

UNIVERSIDADE FEDERAL DE MATO GROSSO
INSTITUTO DE FÍSICA
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA AMBIENTAL

**SWAMP: UMA METODOLOGIA ESPAÇO-TEMPORAL
PARA ANÁLISE DE DADOS AMBIENTAIS USANDO
WAVELETS E SÉRIES TEMPORAIS.**

GRACYELI SANTOS SOUZA GUARIENTI

ORIENTADOR: PROF. DR. JOSIEL MAIMONE DE FIGUEIREDO

Cuiabá, MT
Dezembro de 2018

UNIVERSIDADE FEDERAL DE MATO GROSSO
INSTITUTO DE FÍSICA
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA AMBIENTAL

**SWAMP: UMA METODOLOGIA ESPAÇO-TEMPORAL
PARA ANÁLISE DE DADOS AMBIENTAIS USANDO
WAVELETS E SÉRIES TEMPORAIS.**

GRACYELI SANTOS SOUZA GUARIENTI

Tese apresentada ao Programa de Pós-graduação em Física Ambiental da Universidade Federal de Mato Grosso, como parte dos requisitos para obtenção do Título de Doutora em Física Ambiental.

Orientador: PROF. DR. JOSIEL MAIMONE DE FIGUEIREDO

Cuiabá, MT
Dezembro de 2018

Dados Internacionais de Catalogação na Fonte.

S237s Santos Souza Guarienti, Gracyeli.
SWAMP: UMA METODOLOGIA ESPAÇO-TEMPORAL PARA ANÁLISE DE
DADOS AMBIENTAIS USANDO WAVELETS E SÉRIES TEMPORAIS / Gracyeli
Santos Souza Guarienti. -- 2018
109 f. : il. color. ; 30 cm.

Orientador: Josiel Maimone de Figueiredo.
Tese (doutorado) - Universidade Federal de Mato Grosso, Instituto de Física,
Programa de Pós-Graduação em Física Ambiental, Cuiabá, 2018.
Inclui bibliografia.

1. Periodicidade. 2. precipitação. 3. umidade relativa. 4. temperatura média do ar.
5. SIG WEB. I. Título.

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

Permitida a reprodução parcial ou total, desde que citada a fonte.

UNIVERSIDADE FEDERAL DE MATO GROSSO
INSTITUTO DE FÍSICA
Programa de Pós-Graduação em Física Ambiental

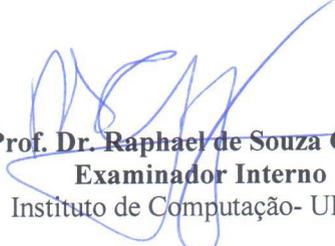
FOLHA DE APROVAÇÃO

TÍTULO: SWAMP: UMA METODOLOGIA ESPAÇO-TEMPORAL PARA ANÁLISE DE DADOS AMBIENTAIS USANDO WAVELETS E SÉRIES TEMPORAIS

AUTORA: GRACYELI SANTOS SOUZA GUARIENTI

Tese de Doutorado defendida e aprovada em 11 de dezembro de 2018, pela comissão julgadora:


Prof. Dr. Josiel Maimone de Figueiredo
Orientador
Instituto de Computação – UFMT


Prof. Dr. Raphael de Souza Gomes
Examinador Interno
Instituto de Computação- UFMT


Prof. Dr. João Basso Marques
Examinador Interno
Programa Nacional de Pós-Doutorado - CAPES


Prof. Dr. Allan Gonçalves de Oliveira
Examinador Externo
Instituto de Computação – UFMT


Prof. Dr. Jonathan Willian Zangeski Novais
Examinador Externo
Universidade de Cuiabá - UNIC

DEDICATÓRIA

Dedico este trabalho a minha filha Marieli e meu esposo César, que com todo amor e carinho, me fizeram ter forças para seguir em frente e ser forte.

AGRADECIMENTOS

- À Deus pelo dom da vida e todas as graças concedidas.
- Aos meus pais Izaildo e Doralina, minha irmã Victória, que mesmo com a distância, sempre me apoiaram incentivaram.
- Ao meu esposo César por todo amor, carinho e dedicação que permitiu que fosse possível a conclusão deste trabalho.
- Ao professor Dr. José de Souza Nogueira (Paraná) grande amigo, que sempre incentivou e ajudou nas horas mais difíceis que precisei.
- A professora Dr.^a Marta Nogueira, pelo apoio e acolhimento.
- Ao professor Dr. Josiel Maimone de Figueiredo, pela orientação dada durante os anos de mestrado e doutorado.
- Ao meu amigo Guilherme Campos, que com sua simplicidade e humildade sempre me ajudou quando precisei.
- Aos meus amigos Thiago Meirelles, Acabias e Vinicius que sempre me apoiaram nesta pesquisa.
- A banca examinadora, Prof. Dr. João, Prof. Dr. Allan, Prof. Dr. Jonathan, Prof. Dr. Raphael pelas observações e contribuições dadas a este trabalho.
- À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, pelo apoio financeiro.
- A todas as pessoas que direta ou indiretamente contribuíram para a realização deste trabalho.

*“Não importa onde você parou,
em que momento da vida você cansou, o
que importa é que sempre é possível e
necessário recomeçar.”*

Carlos Drummond de Andrade

SUMÁRIO

SUMÁRIO	8
LISTA DE FIGURAS	10
LISTA DE EQUAÇÕES.....	14
LISTA DE QUADROS	15
LISTA DE ABREVIACÕES	16
RESUMO.....	17
ABSTRACT	18
1. INTRODUÇÃO	1
1.1 OBJETIVO GERAL.....	2
1.1.1 Objetivos Específicos.....	3
1.2 JUSTIFICATIVA.....	3
1.3 ORGANIZAÇÃO DO TRABALHO	4
2. FUNDAMENTAÇÃO TEÓRICA	5
2.1 MODELOS DE PROCESSOS APLICADOS A DADOS AMBIENTAIS	5
2.2 PRECIPITAÇÃO	10
2.3 PROJETO TRMM	11
2.4 UMIDADE.....	15
2.5 TEMPERATURA MÉDIA DO AR	16
2.6 INMET (INSTITUTO NACIONAL DE METEOROLOGIA)	17
2.7 SÉRIES TEMPORAIS	19
2.7.1 Método de Análise de Séries Temporais	20
2.8 BUSCA POR SIMILARIDADE EM SÉRIES TEMPORAIS	24
2.8.1 Distância Euclidiana	25
2.9 GEOPROCESSAMENTO E SISTEMAS DE INFORMAÇÃO GEOGRÁFICA	27
2.9.1 Traduzindo informações geográficas para o computador	28
2.10 ARQUITETURA DE SIG	29
2.10.1 Representação dos Dados Geográficos.....	31
2.11 SIG WEB	32
2.12 CONCLUSÃO DO CAPÍTULO	34
3. MATERIAIS E MÉTODOS.....	35
3.1 Dados Utilizados.....	35

3.1.1	TRMM.....	35
3.1.2	INMET.....	36
3.1.3	Pontos utilizados.....	38
3.2	Busca por Similaridade.....	39
3.3	Arquitetura.....	40
4.	RESULTADOS E DISCUSSÃO.....	44
4.1	METODOLOGIA SWAMP.....	44
4.1.1	Modelagem.....	46
4.1.2	Tratamento dos dados ambientais.....	50
4.2	ANÁLISE DE PRECIPITAÇÃO COM O SWAMP 1.0.....	55
4.2.1	Análise Espacial Semestral - Séries Temporais Originais.....	55
4.2.2	Ponto A - Análise Semestral - Séries Temporais Originais.....	56
4.2.3	Análise Espacial Anual - Wavelets.....	59
4.2.4	Ponto A - Análise Anual - Wavelets.....	60
4.2.5	Análise Espacial - 16 anos - Wavelets.....	62
4.2.6	Ponto A - Análise 16 anos - Wavelets.....	63
4.3	ANÁLISE DE UMIDADE RELATIVA COM O SWAMP 1.0.....	65
4.3.1	Análise Espacial Semestral - Séries Temporais.....	65
4.3.2	Ponto B - Análise Semestral - Séries Temporais.....	66
4.3.3	Análise Espacial Anual - Wavelets.....	67
4.3.4	Ponto B - Análise Anual - Wavelets.....	68
4.4	ANÁLISE DE TEMPERATURA MÉDIA DO AR COM O SWAMP 1.0.....	70
4.4.1	Análise Espacial Anual - Séries Temporais.....	70
4.4.2	Ponto C - Análise Anual - Séries Temporais.....	71
4.4.3	Análise Espacial Semestral - Wavelets.....	73
4.4.4	Ponto C - Análise Semestral - Wavelets.....	74
4.5	CONCLUSÃO DO CAPÍTULO.....	76
5.	CONSIDERAÇÕES FINAIS.....	78
5.1	CONTRIBUIÇÕES.....	79
5.2	TRABALHOS FUTUROS.....	79
6.	REFERÊNCIAS BIBLIOGRÁFICAS.....	81

LISTA DE FIGURAS

Figura 1: Modelo de Processos para Mineração de Dados, CRISP-DM. Fonte: OLIVEIRA, 2015.	6
Figura 2: Entendimento do Negócio influenciado demais etapas do processo de Mineração de Dados. Fonte: OLIVEIRA, 2015.	8
Figura 3: Etapas do monitoramento ambiental. Fonte: Adaptado de Alcântara, Moresi e Prado (2010).	9
Figura 4: Sensores do satélite TRMM. Fonte: NASA (2018).	14
Figura 5: Pluviômetro, ao centro a antena de comunicação via satélite e à direita o <i>datalogger</i> com sensor de temperatura presos ao mastro. Fonte: INMET, 2018.	18
Figura 6: (a) Totais mensais de passageiros em linhas aéreas internacionais nos EUA, (b) Número anual de lincas capturados em armadilhas no Canadá, (c) Medições anuais de vazões do Rio Nilo em Ashwan, (d) Consumo de gás no Reino Unido no primeiro trimestre de cada ano. Fonte: EHLERS, 2009.	20
Figura 7 - Esquema de etapas de geração de imagens através de SIG. Fonte: Adaptado pela autora.	28
Figura 8: Arquitetura de SIG. Fonte: INPE (2018), modelo adaptado.	29
Figura 9: Arquitetura SIG - Estrutura Dual. Fonte: CÂMARA e RIBEIRO, (2004).	30
Figura 10: Arquitetura SIG - Estrutura Integrada Relacional. Fonte: CÂMARA e RIBEIRO, (2004).	31
Figura 11: Requisição do SIG <i>web</i> . Adaptado (GORNI <i>et al.</i> , 2007).	34
Figura 12: Localização área de estudo dos dados de precipitação do TRMM.	36
Figura 13: Localização área de estudo dos dados de umidade relativa e temperatura média do ar do INMET.	37
Figura 14: Esquema para construção do SWAmP 1.0.	40
Figura 15: Etapas do resultado do desenvolvimento da metodologia SWAmP.	45

Figura 16: Diagrama de Caso de Uso.....	46
Figura 17: Tela para escolha dos pontos para visualização dos resultados das análises.	48
Figura 18: Diagrama de processo de funcionamento do SWAmP 1.0.....	49
Figura 19: Diagrama de tratamento dos dados do SWAmP.....	50
Figura 20: Tabelas com os agrupamentos: semestral, anual e 16 anos de séries temporais e Wavelets. Tabela Ponto.	53
Figura 21: Modelo relacional das tabelas de similaridades do SWAmP 1.0.....	54
Figura 22: Mapa de distância média em quilômetros, do primeiro semestre de 1998 das Séries Temporais.	56
Figura 23: Séries Temporais do Ponto 2568, primeiro semestre 1998; Ponto 1637 primeiro semestre 1999 (Primeira semelhante); Ponto 1452, primeiro semestre 2006(Segunda Semelhante); Ponto 2146, primeiro semestre 2001(Terceira semelhante).	57
Figura 24: Quadro com as informações dos pontos semelhantes ao ponto selecionado pelo usuário.	58
Figura 25: Mapa do SWAMP 1.0 com os pontos semelhantes. Ponto 2568(VERDE); Ponto 1637(AZUL ESCURO); Ponto 1452(AZUL MÉDIO); Ponto 2146(AZUL CLARO).	58
Figura 26: Mapa de distância média em quilômetros de cada ponto, do ano 1998 das Wavelets.	59
Figura 27: Wavelets semelhantes do Ponto 2568, ano de 1998; Ponto 2134, ano de 2001, (primeiro semelhante); Ponto 2197, ano de 2001, (segundo semelhante); Ponto 1984, ano de 2007, (terceiro semelhante).....	60
Figura 28: Quadro com as informações dos pontos semelhantes ao ponto selecionado pelo usuário.	61
Figura 29: Mapa apresentado pelo SWAmP 1.0. Ponto selecionado pelo usuário (VERDE): ano 1998; Primeiro ponto semelhante (AZUL ESCURO): ano 2001; Segundo ponto semelhante (AZUL MÉDIO): ano 2001; Terceiro ponto semelhante (AZUL CLARO): ano 2007.	61

Figura 30: Mapa de distância média em quilômetros, de 16 anos (janeiro de 1998 a dezembro 2013) nas Wavelets.....	62
Figura 31: Wavelets semelhantes do Ponto 2568; Ponto 2273; Ponto 2470; Ponto 2480.	63
Figura 32: Mapa do SWAmP 1.0 com os pontos semelhantes. Ponto selecionado pelo usuario (VERDE); Primeiro ponto semelhante (AZUL ESCURO); Segundo ponto semelhante (AZUL MÉDIO); Terceiro ponto semelhante (AZUL CLARO).	64
Figura 33: Quadro com as informações dos pontos semelhantes ao ponto selecionado pelo usuário.	64
Figura 34: Mapa de distância média em quilômetros, do segundo semestre de 2004 das Séries Temporais.	65
Figura 35: Séries temporais semelhantes. Ponto 1 selecionado pelo usuário, segundo semestre de 2004; Ponto 14 (primeiro ponto semelhante), segundo semestre de 2004; Ponto 27 (segundo ponto semelhante), segundo semestre de 2004; Ponto 22 (Terceiro ponto semelhante), segundo semestre de 2004.....	66
Figura 36: Quadro com as informações dos pontos semelhantes ao ponto selecionado pelo usuário.	67
Figura 37: Mapa com os pontos das Séries temporais semelhantes. Ponto selecionado pelo usuário (VERDE): segundo semestre de 2004; Primeiro ponto semelhante (AZUL ESCURO): segundo semestre de 2004; Segundo ponto semelhante (AZUL MÉDIO): segundo semestre de 2004; Terceiro ponto semelhante (AZUL CLARO): segundo semestre de 2004.	67
Figura 38: Mapa de distância média em quilômetros, ano de 2004 nas Wavelets de umidade relativa.	68
Figura 39: Wavelets do Ponto 1, ano de 2004; Ponto 14, ano de 2004; Ponto 8, ano de 2004; Ponto 27, ano de 2004;.....	69
Figura 40:Quadro com as informações dos pontos semelhantes das Wavelets de umidade relativa.	70

Figura 41: Imagem do SWAmP 1.0 com os pontos semelhantes. Ponto selecionado pelo usuário (VERDE): ano 2004; Primeiro ponto semelhante (AZUL ESCURO): ano 2004; Segundo ponto semelhante (AZUL MÉDIO): ano 2004; Terceiro ponto semelhante (AZUL CLARO): ano 2004.	70
Figura 42: Mapa de distância média em quilômetros de cada ponto, do ano 2011 das Séries Temporais.	71
Figura 43: Séries Temporais do Ponto 22, ano de 2011; Ponto 5, ano de 2011; Ponto 14, ano de 2011; Ponto 8, ano de 2011;.....	72
Figura 44: Quadro com as informações dos pontos semelhantes de temperatura média do ar ano 2011	73
Figura 45: Ponto selecionado pelo usuário (VERDE): ano 2011; Primeiro ponto semelhante (AZUL ESCURO): ano 2011; Segundo ponto semelhante (AZUL MÉDIO): ano 2011; Terceiro ponto semelhante (AZUL CLARO): ano 2011.....	73
Figura 46: Mapa de distância média em quilômetros, primeiro semestre de 2011 nas Wavelets de temperatura média do ar.....	74
Figura 47: Wavelets do Ponto 22, primeiro semestre de 2011; Ponto 5, primeiro semestre de 2011; Ponto 4, primeiro semestre de 2011; Ponto 14, primeiro semestre de 2011.	75
Figura 48: Quadro com as informações dos pontos do primeiro semestre de 2011 de temperatura média do ar de Wavelets.....	75
Figura 49: Mapa com pontos semelhantes ao ponto selecionado pelo usuário. Ponto selecionado pelo usuário (VERDE): primeiro semestre de 2011; Primeiro ponto semelhante (AZUL ESCURO): primeiro semestre de 2011; Segundo ponto semelhante (AZUL MÉDIO): primeiro semestre de 2011; Terceiro ponto semelhante (AZUL CLARO): primeiro semestre de 2011.....	76

LISTA DE EQUAÇÕES

Equação 1 - Decomposição de Wavelets	22
Equação 2 - Wavelets Contínuas.....	22
Equação 3 - Wavelets Contínuas - Inversa.....	22
Equação 4 -Wavelets Morlet	22
Equação 5 - Wavelets Morlet - equação original segundo Jean Morlet.....	23
Equação 6 - Variação da Wavelet Morlet:	23
Equação 7 - Equação de Comprimento	24
Equação 8 - Equação da distância entre dois pontos cartesianos	26

LISTA DE QUADROS

Quadro 1: Descrição Geral dos Instrumentos a Bordo do TRMM.....	13
Quadro 2: Funcionalidades do SWAmP 1.0 detalhadas.....	47
Quadro 3: Pontos e datas utilizados para apresentar os resultados do SWAmP 1.0.	38

LISTA DE ABREVIACES

BPMN	<i>Business Process Model and Notation</i>
CM	<i>Centmetros</i>
DIST	<i>Distancia</i>
GV	<i>Ground Validation</i>
GPI	<i>Goes Precipitation Index</i>
HDF	<i>Hierarchical Data Format</i>
JAXA	<i>Japan Aerospace Exploration Agency.</i>
LIS	<i>Lighting Imaging Sensor</i>
NASA	<i>National Aeronautics and Space Administration.</i>
NetCDF	<i>Network Common Data Form</i>
PR	<i>Precipitation Radar</i>
SGBD	<i>Sistema Gerenciador de Banco de Dados.</i>
SWAMP	<i>SIG Web Ambiental</i>
SIG	<i>Sistema de Informao Geogrfica.</i>
TIROS	<i>Television and InfraRed Observation Satellite</i>
TMI	<i>TRMM Microwave Imager</i>
TRMM	<i>Tropical Rainfall Measuring Mission.</i>
TWC	<i>Transforma de Wavelet Contnua</i>
VIRS	<i>Visible and Infrared Radiometer System</i>

RESUMO

GUARIENTI, G, S, S. SWAMP: UMA METODOLOGIA ESPAÇO-TEMPORAL PARA ANÁLISE DE DADOS AMBIENTAIS USANDO WAVELETS E SÉRIES TEMPORAIS. Cuiabá, 2018. 109 p. Tese (Doutorado) - Pós-Graduação em Física Ambiental, Universidade Federal de Mato Grosso.

Este trabalho apresenta uma nova metodologia espaço-temporal para a manipulação, modelagem e visualização de dados ambientais. Denominada SWAmP (Sistema de informação geográfica via Web para dados Ambientais de Precipitação), essa metodologia é validada com a implementação do *software* SWAmP 1.0, que incorpora as etapas da metodologia em uma interface SIG *web* e utiliza dados do projeto TRMM (*Tropical Rainfall Measuring Mission*), e do INMET (Instituto Nacional de Meteorologia). Os resultados evidenciaram que as etapas de processamento abordados na metodologia SWAmP trazem novas perspectivas para análise de dados ambientais. As funcionalidades providas pelo SWAmP 1.0 são demonstradas com dados de precipitação o(16 anos), umidade relativa média e temperatura da região compreendida entre (-8,-22S e -49,-61W) nos anos de 1998 a 2013.

Palavras-chave: Periodicidade, precipitação, umidade relativa, temperatura média do ar, e SIG WEB.

ABSTRACT

GUARIENTI, G, S, S. **SWAMP: A SPACE-TEMPORAL METHODOLOGY FOR ENVIRONMENTAL DATA ANALYSIS USING WAVELETS AND TEMPORATIVE SERIES.** Cuiabá, 2018, 109 p. Thesis (Doctorate in Environmental Physics) - Physics Institute, Federal University of Mato Grosso.

The work presents a new temporal vision for the manipulation, modeling and visualization of environmental data. Named SWAmP (Web Geographic Information System for Environmental Precipitation Data), this same methodology is validated with an implementation of the SWAmP 1.0 software, which includes the steps of the methodology in a web GIS interface and the use of data from the TRMM (Project of Precipitation Measurement), and INMET (National Institute of Meteorology). They show that, as the imitation steps of mathematical methods, are present in the SWAmP methodology bring new perspectives for the analysis of environmental data. The tests are of type SWAmP 1.0 are demonstrated with data of precipitation (16 years), mean and average environment between (-8, -22S and -49, -61W) in the years 1998 to 2013.

Keywords: Frequency, precipitation, relative humidity, temperature, and GIS WEB.

1. Introdução

O estudo de variáveis ambientais mostra-se de grande importância em muitas áreas científicas, pois compreender seu comportamento e como são executadas as análises em escalas espaço-temporal é importante para caracterizar determinado ambiente e como ele pode ser afetado.

Análises ambientais com dados espaço-temporais, de modo geral, podem auxiliar em áreas como: agricultura no plantio de determinadas culturas, turismo na escolha dos melhores períodos para viagens, defesa civil para previsão e estudo de desastres. As análises são executadas com estudo de uma base dados histórica para gerar uma série temporal dos dados da variável ambiental e analisar seu comportamento.

Para realização das análises, os dados ambientais podem ser coletados de diversas fontes como estações terrestres dispostas de sensores e equipamentos, ou por sensoriamento remoto. Conforme o tipo de equipamento utilizado pode-se obter dados com resoluções temporais a cada segundo, minuto, hora ou 24 horas por dia.

É possível observar que existe um volume de dados que pode até ultrapassar gigabytes de dados, e que pode existir dificuldade na análise dos dados que origina-se pela: quantidade, organização e processo e análise de dados ambientais. Neste contexto, devido à interatividade dos processos, muitas vezes uma tarefa é realizada diversas vezes, e para um novo conjunto de dados não existe uma sequência exata para execução e extração de resultados. Dessa forma é preciso que utilize técnicas computacionais, com um armazenamento eficiente, para facilitar buscas e minimizar possíveis falhas técnicas. Isso nitidamente facilita a análise que pode ser feita quanto a grandes variações que podem ser analisadas quando há várias resoluções temporais ou séries históricas.

Com um sistema computacional, a busca pelo conhecimento acerca das variáveis ambientais como precipitação, temperatura média do ar e umidade relativa e como ocorre sua distribuição temporal e espacial, é uma atividade de suma importância, pois possibilita a modelagem e previsão do tempo, por exemplo.

Por possuir uma distribuição espaço-temporal, as séries temporais de precipitação pluviométrica, temperatura média do ar e umidade relativa, podem ser analisadas com técnicas de comparação e extração de características similares. Assim, o

uso de séries temporais pode ser utilizado para verificar a periodicidade dos dados ambientais e semelhanças.

O estudo de variáveis ambientais exige a utilização de técnicas computacionais para análises, observações e caracterizações, isso possibilita a compreensão de possíveis fenômenos que estão atuando ou que atuaram em algum momento de uma série histórica. O uso de métodos de análises de dados ambientais como Wavelets pode ser utilizado para extrair informações como tempo, frequência e intensidade de energia, trazendo informações além das obtidas analisando somente a série temporal, com uma observação criteriosa do eixo X e Y.

Para possibilitar análises de séries temporais e Wavelets de dados ambientais, utiliza-se um conjunto de ferramentas que permite o tratamento de informações geográficas, chamado Sistemas de Informação Geográfica (SIG). Esses sistemas permitem o tratamento dos dados espaciais e sua manipulação, procurando facilitar buscas, a inferência de informações e visualização.

Um SIG pode ser desenvolvido em ambiente *web* (SIG *web*). Implementado dessa forma, ele está preparado para disponibilizar informações geográficas a seus usuários, ampliando a disponibilidade dos dados. A utilização de SIG *web* para implementar análises ambientais facilita a visualização dos resultados e possibilita diversos tipos de análises para o usuário.

Neste contexto, este trabalho apresenta a metodologia para análises de dados ambientais chamado SWAmP (SIG *Web* Ambiental). Que possibilita a análise de diversos tipos de resultados traz várias possibilidades de visualizações em um sistema nominado SWAmP 1.0, aplicado a dados de precipitação, temperatura média do ar e umidade relativa. Os resultados podem ser utilizados em pesquisas em que exista a necessidade de comparação, visualização espaço-temporal de dados e detecção de padrões semelhantes, auxiliando os pesquisadores que utilizam essas informações em suas pesquisas e tomadas de decisões.

1.1 Objetivo Geral

Desenvolver uma metodologia que aplique técnicas para análises espaço-temporais em dados ambientais.

1.1.1 Objetivos Específicos

Para atingir o objetivo geral deste trabalho, foram definidos os seguintes objetivos específicos:

- Processar e armazenar dados de precipitação do projeto TRMM;
- Processar e armazenar dados de temperatura média do ar e umidade relativa INMET;
- Modelagem computacional do banco de dados com base geográfica;
- Processar séries temporais de dados ambientais com resolução temporal: anual, semestral e 16 anos;
- Processar Wavelets de séries temporais com resolução temporal: de precipitação anual, semestral e 16 anos;
- Realizar detecção de padrões nas séries temporais e Wavelets de dados ambientais com análise de similaridade;
- Modelagem computacional do SWAmP 1.0 (SIG *Web* Ambiental);
- Aplicar técnicas de programação *web*, geoprocessamento e sistemas de informação geográfica para construção da interface do sistema.

1.2 Justificativa

Em dados ambientais existe uma a necessidade de realizar o tratamento e processamento de forma eficaz, para obter uma série de dados com integridade. Para realizar esses procedimentos em uma grande quantidade de dados, uma metodologia pode ser utilizada, para facilitar o processo que pode consistir de várias etapas. O desenvolvimento da metodologia apresentada neste trabalho traz uma sequência de etapas que abstraem para o pesquisador os métodos necessários para chegar ao resultado requerido.

Por outro lado, realizar todas as etapas de uma metodologia não é uma tarefa simples, pois pode exigir conhecimentos avançados sobre o comportamento dos dados e técnicas computacionais, que normalmente é um procedimento complexo e demanda

tempo. Outra questão é a análise dos resultados realizada de forma manual em cada etapa da metodologia.

Para contornar essa situação, e possibilitar também que as análises sejam efetuadas, é possível que seja criado um sistema *web* que permita vários tipos de análises de dados ambientais, simplificando e automatizando todo o processo. Com o sistema os usuários não precisariam se preocupar com a programação dos algoritmos para análise de séries temporais e Wavelets o que resultaria em automatização dos processos de análises de dados ambientais. O uso de um sistema como o desenvolvido neste trabalho resultará em maior consistência nos dados em pesquisas, interpretações e modelagens mais precisas para os dados disponibilizados.

1.3 Organização do Trabalho

Estruturalmente este trabalho está dividido da seguinte forma: o Capítulo 2 está os conceitos de monitoramento de precipitação, geoprocessamento, SIG, projeto TRMM e SIG *web*. No Capítulo 3 está a metodologia desenvolvida, modelagem, e as tecnologias utilizadas para o desenvolvimento. O Capítulo 4 estão os resultados aplicados a alguns pontos do SWAmP 1.0. Por fim, o Capítulo 5 mostra as considerações finais e trabalhos futuros.

2. Fundamentação Teórica

Para um melhor entendimento dos conceitos, ferramentas e linguagens tratadas no decorrer deste trabalho, é necessário que haja uma abordagem conceitual. Portanto este capítulo traz o conceito modelos de processos aplicados a dados ambientais e monitoramentos de precipitação, dos dados de precipitação do projeto TRMM que são utilizados como fonte para o sistema desenvolvido.

Para facilitar o entendimento dos conceitos envolvidos as próximas seções trazem os aspectos de modelos de processos, técnicas de geoprocessamento e SIGs para tratar os dados assim como funcionamento de sistemas *web* para geoprocessamento.

2.1 Modelos De Processos Aplicados a Dados Ambientais

Pode-se definir um processo de várias formas. Davenport (1994) define processo como: "uma ordenação específica das atividades de trabalho no tempo e no espaço, com um começo, um fim, entradas e saídas identificadas...". Um modelo de processos aplicado a dados ambientais pode ser definido como conjuntos de atividades a ser efetivado para alcançar um objetivo como: analisar dados de precipitação, umidade relativa e temperatura média do ar (PRESSMAN, 2005).

Segundo Pressman, (2005), um modelo de processo traz os elementos que são produzidos em cada atividade do processo de análise e os elementos que são necessários para realizar uma atividade para que se alcance o objeto final.

Existem diversos tipos de modelos de processos, a maioria deles se baseia em etapas a serem executadas de forma continua o que inclui loops e interações entre etapas, e um modelo de processos aplicado a dados ambientais também segue os mesmos critérios. (KURGAN; MUSILEK, 2006).

Em Oliveira (2015), Marbán *et al.*(2009) e Mariscal *et al.* (2010) podem ser encontrados outros modelos de processos, alguns modelos são específicos para determinada ferramenta, vendedor ou aplicativo. Tais modelos podem ser utilizados, modificados e utilizados para diversos tipos de dados, como dados ambientais, por exemplo.

O modelo de processo CRISP-DM (*Cross Industry Standard Process for Data Mining*), pode ser utilizado em qualquer aplicação, projeto acadêmico, industrial ou análise de dados ambientais e é ainda independente de fornecedor. Shearer (2000) e Wirth (2000) apresentaram o CRISP-DM oficialmente em seus trabalhos à comunidade científica, composto por 6 etapas, apresentadas na Figura 1 de forma genérica.

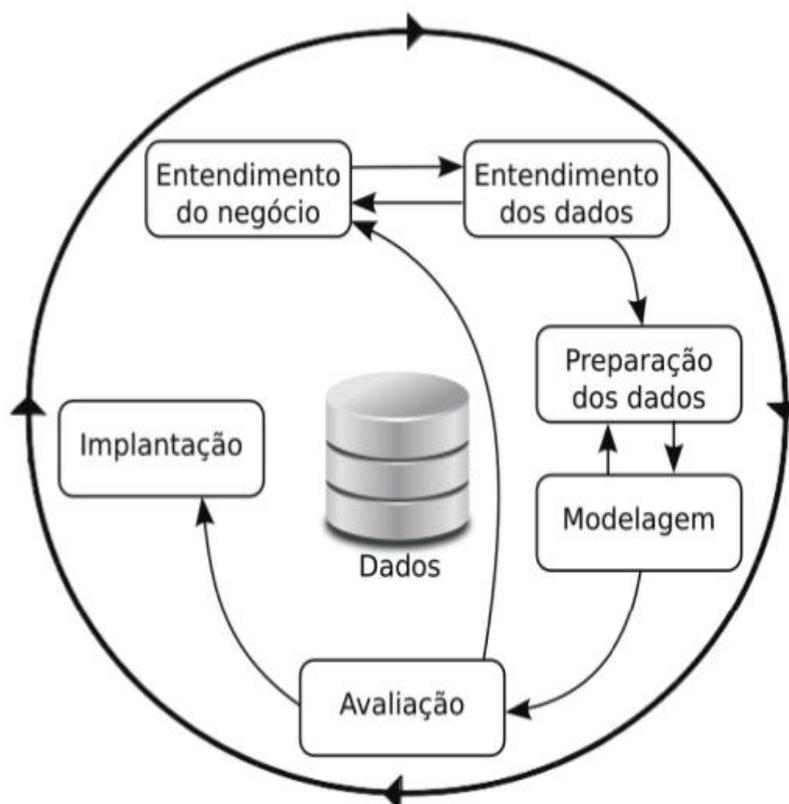


Figura 1: Modelo de Processos para Mineração de Dados, CRISP-DM. Fonte: OLIVEIRA, 2015.

