rangle$ para o momento da quebra das redes. Dessa forma, verificamos que a grande maioria das redes possuem o mesmo valor de $\langle K \rangle$ nesse ponto, em torno de 15. Apenas os meses de junho e agosto se diferenciam com valores de 20 e de 5 respectivamente.

Se considerarmos que 15 é um valor alto de $\langle K \rangle$, valerá a hipótese 1, de que as redes se iniciam com um grande bloco de conexões que vai perdendo atalhos à medida que cresce o critério de correlação e se divide para alguns valores desses critérios.

Por outro lado, se considerarmos 15 um valor baixo de $\langle K \rangle$, valerá a hipótese 2, de que as redes se iniciam com vários pequenos blocos desconexos que vão perdendo atalhos e se dividindo à medida que cresce o critério de correlação.

Mas também podemos analisar os valores de $\langle K \rangle$ para as redes com um mesmo r , o menor deles (0,5), por exemplo. Para os meses de janeiro a abril e para o mês de dezembro, verificaram-se valores de $\langle K \rangle$ acima de 50. Para os outros meses verificaram-se valores abaixo de 40, com destaque para o mês de agosto com $\langle K \rangle$ 12. Os valores estão diretamente relacionados com o perfil de distribuição espacial da chuva nos períodos em questão. Janeiro, Abril e Dezembro são períodos de máximo pluviométrico em várias regiões do Nordeste.

Nesse caso, podemos considerar que 50 é um valor alto de $\langle K \rangle$, mas que 40 é um valor baixo. Assim, as redes de janeiro a abril e de dezembro obedeceriam à regra descrita pela hipótese 1 enquanto as demais redes obedeceriam à regra descrita pela hipótese 2.

5.2.3 Identificação e análise das redes críticas

Verificamos semelhança no padrão de de variação do $\langle l \rangle$ para meses iguais nos 4 anos (1979, 1980, 1981 e 1982), conforme mostram os gráficos relativos ao meses de abril e agosto, exibidos nas Figuras 5.19 e 5.20.

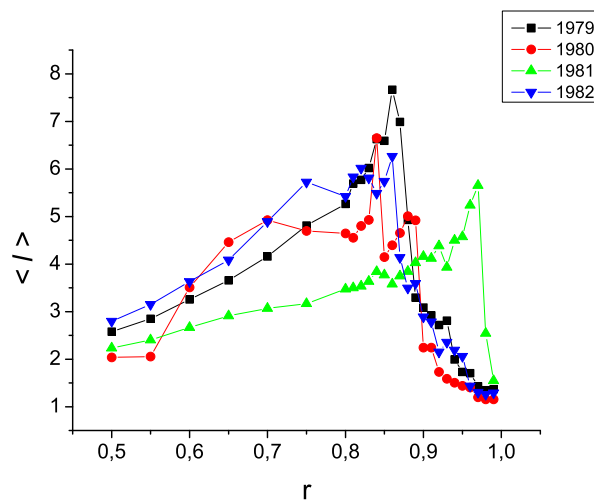


Figura 5.19: Valores do $\langle l \rangle$ para o mês de Abril, nos anos de 1979 a 1982, relacionado com diferentes valores de r

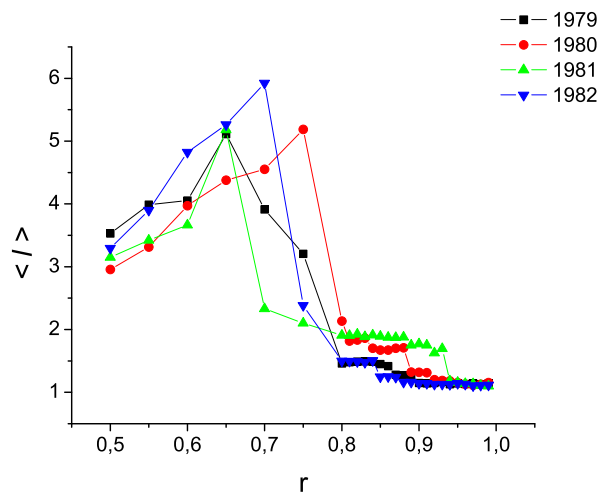


Figura 5.20: Valores do $\langle l \rangle$ para o mês de Abril, nos anos de 1979 a 1982, relacionado com diferentes valores de r

Apoiados nessa semelhança, calculamos a média dos valores de $\langle l \rangle$ associados aos valores críticos de r , em cada mês para os 4 anos estudados, obtendo um valor único de $\langle l \rangle$ para cada mês, apesar de média só ser considerada um valor representativo para um conjunto de dados maior ou igual a 10.