Segue a descrição de cada etapa do CRISP-DM, aplicado a dados ambientais:

- **Entendimento do negócio:** identificam-se os objetivos e necessidades para converter em uma tarefa. Aplicado a dados ambientais, essa etapa corresponde ao entendimento de qual variável climatológica será estudada.
- **Entendimento dos dados:** tem como objetivo descobrir informações nos dados para detectar subconjuntos e formular hipóteses. Em estudos ambientais, nesta etapa os pesquisadores já sabem qual variável climatológica será estudada, por exemplo: a precipitação, e neste caso é feito o entendimento de como esta variável ambiental poderia utilizada.

- Preparação dos dados: é a etapa em que um conjunto de atividades são elaboradas para obter os dados finais e a partir daí será criado e validado o modelo. Em estudos ambientais os dados como os de precipitação poderiam ser agrupados seguindo periodicidades como meses, semestres ou anos, conforme a necessidade do estudo.
- Modelagem: nesta etapa várias técnicas de modelagem podem ser aplicadas, dependente do tipo de dado tratado.
- Avaliação: nesta etapa já existe um modelo construído ou vários, e para seguir para a implantação é preciso que apresente alta qualidade para a perspectiva de análise dos dados ambientais. Caso haja necessidade é possível retornar a primeira etapa de entendimento do negócio
- Implantação: nesta etapa os modelos já validados são implantados e difundidos, por vezes é necessário retornar as etapas anteriores para aprimorar o trabalho quando preciso.

Existe uma conexão apresentada no CRISP-DM entre o Entendimento do negócio e o Entendimento dos dados. Esse conhecimento serve de suporte para a execução também das demais etapas do processo (OLIVEIRA, 2015).

Segundo Oliveira (2015):

o conhecimento da fase de Entendimento do negócio influencia todo o processo, o que gera a necessidade de vinculá-lo com a parametrização de cada etapa de processamento. Entretanto, no modelo tradicional CRISP-DM não fica explícito essa influência, bem como as ferramentas de DM também não refletem a necessidade de interação do Especialista em todo o processo.

Oliveira (2015) apresenta uma modificação do fluxo original e torna mais natural a interação entre as etapas que são incluídas pelas setas azuis como mostra a Figura 2:

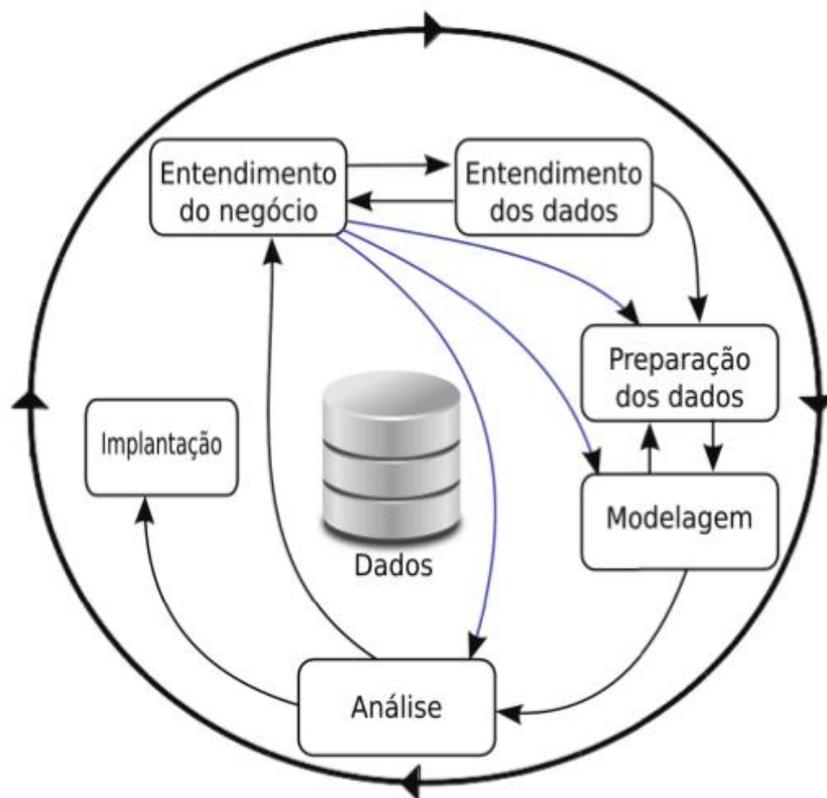


Figura 2: Entendimento do Negócio influenciado demais etapas do processo de Mineração de Dados. Fonte: OLIVEIRA, 2015.

O modelo modificado por Oliveira (2015) é uma excelente abordagem para o modelo CRISP-DM, pois aplicado aos dados ambientais esta modelagem viabiliza o acesso aos processos, pois é possível ter uma interatividade maior entre as etapas.

O modelo de processos CRISP-DM assemelha-se com o monitoramento ambiental, pois possui um ciclo de atividades necessárias para extração de conhecimento e que podem ser aplicado em dados ambientais. A realização do monitoramento ambiental é feita por meio de análises de variáveis climatológicas e seu desenvolvimento durante certo período de tempo. Dessa forma esses dados podem ser medidos em diversos locais em períodos curtos ou longos, isso permite que técnicas de análises espaço-temporais sejam aplicadas de forma que seja realizado o entendimento do comportamento do ambiente e seu monitoramento.

Alcântara, Moresi e Prado (2010), definiram as etapas do monitoramento ambiental que podem ser utilizadas em diversos estudos na área ambiental a fim de

minimizar os impactos e degradação do meio ambiente, possibilitar a conservação, preservação, e recuperação ambiental da região estudada.

Por meio do monitoramento ambiental é realizado o processo de coleta de dados, estudo e acompanhamento dos recursos do meio ambiente, etapas muito similares ao modelo de processo CRISP-DM. Com o objetivo de encontrar e aferir as condições dos recursos em um certo período de tempo e suas tendências (TUCCI e MENDES, 2006).

A Figura 3 apresenta o ciclo das etapas para o monitoramento ambiental.



Figura 3: Etapas do monitoramento ambiental. Fonte: Adaptado de Alcântara, Moresi e Prado (2010).

O conjunto de atividades apresentado por Alcântara, Moresi e Prado (2010) (Figura 3) é composto por 5 etapas, descritas a seguir:

- **Planejamento e Direção:** Nesta etapa é definido qual será a variável ambiental estudada como será realizada a pesquisa. Este procedimento é importante, pois viabiliza o conhecimento do problema e lista as necessidades de monitorar determinado ambiente.
- **Armazenamento e Processamento das Informações:** Nesta etapa é realizado o armazenamento e processamento dos dados.
- **Coleta e Relatórios Adequados:** Nesta etapa os dados armazenados são coletados para gerar relatórios e informações.
- **Análise e Produção:** Nesta etapa as informações coletadas da etapa anterior são analisadas, as hipóteses são testadas e realiza-se a produção de relatórios finais.

- **Disseminação:** Nesta etapa os modelos já validados podem ser utilizados em tomada de decisão ou na elaboração de ações a serem tomadas com base nos procedimentos já executados.

As etapas do monitoramento ambiental servem para auxiliar em pesquisas em que o foco são dados ambientais. O processo é semelhante ao modelo CRISP-DM, porém no caso do monitoramento ambiental as etapas são totalmente voltadas a problemática ambiental.

Nesse sentido, existem diversas variáveis ambientais como temperatura média do ar, umidade e pressão atmosférica, que são empregadas em pesquisas utilizam as etapas do monitoramento ambiental, entre elas a precipitação que é vastamente estudada por diversos pesquisadores, pois afeta diretamente múltiplos setores da sociedade e do meio ambiente (SANTOS *et al.*, 2002).

2.2 Precipitação

A precipitação pode ocorrer pelo encontro de massas de ar sob condições diferentes com interferência orográfica¹ ou não, ou como resultado das variações como vapor d'água, pressão atmosférica e temperatura média do ar.

As chuvas podem ser classificadas em três tipos: chuvas convectivas, as quais possuem características de pequena duração e grande intensidade, com abrangência em pequenas áreas; chuvas frontais, são as que possuem menos intensidade e maior duração, porém uma abrangência em grandes áreas; e chuvas orográficas que apresentam baixa intensidade duração longa e atuação sobre áreas médias (GARCEZ e ALVAREZ, 1998).

O estudo da precipitação e sua distribuição no espaço e tempo são de fundamental importância para caracterização e medidas de sustentabilidade. O

¹ Ocorre quando uma massa de ar carregada de umidade sobe ao encontrar uma elevação do relevo, como uma montanha. O ar mais quente e mais leve, geralmente, mais úmido é empurrado para cima.

gerenciamento sustentável dos recursos naturais está condicionado ao estudo da precipitação, por exemplo (DIAS, 2011).

Dependendo do tipo de estudo, é possível definir diferentes periodicidades de tempo para mensurar a precipitação, ou seja, é possível medir a precipitação a cada 3 horas, diariamente, mensalmente, ou anualmente, por exemplo. Dessa forma, é possível realizar projetos com base em análises do comportamento da precipitação detectando tendências ou alterações no clima local ou regional (SANT' ANNA NETO, 2000).

Os diferentes estudos feitos pelos cientistas que trabalham com monitoramento de precipitação avaliam o comportamento de séries espaço-temporais e os impactos da sociedade no meio ambiente com o objetivo de registrar os dados, prevenir possíveis desastres ambientais (TUCCI e MENDES, 2006). Com a utilização de modelagem para apoiar a gestão de recursos de forma dinâmica com estabelecimento de prevenção ou minimização de desastres naturais.

Porém principalmente para as regiões com baixa disponibilidade de dados pluviométricos, realizar a modelagem e fazer análises com a intenção de minimizar possíveis desastres é complicado. Neste caso os produtos de sensoriamento remoto podem ser uma alternativa para essa questão, pois pode ser uma fonte de dados que mostram concordância entre os índices calculados com dados postos pluviométricos (Yan, et al., 2014; Naumann, et al., 2012; Rhee et al., 2010).

2.3 Projeto TRMM

Com o objetivo de medir a precipitação nos trópicos e servir como fonte de dados para os pesquisadores, a *National Aeronautics And Space Administration* (NASA²) em parceria com a *Japan Aerospace Exploration Agency* (JAXA³) criou o projeto *Tropical Rainfall Measuring Mission* (TRMM) e em 27 de novembro de 1997 no Centro Espacial Tanegashima lançou o satélite TRMM.

² Disponível em < <http://www.nasa.gov/> > Acesso em 24 de janeiro de 2018.

³ Disponível em < <http://www.jaxa.jp/> > Acesso em 24 de janeiro de 2018.

Diversas pesquisas foram realizadas desde seu lançamento, para comparar e verificar a acurácia dos dados estimados pelo satélite TRMM (BARRETT *et al.*, 1994; FRANCHITO *et al.*, 2009; EBERT *et al.*, 1996; FARIAS *et al.*, 2013; HUANG *et al.*, 2013; SMITH *et al.*, 1998; ADLER *et al.*, 2001; NICHOLSON *et al.*, 2003; ROZANTE *et al.*, 2010; FISHER, 2004). Estes trabalhos mostram que o satélite com seus sensores produzem bons resultados (LAYBERRY *et al.*, 2006).

As estimativas produzidas pelo satélite TRMM são consideradas mais confiáveis do que de outros satélites (BARRERA, (2005), NICHOLSON, (2005)). Isso porque existe também o Ground Validation (GV), que coleta dados de radares, pluviômetros e disdrômetros⁴ localizados no solo. Os dados são controlados por qualidade e, em seguida, é gerado os produtos de validação que são comparados com os produtos do satélite TRMM. As quatro principais localizações da GV são Darwin, na Austrália; Houston, Texas; Kwajalein, República das Ilhas Marshall; e, Melbourne, Flórida (NASA, 2018).

Os produtos são produzidos usando técnicas desenvolvidas para controlar cuidadosamente a qualidade dos conjuntos de dados do radar terrestre. Os produtos estimam as taxas de precipitação de superfície, ajustadas por dados de pluviometria de qualidade controlada (NASA, 2018).

Os dados do projeto TRMM são utilizados em estudos climatológicos e servem como base de dados de chuva em áreas de difícil acesso o que possibilita análise espacial e temporal ((COLLISCHONN, 2006); (COLLISCHON *et al.*, 2007); (NÓBREGA *et al.*, 2008), (NÓBREGA, 2008)).

O satélite foi desenvolvido como o objetivo de medir a precipitação e a sua variação na região tropical entre 35°N e 35°S, com órbita oblíqua não heliossíncrona de baixa inclinação e altitude (NÓBREGA, 2008; KUMMEROW *et al.*, 2000). Todavia, devido as atualizações do satélite e/ou dos algoritmos, o satélite passou a medir entre as

⁴Um instrumento que mede a distribuição de tamanhos de gotas de precipitação na atmosfera, ou seja, o número de pequenas gotas, o número de gotículas grandes e onde elas ocorrem.

latitudes de 50° N e 50° S (SCHEEL et al., 2011; ALVES, 2017). Apesar de ser projeto para durar três anos o satélite forneceu dados por 17 anos (JAXA, 2018).

Os sensores disponíveis no satélite estimam da precipitação em locais de difícil acesso, como grandes áreas do oceano o que torna sua utilização bastante usual. Os dados fornecidos pelo satélite TRMM são aplicados em várias pesquisas pelo mundo, destacando-se em contextos que tratam de locais onde não há registro de dados (COLLISCHONN *et al.*, 2007; COLLISCHONN *et al.*, 2008).

O satélite TRMM possui a bordo sensores que possibilitam a estimativa da precipitação são eles: Imageador de Micro-ondas (TMI); Radar de Precipitação (PR) e Sistema de Radiômetro Visível e Infravermelho (VIRS). Possui ainda os sensores para mapeamento de relâmpagos e raios (LIS) e sensor de energia radiante da superfície terrestre e das nuvens (CERES); (TRMM, 2018). O QUADRO 1 disponibiliza a descrição geral dos instrumentos do satélite TRMM e na Figura 4 são ilustrados os instrumentos a bordo do satélite TRMM.

Quadro 1: Descrição Geral dos Instrumentos a Bordo do TRMM.

TMI	É um Imageador de micro-ondas passivo multicanal operando em cinco frequências: 10,65, 19,35, 37,0 e 85,5 ghz em polarização dupla e 22,235 ghz em polarização única. O TMI fornece informações sobre a quantidade de precipitação, a água líquida da nuvem, o gelo da nuvem, a intensidade da chuva e os tipos de precipitação (por exemplo, estratiforme ou convectiva).
PR	É um radar de varredura eletrônica, operando a 13,8 ghz, que mede a distribuição de chuvas em 3-D sobre a terra e o oceano, e define a profundidade da camada da precipitação.
VIRS	Detecta radiação proveniente da Terra em cinco regiões espectrais, variando do visível ao infravermelho, ou 0,63 a 12 micrômetros, fornecendo observações de alta resolução de cobertura das nuvens, o tipo de nuvem e temperatura média

	do ars do topo da nuvem.
LIS	É um instrumento pequeno e altamente sofisticado que detecta e localiza relâmpagos. Ele tem aproximadamente oito polegadas de diâmetro e 14 polegadas de altura, pesa aproximadamente 46 libras e usa cerca de 25 <i>watts</i> de potência.
CERES	Mediu a energia no topo da atmosfera, além de estimar os níveis de energia dentro da atmosfera e na superfície da Terra. Os dados do instrumento CERES foram usados para estudar a energia trocada entre o Sol, a atmosfera, a superfície, as nuvens da Terra e o espaço. Mas o registro de dados disponível é bastante curto, pois esse instrumento operou apenas entre janeiro e agosto de 1998 e março de 2000.

Fonte: adaptado de <https://pmm.nasa.gov/TRMM/trmm-instruments>

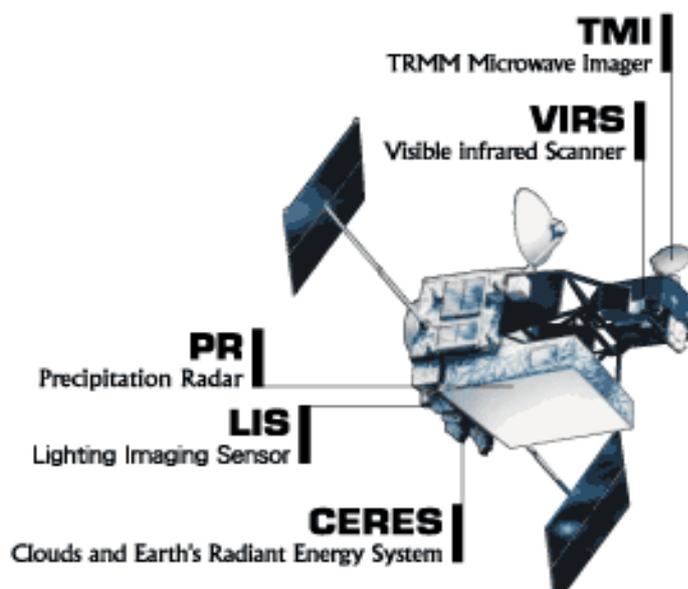


Figura 4: Sensores do satélite TRMM. Fonte: NASA (2018).

Segundo TRMM, (2018) o conjunto de sensores presentes no satélite geram diversos produtos juntamente com a combinação com os algoritmos utilizados para a realização do cálculo. Os produtos que geram estimativas de precipitação são:

- 3B42RT: resolução temporal de três horas.
- 3B42: dados diários.
- 3B43: dados mensais.

Esses produtos tem resolução espacial de 0.25 para qualquer ponto dentro da órbita do satélite.

O conjunto de dados fornecidos pelo projeto em formato *Hierarchical Data Format* (HDF) ou *Network Common Data Form* (NetCDF) integram séries temporais, que auxiliam na estimativa da precipitação e servem como base de dados para os mais diversos tipos de pesquisas ambientais (NÓBREGA, 2008).

2.4 Umidade

O ar atmosférico é constituído entre outros gases, pelo vapor d'água. A temperatura média do ar do ambiente determina seu volume, visto que a capacidade de conter o vapor d'água é função da temperatura média do ar do ar (BUTERA, 1995).

Segundo Tubelis e Nascimento (1992), a umidade relativa do ar pode ser definida, como a relação percentual entre a concentração de vapor de água existente no ar e a sua saturação. Já Ayoade (2002), considera a umidade relativa como a razão entre o conteúdo real da umidade de uma amostra de ar e a quantidade de ar, que o mesmo volume de ar pode ter em uma temperatura média do ar e pressão quando está saturado, podendo ser expressa em forma de porcentagem.

A concentração do vapor d'água é proveniente da superfície da terra, logo ao se afastar deste ela diminui. As interações físicas e fisiológicas com o meio, como vegetais e animais, fazem com que o vapor d'água seja considerado um elemento muito importante no estudo bioclimatológico (OMETTO, 1981).

No verão umidade relativa possui maior a variação diária do que no inverno, de forma análoga à temperatura média do ar do ar. Altos valores são registrados durante a madrugada, quando a temperatura média do ar do ar atinge seu valor mínimo, e mais

baixos no início da tarde, correspondentemente ao máximo da temperatura média do ar de bulbo seco, (BUTERA, 1995).

Segundo Ayoade (2002), o vapor d'água representa uma pequena parcela na atmosfera, apenas 4% de seu volume, pode ser considerado um componente importante na determinação do tempo e do clima, devido aos seguintes fatos:

- Origem de condensação e de precipitação;
- Pode absorver tanto a radiação solar quanto a terrestre e, assim, desempenha o papel de regulador térmico no sistema Terra-atmosfera;
- Exerce efeito sobre a temperatura média do ar do ar;
- Contém calor latente que é importante fonte de energia para a circulação atmosférica e para o desenvolvimento de perturbações atmosféricas.
- Influencia a taxa de evaporação e de evapotranspiração.
- Determina a temperatura média do ar sentida pelo homem, pois atua juntamente com a temperatura média do ar do ar.

Para medir a umidade relativa do ar, os equipamentos utilizados possuem propriedades associadas ao teor de vapor de água contido na atmosfera. Os aparelhos que mensuram esta variável ambiental possuem dois termômetros analógicos, um deles possui seu bulbo em contato com o ar e o outro tem o bulbo envolto com uma gaze umedecida. A gaze perde água de acordo com a concentração de vapor de água no ar, gerando uma temperatura média do ar diferente entre os termômetros. Esta amplitude de temperatura média do ar possibilita medir a umidade relativa. (SENTELHAS, et al., 2007).

2.5 Temperatura média do ar

A temperatura média do ar é um dos principais agentes causadores de modificações do clima local. Sendo a mais importante consequência da radiação solar. O aquecimento da atmosfera próxima à superfície terrestre ocorre a partir do aumento da temperatura média do ar da superfície pelos raios solares. (AYOADE, 2002).

A temperatura média do ar do ar, dentre as demais variáveis ambientais, tem influência direta e significativa sobre muitos processos fisiológicos que ocorrem na natureza. Ela está relacionada com o desenvolvimento e crescimento das plantas, devido ao seu efeito na velocidade das reações químicas e dos processos internos de transporte (PEREIRA et al., 2002).

A temperatura média do ar média do ar possui relação direta com a intensidade de radiação solar recebida e a nebulosidade. As nuvens interceptam o fluxo de radiação infravermelha emitida o que modifica o balanço térmico local. Porém o solo ganha energia solar se o céu estiver limpo, mas perde energia por irradiação mais do que com céu encoberto. Para se determinar a temperatura média do ar do ar em um local depende do ganho de calor, devido ao deslocamento de massas de ar que ocorre através dos ventos, correntes oceânicas e da existência de proximidade de grandes massas de água (BUTERA, 1995).

Segundo Butera (1995), os dados medidos de temperatura média do ar do ar estão disponíveis em diversas formas: valores médios, máximos e mínimos anuais, mensais e diários, etc. Para caracterizar o clima térmico de um local, deve-se utilizar dados provenientes de 11 observações efetuadas em um período de tempo suficientemente longo, para que estas tenham valor estatístico (utilizando no mínimo 3 anos).

2.6 INMET (Instituto Nacional De Meteorologia)

Segundo o ministério da Agricultura (2018), o INMET é Órgão Oficial da Meteorologia do País, responsável pela previsão do tempo. Ele contribui para o com o dados de diversas variáveis ambientais e proporciona subsídios para análises de riscos climáticos e planejamento ambiental.

As estações meteorológicas automáticas possuem um sistema de armazenamento que é composto por um *datalogger*⁵, que faz o registro dos valores observados e salva em um disco de memória. Possui também um sistema de energia, que é independente de

⁵ Processador central de baixo consumo de energia.

energia elétrica, isso inclui baterias e painéis solar. E por fim há um sistema para comunicação direta das estações com o INMET EM Brasília em tempo real (SILVA, 2014)

O INMET fornece dados de: Temperatura Instantânea do Ar, Temperatura Máxima do Ar, Temperatura Mínima do Ar, Umidade Relativa Instantânea do Ar, Umidade Relativa Máxima do Ar, Umidade Relativa Mínima do Ar, Temperatura Instantânea do Ponto de Orvalho, Temperatura Máxima do Ponto de Orvalho, Temperatura Mínima do Ponto de Orvalho, Pressão Atmosférica Instantânea do Ar, Pressão Atmosférica Máxima do Ar, Pressão Atmosférica Mínima do Ar, Velocidade Instantânea do Vento, Direção do Vento, Intensidade da Rajada do Vento, Radiação Solar e Precipitação acumulada (INMET, 2018). A Figura 5 ilustra uma típica estação do INMET onde os dados são coletados pelos instrumentos e sensores que mensuram os valores de cada variável observada. As medidas têm intervalos de minuto a minuto integrando para o período de uma hora para realização de uma transmissão desde 1969(SCHNEIDER, 2012; SILVA, 2014).



Figura 5: Pluviômetro, ao centro a antena de comunicação via satélite e à direita o datalogger com sensor de temperatura presos ao mastro. Fonte: INMET, 2018.

O INMET tem procedimentos de medidas analógicas e digitais. Para medidas analógicas basicamente usa-se medição da resistência de tensão (INMET, 2018). Para medidas digitais usa-se: medida com contador de pulsos, medida da frequência, mensagem de dados seriais. Os dados das variáveis ambientais disponibilizados pelo INMET integram longas séries temporais.

2.7 Séries Temporais

Ao longo da história o desenvolvimento dos estudos relacionados a séries temporais já obteve grandes mudanças, com consideráveis avanços principalmente a partir da década de 1970.

Na ocasião, dois autores Box e Jenkins, desenvolveram uma proposta acadêmica que instigava pesquisadores da época a iniciarem processos de pesquisa que alavancaram uma série de contribuições para a área. A partir deste momento, Séries Temporais passou a ganhar destaque dentro do cenário da Estatística.

Box e Jenkins (1976) definem que existe uma classe de fenômenos naturais, econômicos que quantifica numericamente produzindo uma sequência lógica de dados que são distribuídos no tempo. Essa sequência e ordenação correta dos dados segmentados denominam-se “Serie Temporal”.

Wooldridge (2000) menciona que Série Temporal, é considerada uma sequência de realizações ou observações de uma variável qualquer ao longo do tempo, define que a sequência de pontos precisa estar em ordem sucessiva e pode ocorrer variações uniformes.

Séries temporais podem ser aplicadas à um conjunto de dados estatísticos que mostrem, por exemplo, a temperatura média do ar diária de uma cidade ao longo de um período, ou um número de acidentes ou homicídios de um determinado lugar. Dessa forma, os dados podem ser analisados e mensurados em determinado período.

Os métodos baseados em previsões estatísticas são calculados a partir de dados e conjuntos históricos utilizando uma “série de tempo”, permitindo assim a análise precisa de previsões e estatísticas sobre vários segmentos.

Nogueira (2009) menciona que uma série de tempo é uma série de observações de alguma quantidade de interesse, isto é, utiliza-se de uma variável aleatória em relação ao tempo para calcular dados estatísticos com o intuito de prevenir ou decompor alguns segmentos.

Representação matemática, segundo Nogueira (2009):

$$\{X_1 = x_1, X_2 = x_2, \dots, X_t = x_t\}$$

Onde X é uma variável aleatória de interesse qualquer e i são as observações tomadas através do tempo $1, \dots, 2, \dots, t$.

A Figura 6 ilustra exemplos de séries temporais de quatro tipos de variáveis.

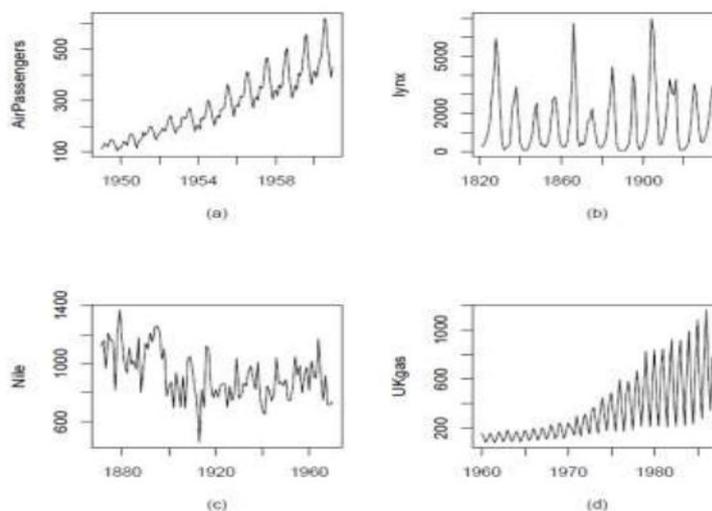


Figura 6: (a) Totais mensais de passageiros em linhas aéreas internacionais nos EUA, (b) Número anual de lincas capturados em armadilhas no Canadá, (c) Medições anuais de vazões do Rio Nilo em Ashwan, (d) Consumo de gás no Reino Unido no primeiro trimestre de cada ano. Fonte: EHLERS, 2009.

A análise de séries temporais tem os seguintes objetivos:

- Encontrar o que gerou determinada série temporal.
- Avaliar tendências da série temporal como comportamento dos dados.
- Realizar previsão por tempo indeterminado (curto ou longo prazo).
- Encontrar periodicidade ou frequência nos dados, com análise espectral.

Identificado o objetivo de análise é possível avaliar qual método é mais adequado para análise.

2.7.1 Método de Análise de Séries Temporais

O método devidamente adequado para análise de uma série temporal irá depender do objetivo de análise. Visto que uma série temporal de qualquer variável que esteja realizando a análise tem propriedades estatísticas e comportamento único, o ideal é que seja identificado e utilizado um modelo adequado (VIEIRA, 2007).

Sendo assim, é possível definir dois domínios em que é possível realizar a análise de séries temporais: do tempo e da frequência. No domínio do tempo existe interesse nos dados que acontecem em determinado momento da série temporal. No domínio da frequência, análise é feita na decomposição em frequências e utilização de equações na série temporal para encontrar informações que não são encontradas em uma simples análise no domínio do tempo (LIMA, 2015).

Um método que pode ser utilizado para análise de séries temporais e que realiza a análise em tempo, frequência e intensidade de energia é a análise de ondaletas, ou *wavelets* (FURON *et al.*, 2008; ROCHA, 2008; LOH, 2003; ALVES, FILHO e SILVEIRA, 2013).

2.7.1.1 Transformada de Wavelet

As Transformadas de Wavelet são definidas como uma abordagem matemática que analisam e modificam funções hierárquicas em diversas áreas de aplicação. Historicamente as “transformadas” possuem um tratamento fundamentado em análise e interpretação de operações que tratam de imagens de modo que possam ser totalmente segmentados, facilitando assim sua compreensão e análise. O termo “Transformada” significa modificar uma representação de um conjunto de dados de maneira que esses possam ser facilmente analisados (DURTRA E MINETTO, 2015).

Wavelets são utilizadas em diversas áreas de estudo, é possível citar, por exemplo: computação numérica, processamento de sinais, análise de imagens, processamento de dados ambientais.

As transformadas Wavelets iniciaram como estudo focado no processamento de imagens, possibilitando a cobertura de deficiências que a análise de Fourier possibilitava apenas em tempo. Na análise de Fourier é possível caracterizar o ciclo diário e fenômenos com escalas de pequenas frequências, mas isso é uma tarefa inviável, pois para sinais de fenômenos não-estacionários que são tipicamente caóticos este tipo de análise não traz bons resultados. Neste contexto é necessário o uso da transformada de Wavelets. Dessa forma, informações de tempo, frequência e escala de intensidade de energia estão presentes na equação (DAUBECHIES,1992).