Tomando a sequência desses valores médios de $\langle l \rangle$, identifica-se como $r_{critico}$ de cada mês, aquele r para o qual ocorre mudança no crescimento do Caminho Mínimo nesta sequência. Para os meses de Novembro, Dezembro, Janeiro, Fevereiro, Março e Abril o valor de $r_{critico}$ está em torno de 0,8, conforme mostram as figuras 5.21 e 5.22. Para os meses de Maio, Junho e Julho, o valor de r_{limite} está em torno de 0,7, como mostra a figura 5.23. E para os meses de Agosto, Setembro e Outubro o valor está em torno de 0,6, como mostra a figura 5.24.

Foi possível agrupar os meses do ano utilizando os valores de $r_{critico}$ como parâmetro, e esses grupos estão diretamente associados aos períodos de máxima e mínima precipitação da região Nordeste, conforme nos diz Chaves(1999), Barbosa(1999) e Sugahara(1982).

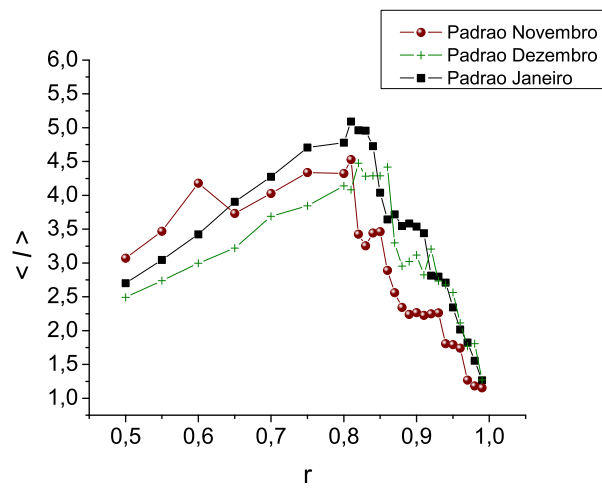


Figura 5.21: Valores do $\langle l \rangle$ para o período de Novembro, Dezembro e Janeiro, definido pelas médias, relacionado com diferentes valores de r

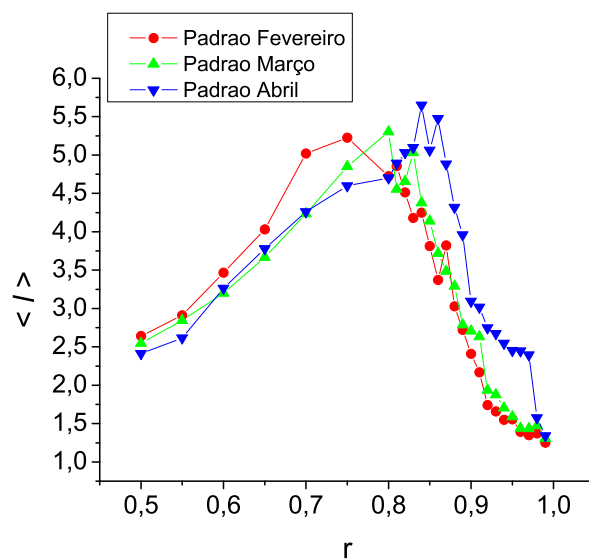


Figura 5.22: Valores do $\langle l \rangle$ para o período de Fevereiro, Março e Abril, definido pelas médias, relacionado com diferentes valores de r