A autora ainda menciona que Wavelets decompõe de um sinal que por meio de funções elementares, criadas a partir de dados escalonáveis de uma função base, denominada de Wavelet mãe, conforme a Equação 1.

Equação 1- Decomposição de Wavelets

$$\psi_{S\tau}(t) = \frac{1}{\sqrt{S}} \psi\left(\frac{t - \tau}{S}\right)$$

Fonte: DAUBECHIES, 1992 - modelo adaptado.

Daubechies (1992) escreve que as Transformadas Wavelets possuem outras características diretas como a transformada Wavelet contínua (CWT) conforme apresenta a Equação 2. Esta por sua vez denomina-se da soma sobre todo o tempo de um sinal multiplicado por funções escaladas e transladadas da função ψ , esta por sua vez gera por meio de um coeficiente um conjunto c que representa o sinal na base Wavelet conforme apresenta a Equação 3.

Equação 2 - Wavelets Contínuas

$$C(S, \tau) = \int_{-\infty}^{\infty} f(t) \psi_{S,\tau}(t) dt$$

Fonte: DAUBECHIES, 1992 - modelo adaptado.

Equação 3 - Wavelets Contínuas - Inversa

$$f(t) = \int_S \int_{\tau} C(S, \tau) \psi_{S,\tau}(t) dt ds$$

Fonte: DAUBECHIES, 1992 - modelo adaptado.

Outro tipo específico de Wavelet está denominado por Wavelet de Morlet, em que pode-se destacar Torrence e Compo (1998) que descrevem a Wavelet de Morlet como uma função periódica cuja amplitude é modulada por uma gaussiana, já o envelope Gaussiano é definido por uma segunda exponencial, enquanto a onda exponencial complexa é a primeira exponencial definida como função de Wavelet básica.

A Wavelet de Morlet foi introduzida para ser utilizada dentro das famílias mais complexas de Wavelets ortogonais conforme Equação 4:

Equação 4 - Wavelets Morlet

$$\psi(t) = \pi^{-0,25} e^{iw_0 t} e^{-1/2t^2}, \text{ para } \omega_0 \geq 5$$

Fonte: TORRENCE e COMPO (1998) - modelo adaptado

Onde:

- ψ - valor da Wavelet para um parâmetro não-adimensional;
- w_0 - frequência (fornece o número de oscilações dentro da própria Wavelet);
- t - período ou escala temporal de análise (adimensional);

Mediante os autores suas variações caracterizadas por Jean Morlet, que foi responsável, em 1942, por introduzir o trabalho das variações perante a comunidade de análises sistêmicas, ficou conhecida como definição de Morlet Wavelet, obtendo a equação variável abaixo:

Equação 5 - Wavelets Morlet - equação original segundo Jean Morlet

$$\psi_{\sigma}(t) = C_{\sigma} \pi^{\frac{-1}{4}} e^{\frac{-1}{2}t^2} (e^{i\sigma t} - K_{\sigma})$$

Fonte: TORRENCE e COMPO (1998) - modelo adaptado

Onde o critério de admissibilidade é definido e pela constante de normalização, obtendo a equação abaixo:

Equação 6 - Variação da Wavelet Morlet:

$$C_{\sigma} = \left(1 + e^{-\sigma^2} - 2e^{\frac{-3}{4}\sigma^2} \right)^{\frac{-1}{2}}$$

Fonte: TORRENCE e COMPO (1998) - modelo adaptado

Estas Wavelets foram conceituadas por Jean Morlet. Autores como Torrence e Compo (1998), Torrence e Webster (1999), Nakken (1998), Reboita (2004), Guedes et al. (2005), Kayano e Blain (2007) e Blain (2010) utilizaram wavelet morlet na análise de séries temporais da precipitação. Geralmente utilizado em sinais geofísicos e dados de ambientais, pois possibilita análises de mudanças de fase e amplitude de séries temporais uma Wavelet complexa como a de Morlet que pode ajudar a capturar o comportamento oscilatório dos dados, dessa forma, a transformada Wavelet revela qual parte do sinal analisado transporta energia e em quais frequências. A Wavelet de Morlet tem uma característica que é o efeito de borda, ou seja, resultado final da análise possui períodos fora da parte superior do cone de influencia deve ser desconsiderados por não

possuírem confiança estatística adequada, dessa forma, é necessário utilizar o cone de influência que é representado por uma linha (BARBOZA *et al.*, 2005).

2.8 Busca por Similaridade em Séries Temporais

Uma série temporal pode ser definida como uma coleção de informações por meio de fenômenos que são explorados ao longo do tempo e organizados de forma cronológica (KIRCHGASSMER; WOLTERS, 2007).

Houve historicamente um grande interesse em séries temporais, e isso provocou uma série de estudos que possuem características em comum, isto é, a busca pela similaridade, comparando as séries temporais como um todo. O conceito é utilizado diretamente na definição do problema, por exemplo, na detecção de motifs, que são padrões. Um motif pode ser definido como o par de subsequências mais semelhante em uma série temporal (MUEEN, 2013).

Furtado (2014) aponta que teoricamente a abordagem de busca por similaridade em séries temporais parece ser simples, todavia tem-se mostrado competitiva em um cenário com vários métodos de classificação complexos, identificando em várias ocasiões alguns problemas decorridos desta classificação. Um dos fatores que complicam tais aplicações está diretamente ligado as Distâncias Comumente, que a medida que é utilizada em séries temporais pode facilmente assumir qualquer valor, perante a equação ($x = (x_1, x_2, \dots, x_n)$ e $y = (y_1, y_2, \dots, y_n)$).

O autor menciona que por meio da aplicação da equação ($x = (x_1, x_2, \dots, x_n)$ e $y = (y_1, y_2, \dots, y_n)$), é facilmente compreendido as séries temporais de comprimento, em que na Equação 7, n e d podem assumir qualquer valor, sendo a distância de Manhattan, com $d = 1$ e a distância euclidiana, com $d = 2$, como segue a equação a Equação 7:

Equação 7 - Equação de Comprimento

$$dist(x, y) = \left(\sum_{k=1}^n |x_k - y_k|^d \right)^{\frac{1}{d}}$$

Fonte: FURTADO, 2014 - modelo adaptado

O mesmo autor ainda relata que diante tais aplicações, as junções necessárias são mais flexíveis, aplicando-se algumas observações baseadas em x no tempo t_x , que é combinada com y em um tempo $t_y \neq t_x$. Sendo assim, compreende-se que é possível calcular uma menor distância obtida através de junções não-lineares observadas de acordo com algumas restrições.

Furtado (2014) define que é preciso estar atento a tais restrições pela busca de similaridade em séries temporais observando restrições como:

- Restrição de fronteira: A junção é realizada entre séries x e toda série y , portanto começa através dos comprimentos destas séries temporais, quando comparadas. Por exemplo: seja $w_1 = (1,1)$ e $w_k(x,y)$ as comparações serão feitas em todas as séries temporais utilizadas.
- Restrição de continuidade: As junções são realizadas através de etapas de maturidade singulares, e estas similaridades são observadas através dos “saltos” de valores expressos na equação.
- Restrição de monotonicidade: As junções aqui são observadas e ordenadas de acordo com suas preservações, obedecendo a uma ordem cronológica de valores. Por exemplo: $w_k = (a,b)$ então $w_{k-1}=(a,b)$ onde $a-a \geq 0$ e $b-b \geq 0$, faz com que os valores comparados nas séries temporais sejam espaçados no tempo.

Encontrar padrões semelhantes em séries temporais é um trabalho que pode ser aplicado em diversas áreas, porém existem diversos tipos de medidas de similaridades que podem ser utilizadas com base nos dados tratados, os trabalhos de Donavam (2001), Foote 1997, Singh e Rajan (2011), Bala, Kumar e Birla (2010), exemplificam diversos tipos de distância e os variados tipos de dados, todos os trabalhos têm em comum é o fato de elencar a Distância Euclidiana como base para séries temporais homogêneas.

2.8.1 Distância Euclidiana

Oliveira (2007) descreve que ao coletar propriedades de indivíduos (séries temporais ou dados), é possível que o conjunto dessas propriedades analisadas possa ser projetado em um espaço métrico e a distância euclidiana dá a proximidade entre os

indivíduos na projeção. Dessa forma considera-se o resultado da distância euclidiana a similaridade entre esses indivíduos estudados.

A distância euclidiana pode ser compreendida como a distância entre duas parcelas (1 e 2), medidas por unidades de indivíduos de duas espécies, compreende-se em (i e j) sendo a raiz quadrada do somatório dos valores quadrados das diferenças entre os valores de i e j, relacionando as variáveis ($k= 1,2,\dots,p$), representada pela Equação 8 (OLIVEIRA, 2007):

Equação 8 - Equação da distância entre dois pontos cartesianos

$$d_{ij} = \sqrt{\sum_{k=1}^p (X_{ik} - X_{jk})^2}$$

Em que: X_{ijk} : representa a característica do indivíduo i; P: representa o número de variáveis na amostra.

Oliveira (2007) aponta que a distância euclidiana é útil para análise e classificação entre pontos mais complexos ou por meio de um conjunto maior de espécies a serem analisadas, por exemplo, quilos, anos, quantidades altas (milhões) conforme expressão, mas o simples fato de alteração das unidades de valor altera completamente o valor do coeficiente, padronizando assim estas variáveis de elementos.

Para dados que não apresentam uma variação significativa em sua complexidade relacionada a cada instância da série ou se não existe uma diferença significativa entre cada grupo de dado, por exemplo: todos os conjuntos de dados possuem o mesmo tamanho 365 dias. Nestes casos a Distância Euclidiana apresenta um excelente desempenho, devido a sua acurácia e tempo de execução (NASCIMENTO, 2016).

Com relação aos dados ambientais existem, na literatura implementações da Distância Euclidiana para análise de similaridade com foco em detecção de padrões em diversas áreas como o trabalho MELO et al. (2006) que trata da determinação de regiões homogêneas quanto à distribuição de frequência de chuvas, o trabalho de Silva et al. (2001) que procurou identificar padrões de chuva no estado Ceará classificando 20 estações em quatro grupos. No trabalho de agrupamento da precipitação mensal de Bacia Hidrográfica (FECHINE, 2008), para encontrar similaridades das projeções de temperatura média do ar média, precipitação e umidade relativa para o Nordeste

Brasileiro (HAMADA et al., 2015), no reconhecimento de padrões de operação de uma turbina a gás da unidade termelétrica, integrante do parque da Petrobrás com o reconhecimento de falhas por dispersão de temperatura média do ar (BARRAGAN et al. 2012), na classificação de séries temporais por similaridade e extração de atributos para identificar insetos (SILVA, 2014).

Como podem ser observados, os trabalhos aqui apresentados descrevem métodos para aplicação da distância Euclidiana em diferentes tipos de dados, porem todos com características semelhantes aos aqui tratados.

2.9 Geoprocessamento e Sistemas de Informação Geográfica

O termo “geoprocessamento” começou a ser utilizado a partir da segunda metade do século XX. É relacionado ao desenvolvimento da ciência espacial. Com o avanço da computação criou-se a possibilidades de armazenar e representar documentos cartográficos em ambiente computacional (CAMARA; DAVIS, 2001). Diante disso os autores escrevem:

[...] denota a disciplina do conhecimento que utiliza técnicas matemáticas e computacionais para o tratamento da informação geográfica e que vem influenciando de maneira crescente as áreas de Cartografia, Análise de Recursos Naturais, Transportes, Comunicações, Energia e Planejamento Urbano e Regional. (CÂMARA; DAVIS, 2001, p. 2).

Geoprocessamento é definido como um conjunto de ferramentas computacionais capazes de realizar análises complexas, ao integrar dados de diversas fontes e ser capaz de criar banco de dados geo-referenciados (CÂMARA; DAVIS, 2001).

O avanço das tecnologias computacionais aplicadas a áreas matemáticas possibilitam a disseminação de cálculos complexos em curtos espaços de tempo, dessa forma, criando novas perspectivas de estudo.

Sistemas de Informação Geográfica (SIG) são definidos por Burrough (1986) como conjunto de métodos e técnicas computacionais, dispostas utilizando uma ferramenta (*software*), capaz de armazenar, analisar, transformar dados em um conjunto de informações espaciais aplicadas ao mundo real.

Silva (1998) condiciona algumas diretrizes para que este conjunto de “ferramentas computacionais” seja considerado como SIG, isto é, estes sistemas

precisam usar recursos computacionais, por meio de um conjunto de dados armazenados (banco de dados), sendo seguro, com informações geo-processadas e referenciadas, livre de erros, contendo funções de análise aplicadas em álgebra.

2.9.1 Traduzindo informações geográficas para o computador

Um dos maiores problemas decorridos da “geoinformação” diz respeito ao entendimento do processamento computacional ligado ao espaço e ao mundo real, surgindo assim o paradigma dos quatro universos (EASTMAN, et. al. 1993).

Este paradigma compreendido por Eastman diferencia em sua essência o universo do mundo real (realidades a serem modeladas e processadas com uso de sistemas), o universo matemático conceitual (tem por definição os cálculos processados das entidades representativas), o universo de representação (contém dados das entidades mapeadas e representadas geometricamente) e por fim o universo de implementação (contendo estruturas e análises utilizando algoritmos de desempenho para estrutura de *hardware* e a quantidade de dados armazenados), ocorrendo assim a decodificação dos dados geográficos para o computador (EASTMAN, et al. 1993).

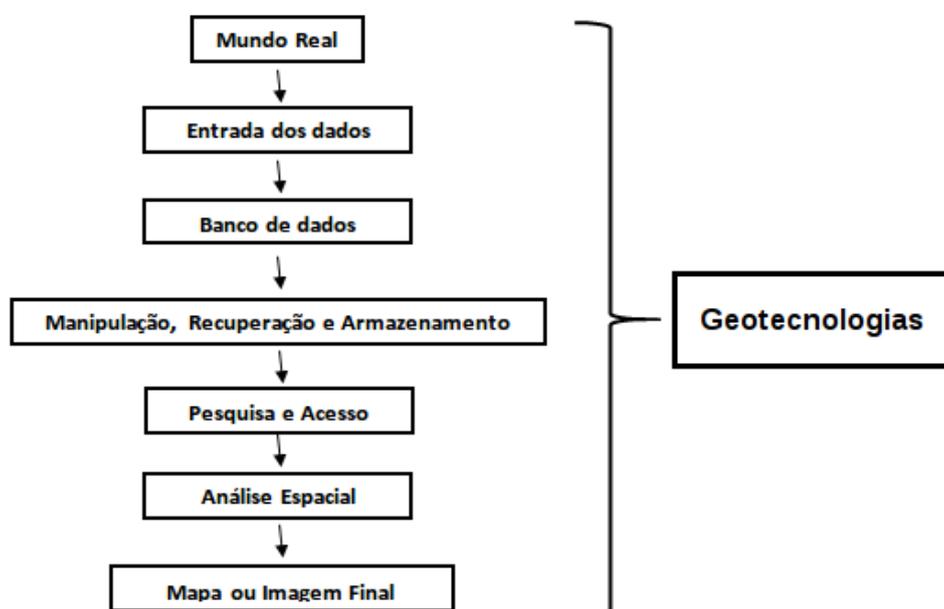


Figura 7 - Esquema de etapas de geração de imagens através de SIG. Fonte: Adaptado pela autora.

Basicamente seguindo o esquema apresentado na Figura 7 as Geotecnologias tem como base algum processo do mundo real, neste caso poderiam ser incluídos aqui mapas, dados de precipitação, dados de umidade e outros dados georreferenciados. Na

sequência é feito a entrada de dados, em um banco de dados, onde é possível manipular, armazenar e recuperar os dados com base no conteúdo.

O esquema de Geotecnologia também prevê como próximo passo a pesquisa e acesso de usuários aos dados e informações ao sistema, assim como análise espacial, com a visualização de um ou vários mapas com imagens resultantes de suas buscas. Eastman, et al. (1993), descreveu o esquema que utilizado até hoje em arquiteturas de SIG.

2.10 Arquitetura de SIG

Por meio da empregabilidade da tecnologia matemática computacional, é possível notar diversos avanços em SIG e conseqüentemente nas arquiteturas SIG. Um SIG precisa ter características computacionais para que o seu emprego seja realmente aplicado as informações geográficas (SILVA, 1998). A Figura 8 ilustra a estrutura de como deve ser a arquitetura de um SIG.

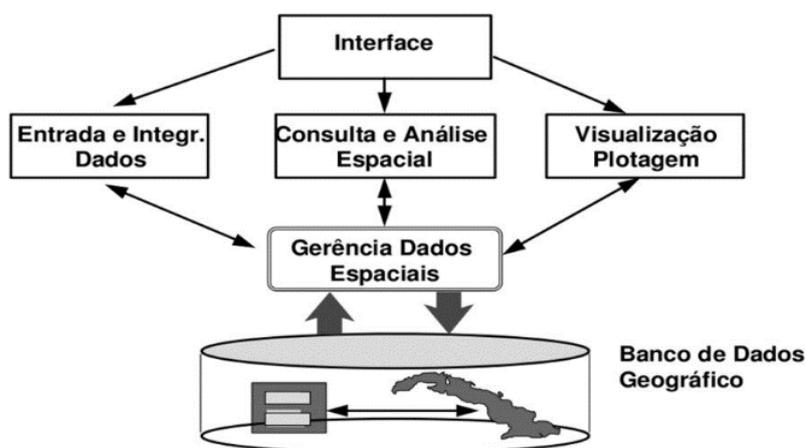


Figura 8: Arquitetura de SIG. Fonte: INPE (2018), modelo adaptado.

O Instituto Nacional de Pesquisas Espaciais (INPE) define características básicas de um SIG ilustrado pela Figura 8:

- Interface com o usuário: corresponde ao meio que o usuário utiliza para interagir com o SIG.
- Entrada e integração dos dados: representa os dados espaciais inseridos no banco de dados geográfico (BDG).
- Funções de processamento gráfico e de imagens, visualização de plotagem.

- Armazenamento e Recuperação de dados.

Um SIG precisa necessariamente oferecer por meios de seus mecanismos computacionais condições para que haja processamento de dados espaciais de modo que ao utilizar um BDG possa ser disponibilizada uma cópia, quando for necessário (ZEILHOFER, 2003).

Câmara e Ribeiro (2004) destacam que os SIGs possuem duas grandes estruturas Dual e Integrada, podendo estender para arquiteturas diferentes baseadas no tipo de serviço que o sistema irá fornecer.

Arquitetura Dual: esta estrutura é baseada na implementação de um Sistema Gerenciador de Banco de Dados Relacional (SGDBR), em que por meio desta estrutura seria possível armazenar os componentes convencionais relativos aos objetos espaciais. O gerenciamento de banco de dados utilizando os recursos básicos (inclusão, edição, exclusão e consultas) é possível controlar tais componentes equacionando as questões de otimização das consultas, transações, proporcionando ao usuário controle, integridade dos dados, confiabilidade e segurança dos dados.

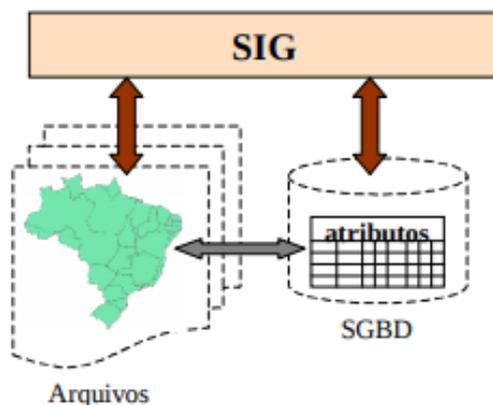


Figura 9: Arquitetura SIG - Estrutura Dual. Fonte: CÂMARA e RIBEIRO, (2004)

Arquitetura Integrada: já nesta estrutura, todos os dados relativos ao processamento das equações baseadas em gráficos ou imagens, ficam dispostos no banco com o uso de tabelas contendo valores alfanuméricos. Como o banco possui estrutura relacional e a arquitetura também, é formado um conjunto de chaves primárias e secundárias de modo que identifiquem os valores dos objetos no banco. Assim, é possível fazer o gerenciamento da implementação de recursos analíticos utilizando geoprocessamento disponibilizado juntamente com as ferramentas computacionais,

produzindo ganhos notáveis relacionados a robustez, performance, estabilidade, porém pode apresentar problemas baseados na consulta em SQL (*Structured Query Language*).

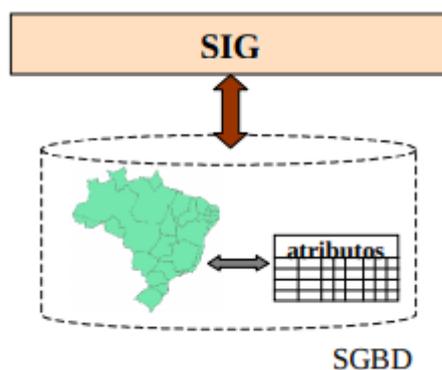


Figura 10: Arquitetura SIG - Estrutura Integrada Relacional. Fonte: CÂMARA e RIBEIRO, (2004).

A escolha de qual arquitetura utilizar irá depender do tipo de dado e aplicação que será construída.

2.10.1 Representação dos Dados Geográficos

Segundo Barcellos (2002) o que diferencia dados geográficos de outros dados, é justamente sua composição espacial. A representação de dados geográficos funciona como “representações da superfície terrestre em conjunto com localização espacial”.

Para realizar consultas no BDG, utiliza-se também a linguagem SQL (*Structured Query Language*). Entretanto, como se trata de dados com referência geográfica, com características especiais, é necessário utilizar uma extensão junto ao SGBD, de maneira que ele possa interpretar e manipular os dados espaciais adequadamente.

A partir desta coleta de informações baseadas em coordenadas é possível a representação dos dados geográficos, fazer armazenamento em uma base concreta e segura, processar as informações e disponibilizar utilizando várias ferramentas, inclusive por meio de uma rede *SIG Web*.

Ao se manipular dados com essas características é preciso considerar relações topológicas como adjacência, pertinência, intersecção, cruzamento dos dados e a proximidade deles. Todas essas informações são manipuladas em duas grandes classes: Vetorial e Matricial, onde:

Classe Vetorial: O armazenamento ocorre pelo modelo vetorial, possuindo a localização de atributos gráficos, sendo representativo com o uso de coordenadas geográficas representado com: arcos, pontos, linhas, etc.

Classe Matricial: A representação acontece utilizando colunas e linhas, aplicando-se valor das células mediante os atributos analisados, apontando estatísticas precisas sobre o objeto em análise.

Para se trabalhar com a manipulação de dados geográficos para diferentes ferramentas computacionais existe a necessidade definir um conjunto de normativas e especificações técnicas, e para o caso dos dados geográficos existe a organização OGC (*Open Gis Consortium*).

O OGC trabalha em um processo para desenvolver padrões para manipulação e serviços geográficos. É uma organização internacional formada por organizações comerciais, governamentais, não lucrativas e instituições de pesquisa (OGC, 2014). A OGC define padrões que permitem a representação de dados espaciais. Os principais padrões estabelecidos são:

- *Web Feature Service (WFS)*: uma de suas função e permitir o acesso e manipulação de dados geográficos na *web* utilizando chamadas independentes de plataforma.
- *Web Coverage Service (WCS)*: essa norma define requisitos como recuperar dados geoespaciais com representação de espaço.
- *Web Map Service (WMS)*: permite implementar uma interface HTTP, para solicitar imagens de mapas a partir dos dados geográficos armazenados no banco de dados. Exibe os mapas no página *web*.

Os SIG's implementados para *web* utilizam as especificações oferecidas pelo OGC, para que haja a padronização, o compartilhamento e a integração dos dados para diferentes sistemas.

2.11 SIG Web

SIG *Web* são sistemas baseados na arquitetura SIG, porém com características para ambiente virtual (*online*). Estas plataformas de gerenciamento armazenam,

processam, analisam dados geográficos espaciais por meio de plataformas que são manipuladas por usuários, baseados em entrada de dados convencional (CÂMARA et al., 1996)

Esta característica para sistemas de informação geográfico baseado em plataformas virtuais permite que milhões de usuários compartilhem da mesma base de dados para estudos proporcionando um ganho em robustez no banco, além de rapidez pela qual as informações são processadas fornecendo aos usuários dados precisos para seus estudos (SCHIMIGUEL et al., 2005).

Sistemas baseados nesta estrutura contam com uma equipe técnica que é responsável pelo “gerenciamento” dos dados espaciais (bases cartográficas), permitindo que tais cálculos sejam disponibilizados em ambiente desktop podendo ou não ter características espaciais.

Estes sistemas facilitam aos operadores o acesso, pois basta ter um navegador e de acordo com as políticas das empresas envolvidas um sistema de autenticação de usuários que permite não só o acesso, mas a disseminação das informações coletadas como um todo, promovendo o ensino e aprendizagem colaborativa pelos profissionais envolvidos (GORNI et al, 2007).

De acordo com Pires e Medeiros (1997) é possível utilizar recursos tecnológicos *Web*, disponíveis para ambientes SIG, e com isso proporcionar não só o controle geral de dados geográficos, mas também a possibilidade de aplicação da estrutura para Visão, Base de Conhecimento e Comunicação, criando assim novos paradigmas para esta concepção.

De acordo com Zeilhofer (2015) sistemas concentrados em ambiente SIG *Web*, permitem que por meio da inserção de dados de precipitação, a análise espacial processada, armazenada e disponibilizada através de mapas, estruturando assim tal arquitetura para *web*. Normalmente tais aplicações possuem um servidor central que recebe os dados das estações cliente, processa dentre os serviços disponíveis e posteriormente fornece através de mapas estatísticos as informações desejadas pela equipe técnica.

Na Figura 11 é possível identificar uma requisição em um SIG *web*. O navegador é o cliente que “conversa” com o servidor de *web*, este requisita os dados do servidor de

mapas que interpreta o que foi solicitado, buscando as informações requeridas pelo cliente no Banco de Dados Geográfico (BDG), por meio das especificações da OGC (WFS, WMS ou WCS) e então produz uma saída e esta é retornada ao cliente.

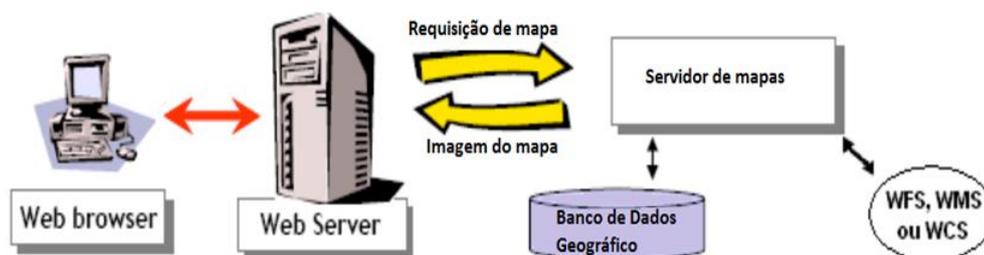


Figura 11: Requisição do SIG web. Adaptado (GORNI *et al.*, 2007).

A utilização de SIG *web* é imensa, e colabora com o monitoramento ambiental de diversas variáveis ambientais. Dentre essas variáveis está à precipitação, que é objeto de estudo deste trabalho.

2.12 Conclusão do Capítulo

As séries temporais podem ter seu comportamento analisado para verificar a dinâmica dos dados ou identificar padrões. Em muitos casos analisar somente a série temporal não é suficiente, pois é preciso um estudo mais elaborado para entender o comportamento dos dados, como uma análise de tempo, frequência e intensidade de energia dos dados. Quando a série temporal é do tipo não estacionária, como a precipitação, existe a possibilidade de verificar a periodicidade nos dados com uma visualização em tempo, escala e intensidade de energia, por meio da análise de Wavelets. Dessa forma é possível realizar a análise espectral que disponibiliza mais informações de uma série temporal além da possibilidade de verificação de periodicidade nos dados.

Para realização de análise climática, a possibilidade de utilizar comparações em séries temporais e Wavelets em um sistema com as geotecnologias que fornecem meios para tratar dados georreferenciados é uma solução para facilitar o acesso, a visualização e manejo as informações. Para ampliar a disponibilidade das informações neste tipo de sistema, também foi discutido neste capítulo sobre o SIG *web* que permite que conteúdos com informações geográficas sejam disponibilizadas na *web*, o que possibilita a expansão de publicação dos dados.

Por fim, neste capítulo foram apresentados os conceitos teóricos e as problemáticas referentes ao sistema aqui apresentado e nos próximos capítulos serão apresentados: a metodologia, resultados e conclusão.

3. Materiais e Métodos

Neste trabalho foi realizado desenvolvimento da metodologia SWAmP para produção e disponibilização de análises espaço-temporais de dados ambientais, que possibilita o processamento de dados de precipitação, umidade relativa e temperatura média do ar em várias resoluções de temporais. As várias etapas da metodologia SWAmP possibilitaram a concepção do SIG *WEB* SWAmP 1.0 que proporciona ao usuário a visualização e realização de buscas por similaridade em séries temporais e Wavelets nos dados ambientais.

A metodologia SWAmP, pode ser implementada em qualquer tipo de dado ambiental com a estrutura utilizada neste trabalho. Os dados e tecnologias selecionados para elaboração do SWAmP1.0 são comentados neste capítulo.

3.1 Dados Utilizados

Os dados utilizados neste trabalho foram obtidos de duas fontes de dados diferentes: Instituto Nacional de Meteorologia (INMET, 2018), satélite *Tropical Rainfall Measuring Mission* (TRMM) (TRMM, 2015).

3.1.1 TRMM

Foram utilizados dados diários de precipitação fornecidos pelo projeto TRMM produto 3B42, pré-processados por Guarienti (2015). O período de estudo corresponde janeiro de 1998 a dezembro 2013, pois o projeto TRMM forneceu dados até 2014⁶.

Os dados diários de 16 anos inseridos no banco de dados totalizam 15 milhões de registros de uma área que possui 2597. Os dados disponibilizados por Guarienti (2015), que inicialmente estavam em formato NetCDF, após processamento, incluído ao

⁶ <https://earthdata.nasa.gov/trmm-to-gpm> <Acesso em novembro de 2018>

banco de dados, uma tabela com um identificador (id), um campo de Latitude (lat), um campo de Longitude (lon) e o campo de dado de precipitação no formato diário (dado) e data. A Figura 12 ilustra os pontos utilizados com os dados de precipitação do projeto TRMM, os pontos possuem distância aproximada de 30 quilômetros.