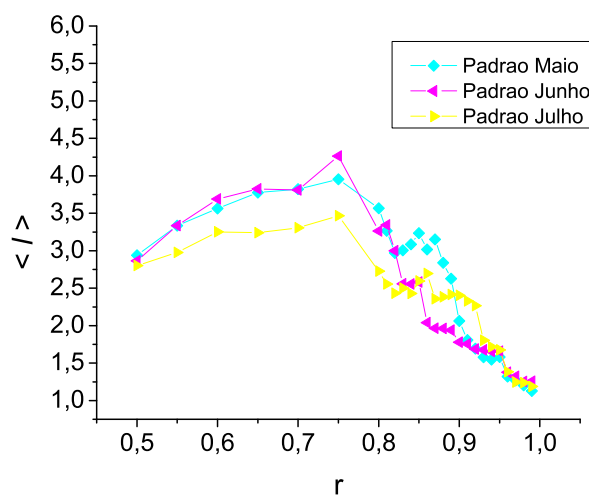


Figura 5.23: Valores do $\langle l \rangle$ para o período de Maio, Junho e Julho, definido pelas médias, relacionado com diferentes valores de r

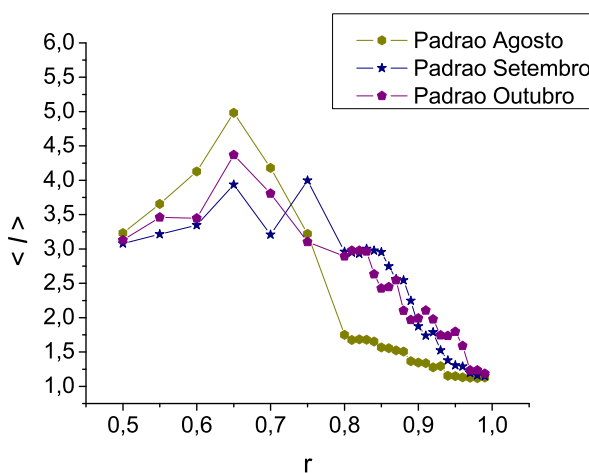


Figura 5.24: Valores do $\langle l \rangle$ para o período de Agosto, Setembro e Outubro, definido pelas médias, relacionado com diferentes valores de r

Fica evidenciado que os valores dos $r_{critico}$ variam de acordo com a distribuição das chuvas no Nordeste. Os meses com $r_{critico}$ semelhantes (Novembro, Dezembro, Janeiro; Fevereiro, Março Abril; Maio, Junho, Julho; Agosto, Setembro, Dezembro) também representam períodos de máxima e mínima pluviometria nas diferentes regiões do Nordeste, conforme foi visto na seção 2.1. Este resultado mostra que o cálculo do $r_{critico}$ pode ser usado para avaliar sazonalidade, períodos de máxima precipitação ou outras características climáticas não abordadas em nosso trabalho.

Após essa sequência de passos, foram identificados os $r_{criticos}$ mensais, conforme mostra a Tabela 5.2. Chamamos as redes construídas seguindo a associação entre $r_{critico}$ e o mês do ano de *redes críticas*.

MESES	$r_{critico}$
Janeiro	0,81
Fevereiro	0,75
Março	0,8
Abril	0,84
Maio	0,75
Junho	0,75
Julho	0,75
Agosto	0,65
Setembro	0,65
Outubro	0,65
Novembro	0,81
Dezembro	0,83

Tabela 5.2: Valores do $r_{critico}$ para cada mês do ano.

5.3 Análise Topológica das redes críticas.

Uma vez identificado o $r_{critico}$ para cada mês do ano, foi feita a análise dos índices caracterizadores de redes complexas utilizando estes valores como critérios de correlação.

Semelhante ao que foi feito com as redes geradas pelo método *LE*, e explicitado na Seção 4.2, comparamos valores do *Caminho Mínimo Médio* ($\langle l \rangle$) e do *Diâmetro* (D) de diferentes redes.

Utilizamos o $r_{critico}$ de cada mês como critério de correlação. Os exemplos mostrados a seguir, na figura 5.25, referem-se à comparação entre os meses de Janeiro e Agosto, por serem típicos de Verão e Inverno, com $r_{criticos}$ iguais a 0,81 e 0,65, respectivamente.

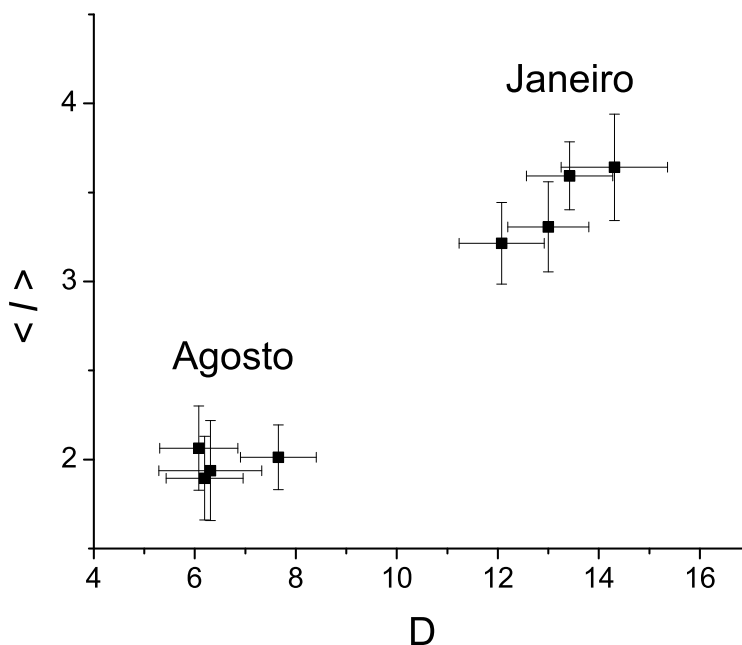


Figura 5.25: Relação entre $\langle l \rangle$ e D dos meses de Janeiro e Agosto dos anos de 1979 a 1982

É nítida a diferença entre os índices dos dois meses. Janeiro apresenta $\langle l \rangle$ e D maiores que Agosto, refletindo comportamento da distribuição pluviométrica da região Nordeste nestes meses e conseqüentemente nas estações chuvosa (verão/outono) e de seca (inverno/primavera).

Em Agosto, mês típico de inverno, ocorrem poucas chuvas no Nordeste, a maioria delas na área costeira. As redes de correlação, portanto, captam uma quantidade menor de estações, apresentam menos arestas, e conseqüentemente o diâmetro tende a ser menor.

Conforme explicitado na seção 3.3.2, como as arestas representam a existência de alta correlação na variação da quantidade de chuva das estações pluviométricas, redes muito conectadas representam períodos em que a variação da quantidade de chuva foi parecida na maioria das estações, ou seja, representam períodos de chuva mais igualmente distribuída na região. Analogamente, redes com poucas conexões

representam períodos em que a variação da quantidade de chuva foi diferente na maioria das estações, períodos de grande variabilidade pluviométrica na região.

Também comparamos os valores de $\langle l \rangle$ com os de $\langle C \rangle$. Na figura 5.26 fica nítida a relação entre estes dois índices e o período do ano: o mês de janeiro é caracterizado por valores de $\langle C \rangle$ mais altos que o mês de agosto.

Considerando que $\langle C \rangle$ está associado à completude da rede, verifica-se no gráfico que as redes de Janeiro são mais completas que as de Agosto. A completude está associada ao grau dos vértices que compõem a rede e também pode ser verificada pelo K , no gráfico da Figura 5.27. Neste gráfico verifica-se que Janeiro também possui valores de K maiores que Agosto.

Os valores descritos informam que existem mais conexões nas redes de Janeiro que nas de Agosto, o que significa que as localidades estudadas têm suas chuvas mais correlacionadas no primeiro mês que no segundo. Os resultados indicam, portanto, que é possível verificar o grau de distribuição pluviométrica de uma região em um determinado período através da análise de índices como $\langle C \rangle$ e K .