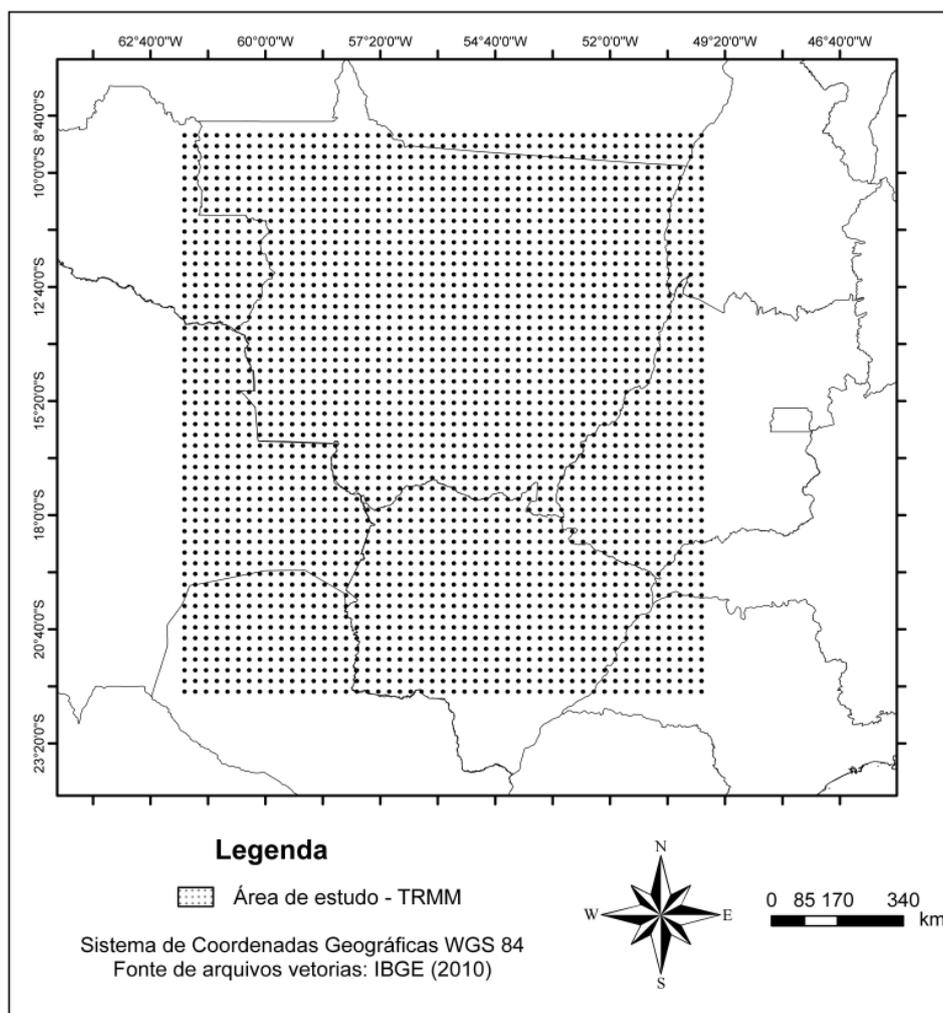


Figura 12: Localização área de estudo dos dados de precipitação do TRMM.

Além dos dados de precipitação do projeto TRMM também foram utilizados dados de umidade relativa e temperatura média do ar.

3.1.2 INMET

Foram utilizados dados de séries temporais obtidas do INMET de 28 estações, cada uma representando uma estação meteorológica automática. As estações selecionadas estão localizadas na região centro-oeste, para coincidir com os dados de precipitação da grade do TRMM. As estações meteorológicas que não apresentaram

dados foram excluídas, e os que tinham um número de dados com falhas ou *outliers* foram utilizado o Framework desenvolvido por VENTURA (2015). As estações utilizadas são identificadas pelo código do INMET, latitude, longitude, campo de dado e data.

Foram utilizados dados de Temperatura média do ar Média (°C) e Umidade Relativa Média (%), do período correspondente a janeiro de 1998 a dezembro 2013. Na Figura 13 são mostradas as estações selecionadas para este estudo.

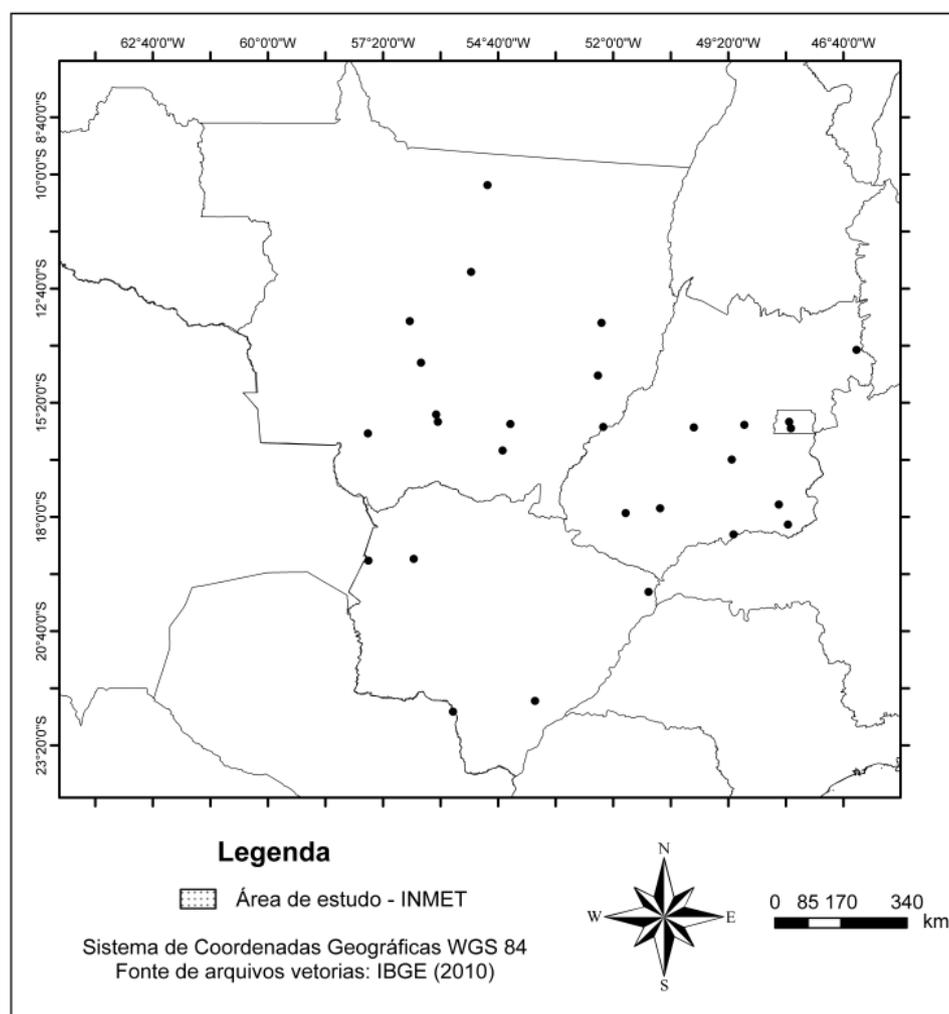


Figura 13: Localização área de estudo dos dados de umidade relativa e temperatura média do ar do INMET.

A Tabela 1 traz a descrição com os nomes das estações terrestres que possuem sensores e instrumentos para mensurar e disponibilizar dados do INMET.

Estação meteorológica

Estação meteorológica

Estação meteorológica

Aragarças	Ivinhema	Pirenópolis
Brasília	Jataí	Ponta porá
Cáceres	Matupá	Posse
Canarana	Nhumirim (nhecolândia)	Poxoréo
Catalão	Nova Xavantina	Rio verde
Corumbá	Padre Ricardo Remeter	Roncador
Cuiabá	Paranaíba	Rondonópolis
Gleba celeste	Diamantino	São Jose Do Rio Claro
Goiânia	Formosa	Itumbiara
Ipameri		

O SWAmP 1.0, possibilita a detecção de padrões similares em dados de umidade relativa e temperatura média do ar, a fim de disponibilizar uma ferramenta útil para pesquisadores que necessitem comparar dados, e, para isso, a interface deve ser simples e intuitiva.

3.1.3 Pontos utilizados

O Quadro 3 apresenta as coordenadas dos pontos utilizados para demonstração neste trabalho, ou seja, para apresentar o resultado em Wavelets e séries temporais dos dados de precipitação, umidade relativa e temperatura média do ar nas resoluções temporais: semestrais, anual e 16 anos foram utilizados os dados dos pontos apresentados no Quadro 3.

Quadro 2: Pontos e datas utilizados para apresentar os resultados do SWAmP 1.0.

Coordenadas	Análise anual	Análise semestral	Organização do Trabalho	Fonte dos Dados
-------------	---------------	-------------------	-------------------------	-----------------

-9.125, -49.875	1998	Primeiro semestre 1998	Ponto A	Precipitação TRMM
-18.98, -56.619999999999997	2004	Segundo semestre 2004	Ponto B	Umidade Relativa INMET
-14.40, -56.450000000000003	2011	Primeiro semestre 2011	Ponto C	Temperatura média do ar INMET

Após o processamento e verificação de dados com falhas, o ano de 2004 ficou 23 pontos, e o ano de 2011 com 18 pontos de séries temporais e Wavelets. Os dados do TRMM de 1998 não possuem falhas ou pontos excluídos.

Todos os pontos, anos e semestres dos dados foram escolhidos de forma aleatória para apresentar alguns resultados da metodologia que culminou na implementação do sistema SWAmP 1.0, que possui para dados de precipitação de 2597 pontos, 16 possibilidades de anos e 32 possibilidades para análises semestral de ponto, além da análise de 16 anos pontual. Para dados de umidade relativa e temperatura média do ar 28 pontos com dados do período de 1998 a 2013 com análise semestral e anual.

3.2 Busca por Similaridade

A busca por similaridade foi implementada utilizando a linguagem Python para a comparação entre duas séries temporais e entre duas Wavelets separadamente. O programa foi elaborado utilizando como parâmetro a distância Euclidiana realização dos cálculos. Os dados utilizados foram separados nas resoluções temporais: semestral, anual ou 16 anos, para que o usuário tenha diferentes visualizações dos dados.

Para a busca foi passado para o programa uma série temporal equivalente a resolução temporal trabalhada, e foi realizada a comparação em todas as séries temporais contidas no banco de dados, pegando o menor valor de distância de comparação de distância Euclidiana, ou seja, quanto mais próximo de zero, maior a similaridade entre as séries temporais. O mesmo processo foi realizado com as Wavelets,

mas neste caso foram utilizados os dados de dentro do cone de influência, visto que os períodos fora desta área devem ser desconsiderados devido ao efeito de borda, característico da análise de Wavelets. Na análise do espectro de energia, o eixo x mostra o comprimento temporal, ou seja, a resolução temporal (diária, mensal, anual), e o eixo y representa os períodos presentes na série; a escala de cores indica a energia associada a cada período da série.

3.3 Arquitetura

Essa seção descreve as ferramentas, tecnologias e linguagens de programação que culminou na elaboração do SWAmP 1.0, o sistema para processamento e detecção de padrões espaço-temporal aplicado em *Wavelets* e séries temporais de dados ambientais, utilizando a análise de similaridade apresentado neste trabalho. A Figura 14 ilustra o esquema com as etapas da implementação.

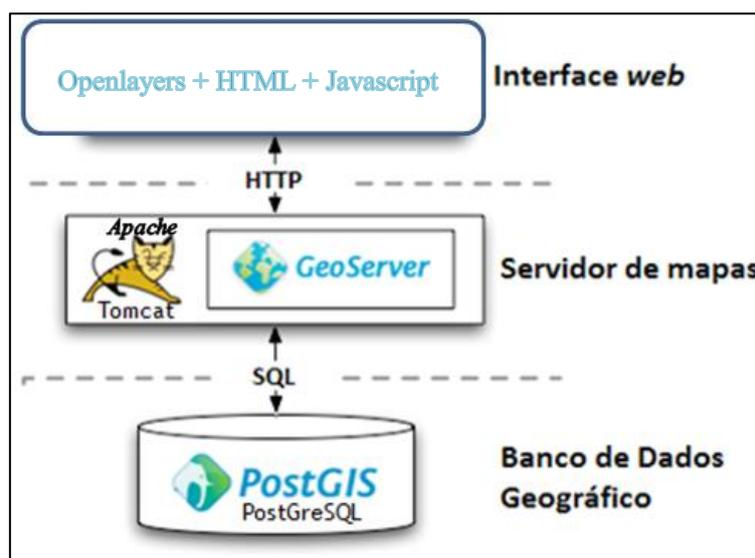


Figura 14: Esquema para construção do SWAmP 1.0.

Seguindo o esquema apresentado na Figura 14, a primeira etapa foi a escolha do Sistema Gerenciador de Banco de Dados (SGBD), que para o desenvolvimento do SWAmP 1.0, foi escolhido PostgreSQL⁷ versão 9.5.2, que possui código aberto, confiabilidade e integridade na transação de dados (CARVALHO, 2017).

⁷ Disponível em < www.postgresql.org >

O SGBD armazena as Séries Temporais e Wavelets de precipitação nas resoluções temporais: semestral, anual e da série histórica completa de dezesseis de dados de precipitação do TRMM, bem como as distâncias de similaridades de todos os pontos com todos os pontos, também armazena os dados com resolução temporal anual e semestral de temperatura média do ar e umidade relativa do INMET. O PostGis⁸ é uma extensão espacial que adiciona ao PostgreSQL o suporte a objetos geográficos, possibilitando que ele seja utilizado como um banco de dados espacial, dessa forma, oferece suporte ao uso de índices espaciais, tratamento de informações geográficas, importação e exportação de dados geográficos. Para realização da busca espacial, o SWAmP 1.0, utiliza a função de distância do PostGis, para excluir da busca os pontos próximos em um raio de 100 km, afim de que os pontos selecionados não sejam os vizinhos imediatos.

Para isso foi necessário realizar uma sequência de passos que inclui a criação de um índice do tipo GIST⁹, e utilização da busca geográfica ST_DWithin. O índice do tipo GIST foi criado para que as buscas fossem realizadas utilizando a medida de distância por quilometragem, dessa forma foram criados os resultados de uma função em vez de uma coluna armazenada. Após realizar a configuração do PostGis, foi implementada a busca geográfica.

Para hospedagem foi utilizado o servidor *Web* Apache/Tomcat8¹⁰, e para realizar as funcionalidades de SIG *Web* foi utilizado o servidor de mapas Geoserver 2.11¹¹ que é um servidor de código aberto para compartilhar dados geoespaciais, e segue os padrões da OGC. A principal função do GeoServer é facilitar o uso dos dados geoespaciais e

⁸ Disponível em < <http://postgis.refractor.net/> >

⁹ (GiST) Generalized Search Tree é uma estrutura de dados que pode ser usada para construir quase todo tipo de árvore de busca, normalmente utilizada para dados geográficos.

¹⁰ Disponível em < <http://tomcat.apache.org/> >

¹¹ Disponível em < <http://geoserver.org/> >

permitir a disseminação de informações geográficas de maneira interoperável (Oliveira, 2008).

A fim de publicar os dados no sistema *web*, é necessário configurar o GeoServer para que ele se comunique com o banco de dados, visto que Séries Temporais e Wavelets de dados ambientais estão armazenados nele. É preciso informar ao GeoServer o nome da tabela que se deseja disponibilizar, o nome do banco de dados e as informações de acesso. Após a configuração do GeoServer, é feita a construção da interface para o usuário.

Para realizar a inclusão do mapa de biomas e a camada dos pontos no SWAmP 1.0, foi utilizado o Openlayers¹² versão 4 que é uma API (*Application Programming Interface*) desenvolvida em JavaScript, responsável por gerar mapas dinâmicos para *web*. Dessa forma é possível realizar a conexão do Geoserver com o Openlayers para exibição das camadas no mapa na página *web*.

No desenvolvimento da interface foi utilizado HTML5 (*HiperText Markup Language*), que permite a modelagem de páginas dinâmicas e atrativas. Foi utilizado também o CSS (*Cascading Style Sheets*), que possibilitou a melhoria na aparência do site.

Para gerar gráficos de séries temporais, foi utilizado o D3(*Data-Driven Documents*)¹³, que é uma biblioteca JavaScript para visualizações de dados em *Web Sites*. E para gerar os gráficos Wavelets, foi utilizado um script desenvolvido em linguagem Python que recebe os dados de consulta no banco, gera Wavelets e salva no servidor para futuras consultas. Para realizar a conexão e consultas no Banco de Dados foi utilizado PHP versão 7¹⁴.

A junção de todas as etapas e processos descritas neste tópico de resultados da metodologia SWAmP permitem a realização, aplicação e utilização em outros tipos de dados, o que possibilita a produção do conhecimento, inerente ao conceito de

¹² Disponível em <<https://openlayers.org/>>

¹³ Disponível em <<https://d3js.org/>>

¹⁴ Disponível em <<https://secure.php.net/>>

metodologia. Todas as linguagens e tecnologias aqui apresentadas nas etapas da metodologia SWAmP permitiram o desenvolvimento do SWAmP 1.0 para a realização e aplicação dos resultados nos dados de precipitação, temperatura média do ar e umidade relativa. Os códigos para implementação do SWAmP 1.0, estão disponíveis em <https://github.com/gracyeli/SWAmP1.0>.

Sendo assim, é possível acessar o SWAmP 1.0 pelo link: <http://geoserver2.ic.ufmt.br/sig/index.html> para realizar testes em qualquer ponto da grade e escolher entre a análise Wavelet ou Série Temporal, semestre, ano ou análise de 16 anos.

4. Resultados e Discussão

O objetivo principal deste estudo, foi o desenvolvimento da metodologia SWAmP para análise de similaridades espaço-temporal que envolve: obtenção dos dados, modelagem com diferentes resoluções, construção de algoritmos para proporcionar análise de similaridades, construção do banco de dados geográfico, construção da infraestrutura *web*, construção da interface *web* e toda a interatividade que o sistema pode proporcionar ao usuário para realizar suas análises, como visto no Capítulo 3 Materiais e Métodos, portanto este capítulo traz os respectivos resultados.

4.1 Metodologia SWAmP

A metodologia SWAmP desenvolvida e implementada no SWAmP 1.0 utiliza as etapas fundamentadas pelo CRISP-DM, baseada no modelo modificado por Oliveira (2015) que aplicado aos dados de precipitação, umidade relativa e temperatura média do ar (séries temporais e Wavelets) possibilita maior interação entre as etapas.

A metodologia SWAmP é um conjunto de ações, tecnologias e processos empregados para a pesquisa de dados ambientais, que proporciona ao pesquisador ou usuário, uma ferramenta que possibilita a visualização da análise espaço-temporal de dados de precipitação do projeto TRMM, temperatura média do ar e umidade relativa do INMET. Seu resultado possui sequência de etapas que possuem critérios para seu funcionamento.

O resultado do desenvolvimento da metodologia SWAmP proporciona ao usuário final um sistema que possibilita vários tipos de visualizações, análises de similaridade em Wavelets e séries temporais. Uma ferramenta que pode ser utilizada para modelar o processo de funcionamento de sistemas é o padrão de modelagem de processos de negócio, *Business Process Model and Notation* (BPMN) que fornece uma notação gráfica para especificação de processos (ALLWEYER, 2010; JESTON e NELIS, 2006). A Figura 15 ilustra as etapas da metodologia SWAmP que permitem a reprodução e aplicação da metodologia com outros tipos de dados e tecnologias.

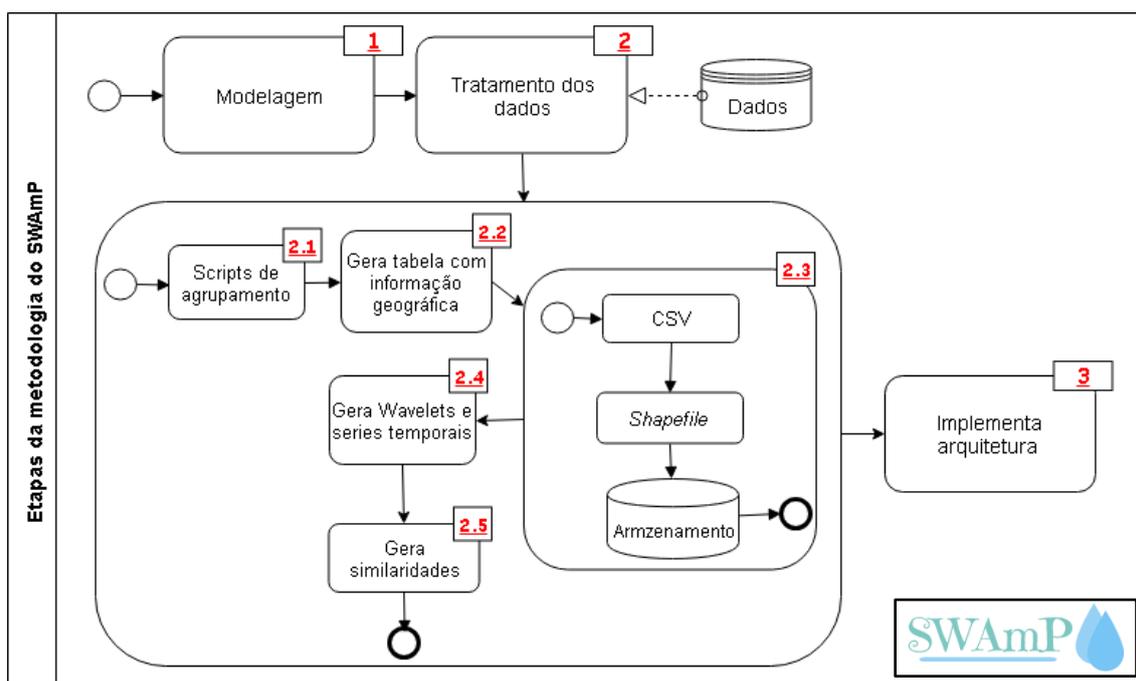


Figura 15: Etapas do resultado do desenvolvimento da metodologia SWAmP

Conforme ilustrado pela Figura 15, a primeira etapa do resultado do desenvolvimento da metodologia SWAmP, é a Modelagem dos dados ambientais, onde é feita toda a análise de como deve ser implementada em conjunto com o tipo de dado. A Modelagem dos dados deve envolver aspectos relacionados com os tipos de dados e a manipulação que se deseja fazer com os mesmos. Assim, os aspectos espaço-temporais devem ser envolvidos, relacionando a possibilidade de armazenamento, processamento e visualização.

A 2ª etapa refere-se ao Tratamento dos dados que envolve as adaptações e uniformizações dos dados para serem utilizados nas etapas a seguir. Essa etapa compreende de uma sequência de ações que começa com o subprocesso 2.1 que corresponde aos scripts de agrupamento onde é realizada a junção dos dados para a resolução temporal a ser trabalhada. A próxima etapa apresentada na Figura 15 corresponde ao subprocesso 2.2, que é Gerar a tabela com a informação geográfica. Este é um subprocesso que possui as etapas apresentadas no subprocesso 2.3 para sua realização, de modo geral, inclui a criação de um arquivo CSV com a latitude e longitude dos pontos, depois a produção de um *Shapefile* e por fim o armazenamento do *Shapefile* no SGBD com uma extensão para dados geográficos. Desse modo, a próxima etapa é o subprocesso 2.4 que corresponde a ação de gerar as series temporais e Wavelets dos dados. Até esta etapa da metodologia SWAmP, é gerado um banco de

dados com as tabelas das séries temporais, Wavelets juntamente com a informação geográfica. O subprocesso 2.5 trata da análise de similaridade, ou seja, essa etapa gera os scripts para realização da busca por semelhança das séries temporais e Wavelets implementadas no banco de dados. A etapa de Tratamento de dados é descrita de forma detalhada na secção 4.1.2.

O 3º processo da metodologia SWAmP corresponde a implementação da arquitetura. Neste processo, é necessário a implementação de um servidor *web*, juntamente com um servidor de mapas para a disponibilização dos mapas com os pontos trabalhados para visualização e *web*, outras tecnologias e linguagens de programação para *web* são necessárias para sua implantação, por exemplo, é necessário uma biblioteca específica para se trabalhar com mapas, e também é necessário que os scripts façam conexão com o banco de dados geográfico e com o servidor de mapas.

As próximas seções descrevem detalhadamente cada uma das etapas da metodologia SWAmP assim como as tecnologias, linguagens de programação utilizadas em cada fase.

4.1.1 Modelagem

O primeiro resultado da metodologia do SWAmP é o diagrama de caso de uso representado na Figura 16 apresenta as funcionalidades do SWAmP 1.0.

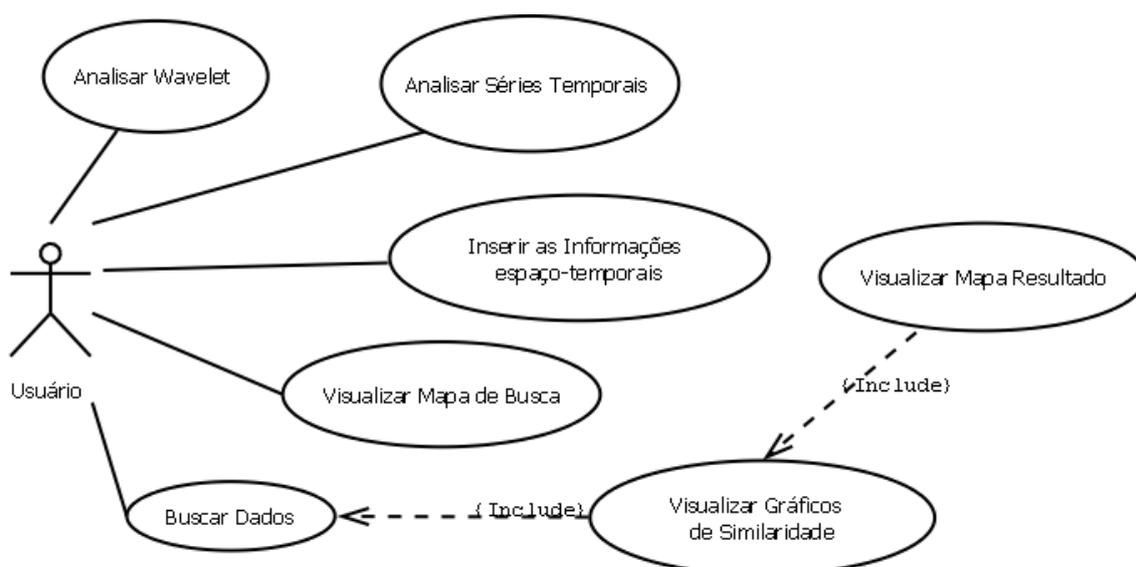


Figura 16: Diagrama de Caso de Uso.

O Quadro 2 descreve detalhadamente cada funcionalidade do SWAmP 1.0 inerente ao usuário que pode escolher entre realizar análise em séries temporais ou Wavelets.

Quadro 3: Funcionalidades do SWAmP 1.0 detalhadas.

Analisar Wavelet	O usuário faz análise de Wavelets.
Analisar Séries Temporais	O usuário escolhe realizar análise em séries temporais.
Inserir Informações Espaço-Temporais	O usuário deve escolher escala temporal que deseja realizar a detecção de padrões: semestral, anual ou 16 anos. E escolher entre os anos 1998 a 2013, seus semestres ou a análise de todos os anos em conjunto.
Visualizar Mapa de Busca	Um mapa é apresentado ao usuário com os pontos disponíveis para consulta.
Buscar Dados	O usuário preenche os dados ou utilizar o mapa e realizar a busca. Em seguida o sistema recebe a solicitação e efetua a pesquisa no banco de dados.
Visualizar Gráficos de Similaridade	O sistema possibilita a visualização dos gráficos de Wavelets em diferentes cores seguindo o grau de similaridade. Para análise de séries temporais, o usuário também pode visualizar os gráficos de resultados que são apresentados com cores diferentes seguindo o grau de similaridade.
Visualizar Mapa Resultado	Para a análise espacial é apresentado um mapa com o com os pontos nas mesmas cores das Wavelets e Séries Temporais, para que o

	usuário entenda qual ponto é de determinada informação.
--	---

A Figura 17 mostra a Tela para seleção dos dados de umidade relativa, nesta tela o usuário pode escolher entre clicar sobre o ponto azul no mapa ou digitar a Latitude e Longitude do ponto que deseja realizar a busca, e assim clicar sobre o botão buscar. Esta ação proporciona ao usuário maior flexibilidade, pois permite a digitação dos dados do ponto no formulário ou a seleção do ponto com mouse. A possibilidade de realizar a busca em diferentes periodicidades de tempo também é relevante para diferentes pesquisas, dessa forma, esse requisito também foi executado neste trabalho, sendo utilizados os espaços temporais: semestral, anual e 16 anos.

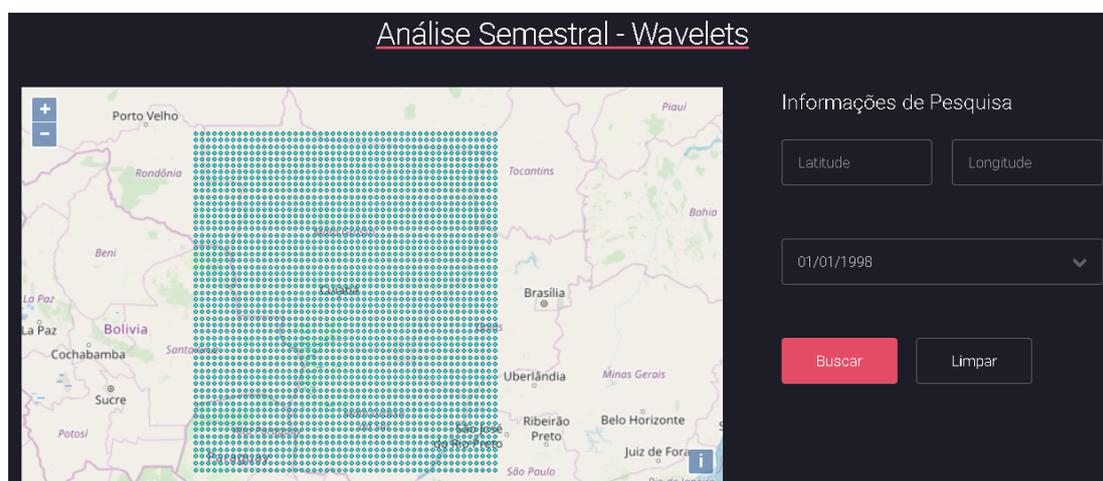


Figura 17: Tela para escolha dos pontos para visualização dos resultados das análises.

A necessidade de guardar os dados de uma busca é um requisito relevante, considerando a necessidade do usuário em trabalhar com esses dados em outros programas. Após realizar uma busca, o SWAmP 1.0 mostra como resultado, o gráfico (Wavelet ou Série Temporal) selecionado pelo usuário na cor VERDE, e mostra as Wavelets ou Séries Temporais similares pelo grau de similaridade em tons de azul:

- 1º Grau de similaridade: azul escuro.
- 2º Grau de similaridade: azul médio.
- 3º Grau de similaridade: azul claro.

Outra questão importante é a análise espacial. Para isso, o SWAmP 1.0 oferece visualização na mesma ordem de cores apresentadas nos gráficos, isso possibilita que o

usuário verifique geograficamente onde estão as Wavelets ou Séries Temporais estão dispostas. Como resultado, este capítulo também apresenta uma análise espacial dos pontos, que mostra geograficamente a distância média de todos os pontos no ano selecionado.