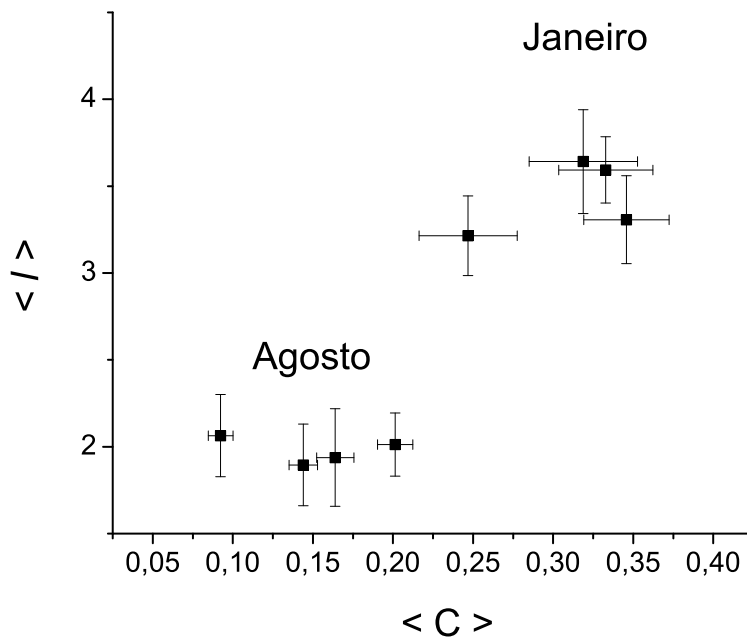


Figura 5.26: Relação entre $\langle C \rangle$ e l dos meses de Janeiro e Agosto dos anos de 1979 a 1982

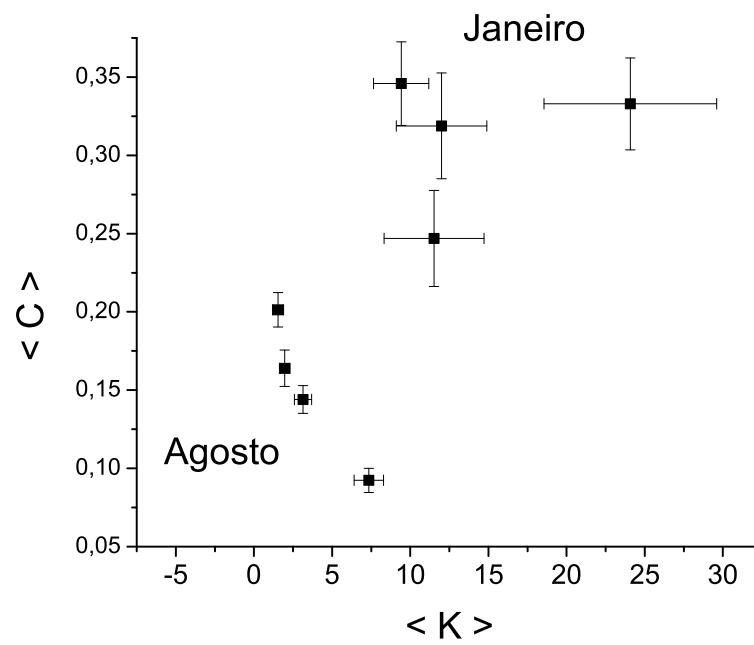


Figura 5.27: Relação entre $\langle C \rangle$ e K dos meses de Janeiro e Agosto dos anos de 1979 a 1982

6 CONCLUSÕES

Considerando o comportamento complexo da distribuição pluviométrica do Nordeste brasileiro (alta variabilidade espacial e temporal da precipitação) desenvolvemos uma metodologia de análise de dados pluviométricos baseada em redes complexas com relacionamentos segundo critérios de *Limitação Espacial* e de *Correlação Linear estatística*. A construção dessas redes possibilitou analisar os dados pluviométricos sob a ótica de diferentes Sistemas Complexos.

A análise topológica e estatística das redes, permitiu identificar padrões relacionados com a sazonalidade. Redes relativas a períodos com chuva mais distribuída ao longo do território (meses de verão e outono) possuem características diferentes das relativas a períodos de chuva mais localizada na região litorânea (meses de inverno e primavera). Estes resultados corroboram para os obtidos por Santana(2005), entretanto a maior base de dados traz maior confiança aos resultados.

Para as redes construídas pelo método *Limitação Espacial*, essa sazonalidade é melhor verificada pelos índices *Caminho Mínimo Médio* e *Diâmetro*.

Para as redes construídas pelo método *Correlação Linear Estatística*, verificou-se a existência de comportamento crítico dependente do critério de correlação escolhido. O aumento do critério de correlação produz redes com cada vez menos conexões, e conseqüentemente com alterações em sua topologia. Até determinado valor de correlação as redes perdem arestas, mas o valor do caminho mínimo médio cresce, o que indica que as arestas perdidas são “atalhos” e não provocam quebra da rede. Entretanto, a partir do momento em que o valor do caminho mínimo médio começa a diminuir, as redes começam a se dividir. Esse momento de divisão é chamado momento crítico, pois as redes começam a perder características de sua topologia. À correlação correspondente a essa divisão chamamos *correlação crítica*.

Essa informação de criticalidade foi importantíssima para guiar a análise dos índices clássicos de caracterização de redes. Entende-se que o momento de criticalidade é aquele em que a rede traz mais informações sobre o fenômeno estudado e isso motivou o uso desse critério para definir que valor de correlação utilizar para construir as redes em cada período do ano.

Semelhante aos resultados obtidos com o método de *Limitação Espacial*, o estudo das redes críticas também apontou variabilidade sazonal em sua topologia. O *Coefficiente de Aglomeração Médio* e o *Caminho Mínimo Médio* revelaram características diferentes para estações com chuva mais distribuída (verão e outono) e para os períodos de chuva mais localizada (inverno e primavera). De forma análoga, o estudo do *Grau Médio* e do *Diâmetro* também demonstrou essas características sazonais.

Por fim, o estudo da distribuição dos tamanhos das conexões (distâncias entre as

estações que compõem arestas) para diferentes escalas, revelou características fractais na topologia das redes de chuva estudadas. O estabelecimento de um parâmetro de comparação formado pelas redes completas (uma rede completa referente aos dados de chuva e outra referente a um reticulado de mesmo número de arestas) permitiu inferir que as conexões de nossas redes não são homogêneas. Cada região apresenta características de conexão própria, resultado apontado pela diferença entre os índices fractais dessas redes com os do reticulado. Também estão longe de serem completas, apontados por comparação análoga com as redes completas.

Além disso, a análise da variação dos índices fractais das redes para dados da região Sul do Nordeste (Bahia) e da região Norte do Nordeste (demais estados da região) apontou valores bem próximos para as diferentes regiões. O que indica que o fenômeno associado a este índice deve ter atuação semelhante nas duas regiões. Essas informações permitem identificar um novo tipo de comportamento fractal na variabilidade pluviométrica do Nordeste.

Por se tratar de um trabalho metodológico, concentramos esforços na associação de métodos e padrões estatísticos das redes complexas com padrões estatísticos dos dados de pluviometria, mas reconhecemos que o maior desafio é identificar os números e equações aqui descritos na natureza. Associar cada índice de redes ou a dimensão fractal a características de relevo, temperatura, maritimidade e continentalidade. Para essa nova etapa será preciso acesso a uma base de dados de temperatura, vegetação, umidade e outras variáveis fisiográficas de igual periodicidade (diária) período que a base de dados de chuva que utilizamos para este trabalho.

O certo é que a identificação destes padrões, a capacidade de extrair das redes características sobre variação de chuva sazonal da região Nordeste, são um grande estímulo a investigar as possibilidades que esta teoria, e os Sistemas Complexos em geral, têm a oferecer para o estudo dos complexos fenômenos ambientais. Fica o convite a novos aventureiros desse mar de complexidade e ordem ainda inexplorado.

REFERÊNCIAS

- ABE, S.; SUZUKI, N. Scale-free network of earthquakes. **Europhysics Letters**, [S.l.], v.65, n.4, p.581–586, 2004.
- ALBERT, R.; JEONG, H.; BARABÁSI, A.-L. Diameter of the world wide web. **Nature**, [S.l.], v.401, p.130 – 131, 1999.
- ARAGÃO, J. O. R. **Um estudo das estruturas das perturbações sinóticas do Nordeste do Brasil**. 1975. Dissertação (Mestrado em Meteorologia) – Instituto Nacional de Pesquisas Espaciais.
- BARABÁSI, A.-L.; ALBÉRT, R. Statistical Mechanics of Complex Network. **Reviews of Modern Physics**, [S.l.], v.74, p.47–85, 2002.
- BARBOSA, H. A. **Análise espaço-temporal de índice de vegetação AVHRR/NOAA e precipitação na região Nordeste do Brasil em 1982-85**. 1999. 170p. Dissertação (Mestrado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos.
- BATAGELI, V.; MRVAR, A. **Pajek - Program for large Network Analysis**. Último acesso em 11 de novembro de 2007.
- CALDEIRA, S. **Entendendo o processo de escrita de textos usando Redes Complexas**. 2005. 139p. Dissertação (Mestrado em Modelagem Computacional) – Fundação Visconde de Cairú.
- CASTRO, F. R. **Redes de migração no Brasil: caracterização e modelagem através de redes complexas**. 2007. Mestrado Interdisciplinar em Modelagem Computacional – Centro de Pesquisa e Pós-Graduação da Fundação Visconde de Cairú.
- CHARNEY, J. G. Dynamics of desert and drought in the Sahel. **Quartely Journal Royal Meteorological Society**, [S.l.], n.101, p.193 – 202, 1975.
- CHAVES, R. R. **Variabilidade da precipitação na região sul do Nordeste e sua associação com padrões atmosféricos**. 1999. 161p. Dissertação (Mestrado em Meteorologia) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos.
- EGUILUZ, V. M.; CHIALVO, D.; CECCHI, G. A.; BALIKI, M.; APKARIAN, A. V. Scale-Free Brain Functional Networks. **Physical Review Letters**, [S.l.], 2005.
- FEDER, J. **Fractals**. New York: [s.n.], 1988.