Oferece ainda como resultado, informações de Tempo, Latitude e Longitude em um Quadro, com a mesma disposição de cores para facilitar a análise de todas as informações. Após a modelagem das funcionalidades do SWAmP 1.0, que corresponde ao que o sistema deve ter, foi elaborado utilizando o modelo de negócios BPMN, as etapas de como o usuário deve utilizar o sistema. O esquema de utilização do SWAmP 1.0 é apresentada na Figura 18.

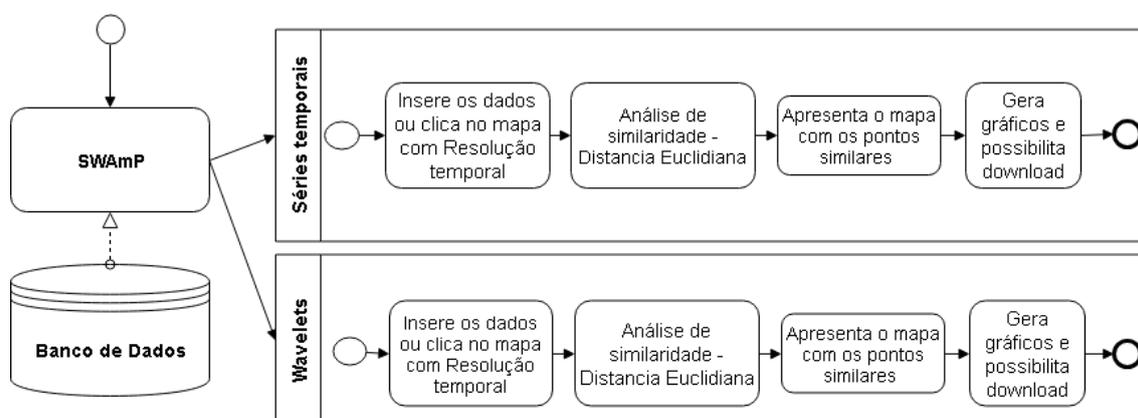


Figura 18: Diagrama de processo de funcionamento do SWAmP 1.0.

Como apresentado na Figura 18, a aplicação solicita ao Banco de dados as informações, e, o usuário pode escolher entre fazer a análise de similaridade de Wavelets ou Séries temporais. Ao escolher uma das opções o usuário pode inserir os dados no formulário com a Resolução Temporal (Semestral, Anual, 16 anos) e realizar a análise de similaridade.

Após buscar os dados, é apresentado o mapa com os pontos similares, os gráficos e as informações geográficas para análise, além da possibilidade do usuário realizar o *download* das informações solicitadas.

O diagrama BPMN possibilita a visualização de utilização do sistema. Após realizar a modelagem do sistema e moldar como deve ser sua utilização, a próxima etapa para o desenvolvimento da metodologia SWAmP para detecção de padrões é o tratamento dos dados.

4.1.2 Tratamento dos dados ambientais

No presente trabalho foi desenvolvido dez scripts com a linguagem de programação Python versão 3 para manipulação e tratamentos dos dados. A etapa de Tratamento de dados segue os parâmetros apresentados na Figura 15 da metodologia SWAmP, dessa forma a Figura 19 ilustra o diagrama BPMN que descreve de forma detalhada cada etapa.

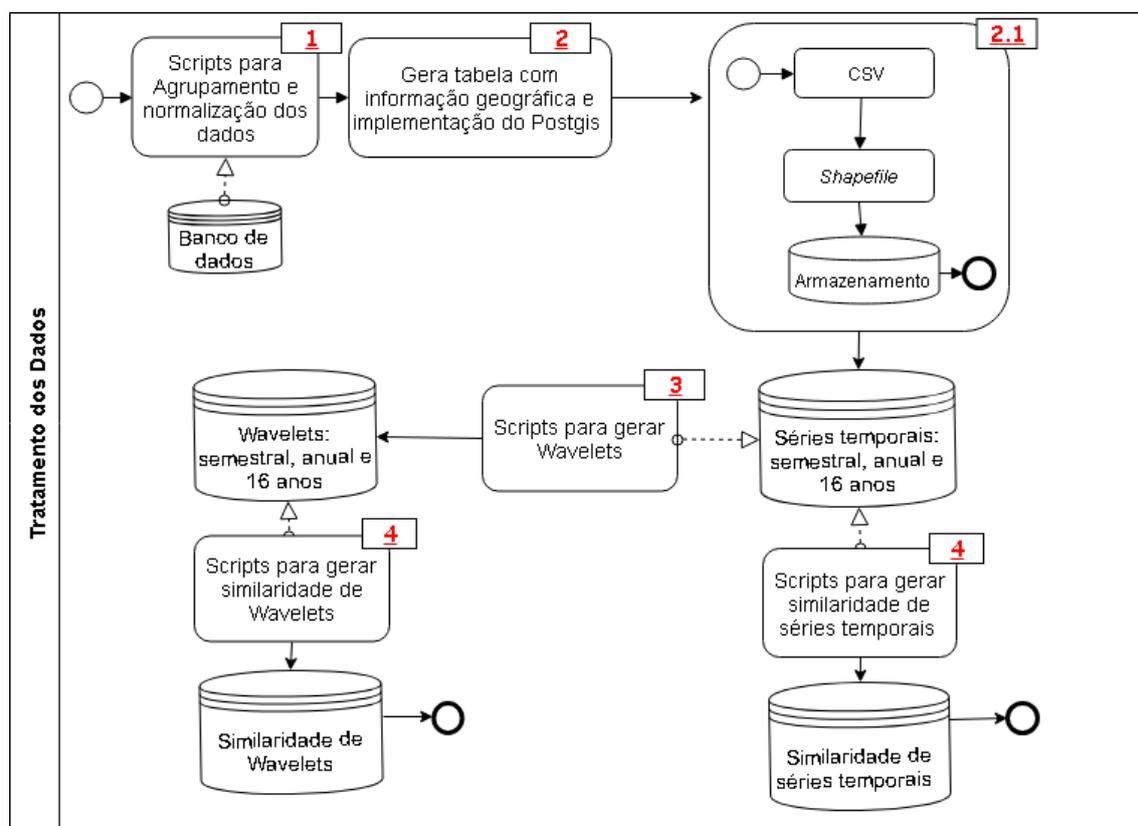


Figura 19: Diagrama de tratamento dos dados do SWAmP.

Sendo assim, a 1ª etapa foi realizar o agrupamento e normalização dos dados, conforme ilustrado no diagrama BPMN (Figura 19). Com os scripts elaborados em Python e nomeados: GeraSemestral.py, GeraAnual.py e GeraTotal.py, foram agrupados os dados ambientais respectivamente em semestral, anual e 16 anos, ao término de cada script, os dados agrupados foram armazenados no banco de dados nas tabelas: Semestral, Anual e Total (16 anos). Assim como os dados de umidade relativa e temperatura média do ar que utilizaram os mesmos scripts para a realização dos agrupamentos nas resoluções temporais: Semestral e Anual.

A 2ª etapa do tratamento dos dados ilustrada na Figura 19 pelo diagrama BPMN foi o processo para desenvolver a tabela chamada ponto, que contém os dados de latitude, longitude e a geometria. Esse passo foi necessário, facilitar o processo de busca e normalização dos dados.

Para gerar a tabela ponto com a informação geográfica (campo geometria) foi realizado um conjunto de ações apresentadas na Figura 19 que correspondem ao subprocesso 2.1.

As ações do subprocesso 2.1 baseiam-se na extração dos dados originais, no armazenamento desses dados em um banco de dados, criação do arquivo CSV¹⁵, que contém a informação da latitude e longitude de cada ponto, a geração do *shapefile*¹⁶ a partir do arquivo CSV e então a importação do *shapefile* no BDG. Para gerar o *shapefile* com a marcação dos pontos, para isso utilizou-se a ferramenta Quantum GIS¹⁷ que possibilita a conversão de dados em formato CSV (latitude, longitude) para *shapefile*. O *shapefile* gerado foi importado para BDG, com o comando Shp2pgsql que permite adicionar dados geográficos ao BDG.

Ao importar o *shapefile*, o PostGis cria automaticamente um atributo que contém a especificação geométrica do objeto, que indica o tipo da representação, se este é do tipo ponto, linha ou polígono, e atribui a ele uma coordenada espacial seguindo os padrões OGC. Após inserir o *shapefile* no BDG, é possível realizar quaisquer operações geométricas/geográficas na tabela no BDG.

A 3ª etapa do tratamento dos dados apresentada no diagrama BPMN (Figura 19) consiste na criação dos scripts e banco de dados de Wavelets. Para gerar as Wavelets foi

¹⁵ *Comma-Separated Values* (CSV) é um formato de arquivo que armazena dados tabelados, separados por um delimitador, geralmente vírgulas e a quebra de linha para separar os valores.

¹⁶ Formato de arquivo contendo dados geoespaciais vetoriais. Descreve geometria como pontos, linhas e polígonos, além dos dados relacionados aos componentes geométricos.

¹⁷ Aplicação de código aberto que proporciona visualização de dados, edição e recursos de análise em SIG. Disponível em <<http://www.qgisbrasil.org/>> Acesso 20 de setembro de 2016.

adaptado o script¹⁸ disponibilizado por Paos e Colorado, utilizando a linguagem de programação Python versão 3. Para realizar a conexão dos scripts que geram as Wavelets e base de dados (semestral, anual e 16 anos) foi utilizado a biblioteca Psycopg versão 2, juntamente com as bibliotecas: NumPy, Matplotlib e SciPy usadas para auxiliar nos cálculos e plots das Wavelets que são geradas a cada solicitação do usuário.

A Figura 20 apresenta o Diagrama do modelo relacional que mostra modelagem do Banco de Dados com as 3 primeiras etapas do tratamento dos dados apresentados, e com as tabelas originais de precipitação do TRMM, de umidade relativa e temperatura do INMET. É possível observar que neste momento foram criadas 3 tabelas com as periodicidades: Semestral, Anual e 16 anos. E como atributos dentro delas existem vetores¹⁹ chamados:

- SériesT: armazenam as séries temporais
- wavelet: armazenam as Wavelets
- Data: armazena a data do evento.

E a tabela Ponto é uma tabela que armazena a informação geográfica, possui:

- Id: um identificador.
- lat: um campo de Latitude.
- lon: um campo de Longitude.
- geom: um campo de geometria.

Após o agrupamento dos dados nas resoluções temporais: semestral, anual e 16 anos, foi implementado o reconhecimento de padrões de semelhança utilizando o cálculo de distância Euclidiana com os dados das séries temporais e Wavelets separadamente. Esse passo corresponde a 4ª etapa do tratamento dos dados da metodologia SWAmP implementada no SWAmP 1.0. Sendo calculada a distância de cada ponto em seu específico agrupamento (semestral, anual, 16 anos) com todos outros pontos de mesmo agrupamento (séries temporais e Wavelets).

¹⁸ paos.colorado.edu/research/wavelets/software.html

¹⁹ É um conjunto de variáveis do mesmo tipo acessíveis com um único nome. Armazenadas de forma contínua e ocupando as posições de forma fixas.

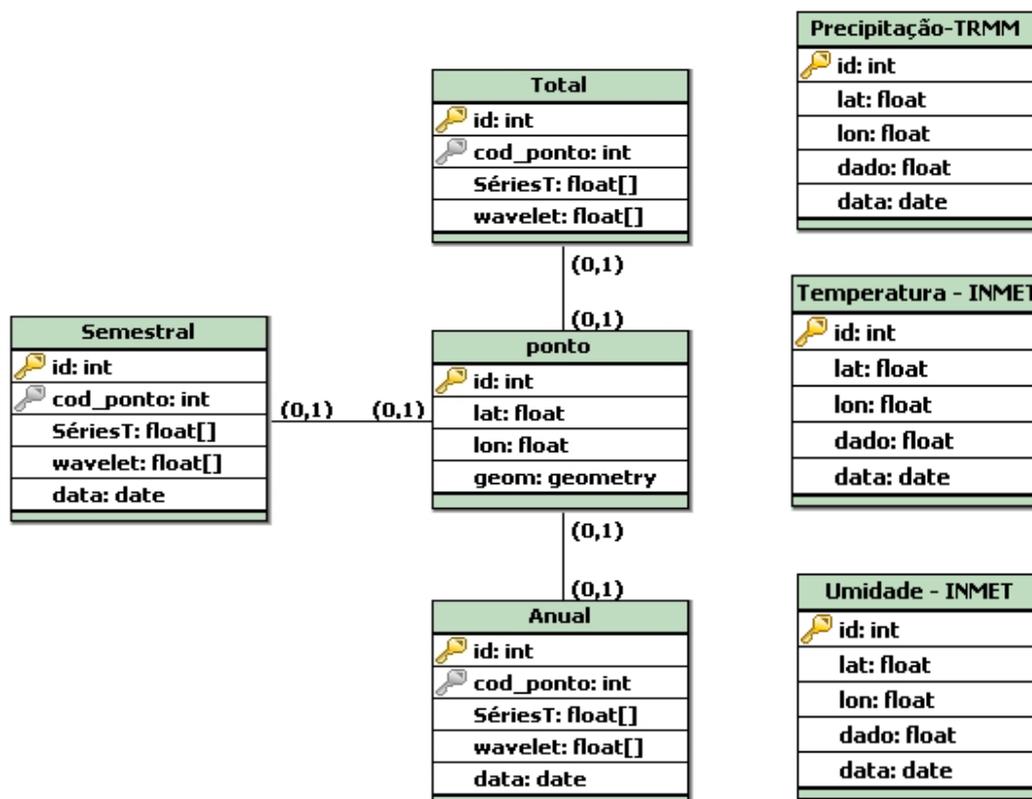


Figura 20: Tabelas com os agrupamentos: semestral, anual e 16 anos de séries temporais e Wavelets. Tabela Ponto.

Na etapa 4, foram criados os scripts e utilizado o cálculo de distância Euclidiana para análise de similaridade das séries temporais originais nominados:

- GeraDistanciaSemestreS.py
- GeraDistanciaAnualS.py
- GeraDistanciaTotalS.py

Para calcular a similaridade entre as Wavelets foram desenvolvidos e utilizados os seguintes scripts:

- GeraDistanciaSemestreW.py
- GeraDistanciaAnualW.py
- GeraDistanciaTotalW.py

Ao se ter o resultado das semelhanças nas periodicidades semestral, anual e 16 anos, foram selecionados os trezentos pontos com maior similaridade que o usuário poderia selecionar para ser analisado e armazenado no banco de dados, ou seja, para cada ponto armazenado no banco de dados existem trezentos similares para os dados de

precipitação, porém para os dados de umidade relativa e temperatura média do ar foram armazenadas as 10 primeiras, pelo fato de possuir apenas 28 estações.

A Figura 21 ilustra o diagrama do modelo relacional das tabelas criadas, e seus relacionamentos.

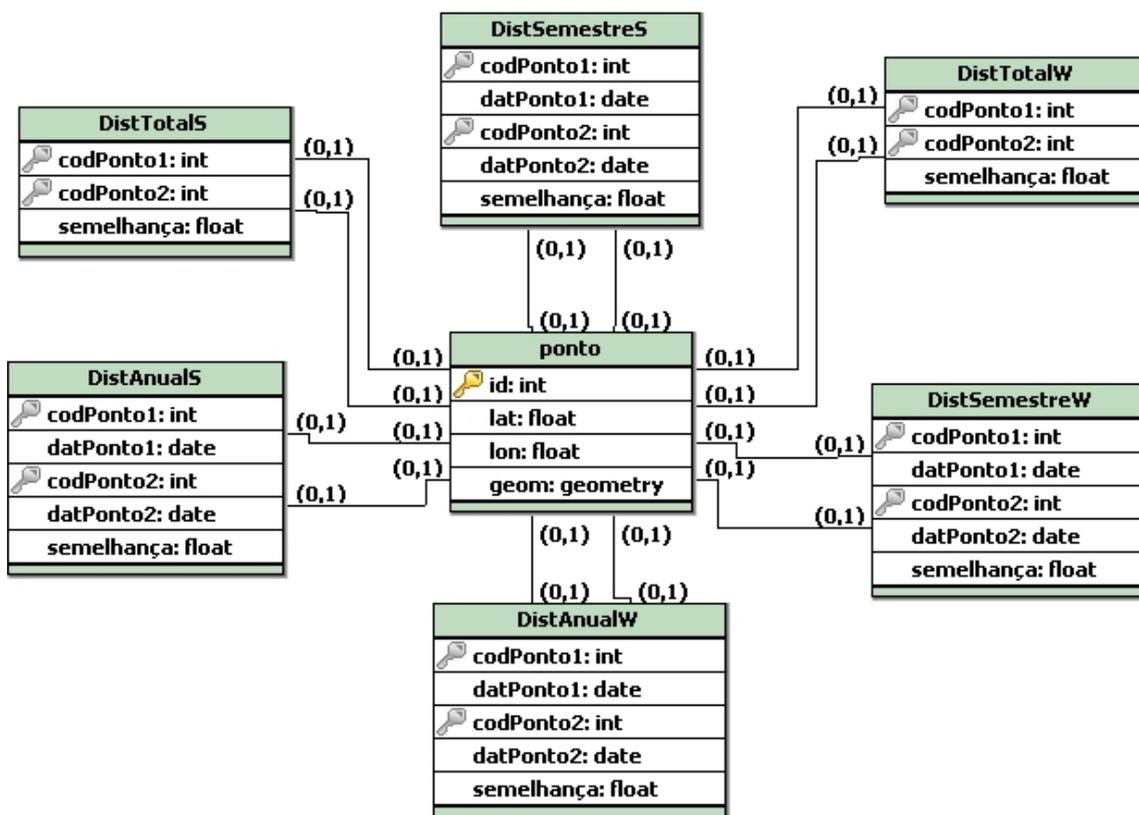


Figura 21: Modelo relacional das tabelas de similaridades do SWAmP 1.0.

As tabelas de similaridades estão todas conectadas com a Tabela Ponto, para que não aconteça de maneira alguma perda da informação geográfica. Os atributos apresentados da Figura 21, nominados:

- semelhança: float²⁰ - armazena o valor de similaridade.
- codPonto1:int²¹ - armazena o código do ponto analisado.

²⁰ Tipo de dado para números de ponto flutuante, um número que tem um ponto decimal.

²¹ Tipo de dado para números inteiros.

- `datPonto1:date` - armazena a data correspondente ao ponto analisado.
- `codPonto2:int` - armazena o código do ponto semelhante.
- `datPonto2:date` - armazena a data correspondente ao ponto semelhante.

O tratamento e armazenamento dos dados ambientais da metodologia SWAmP possibilitam o manejo dos dados em um SGBD, porém, a visualização dos resultados é possível no sistema SWAmP 1.0.

4.2 Análise de Precipitação com o SWAmP 1.0

Foram realizados testes em diversos pontos dentro da grade de 2597 pontos de precipitação, para verificar a similaridade e a acurácia dos resultados da metodologia que culminou na implementação do sistema SWAmP 1.0. Visto que mostrar o resultado de todos os pontos em todos os semestres e anos (janeiro de 1998 a dezembro 2013) neste trabalho não é viável devido ao volume de pontos, datas e dados foi selecionado um ponto aleatório com data aleatória, e aqui apresentado o resultado de Wavelets e Séries Temporais de precipitação.

4.2.1 Análise Espacial Semestral - Séries Temporais Originais

A análise espacial realizada na discussão dos resultados deste trabalho oferece ao leitor uma visão geográfica de dos pontos utilizados para apresentação do SWAmP 1.0.

O primeiro estudo espacial aqui apresentado trata da análise espacial em Séries Temporais com resolução temporal semestral, sendo utilizado o primeiro semestre de 1998 nos dados de precipitação conforme exemplificado no Quadro 3 do Capítulo 3.

Os pontos estão em um espaço que contem 2597 pontos, para cada ponto foi analisado seus três pontos semelhantes espacialmente para verificar qual a média em quilômetros a distancia selecionado. Para análise do primeiro semestre de 1998, apresentada na Figura 22, verifica-se que os pontos tem uma variação de distância entre 114 quilômetros a 1643 quilômetros para o ponto com similaridade de quilômetros média mais distante, porém em sua maioria os pontos estão a uma distância a baixo de 300 quilômetros.

série temporal observa-se que o comportamento é semelhante, pelo fato, de que quando no eixo X ocorre a precipitação na série temporal selecionada pelo usuário, na série temporal similar também ocorre em volume menor, porém na mesma escala de dias.

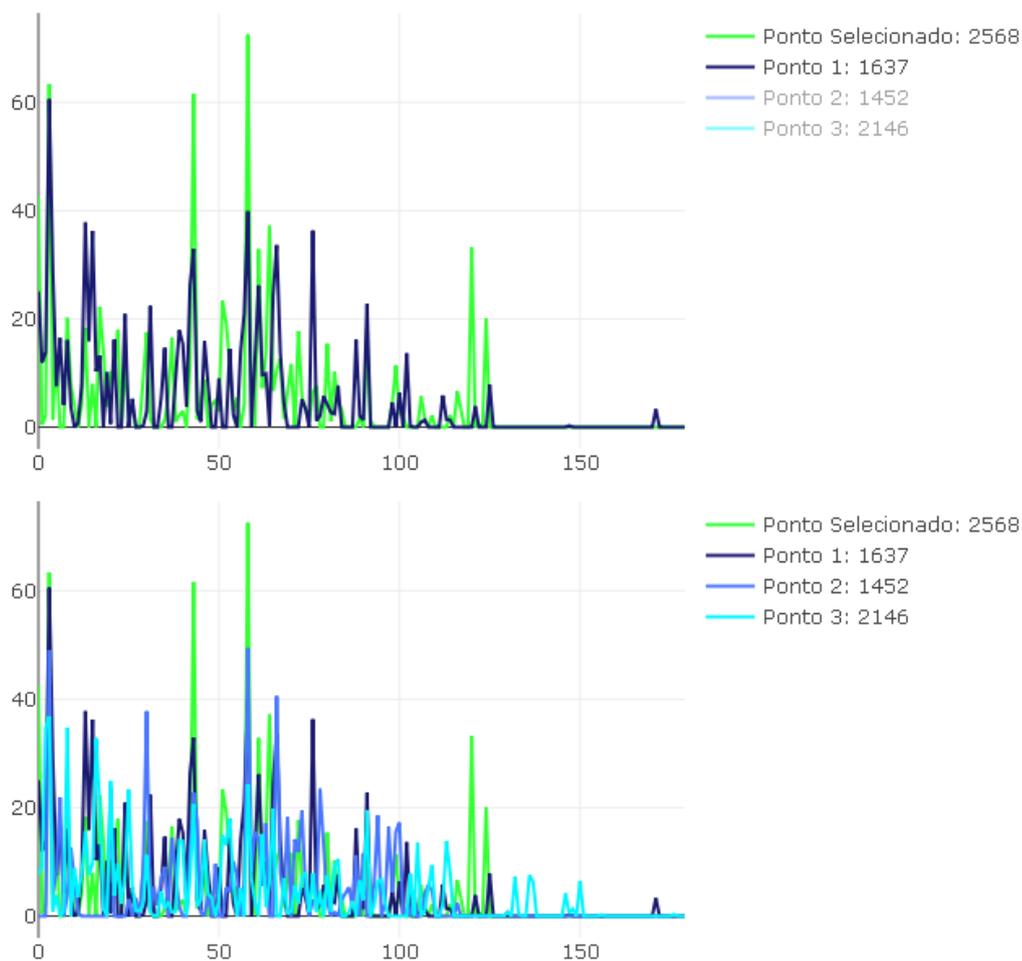


Figura 23: Séries Temporais do Ponto 2568, primeiro semestre 1998; Ponto 1637 primeiro semestre 1999 (Primeira semelhante); Ponto 1452, primeiro semestre 2006 (Segunda Semelhante); Ponto 2146, primeiro semestre 2001 (Terceira semelhante).

A Figura 24 traz o Quadro que o SWAmP 1.0 fornece ao usuário, que mostra que os semestres semelhantes ao primeiro semestre de 1998:

- O primeiro ponto semelhante é do primeiro semestre de 1999.
- O segundo ponto semelhante é do primeiro semestre de 2006.
- O terceiro ponto semelhante é do primeiro semestre de 2001.

Codigo do Ponto	Data	Latitude	Longitude
2568	01-01-1998	-9.125000000000000	-49.875000000000000
1637	01-01-1999	-13.875000000000000	-55.625000000000000
1452	01-01-2006	-14.875000000000000	-49.875000000000000
2146	01-01-2001	-11.375000000000000	-56.375000000000000

Figura 24: Quadro com as informações dos pontos semelhantes ao ponto selecionado pelo usuário.

O mapa apresentado na Figura 25 mostra os pontos semelhantes para visualização do usuário, e a distancia desses pontos é de: 754.2 quilômetros para Azul escuro; 636.0 quilômetros para Azul médio; 818.1 quilômetros para Azul claro.

Observa-se que o ponto com maior semelhança (Azul escuro), está no bioma Amazônia diferentemente dos outros dois pontos semelhantes que se encontram no bioma Cerrado como o ponto selecionado. Os regimes de chuvas no Brasil depende de uma sequência de fatores, um deles pode ser considerado o tipo de bioma, neste caso o ponto com maior semelhança encontra-se em um bioma vizinho ao ponto selecionado.

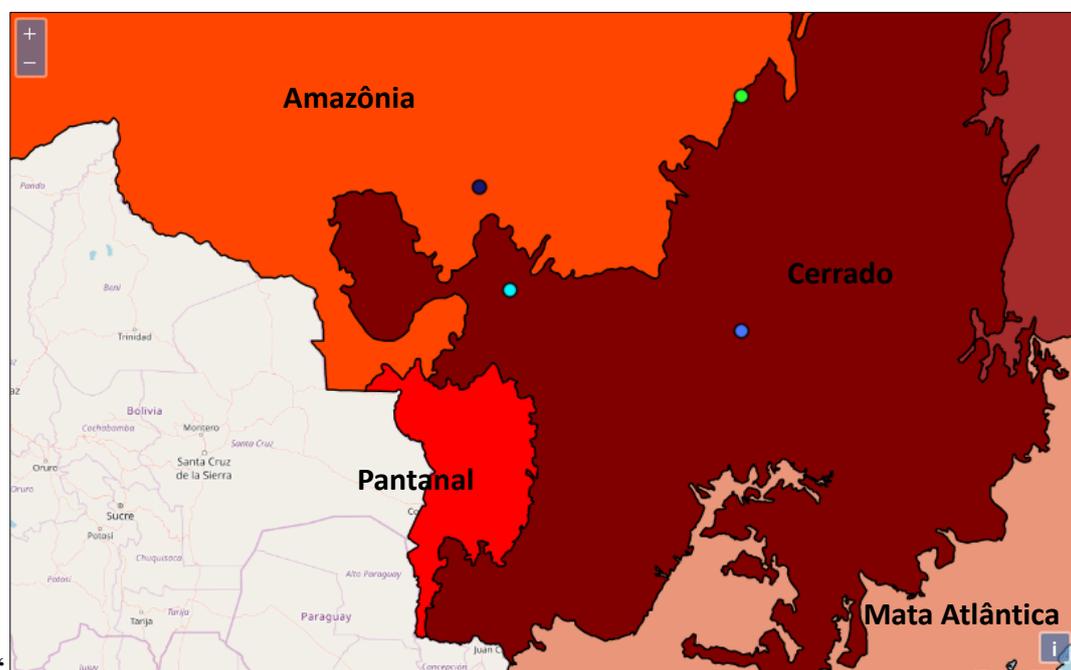


Figura 25: Mapa do SWAMP 1.0 com os pontos semelhantes. Ponto 2568(VERDE); Ponto 1637(AZUL ESCURO); Ponto 1452(AZUL MÉDIO); Ponto 2146(AZUL CLARO).

O ponto A faz parte do grupo de pontos que tem séries temporais semelhantes com distancias maiores de 300 quilômetros.

4.2.3 Análise Espacial Anual - Wavelets

A análise espacial realizada com resolução temporal anual segue os mesmos conceitos da análise espacial efetuada com resolução temporal semestral. O Quadro 3, mostra que para a realização da análise anual apresentada nos resultados e discussões deste trabalho se utilizou o ano de 1998 escolhido de forma aleatória, para tanto, a análise geográfica dos pontos semelhantes espacialmente apresenta o mesmo ano.

A Figura 26 ilustra o mapa com a média dos quilômetros dos três pontos semelhantes do ano de 1998 de todos os pontos da grade de dados utilizada neste trabalho. Observa-se ao analisar o mapa, que grande parte dos pontos possuem uma média de quilômetros próximos. Isso significa que em sua maioria os pontos semelhantes estão a uma distância em torno de 200 quilômetros. É possível observar também que existem pontos com dados semelhantes em locais com distancias em torno de 800 quilômetros, porém no caso da resolução temporal anual, e no ano de 1998, não houve muitos pontos.

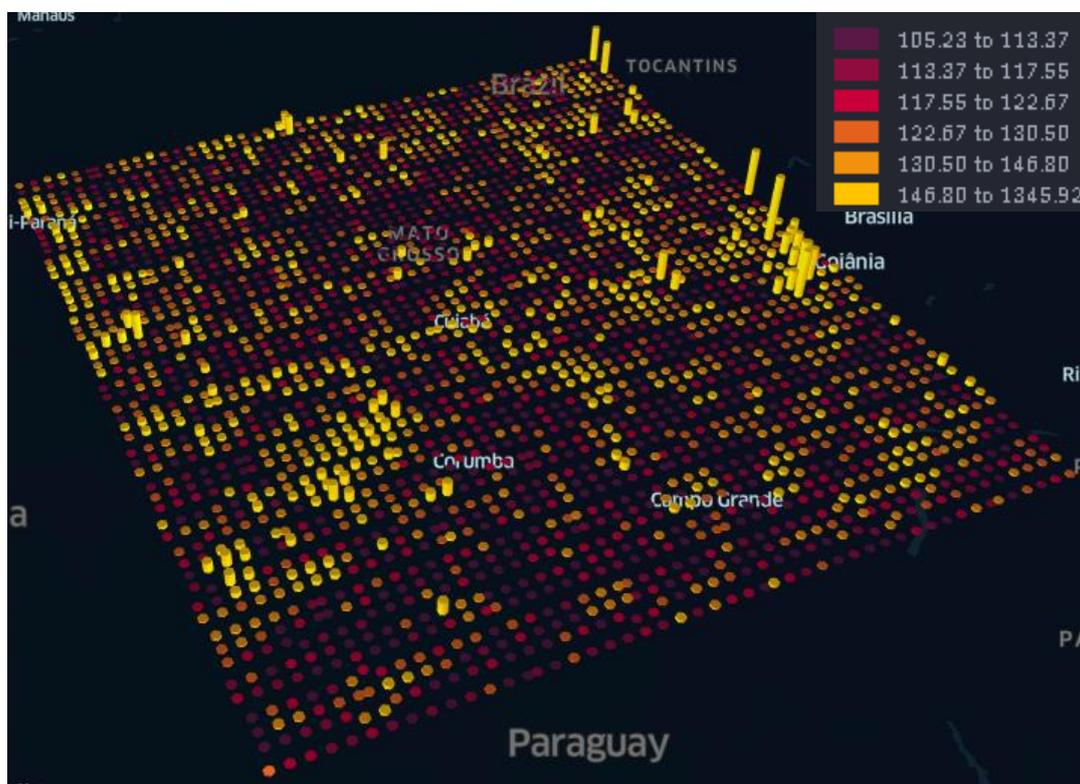


Figura 26: Mapa de distância média em quilômetros de cada ponto, do ano 1998 das Wavelets.

Após a realização da análise espacial do mapa que permite uma visualização geral de como os dados estão dispostos espacialmente em questão de distância

geográfica, a próxima etapa é a análise que o SWAmP 1.0 oferece ao usuário ilustrando os resultados dos pontos.

4.2.4 Ponto A - Análise Anual - Wavelets

Conforme explicitado no início deste capítulo, foi realizada a análise de similaridade do espectro de energia das Wavelets, cujo ponto e data escolhidos aleatoriamente foram (-9.125, -49.875) do período do ano de 1998 referente ao ponto para os dados de precipitação do projeto TRMM. A Figura 27 apresenta o resultado da busca por padrões em Wavelet aplicado ao ponto selecionado pelo usuário. Com o resultado da busca identifica-se que este ponto tem uma Wavelet com alta intensidade de energia na escala temporal entre os dias 50 e 70 no eixo X ambas com frequência de 14 dias no eixo Y. O Período de seca também é bem definido pelo ciclo de baixa intensidade de energia no dia 150 a 280 para todas Wavelets.

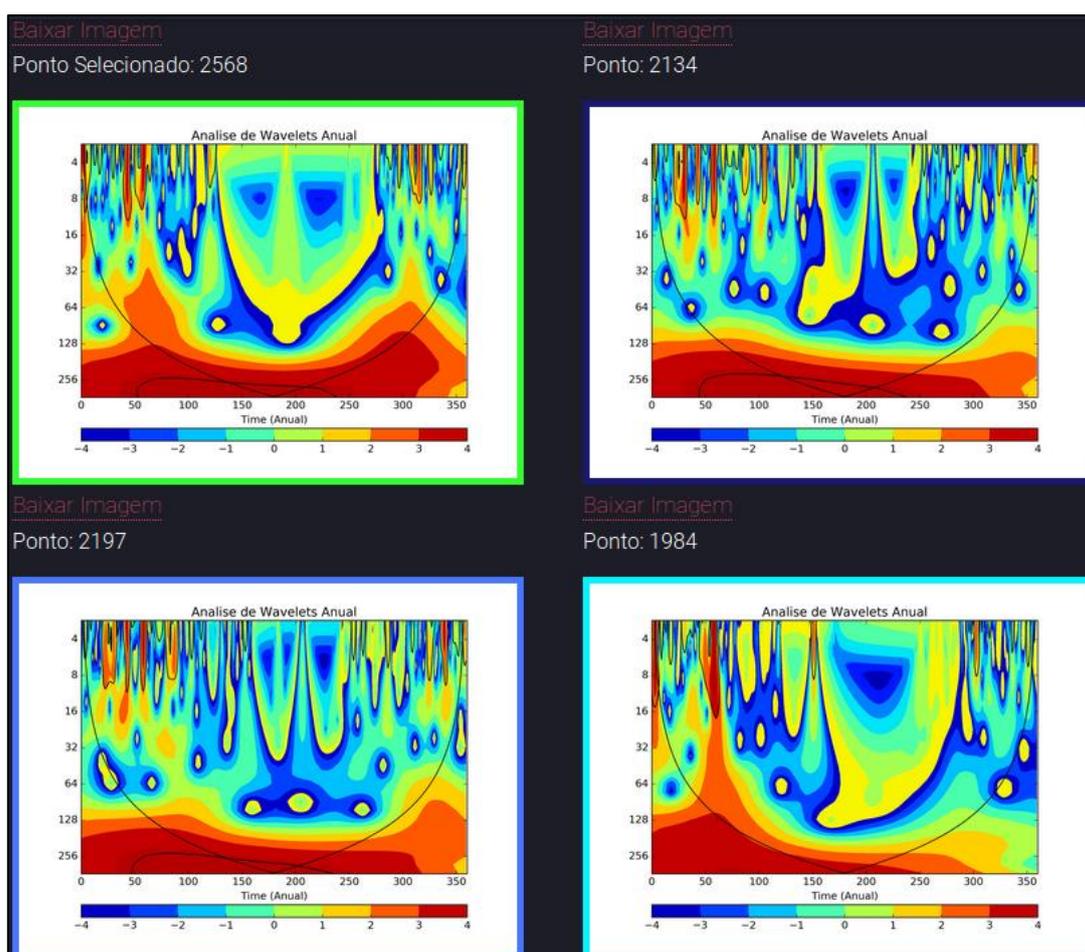


Figura 27: Wavelets semelhantes do Ponto 2568, ano de 1998; Ponto 2134, ano de 2001, (primeiro semelhante); Ponto 2197, ano de 2001, (segundo semelhante); Ponto 1984, ano de 2007, (terceiro semelhante).

A Figura 28 apresenta o Quadro que traz as informações de latitude e longitude, além de qual ano foi semelhante ao ano de 1998, que neste caso os anos semelhantes foram:

- O primeiro ponto semelhante é do ano de 2001.
- O segundo ponto semelhante é do ano de 2001.
- O terceiro ponto semelhante é do ano de 2007.

Codigo do Ponto	Data	Latitude	Longitude
2568	01-01-1998	-9.125000000000000	-49.875000000000000
2134	01-01-2001	-11.375000000000000	-59.375000000000000
2197	01-01-2001	-11.125000000000000	-59.375000000000000
1984	01-01-2007	-12.125000000000000	-50.375000000000000

Figura 28: Quadro com as informações dos pontos semelhantes ao ponto selecionado pelo usuário.

A Figura 27 ilustra o mapa com a relação dos pontos para que o usuário visualize onde estão seus resultados. A distância do ponto selecionado (Verde) para os pontos semelhantes: Azul escuro é de 859.7 quilômetros; Azul médio é de 839.8 quilômetros; e Azul claro é de 336.3 quilômetros. É possível observar que os pontos semelhantes apesar da distância estão localizados em biomas que possuem características semelhantes ao ponto analisado.

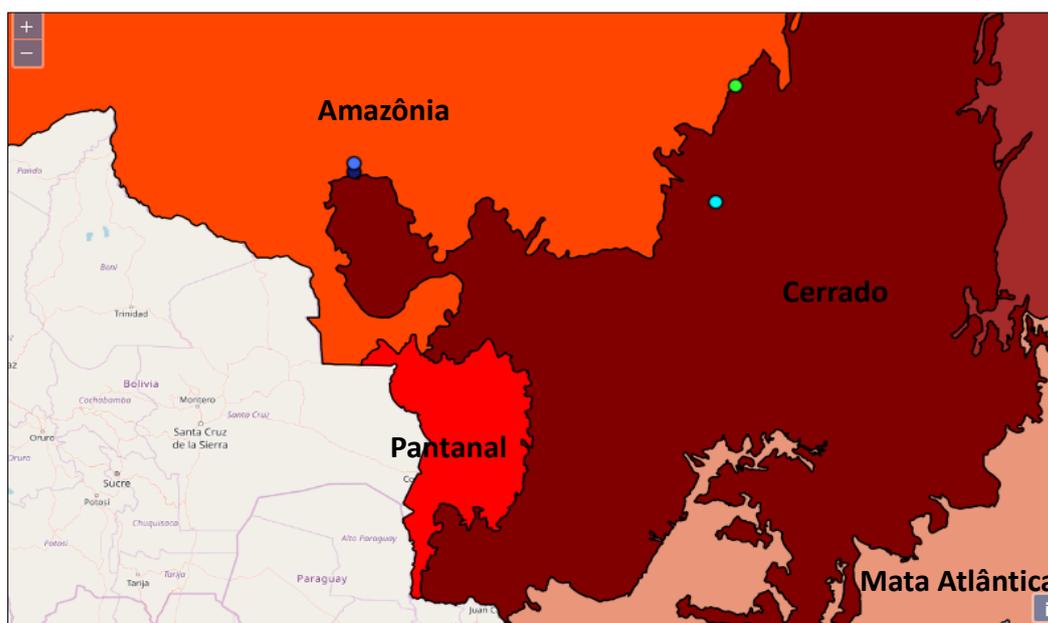


Figura 29: Mapa apresentado pelo SWaMP 1.0. Ponto selecionado pelo usuário (VERDE): ano 1998; Primeiro ponto semelhante (AZUL ESCURO): ano 2001; Segundo ponto semelhante (AZUL MÉDIO): ano 2001; Terceiro ponto semelhante (AZUL CLARO): ano 2007.

Os resultados apresentados neste trabalho servem como base para trabalhos futuros, onde é possível realizar outras buscas no SWAmP 1.0 e inferir também dados de outras variáveis ambientais para completa-las. O sistema aqui apresentado possibilita ao usuário uma grande possibilidade de análises e tomadas de decisões

4.2.5 Análise Espacial - 16 anos - Wavelets

A análise espacial realizada com a resolução temporal de 16 anos utiliza os dados de quilômetros dos três pontos semelhantes de todos os pontos da grade de dados utilizada neste trabalho. Todos os pontos se encontram em uma média de 200 quilômetros do ponto selecionado pelo usuário. Na Figura 30 é possível verificar a distribuição dos pontos no mapa, que varia de 104 quilômetros a 253.72 para o ponto com dados mais distantes do ponto selecionado pelo usuário, ou seja, neste caso os pontos estão relativamente próximos.

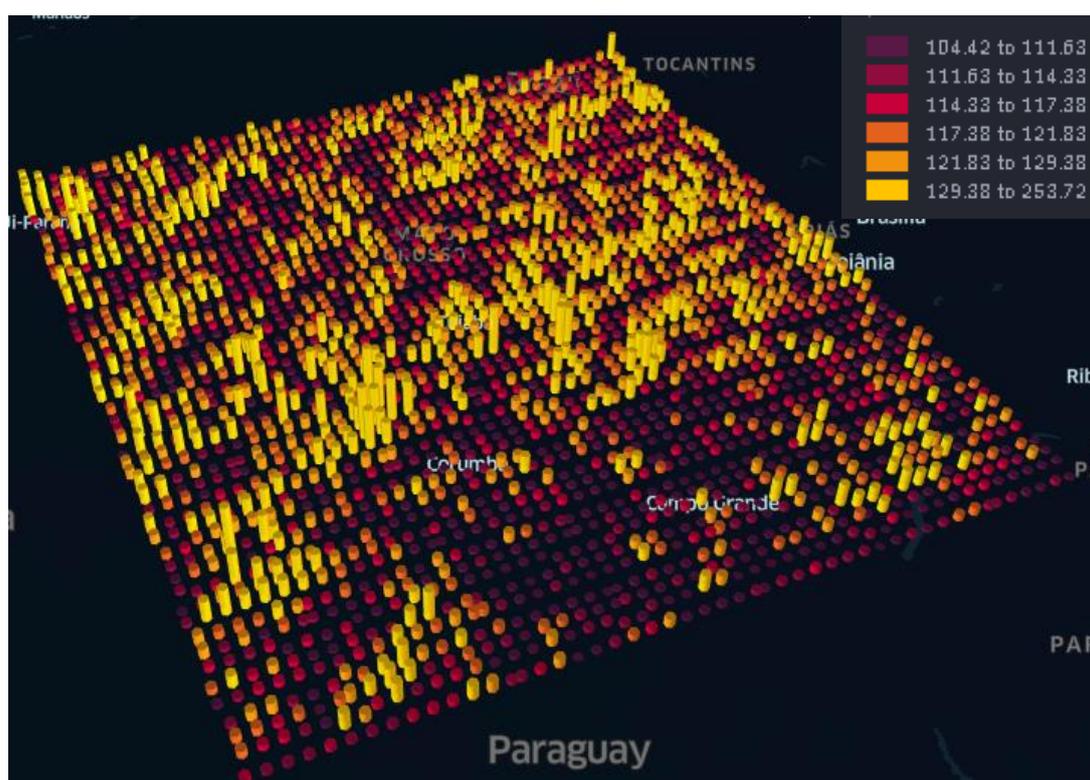


Figura 30: Mapa de distância média em quilômetros, de 16 anos (janeiro de 1998 a dezembro 2013) nas Wavelets.

Após a análise espacial, que apresenta a média geral em quilômetros de todos os pontos semelhantes na resolução temporal de 16 anos, foi selecionado para apresentação do resultado de precipitação um ponto com resolução temporal com 16 anos de dados.

4.2.6 Ponto A - Análise 16 anos - Wavelets

O resultado do ponto A com coordenadas: (-9.125, -49.875) com resolução temporal de 16 anos do período de 1998 a 2013 é ilustrado pela Figura 31 que revela que no eixo X existe um domínio temporal do período dos dias de 500 a 1000, de 1000 a 2000, de 3500 a 4000, de 4000 a 5000, sendo analisado no eixo Y correspondente a uma intensidade de energia com frequência de 360 dias para esses períodos.

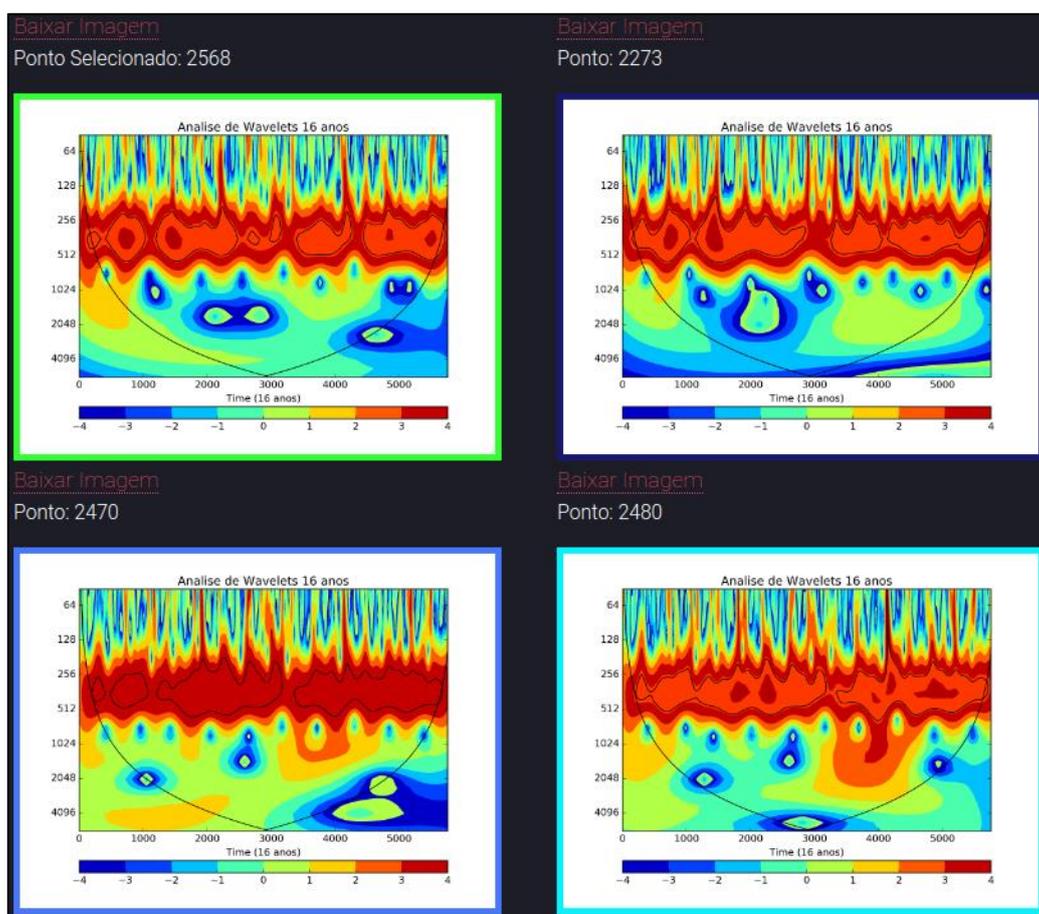


Figura 31: Wavelets semelhantes do Ponto 2568; Ponto 2273; Ponto 2470; Ponto 2480.

O mapa ilustrado pela Figura 32 traz os pontos na mesma sequência de cores, para que o usuário possa visualizar no SWAmP 1.0 onde estão seus resultados, e para que o usuário possa utilizar suas informações em outros sistemas, o SWAmP 1.0 possibilita a análise em um quadro na Figura 33, onde encontra-se a latitude e longitude do ponto selecionado e dos pontos resultantes. A distância do ponto selecionado (Verde) para os pontos semelhantes é de:

- 185.18 quilômetros para o ponto Azul escuro;

- 173.80 quilômetros para o ponto Azul médio;
- 148.02 quilômetros para o ponto Azul claro.

Para a verificação do resultado da análise de 16 anos apresentada pelo mapa da Figura 32, observa-se que pelo conjunto conter um grande volume de dados (16 anos de dados diários), os pontos estão relativamente próximos, mostrando que quanto maior o volume de dado, mais próximo estarão os pontos e logicamente no mesmo Bioma ou em Bioma vizinho como ilustrado pela Figura 32.

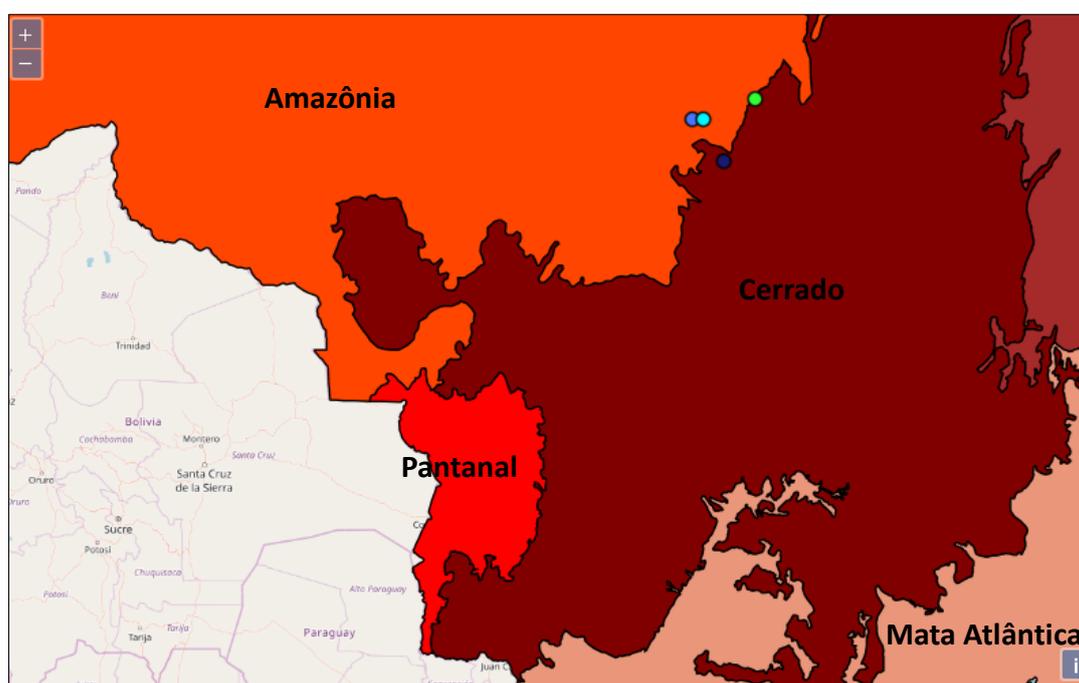


Figura 32: Mapa do SWAmP 1.0 com os pontos semelhantes. Ponto selecionado pelo usuário (VERDE); Primeiro ponto semelhante (AZUL ESCURO); Segundo ponto semelhante (AZUL MÉDIO); Terceiro ponto semelhante (AZUL CLARO).

Latitude	Longitude
-9.125000000000000	-49.875000000000000
-10.625000000000000	-50.625000000000000
-9.625000000000000	-51.375000000000000
-9.625000000000000	-51.125000000000000

Figura 33: Quadro com as informações dos pontos semelhantes ao ponto selecionado pelo usuário.

4.3 Análise de Umidade Relativa com o SWAmP 1.0

Este trabalho também apresenta os resultados da metodologia SWAmP que culminou com o sistema SWAmP 1.0. Esta seção traz o resultado do sistema para os dados de umidade relativa do ponto (-18.98, -56.619999999999997), que foi realizado comparações com os 28 pontos da grade do período de 1998 a 2013 com análise semestral e anual.

4.3.1 Análise Espacial Semestral - Séries Temporais

O uso de análises espaciais como a apresenta Figura 34 é importante para representar espacialmente qual a distância média dos pontos semelhantes em quilômetros selecionados pelo usuário no ano analisado, neste caso, o ano de 2004.

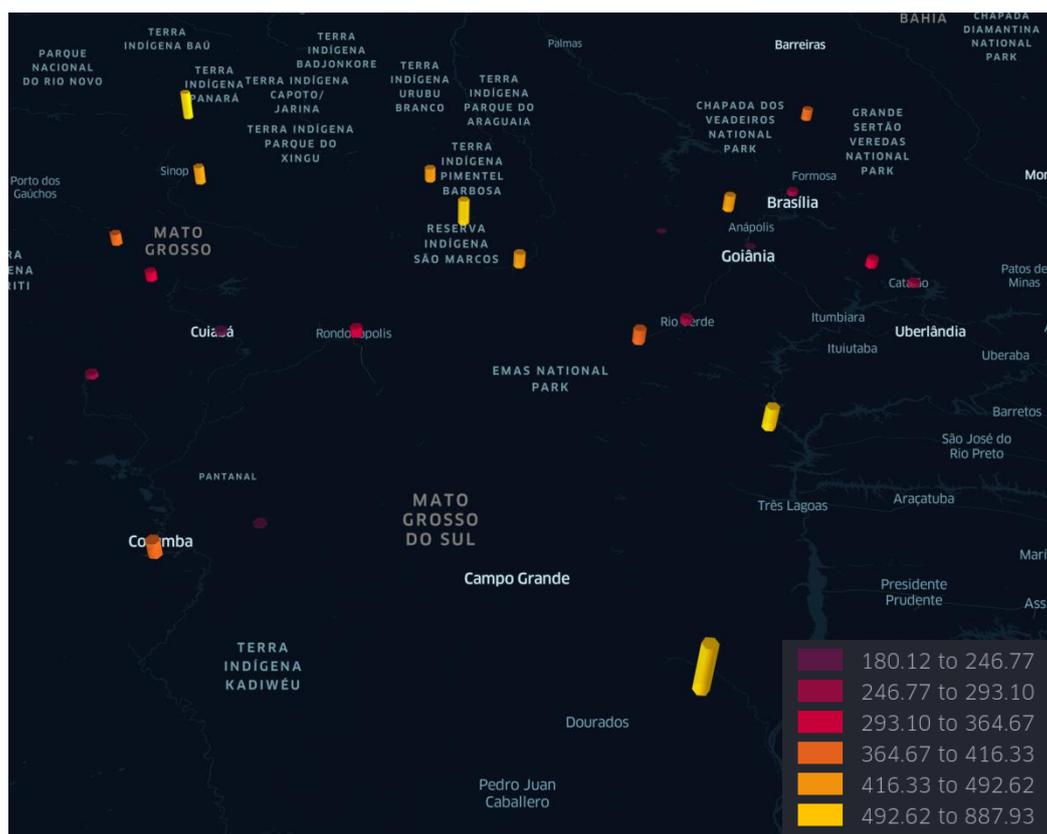


Figura 34: Mapa de distância média em quilômetros, do segundo semestre de 2004 das Séries Temporais.

Com a visualização ilustrada pela Figura 34, observa-se que os pontos que possuem as três primeiras séries temporais semelhantes estão a uma distância média em quilômetros entre 180km a 492km, havendo alguns pontos que possuem séries temporais semelhantes que estão a uma distância de até 887km. Isso significa que pode

ser encontrado pontos que podem estar em biomas diferente, por exemplo, ou podem ser estudados em conjunto com outro dado.

4.3.2 Ponto B - Análise Semestral - Séries Temporais

A análise do comportamento da umidade relativa é importante, uma vez que possibilita detectar tendências ou alterações no clima, em escalas local, regional ou global. A Figura 35, mostra que o segundo semestre do ano de 2004 a umidade relativa média, esteve em torno de 60% a 90% do mês de julho a setembro aproximadamente, e, e depois caiu para 40% em todas as séries temporais no período de aproximadamente um mês, e até o final do ano todas as séries temporais tiveram um comportamento semelhante entre 60% a 85%.

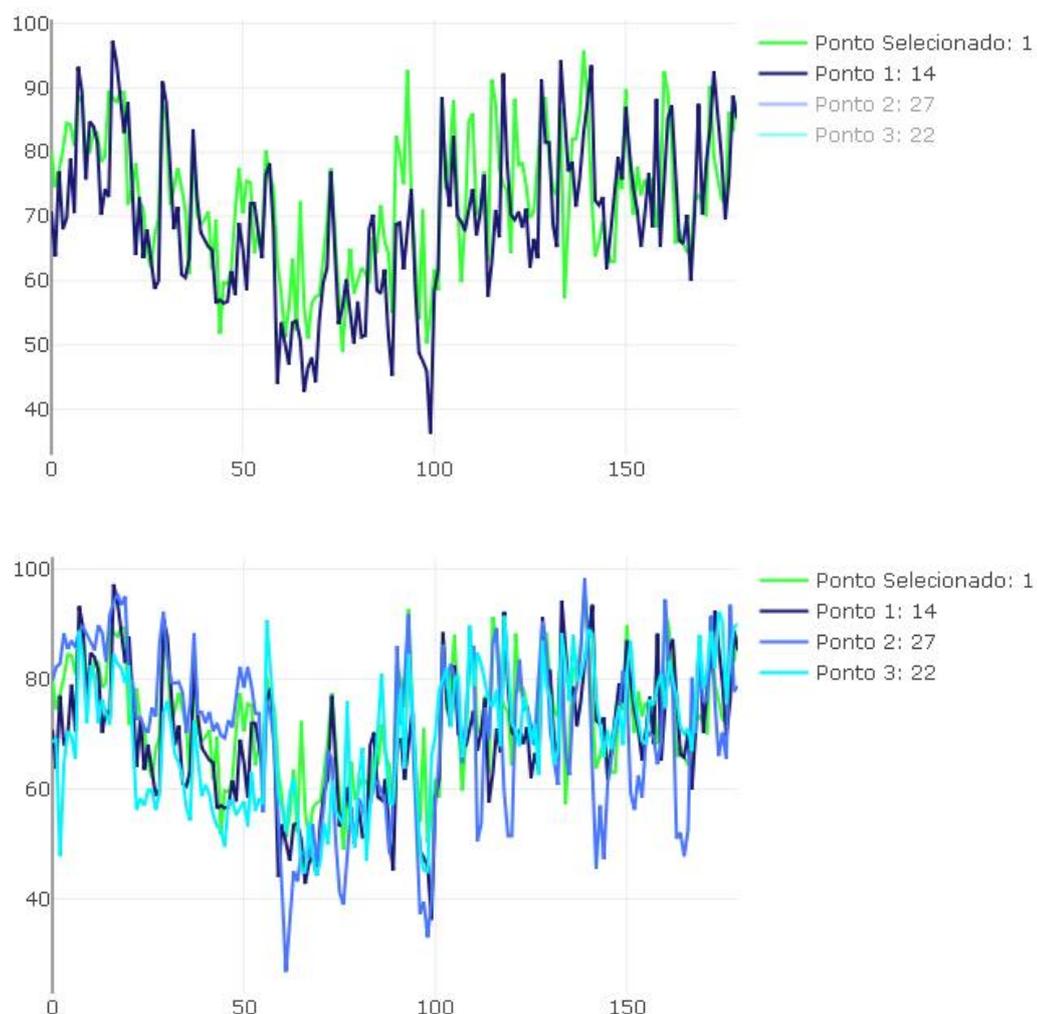


Figura 35: Séries temporais semelhantes. Ponto 1 selecionado pelo usuário, segundo semestre de 2004; Ponto 14 (primeiro ponto semelhante), segundo semestre de 2004; Ponto 27 (segundo ponto semelhante), segundo semestre de 2004; Ponto 22 (Terceiro ponto semelhante), segundo semestre de 2004.

. A Figura 36, traz o Quadro com as informações dos pontos, a distância dos pontos selecionados é de 110.6 para o azul médio, 359.1 para o azul escuro, 110.6 para azul claro. A Figura 37 traz o mapa com os biomas onde é possível observar que o ponto selecionado está no bioma pantanal e um ponto semelhante está no mesmo bioma e os outros dois estão no bioma cerrado

Codigo do Ponto	Data	Latitude	Longitude
1	01-07-2004	-18.980000000000000	-56.619999999999997
14	01-07-2004	-15.779999999999999	-56.060000000000002
27	01-07-2004	-19.020000000000000	-57.670000000000002
22	01-07-2004	-14.400000000000000	-56.450000000000003

Figura 36: Quadro com as informações dos pontos semelhantes ao ponto selecionado pelo usuário.

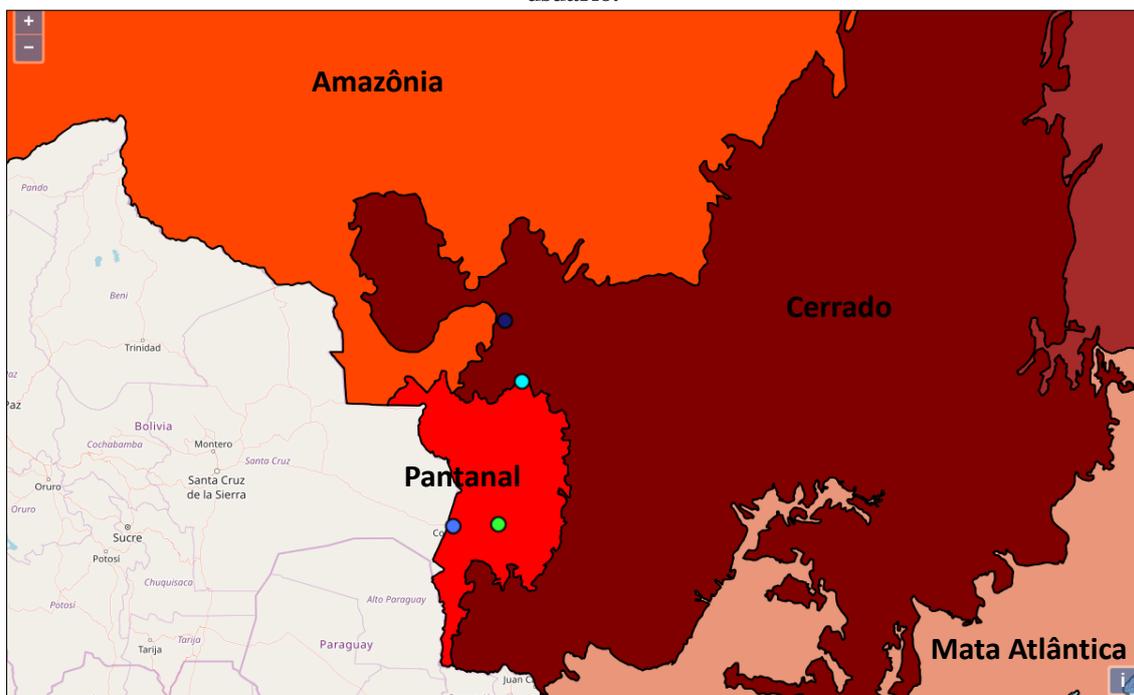


Figura 37: Mapa com os pontos das Séries temporais semelhantes. Ponto selecionado pelo usuário (VERDE): segundo semestre de 2004; Primeiro ponto semelhante (AZUL ESCURO): segundo semestre de 2004; Segundo ponto semelhante (AZUL MÉDIO): segundo semestre de 2004; Terceiro ponto semelhante (AZUL CLARO): segundo semestre de 2004.