- FILHO, M. F. G. **Um estudo sobre a influência do albedo diferencial e da orografia na circulação atmosférica:** uma aplicação para o nordeste brasileiro. 1979. Dissertação (Mestrado em Meteorologia) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos.
- FONSECA, J. S. da; MARTINS, G. A. **Curso de Estatística.** 6.ed. Editora ATLAS: [s.n.], 1996. 320p.
- GLEICK, J. **Caos:** a criação de uma nova ciência. 1.ed. Rio de Janeiro: [s.n.], 1987.
- KOUSKY, V. E. Frontal influences on Northeast Brazil. **Monthly Weather Review**, [S.l.], n.107, p.1140 – 1153, 1979.
- EUCLIDES, P. (Ed.). **Aspectos teóricos da computação.** IMPA - Rio de Janeiro: [s.n.], 1979.
- MARKHAM, C. G. Apparen periodicities in rainfall at Fortaleza, Cará, Brazil. **Journal of Applied Meteorology**, [S.l.], n.13, p.176 – 179, 1972.
- MILGRAM, S. The Small World Problem. **Psychology Today**, [S.l.], p.60 – 67, 1967.
- MIRANDA, J. G. V. **Análise Fractal de Reescalonamento temporal para chuvas.** 1997. 80p. Dissertação (Mestrado em Física) – Universidade Federal da Bahia, Salvador.
- MIRANDA, J. G. V.; ANDRADE, R. F. S.; SILVA, A. B. da; SILVA FERREIRA, C. da; GONZÁLEZ, A. P.; LÓPEZ, J. L. C. Temporal and spatial persistence in rainfall records from Northeast Brazil and Galicia (Spain). **Theory Application Climatology**, [S.l.], v.77, n.1, p.113–121, 2004.
- NEWMAN, M. E. J. The Struture and function of Complex Network. **Society for Industrial and Applied Mathematics**, [S.l.], v.45, n.2, p.167 – 256, 2003.
- SANTANA, C. N. de. **Análise da pluviometria do Nordeste brasileiro segundo modelagem em redes.** 2005. 55p. Monografia (curso de Bacharelado em Ciência da Computação) – Universidade Federal da Bahia, Salvador.
- SPIEGEL, M. R. **Estatística:**resumo da teoria, 875 problemas resolvidos, 619 problemas propostos;tradução de pedro cosentino. 13.ed. São Paulo : Editora McGraw Hill do Brasil Ltda.: [s.n.], 1977. 580p.
- STRANG, D. M. G. **Análise climatológica das normais pluviométricas do Nordeste brasileiro.** 1972. 70p. Relatório técnico – Centro Técnico Aeroespacial, São José dos Campos.
- SUGAHARA, S. **Um estudo sobre a viabilidade de prever chuvas no Nordeste do Brasil, com antecedência de uma a nove meses.** 1982. Dissertação (Mestrado em Meteorologia) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos.
- SZWARCFITER, J. L. **Grafos e algoritmos computacionais.** Rio de Janeiro, Campus: [s.n.], 1984.

YAMAZAKI, Y.; RAO, V. B. Tropical cloudiness over South Atlantic Ocean. **Journal Meteorological Society Japan**, [S.l.], v.55, p.204 – 207, 1977.