4.3.3 Análise Espacial Anual - Wavelets

A análise espacial com resolução temporal anual utilizou os dados de quilômetros dos três pontos semelhantes de séries temporais de umidade relativa de cada ponto da grade dados utilizada neste trabalho do ano de 2004 conforme explicitado no Quadro 3 do Capítulo 3 no Capítulo 3.

A Figura 38 ilustra o mapa com os pontos, que possuem séries temporais semelhantes que estão distribuídos na resolução temporal anual, do ano de 2004, em sua maioria com distâncias médias de 200 a 300 quilômetros, com pontos nos extremos do mapa que possuem distancias maiores variando entre 359 a 645 quilômetros. Isso significa que as séries temporais semelhantes tem um padrão de similaridade geográfica com dados próximos.

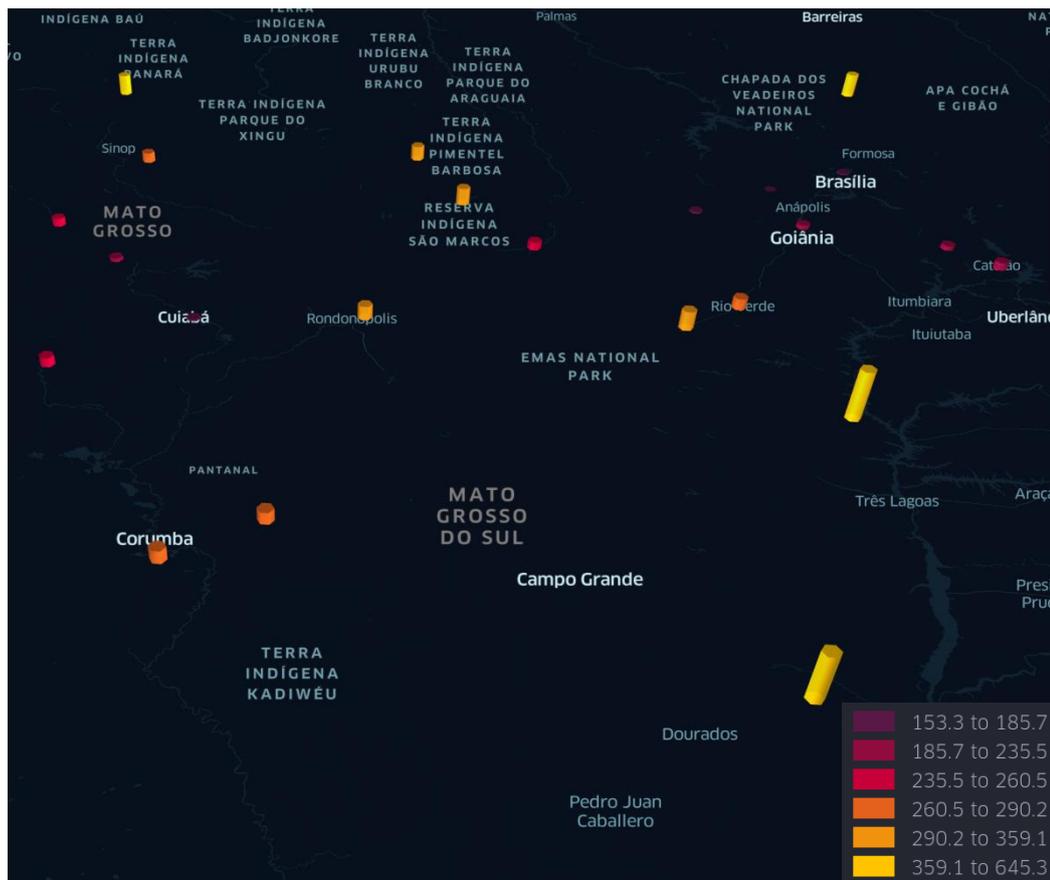


Figura 38: Mapa de distância média em quilômetros, ano de 2004 nas Wavelets de umidade relativa.

4.3.4 Ponto B - Análise Anual – Wavelets

O segundo resultado corresponde ao ponto B com coordenadas: (-18.98, -56.619999999999997) do ano de 2004. A Figura 39 apresenta o resultado da similaridade das Wavelets. Para tanto, em ambas as Wavelets, é possível observar a periodicidade de 14 dias do período que corresponde aos dias no eixo X de alta intensidade de energia: 250 e 300. Isso significa que neste período houve um índice alto de percentual de umidade relativa. Observando os ciclos apresentados pelas Wavelets na Figura 39 fica claro que a umidade relativa média nos pontos demonstrados, evidenciam

a sazonalidade bem definida dos valores de umidade relativa média do dia 1 ao dia 365 do ano de 2004.

A Figura 40 ilustra o Quadro que o SWAmP 1.0 traz para o usuário analisar quais são os anos semelhantes ao ano que ele escolheu para ser analisado em seu ponto, neste caso, o ano semelhante foi o 2004 para todos os pontos semelhantes.

Para finalizar a análise, o SWAmP 1.0 apresenta o mapa com a ilustração dos pontos para que o usuário verifique onde estão, de forma interativa, seus resultados na Figura 41. A distância do ponto selecionado (Verde) para os pontos semelhantes é de:

- 359.2 quilômetros para o ponto Azul escuro;
- 343.8 quilômetros para o ponto Azul médio;
- 110.7 quilômetros para o ponto Azul claro.

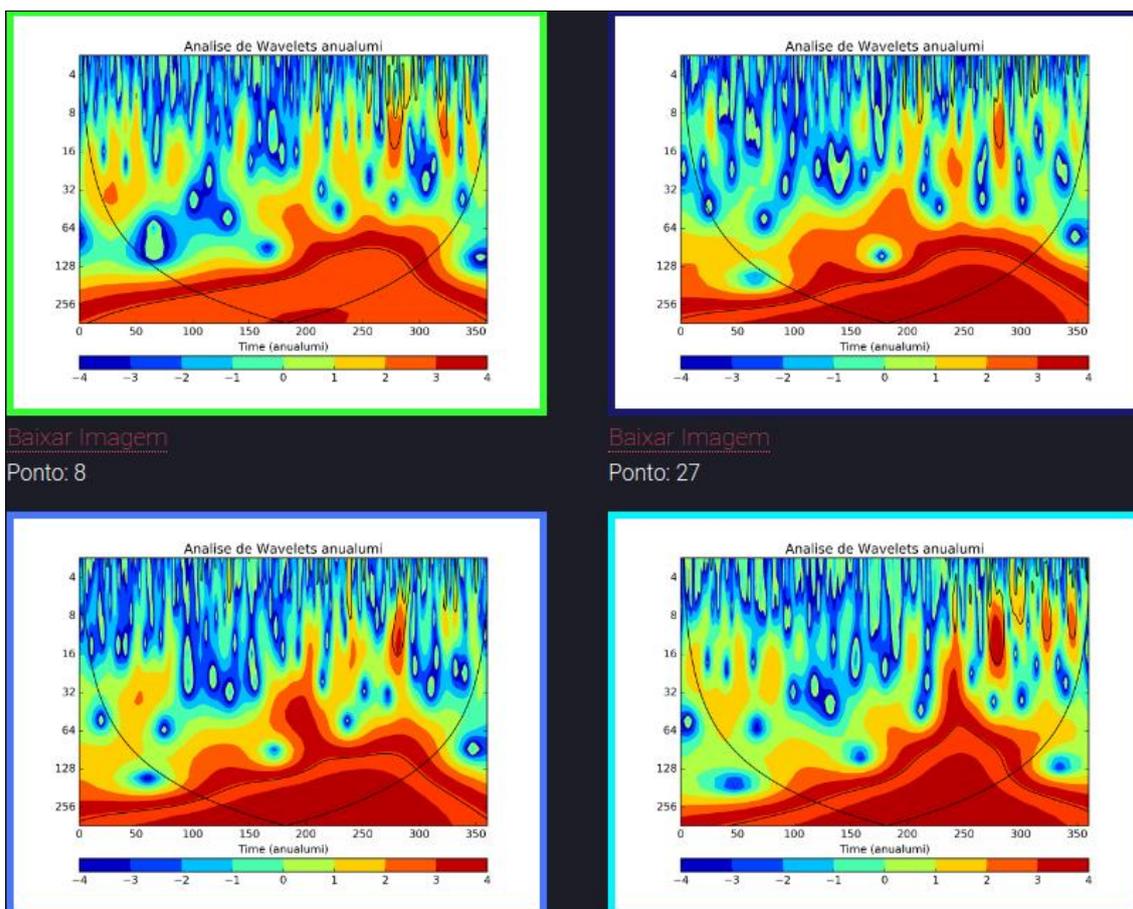


Figura 39: Wavelets do Ponto 1, ano de 2004; Ponto 14, ano de 2004; Ponto 8, ano de 2004; Ponto 27, ano de 2004;

Codigo do Ponto	Data	Latitude	Longitude
1	01-01-2004	-18.980000000000000	-56.619999999999997
14	01-01-2004	-15.779999999999999	-56.060000000000002
8	01-01-2004	-16.050000000000001	-57.680000000000000
27	01-01-2004	-19.020000000000000	-57.670000000000002

Figura 40: Quadro com as informações dos pontos semelhantes das Wavelets de umidade relativa.

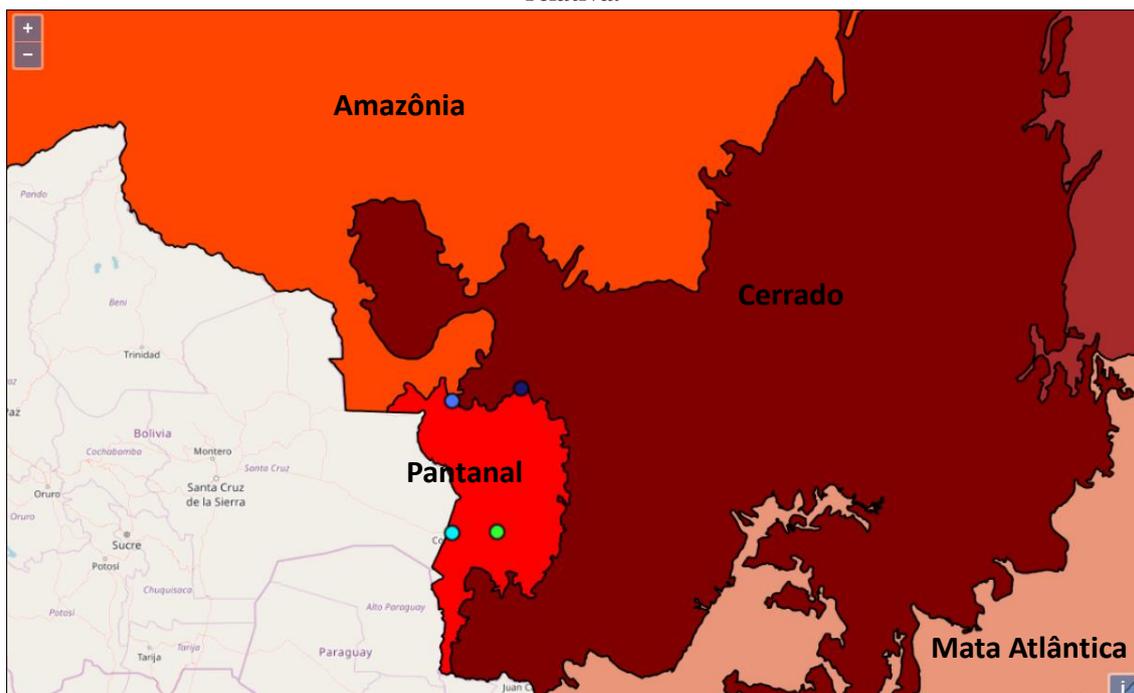


Figura 41: Imagem do SWAmP 1.0 com os pontos semelhantes. Ponto selecionado pelo usuário (VERDE): ano 2004; Primeiro ponto semelhante (AZUL ESCURO): ano 2004; Segundo ponto semelhante (AZUL MÉDIO): ano 2004; Terceiro ponto semelhante (AZUL CLARO): ano 2004.

4.4 Análise de Temperatura média do ar com o SWAmP 1.0

Foi também selecionado um ponto aleatório para apresentar aqui o resultado do SWAmP 1.0 para os pontos com dados de temperatura média do ar do ano de 2011 com análise semestral e anual.

4.4.1 Análise Espacial Anual - Séries Temporais

Para a análise espacial com resolução temporal anual com as séries temporais originais de temperatura média do ar, utilizou-se os dados de quilômetros dos três pontos semelhantes, de cada ponto da grade dados utilizadas neste trabalho do ano de 2011.

A Figura 42 ilustra o mapa com os pontos, onde é possível verificar que os pontos estão distribuídos em sua maioria a uma distância média de 360 quilômetros, os pontos que possuem distâncias superiores chegam a 800 quilômetros estão nos extremo superior do mapa. Isso significa que as séries temporais originais de temperatura média do ar tem sua característica para este ano, ter dados semelhantes a uma distância média abaixo de 400 quilômetros, possuindo apenas pontos nos extremos do mapa que tem séries temporais de temperatura média do ar com distâncias de aproximadamente 800 quilômetros.

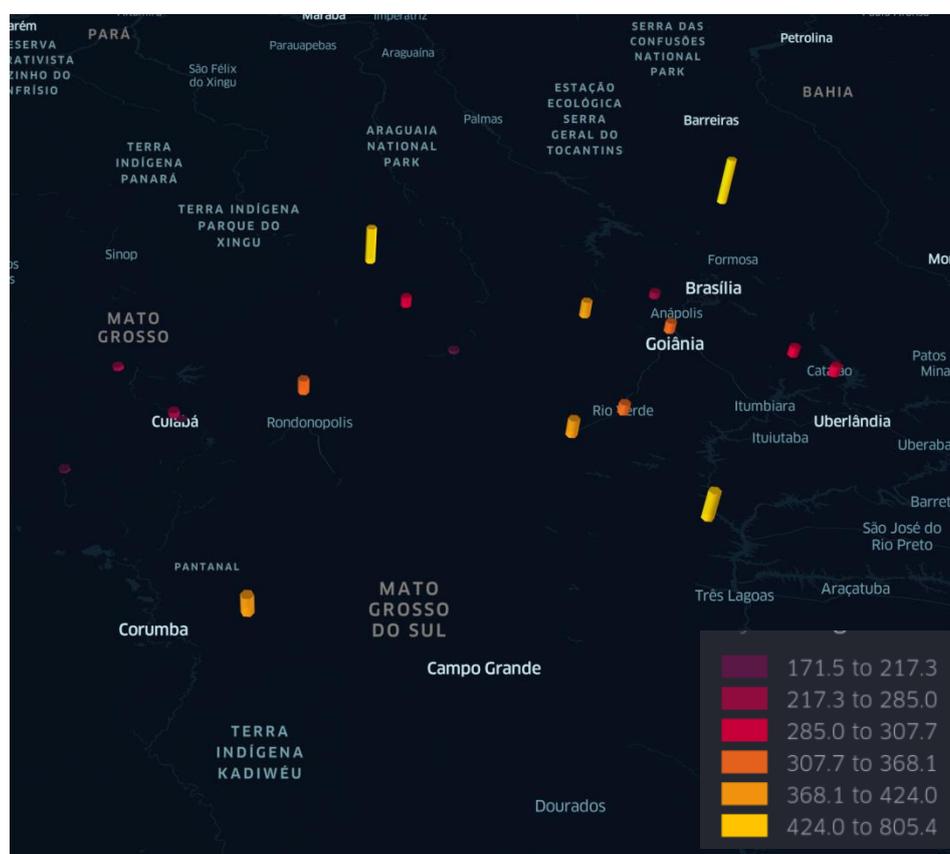


Figura 42: Mapa de distância média em quilômetros de cada ponto, do ano 2011 das Séries Temporais.

4.4.2 Ponto C - Análise Anual - Séries Temporais

O resultado correspondente as coordenadas (-14.40, -56.450000000000003) é apresentado na Figura 43 cujo ano escolhido de forma aleatória foi o de 2011. Observando as Figuras, as séries temporais apresentam comportamento, desempenho temporal e tendência comum entre elas, que fica evidente ao verificar o eixo X e Y, onde todas as séries temporais de temperatura média do ar

mostram que a média de temperatura média do ar para os locais apresentados pelo mapa fica em torno de 27° a 33° do início do ano até o dia 100, após esse período existe uma queda de temperatura média do ar, que aproximadamente no meio do ano de 2011 chega a registrar 15° graus para as séries temporais, após esse período, a temperatura média do ar se eleva novamente para a média de 27° a 33° até o final do ano.

A Figura 44 ilustra os resultados da análise temporal na resolução anual, indicando que o ano semelhante ao ponto selecionado é 2011 para todos os pontos similares. A ilustração referente a latitude e longitude dos pontos no mapa são apresentados na Figura 45, para que o usuário possa analisar de forma visual onde estão localizados os pontos obtidos em sua busca. A distância do ponto selecionado (Verde) para os pontos semelhantes: Azul escuro é de 273 quilômetros; Azul médio é de 158.6 quilômetros; e Azul claro é de 225.8 quilômetros. Todos os pontos estão localizados no bioma Cerrado.

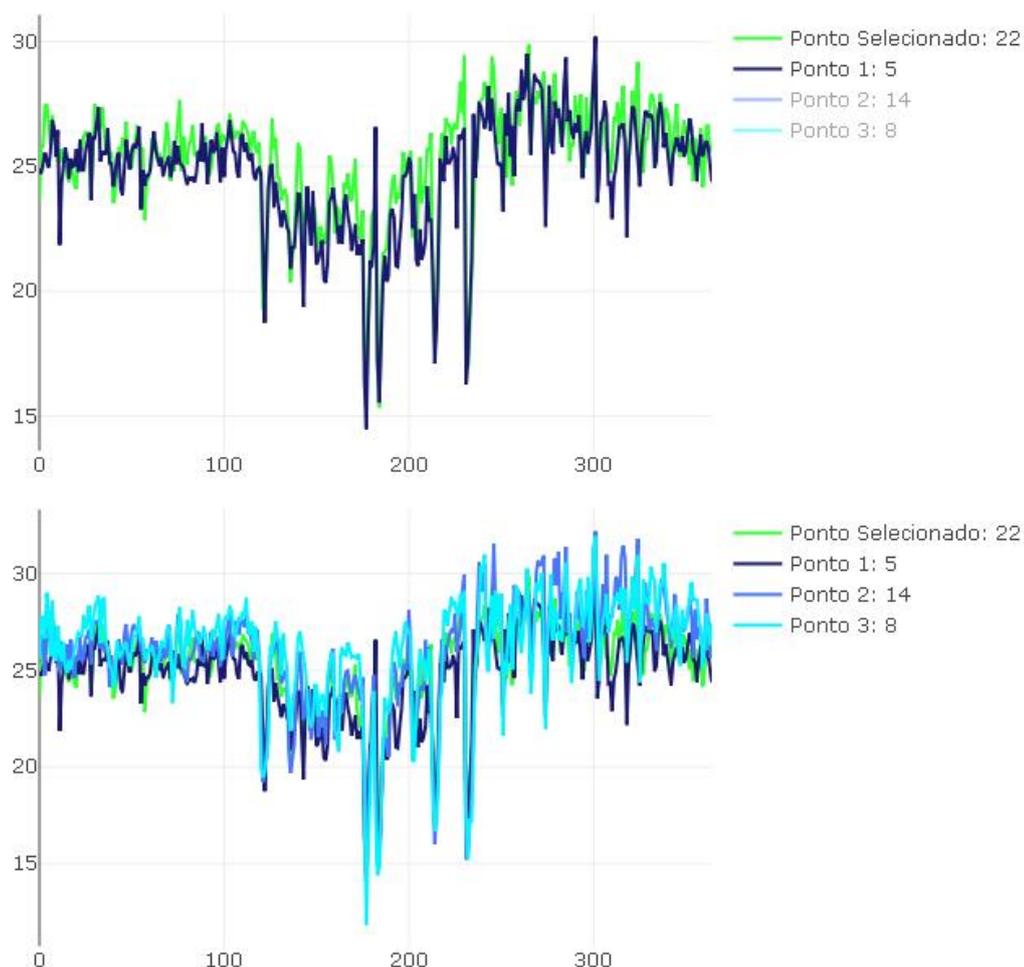


Figura 43: Séries Temporais do Ponto 22, ano de 2011; Ponto 5, ano de 2011; Ponto 14, ano de 2011; Ponto 8, ano de 2011;

Codigo do Ponto	Data	Latitude	Longitude
22	01-01-2011	-14.400000000000000	-56.450000000000003
5	01-01-2011	-15.830000000000000	-54.380000000000003
14	01-01-2011	-15.779999999999999	-56.060000000000002
8	01-01-2011	-16.050000000000001	-57.680000000000000

Figura 44: Quadro com as informações dos pontos semelhantes de temperatura média do ar ano 2011

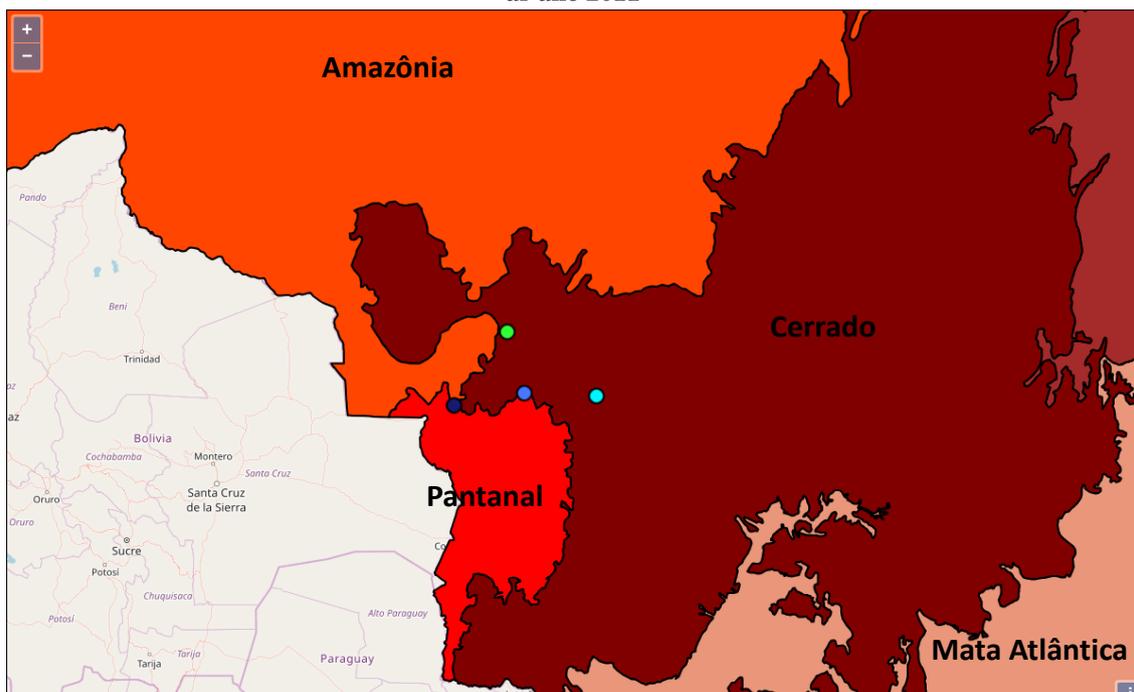


Figura 45: Ponto selecionado pelo usuário (VERDE): ano 2011; Primeiro ponto semelhante (AZUL ESCURO): ano 2011; Segundo ponto semelhante (AZUL MÉDIO): ano 2011; Terceiro ponto semelhante (AZUL CLARO): ano 2011.

4.4.3 Análise Espacial Semestral - Wavelets

A análise espacial em séries temporais com resolução temporal semestral de Wavelets de temperatura média do ar do primeiro semestre de 2011 segue os mesmos preceitos das análises espaciais anteriores com a média dos quilômetros dos três pontos semelhantes de todos os pontos da grade de dados de temperatura média do ar de 28 pontos, ilustrado pela Figura 46, que mostra que a maioria dos pontos possui séries temporais semelhantes a uma distância média de 158 a 499 quilômetros.

A análise de onde estão estes dados no mapa, ilustrada pela Figura 46 caracterizando em ordem de grandezas de cores onde estão os pontos com séries temporais mais distantes do ponto selecionado. Neste mapa, observa-se que os pontos estão distribuídos no mapa e todos corroboram a distância média de 300 quilômetros,

os pontos estão no mesmo bioma (Cerrado) e estão a uma distância de 158.3 km para o ponto claro, 139 km para o ponto azul médio, 273 km para o ponto azul escuro.

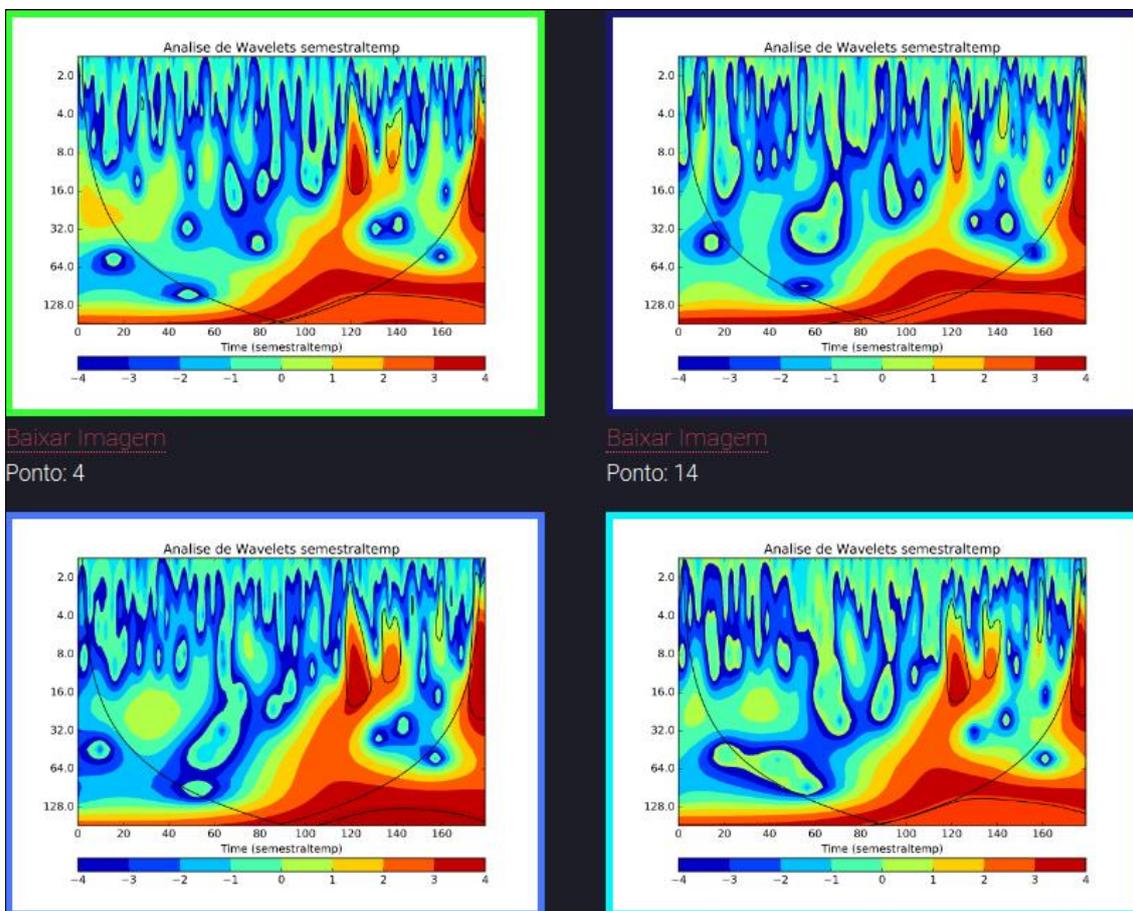


Figura 47: Wavelets do Ponto 22, primeiro semestre de 2011; Ponto 5, primeiro semestre de 2011; Ponto 4, primeiro semestre de 2011; Ponto 14, primeiro semestre de 2011.

Codigo do Ponto	Data	Latitude	Longitude
22	01-01-2011	-14.400000000000000	-56.450000000000003
5	01-01-2011	-15.830000000000000	-54.380000000000003
4	01-01-2011	-15.609999999999999	-56.100000000000001
14	01-01-2011	-15.779999999999999	-56.060000000000002

Figura 48: Quadro com as informações dos pontos do primeiro semestre de 2011 de temperatura média do ar de Wavelets.

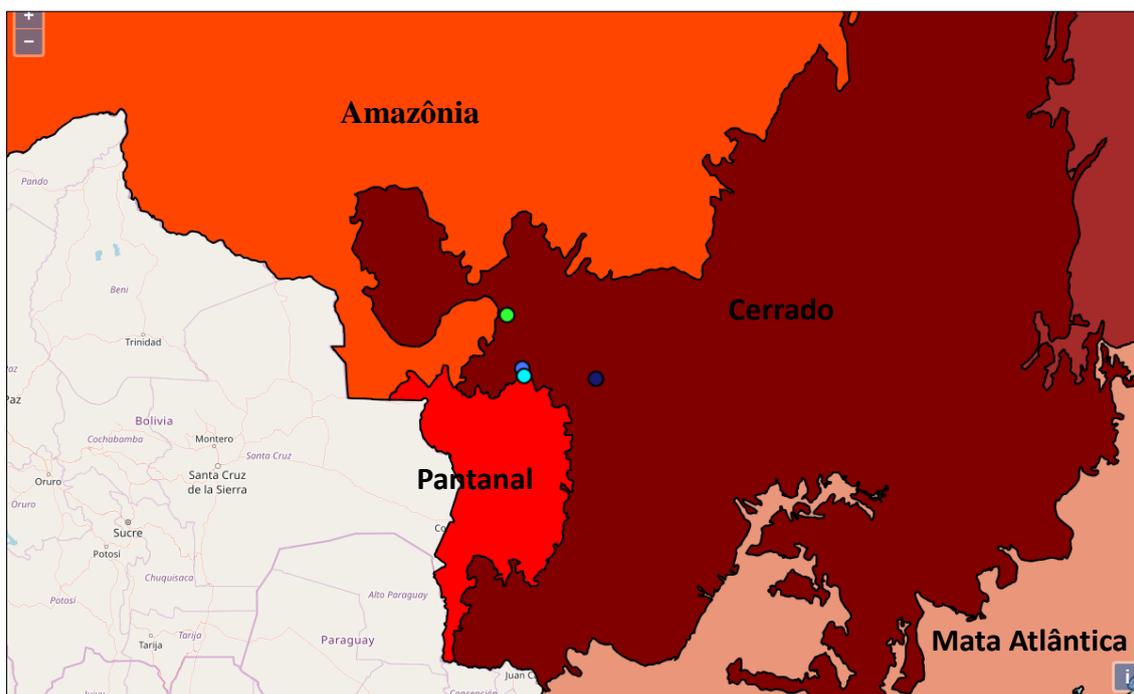


Figura 49: Mapa com pontos semelhantes ao ponto selecionado pelo usuário. Ponto selecionado pelo usuário (VERDE): primeiro semestre de 2011; Primeiro ponto semelhante (AZUL ESCURO): primeiro semestre de 2011; Segundo ponto semelhante (AZUL MÉDIO): primeiro semestre de 2011; Terceiro ponto semelhante (AZUL CLARO): primeiro semestre de 2011.

4.5 Conclusão do Capítulo

Os resultados apresentados neste capítulo para a análise de similaridade em Wavelets nas resoluções temporais: semestral, anual e 16 anos mostram que foram encontrados períodos de dias de alta intensidade de energia e que existem domínios temporais semelhantes nas Wavelets encontradas pelo SWAmP 1.0. Tais resultados podem ser utilizados para uma análise em 3 dimensões (tempo, frequência e intensidade de energia) das quatro imagens apresentadas para cada periodicidade nos dados de precipitação, temperatura média do ar e umidade relativa.

Com relação a análise espacial realizada nas Wavelets, os pontos com resolução temporal semestral e anual verificou-se as Wavelets semelhantes tem a uma distância média de 200 quilômetros, entretanto, existem pontos nos extremos do mapa que possuem Wavelets com distâncias maiores chegando a 1300 quilômetros. Sobre a análise espacial nas Wavelets com resolução temporal de 16 anos ficou evidente que a distância média estava a baixo de 220 quilômetros, para as Wavelets semelhantes.

As análises realizadas no SWAmP 1.0 apresentadas neste trabalho com as séries temporais, com resolução temporal: semestral, anual e 16 anos, mostram que a

metodologia que culminou com a implementação do sistema traz as séries temporais semelhantes para os pontos selecionados, e permite a comparação em diferentes escalas. Mostrou também a possibilidade de encontrar pontos com séries temporais similares com distâncias superiores a 1000 quilômetros como no caso da análise espacial semestral de precipitação. No entanto, para as análises espaciais realizadas nas resoluções temporais semestral, anual, de séries temporais de umidade relativa e temperatura média do ar mostram que a distância média dos pontos semelhantes selecionados pelo usuário se encontra entre 200 e 300 quilômetros.

Por fim, os resultados apresentados neste tópico, mostram diferentes pontos e vários resultados que podem ser utilizados para trabalhos futuros em análises ambientais.

5. Considerações Finais

Neste trabalho foi desenvolvida a metodologia SWAmP para tratamento processamento, e armazenamento de dados ambientais. Os dados foram modelados em um banco de dados geográfico e processados para gerar Wavelets e séries temporais com resolução temporal: anual, semestral e 16 anos do período de janeiro 1998 a dezembro 2013. Para possibilitar a visualização dos resultados em uma interface *web* e permitir a interoperabilidade de diversos usuários aos dados e resultados, a metodologia desenvolvida neste trabalho, que possibilita detecção de padrões semelhantes nas séries temporais e Wavelets de dados ambientais culminou em um SIG *web* denominado SWAmP 1.0.

A abordagem de busca por similaridade usando a Distância Euclidiana mostrou-se eficiente em encontrar padrões semelhantes em ambos os tipos de dados (nas séries temporais e *Wavelets*). O uso de técnicas de busca por similaridade permitiu que locais diferentes pudessem ter seus comportamentos temporais comparados no SWAmP 1.0 com a integração de diferentes tecnologias.

A análise espacial possibilitou uma visibilidade sobre os dados apresentados neste trabalho, de modo geral, pode-se concluir que em sua maioria os pontos que possuem dados semelhantes estão a uma distância média de 200 a 300 quilômetros, existindo alguns pontos com dados semelhantes com distâncias superiores que poderiam ser estudados juntamente com outras variáveis ambientais ou um estudo local.

Os resultados apresentados neste estudo contribuem em várias pesquisas que podem necessitar de processamento e análise de séries temporais e Wavelets de precipitação, umidade relativa e temperatura em espaços temporais: semestral, anual e de 16 anos o que possibilita a visualização de informações de dados em diferentes escalas temporais de forma intuitiva espacial e temporal.

A metodologia SWAmP pode ser empregada para o processo de análise ambiental, conforme demonstrado com a implementação do sistema SWAmP 1.0 com: a comparação dos dados, agrupamento de dados semelhantes, análises e tomada de decisão baseada em conhecimento, para servir como base para os pesquisadores que necessitem informações e tomadas de decisão.

5.1 Contribuições

De modo geral, a metodologia SWAmP e o SWAmP 1.0 fornece para os pesquisadores e usuários, uma ferramenta que pode auxiliar em pesquisas em instituições que necessitam analisar dados de series temporais ou Wavelets para permitir a análise ambiental com funcionalidade espaço-temporal. Apresentando as seguintes contribuições específicas:

- Implementação de um SIG *Web* denominado SWAmP 1.0 que possibilita visualização e análise do mapa e séries temporais e *Wavelet*.
- Processamento em diferentes resoluções temporais: semestrais, anuais e 16 anos, que permite a visualização dos dados ambientais.
- Integração entre a análise de similaridade e *Wavelets*, que possibilita uma análise auxiliar à análise realizada comumente sobre as séries temporais.
- Modelagem do BDG, com os dados de precipitação do projeto TRMM, umidade relativa e temperatura média do ar do INMET que proporciona a busca por dados em diferentes periodicidades e com busca espacial.
- Processamento espaço-temporal do comportamento da precipitação utilizando séries temporais e *Wavelets*, que acrescenta aos processos de tomada de decisão um conjunto novo de informações sobre dados ambientais.
- Processamento e visualização dos resultados.

5.2 Trabalhos Futuros

Considerando que o SWAmP 1.0 foi implementado e todas as etapas impostas nos objetivos deste trabalho foram concluídas, o mesmo se encontra pronto para uso. Todavia, ainda existem possibilidades para que seja aperfeiçoado em trabalhos futuros e otimizado da seguinte forma:

- Inclusão de outros tipos de análises ao SWAmP 1.0 para que o mesmo se torne mais robusto. A análise Fourier poderia trazer diferentes resultados para análise de séries temporais, e a implementação de outros tipos de busca por similaridade como a de Manhattan, Cosseno ou Correlação

- Inclusão de filtros dinâmicos na busca temporal, para alterar as resoluções temporais, ou seja, o usuário poderia alterar a busca para o período necessário.
- Aperfeiçoamento o processo de gerar o banco dados das séries temporais e Wavelets. Essa função tornaria o sistema mais robusto, visto que para desenvolver o banco de dados do SWAmP 1.0 ele foi modelado nas resoluções temporais tratadas no presente trabalho.
- Possibilitar diferentes buscas geográficas pela interface, para que o usuário escolha sua delimitação por quilômetros.
- Implementação de mais funcionalidades de SIG

6. Referências Bibliográficas

ADLER, C.; PETTY, G.; MORISSEY, M.; GOODMAN, H M. 2001. *Intercomparison of global precipitation products: The third Precipitation Intercomparison Project (PIP-3)*. Bull. Amer. Meteor. Soc., 82, 1377-1396.

ALCÂNTARA, A.; MORESI A. D. E.; PRADO A. H. Cenários prospectivos, monitoração ambiental e metadados. DataGramZero - Revista de Ciência da Informação - v.11, n.1, 2010.

ALLWEYER, T. (2010). BPMN 2.0 Introduction to the Standard for Business Process Modeling. Norderstedt, MA: Books on Demand.

ALVES B. C. C., FILHO, F. A. S., SILVEIRA, C. S. Análise de Tendências e Padrões de Variação das Séries Históricas de Vazões do Operador Nacional do Sistema (ONS); RBRH - Revista Brasileira de Recursos Hídricos Volume 18 n.4, 2013.

ALVES, K. S. S. Validação da precipitação estimada pelos produtos 3B42 e 3B43 do TRMM sobre a reserva biológica Jaru - RO. Cuiabá, MT. 70 f. Dissertação (Mestrado em Física Ambiental) - Instituto de Física, Universidade Federal de Mato Grosso, 2017.

AYOADE, J.O. Introdução à Climatologia para os Trópicos. 8. ed. Rio de Janeiro: Bertrand Brasil, 2002. 332p, 1v

BALA A., KUMAR A., BIRLAN., / VOICE COMMAND RECOGNITION SYSTEM BASED ON MFCC AND DTW. International Journal of Engineering Science and Technology Vol. 2 (12), 2010, 7335-7342

BARBOZA, E. B. M. ROSA, R. R., VIJAYKUMAR, N. L., BOLZAN, M. J. A., TOMASELLA, J. Caracterização por Ondoletas de Processos Físicos Não-Lineares. Congresso Nacional de Matemática Aplicada e Computacional. São Paulo, SP, Brasil, 2005.

BARCELLOS, Christovam de Castro et al. Situação atual do geoprocessamento e da análise de dados espaciais em saúde no Brasil. 2002.

BARRAGAN JOAO FRANCISCO MONTEIRO (UFBA), CRISTIANO HORA DE OLIVEIRA FONTES (UFBA), OTACILIO JOSE PEREIRA (UFBA), SERGIO TORRES SA BARRETTO (PETROBRAS), LUCIANA DE ALMEIDA PACHECO

(PETROBRAS), Análises De Métricas De Similaridades Em Séries Temporais Para Reconhecimento De Padrões,. XXXII ENCONTRO NACIONAL DE ENGENHARIA DE PRODUCAO, 2012.

BARRERA, D. F., *Precipitation estimation with the hydroestimator technique: its validation against raingauge observations*. In: Proceedings of the VII IAHS Meeting, Foz do Iguaçu, April 3-9, 2005.

BARRETT, E. C. *et al. The first WetNet Precipitation Intercomparison Project: Interpretation of results*. Remote Sens. Rev., 11, 1994.

BATISTA, G. E.; WANG, X. ; KEOGH, E. J.. A complexity-invariant distance measure for time series. In: SDM, volumen 11, p. 699-710, 2011.

BLAIN, G. C. Precipitação pluvial e temperatura média do ar do ar no Estado de São Paulo: periodicidades, probabilidades associadas, tendências e variações climáticas. Tese. Departamento de Agronomia, Universidade de São Paulo, Piracicaba, 2010.

BOX, G. E. P.; JENKINS, G. M. Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day, 2ª edição, 1976;

BUTERA, F. M. Architettura e ambiente: manuale per il controllo della qualità termica, luminosa e acustica degli edifici. Milano: Etaslibri, 1995.

C MARA, G.; DAVIS, C. Introdução à ciência da geoinformação. São José dos Campos: INPE, 2001. p. 2-6.

C MARA, Gilberto e RIBEIRO, Gilberto Queiroz Introdução da Ciência da Geoinformação, São Jose dos Campos, INPE 2004.

CÂMARA, G.; CASANOVA, M.; HEMERLY, A.; MAGALHÃES, G.; BAUZER-MEDEIROS, C., 1996, Anatomia de Sistemas de Informação Geográfica, Curitiba, Sagres.

CÂMARA, G; Representação computacional de dados geográficos. Disponível em: <<http://www.dpi.inpe.br/livros/bdados/cap1.pdf>>.

CARVALHO, V. PostgreSQL: Banco de dados para aplicações *web* modernas. Editora Casa do Código, 2017.

COLLISCHONN, B. 2006. Uso de precipitação estimada pelo satélite TRMM em modelo hidrológico distribuído. Dissertação.. Universidade Federal do Rio Grande do Sul.

COLLISCHONN, B.; ALLASIA, D.; COLLISCHONN, W.; TUCCI, C.E.M. Desempenho do satélite TRMM na estimativa de precipitação sobre a bacia do Paraguai superior. *Revista Brasileira de Cartografia*, 59. 2007.

COLLISCHONN, B.; COLLISCHONN, W.; TUCCI, C. E. M. (2008). *Daily hydrological modeling in the Amazon basin using TRMM rainfall estimates Journal of Hydrology*, 360(1-4):207-216doi:10.1016/j.jhydrol.2008.07.032.

DAUBECHIES, Ingrid: CBMS-NSF Regional Conference Series in Applied Mathematics - Ten Lectures on Wavelets - 1992 - ISBN: 978-0-89871-274-2

DAVENPORT, T. H. Reengenharia de Processos. Rio de Janeiro: Campus, 1994.

DIAS, R. Gestão ambiental: responsabilidade social e sustentabilidade. 2. ed São Paulo: Atlas, 2011.

DUTRA, Ricardo da Silva e MINETTO, Rodrigo: Transformadas Wavelets: Teorias e Aplicações em Análise de Imagens Digitais. Departamento de Informática - UFPR - 2015.

EASTMAN, R.; KYEM, P.A.K.; TOLEDANO, J.; JIN, W. Explorations in Geographic Information Systems Technology. Gis and decision making. UNITAR, Worcester, M.A. 1993.

EBERT, E. E., MANTON, M.J.; ARKIN, P.A.; ALLAM, R.J.; HOLPIN, G. E.; GRUBER, A. *Results from the GPCP Algorithm. Intercomparison Programme*. Bull. Amer. Meteor. Soc., 77, 2875-2887. 1996.

EHLERS, R. S. Análise de Séries Temporais. Curso de Séries Temporais. Agosto de 2009 e 2011.

FARIAS, S. E. M.; ARANTES, A. E.; FERREIRA JUNIOR, L. G.; SANO, E. E.; BARROS, J. R. Avaliação das Estimativas de Precipitação do satélite TRMM para o Estado de Goiás: uma abordagem climatológica. Anais XVI Simpósio Brasileiro de Sensoriamento Remoto - SBSR, Foz do Iguaçu, PR, Brasil, p. 1938 - 1944, 2013.

FECHINE, J. A. L.; GALVÍNCIO, J. AGRUPAMENTO DA PRECIPITAÇÃO MENSAL DA BACIA HIDROGRÁFICA DO RIO BRIGIDA-PE, ATRAVÉS DA MULTIVARIADA. Revista Brasileira de Geografia Física Recife-PE. V. 01 n. 01, Mai/Ago, 2008, 39-46 D. 39

FISHER, B.L. *Climatological Validation of TRMM TMI and PR Monthly Rain Products over Oklahoma*. J. Appl. Meteor., 43, 519-535, 2004.

FOOTE J. A Similarity Measure for Automatic Audio Classification. Technical Report SS-97-03. Compilation copyright © 1997.

FRANCHITO, S. H.; RAO, V. B.; VASQUES, A. C.; SANTO, C. M. eCONFORTE, J. C. Validation of TRMM precipitation radar monthly Rainfall estimates over Brazil. Journal of Geophysical Research: Atmospheres, v. 114, n. D2, 2009.

FURON, A. C.; RIDDLE, C.W.; SMITH, R.; WARLAND, J.S. *Wavelets analysis of wintertime and spring thaw CO₂ and N₂O fluxes from agricultural fields*. Agricultural and Forest Meteorology, v.48, p.305-317, 2008.

FURTADO, Diego Silva. Classificação de series temporais por similaridade e extração de atributos com aplicação na identificação automática de insetos. Dissertação de Mestrado em Ciências de Computação e Matemática Computacional 0 USP - São Carlos SP - 2014.

GARCEZ, L.N. & ALVAREZ, G.A. Hidrologia. 2.ed. São Paulo, Edgard Blücher, 1988. 291p.

GOMES, J. M.; VELHO, L. Computação Visual: Imagens. Rio, SBM, 1995.

GORNI, D.; GIANNOTTI M.; KNOPIK A.; BRITO P.; RODRIGUES M., Open source *web GIS Sistema de Informação Geográfica de Expedições*. Simpósio Brasileiro de Sensoriamento Remoto, Florianópolis, Brasil 21-26 abril 2007, INPE, p. 3953-3960.

GORNI, Daniel et al. Open source Web GIS-Sistema de Informação Geográfica de Expedições. Anais XIII Simpósio Brasileiro de Sensoriamento Remoto, INPE, p. 3953-3960, 2007.

GUARIENTI, G, S, S. Desenvolvimento de uma técnica computacional de processamento espaço-temporal aplicada em séries de precipitação. Cuiabá, 2015. 121p.

Dissertação (Mestrado) - Pós Graduação em Física Ambiental, Universidade Federal de Mato Grosso.

GUEDES, R. L.; ANDREOLI, R.V.; KAYANO, M.T.; OYANA, M.D.; ALVES, M.A.S. Série temporal de precipitação mensal de Fortaleza, Brasil: Comparação entre observações e dados de reanálise do NCEP/NCAR. *Revista Brasileira de Meteorologia*, São José dos Campos, v.20, n.1. p.83-92, 2005.

GUNOPULOS, D. ; DAS, G.; MANNILA, H.. Finding similar time series In: *PRINCIPLES OF DATA MINING AND KNOWLEDGE DISCOVERY*, p. 88-100. Springer, 1997.

HAMADA E., MAIA A. H.N., LUIZ A. J. B.. III Similaridade das Projeções de Temperatura média do ar Média, Precipitação e Umidade Relativa para o Nordeste Brasileiro Seminário da Rede AgroHidro e I Workshop do projeto, 2015.

HEUSER, Carlos Alberto. Projeto de Banco de Dados. Porto Alegre: Sagra Luzzatto, 1998.

HUANG, Y.; CHEN, S.; CAO, Q.; HONG, Y.; WU, B.; HUANG, M. e YANG, X. Evaluation of version-7 TRMM multi-satellite precipitation analysis product during the Beijing extreme heavy rainfall event of 21 July 2012. *Water*, v.6, n.1, p. 32-44, 2013.

INPE disponível em <http://www.inpe.br/>, acessado em 06 de abril de 2018.

INMET disponível em http://www.inmet.gov.br/portal/css/content/topo_iframe/pdf/Nota_Tecnica-Rede_estacoes_INMET.pdf acessado em 13 de dezembro de 2018.

JAXA; < <http://www.jaxa.jp/>> Acesso em 25 de janeiro de 2018.

JESTON, J.; NELIS, J. BPM Implementations: Does One Approach Fit All? *BP Trends*, 2006b.

KIRCHGASSMER, G. e WOLTERS, J. (2007), *Introduction to Modern Time Series Analysis*, Springer-Verlag Berlim Heidelberg.

KUMMEROW, C. *et al. The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit. Journal of Applied Meteorology*, v. 39, n. 12, p.1965-1982, 2000.

KURGAN, L. A.; MUSILEK, P. A survey of knowledge discovery and data mining process models. *The Knowledge Engineering Review*, v. 21, p. 1-24, 3 2006. ISSN 1469-8005. Disponível em: http://journals.cambridge.org/article_S0269888906000737.

KUROSE J. F.; ROSS K. W., *Redes de Computadores e a Internet - Uma abordagem Top-Down*. Edição 5 Editora Pearson, 2010.

LAYBERRY, R.; KNIVETON, D. R.; *et al.* 2006. *Daily Precipitation over Southern Africa: A new resource for climate studies*. *Journal of Hydrometeorology*. 7, 149-159.

LIMA, ERONDINA AZEVEDO . *Modelagem de séries temporais em dados de temperatura média do ar do ar por semivariografia univariada: interpolação spline e validação cruzada / Erondina Azevedo de Lima*. - 2015.

LOH, R. H. *Time Series Forecast With Neural Network and Wavelet Techniques*. The University of Queensland. Note Class, 49p, 2003.

MARBÁN, O.; MARISCAL, G.; SEGOVIA, J. Data mining and knowledge discovery in real life applications. In: PONCE, J.; KARAHOCA, A. (Ed.). [S.l.]: InTech, 2009. cap. A Data Mining and Knowledge Discovery Process Model.

MARISCAL, G.; MARBÁN, O.; FERNANDÉZ, C. A survey of data mining and knowledge discovery process models and methodologies. *The Knowledge Engineering Review*, v. 25, n. 2, p. 137-166, 2010.

MELO JUNIOR, Júlio C. F. de; SEDIYAMA, Gilberto C.; FERREIRA, Paulo A. and LEAL, Brauliro G.. Determinação de regiões homogêneas quanto à distribuição de frequência de chuvas no leste do Estado de Minas Gerais. *Rev. bras. eng. agríc. ambient.* [online]. 2006, vol.10, n.2 [cited 2018-04-23], pp.408-416. Available from: <http://www.scielo.br/scielo.php?script=sci_arttextepid=S1415-43662006000200023&lng=en&rm=iso>. ISSN 1415-4366. <http://dx.doi.org/10.1590/S1415-43662006000200023>.

Vieira, Guilherme Ernani. *Análise de métodos de previsão de demanda baseados em séries temporais em uma empresa do setor de perfumes e cosméticos / Albino Mileski Junior; [orientador]*, -- 2007.

MUEEN, A. (2013), Enumeration of time series motifs of all lengths. *International Conference on Data Mining*, pag. 547-556

NAKKEN, M. *Wavelet analysis of rainfall-runoff variability isolating climatic from anthropogenic patterns*. Environmental Modelling e Software, Camberra, v.14, p.283-295, 1998.

NASA. Disponível em: <https://pmm.nasa.gov/TRMM/>. Acesso em janeiro 2018.

NASCIMENTO, J. L. A.; Medidas de Similaridade entre Séries Temporais / ; orientador: Eduardo Sany Laber. — 2016.

Naumann, G.;Barbosa, P.; Carrao, H.; Singleton, A.; Vogt, J.; Monitoring Drought Conditions and Their Uncertainties in Africa Using TRMM Data. American Meteorological Society. DOI: 10.1175/JAMC-D-12- 0113.1. 2012

NICHOLSON, S. E.; *et. al.* 2003. *Validation of TRMM and Other Rainfall Estimates with a High-Density Gauge Dataset for West Africa. Part II: Validation of TRMM Rainfall Products*. J. Appl. Meteor., 42, 1355-1368.

NICHOLSON, S., 2005. *On the question of the “recovery” of the rains in the West African Sahel*. Journal of Arid Environments 63, 615-641.

NÓBREGA, R S; SOUZA, E, P; SOUSA, F, A, S; ANÁLISE DA UTILIZAÇÃO DE DADOS DO SATÉLITE TRMM EM UM MODELO HIDROLÓGICO SEMIDISTRIBUÍDO NA BACIA DO RIO JAMARI (RO). Revista Brasileira de Geografia Física. Recife-PE. V. 01, n. 01, Mai/Ago, 2008, 47-61.

NÓBREGA, R. S. Modelagem de impactos do desmatamento nos recursos hídricos da bacia hidrográfica do Rio Jamari (RO) utilizando dados de superfície e do TRMM. Tese. Universidade Federal de Campina Grande, Campina Grande, 2008.

NÓBREGA, R. S., SOUZA, E. P., GALVÍNCIO, J. D, Análise da Estimativa de Precipitação do TRMM em uma Sub-Bacia da Amazônia Ocidental. Revista de Geografia. Recife: UFPE - DCG/NAPA, v. 25, nº1, jan/abr.2008.

NOGUEIRA, Fernando. Modelos de Previsão: Universidade Federal de Juiz de Fora, Brasil, 2009.

OLIVEIRA, A. G. MiMi: Plataforma computacional para mineração de dados micrometeorológicos. Cuiabá, 2015, 84f. Tese (Doutorado em Física Ambiental)- Instituto de Física, Universidade Federal de Mato Grosso.

- OLIVEIRA, de. Raphael Borges, Utilização da classificação por distância euclidiana no mapeamento dos focos de arenização no setor sul da alta bacia do Rio Araguaia, Instituto de Estudos Sócio Ambiental - Universidade Federal de Goiás, 2007.
- OLIVEIRA, E. F. 2008. *Geo na web: As novidades da Internet Geográfica*. Periódico, edição 55 - InfoGeo, Curitiba, 2008.
- OMETTO, J. C. *Bioclimatologia Vegetal*. São Paulo SP: Agronômica Ceres, 1981.440p, 1v.
- PEREIRA, A. R.; ANGELOCCI, L. R.; SENTELHAS, P. C. *Agrometeorologia, Fundamentos e aplicações práticas*. Piracicaba SP, 2002, 478p.
- PIRES, Fátima; MEDEIROS, Cláudia B. *Um ambiente computacional para modelagem de aplicações geográficas*. Campinas: Unicamp, Instituto de Computação, 1997.
- PRESSMAN, R. *Software Engineering: A Practitioner's Approach*. [S.l.: s.n.], 2005.
- REBOITA, M. S. *Elementos da variabilidade climática no extremo sul do Brasil, no período de 1990 a 2001*. 2004. 211f. Dissertação (Mestrado em Engenharia Oceânica), Universidade Federal do Rio Grande do Sul, Rio Grande, 2004.
- RHEE, J.; IM, J.; CARBONE, G.J.; *Monitoring agricultural drought for arid and humid regions using multi-sensor remote sensing data*. *Remote Sensing of Environment*, vol. 114, p. 2875-2887.2010.
- ROBERT E. DONOVAN. *A New Distance Measure for Costing Spectral Discontinuities in Concatenative Speech Synthesizers*, Fourth ISCA ITRW on speech Synthesis (SSW-4). 2001.
- ROCHA, V. B. *Uma abordagem de wavelet aplicada à combinação de previsões: Uma análise teórica e experimental*. 154f. Dissertação (Métodos Numéricos em Engenharia) - Universidade Federal do Paraná, Curitiba, 2008.
- ROZANTE, J. R.; MOREIRA, D. S.; GONÇALVES, L. G. G. e Vila, D. *Combining TRMM and surface observation of precipitation: technique and validation over South America*. *Weather and Forecasting*, p. 885-894, 2010.

SANT' ANNA NETO, J. L. As chuvas no estado de São Paulo: a variabilidade pluvial nos últimos 100 anos. Variabilidade e mudanças climáticas: implicações ambientais e socioeconômicas. Maringá: Eduem, 2000. Capítulo 8, p. 95-120.

SANTOS, A. C.; ANDRADE, A.P.; LIMA, J.R.S. *et al.* Variabilidade temporal da precipitação pluvial: nível de nitrogênio no solo e produtividade de cultivares de girassol. *Ciência Rural*, v.32, n.5, p.757-764, 2002;

SANTOS, N. B. R.; Distribuições espaciais e temporais de temperatura média do ar de superfície no bioma cerrado: uma análise integrada a partir de dados orbitais de resolução moderada; Tese (Programa Multidisciplinar de Doutorado em Ciências Ambientais) - Universidade Federal De Goiás, Goiânia, 2010.

SCHEEL, M. L. M.; ROHRER, M.; HUGGEL, C.; SANTOS VILLAR, D.; SILVESTRE, E. e HUFFMAN, G. J. Evaluation of TRMM Multi-satellite Precipitation Analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution. *Hydrology and Earth System Sciences*, v. 15, n. 8, p. 2649-2663, 2011.

SCHIMIGUEL, J; BARANAUSKAS, M. C.; MEDEIROS, C. B. Usabilidade de aplicações SIG *web* na perspectiva do usuário: um estudo de caso. Disponível em: <http://www.ip.pbh.gov.br/ANO8_N1_PDF/ANO8N1_Juliano.pdf>. Acesso em 27 de janeiro de 2014.

SCHIMIGUEL, Juliano; BARANAUSKAS, Maria Cecilia Calani; MEDEIROS, Claudia Bauzer. Usabilidade de Aplicações SIG *Web* na Perspectiva do Usuário: um Estudo de Caso. In: *GeoInfo*. 2005.

SCHNEIDER, Marcelo. Detecção de tendências climáticas na região norte-nordeste do estado de São Paulo no período de inverno e pré- estação chuvosa, através da análise de séries temporais e modelagem numérica, 2012. 141 f. Tese de Doutorado – Programa de Pós-Graduação em Ciência Ambiental (PROCAM) Universidade de São Paulo.

SENTELHAS, P. C.; NASCIMENTO, A. L. C. Variação sazonal da relação entre o saldo de radiação e a irradiância solar global. *Revista Brasileira de Meteorologia*, v.18, n.1, p.71-77, 2003.

- SHEARER, C., *The CRISP-DM model: the new blueprint for data mining*, J Data Warehousing (2000); 5:13—22.
- SILVA J. X. O que é Geoprocessamento? Rio de Janeiro, Revista do CREA-RJ, Out/Nov 2009.
- SILVA, A.B. 1998. Sistemas de Informações Geo-referenciadas: uma introdução. No prelo. Universidade Estadual de Campinas (UNICAMP), Campinas, SP.
- SILVA, F. S. Classificação De Séries Temporais Por Similaridade E Extração De Atributos Com Aplicação Na Identificação Automática De Insetos. Dissertação De Ciências E Matemática E De Computação ICMC-USP, 2014.
- SILVA, J. B. DA; CUNHA, F. B.; GAVIÃO NETO, W. P. Modelagem das chuvas trimestrais por regiões homogêneas no Estado do Ceará. Revista Brasileira de Agrometeorologia, Santa Maria, v.9, n.2, p.317-24, 2001.
- SINGH S. E.G.; RAJAN V. Quantization Approach for Speaker Recognition using MFCC and Inverted MFCC. International Journal of Computer Applications (0975 - 8887) Volume 17- No.1, March 2011.
- SMITH, E. A.; *et al.* 1998. *Results of WetNet PIP-2*. J. Atmos. Sci., 55, 1483-1536.
- TORRENCE, C.; COMPO, Y. G. P. A practical guide to wavelet analysis. Bulletins of American Meteorological Society, v.79, p.61-78, 1998.
- TORRENCE, C; COMPO, Y.G.P. A practical guide to wavelet analysis. Bulletins of American Meteorological Society, Boston, v.79, p.61-78, 1998.
- TORRENCE, C; WEBSTER, P. *Interdecadal changes in the ENSO-monsoon system*, Journal of Climate, Washington, v.12, p.2679-2690, 1999.
- TRMM. *Mission Overview*. Disponível em: <https://pmm.nasa.gov/TRMM/trmm-instruments> Acesso em 18 janeiro de 2018.
- TUBELIS, A.; NASCIMENTO, F. J. L. do. Meteorologia descritiva; fundamentos e aplicações brasileiras. Editora: Nobel, ano: 1992.
- TUCCI, C.E.M; MENDES, C.A. Avaliação Ambiental Integrada de Bacia Hidrográfica. Ministério do Meio Ambiente. ISBN 85-7738-047-5. 2006.

WIRTH, R., *CRISP-DM: Towards a standard process model for data mining*. In: Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining, 2000.

WOOLDRIDGE, Jeffrey M. *Introductory Econometrics: A Modern Approach*, 2000, South-Western College Publishing, Cap. 1 pág. 8.

YAN,N.; WU, B.; CHANG, S.; BAO, X.; Evaluation of TRMM Precipitation Product for Meteorological Drought Monitoring in Hai Basin. 35th International Symposium on Remote Sensing of Environment (ISRSE35), doi:10.1088/1755-1315/17/1/012093. 2014.

ZEILHOFER, P.; LIMA, E. B. R.; SANTOS, F. M.; RIGO JÚNIOR, L. O. Um ambiente SIG para modelagem integrada da qualidade da água utilizando Qual2e. *Caminhos de Geografia*, v.8, n.10, p.107-125, 2003.

ZEILHOFER, Peter. *Sistemas para registro e manipulação de dados de precipitação de chuva do projeto TRMM*, 2015 - Universidade Federal de Mato Grosso, Cuiabá.