



Graduação em Ciência da Computação

Vinicius Oliveira Folha

**VTMV: UMA FERRAMENTA INTERATIVA DE VISUALIZAÇÃO
PARA A ANÁLISE DE TRAJETÓRIAS MULTIDIMENSIONAIS**

Trabalho de Graduação



Universidade Federal de Pernambuco
secgrad@cin.ufpe.br
www.cin.ufpe.br/~secgrad

RECIFE
2018



Universidade Federal de Pernambuco
Centro de Informática
Graduação em Ciência da Computação

Vinicius Oliveira Folha

**VTMV: UMA FERRAMENTA INTERATIVA DE VISUALIZAÇÃO
PARA A ANÁLISE DE TRAJETÓRIAS MULTIDIMENSIONAIS**

Trabalho apresentado ao Programa de Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Bacharel em Ciência da Computação.

Orientador: *Paulo Salgado Gomes de Mattos Neto*

RECIFE
2018

Agradecimentos

Agradeço primeiramente a Deus por ter me dado saúde força para superar as dificuldades. Aos meus pais Aessio e Rosicleide, por serem exemplo de honestidade e dedicação. Por meio do seu amor, compreensão e incentivo foram fundamentais no decorrer dessa longa e exaustiva jornada, porém gratificante. As minhas irmãs, Thaisa e Regina, que fizeram compreender que o futuro é alcançado a partir do frequente empenho no presente.

A Universidade Federal de Pernambuco e seu corpo docente, pela oportunidade de fazer curso. Meus agradecimentos seguem aos professor Nivan Ferreira e Paulo Salgado, pela orientação, apoio e confiança durante essa caminhada. Agradeço também aos meus amigos que fizeram dessa graduação, uma jornada mais leve e prazerosa, sempre me apoiando em todos os momentos.

*A verdadeira viagem de descobrimento
não consiste em procurar novas paisagens,
mas em ter novos olhos.*

—MARCEL PROUST

Resumo

Devido ao enorme avanço tecnológico na aquisição de localização, juntamente à evolução dos dispositivos móveis, houve uma grande explosão no volume de dados espaço-temporais. Esse aumento da quantidade de dados possibilitou que investigações de objetos em movimento, como pessoas, veículos, animais e materiais fossem desenvolvidas. Essa análise orientada aos dados auxilia na obtenção do conhecimento proporcionando assim melhorias significativas na qualidade de vida dos seres humanos. Embora sejam vastos os estudos referentes às trajetórias, eles são segmentados e focados no aspecto espaço-temporal e em áreas particulares. Essa limitação dificulta que os pesquisadores investiguem as trajetórias de domínios diversos e que tenham outros fatores além do espaço e tempo para serem estudados. Apesar da relevância desses fatores, muitas vezes são ignorados, o que pode acabar ocultando padrões. Este trabalho propõe sanar essas lacunas presentes na literatura através da criação da ferramenta Visualização de Trajetórias Multidimensionais (VTMV), que é uma proposta Web interativa com foco na análise das trajetórias de diversos domínios e nas variáveis que vão além das espaço-temporais. Para este propósito, a VTMV faz uso de diferentes agregações e representações visuais para analisar os aspectos geográficos em função do tempo, assim como as distintas propriedades presente nos dados. A ferramenta é testada em alguns estudos de caso, que possibilitou certificar a eficiência e a capacidade do VTMV na exploração de trajetórias e suas dimensões em diferentes domínios, logo sendo possível revelar padrões e tendências ocultas na grande massa de dados.

Palavras-chave: Visualização de dados, trajetórias multidimensionais, Dados espaço-temporais, exploração visual

Abstract

Due to the enormous technological advance in the acquisition of location, as well as the evolution of mobile devices, a great explosion in the volume of spatiotemporal data happen. This increase in the amount of data made it possible for investigations of moving objects such as people, vehicles, animals, and materials to be developed. This data-driven analysis helped to gain knowledge, thus providing significant improvements in the quality of life of human beings. Although trajectory studies are vast, they are segmented and focused on the spatiotemporal aspect and on some particular areas. This makes it difficult for researchers to investigate the trajectories of different domains and that have other factors to be studied. Despite the relevance of these factors, they are often ignored, which may result in hiding some patterns. This paper addresses the literature gap by creating the Visualização de Trajetórias Multidimensionais tool (VTMV), which is an interactive Web proposal focused on the analysis of the trajectories of several domains and on the variables that go beyond space-time. For this purpose, VTMV makes use of different aggregations and visual representations with the goal to analyze the geographical aspects as a function of time, as well as the different properties present in the data. The tool is tested in some case studies, which made it possible to certify the efficiency and capacity of VTMV in the exploration of trajectories and their dimensions in different domains, thus revealing hidden patterns and tendencies in the large mass of data.

Keywords: Data visualization, multidimensional trajectories, spatial-temporal data, visual exploration

Listas de Figuras

2.1	Mapa de Charles Minard - Campanha de Napoleão na Rússia	13
2.2	Sites das ferramentas Web de Trajetórias	14
2.3	Ferramentas de pesquisas feitas sobre trajetórias	14
4.1	Comparando os diferentes <i>Graphical Map Tiles</i>	21
4.2	A interface do VTMV - Visualização de Trajetórias Multidimensionais	22
4.3	A visualização do mapa com a análise da variável <i>wind</i>	23
4.4	A visualização do histograma com a análise da variável <i>wind</i>	24
4.5	A visualização do gráfico de linhas com a análise da variável <i>wind</i>	25
4.6	A visualização do gráfico de Gantt com múltiplos filtros.	26
5.1	Intervalo que as incoerências no valores do atributo <i>Wind</i>	28
5.2	Intervalo que as incoerências no valores do atributo <i>Pressure</i>	28
5.3	Os períodos nos quais apresentam um baixo número de ocorrência no ano.	29
5.4	Relação de dependência entre as grandezas <i>Pressure</i> e <i>Wind</i>	29
5.5	As trajetórias desses dados tendem a iniciar no mesmo região do gráficos.	30
5.6	Picos e vales presente na grandeza <i>Wind</i> ao longo do tempo.	30
5.7	Os furacões que apresentam comportamento similares quando tem média a média do <i>Wind</i> acima de 50 nós.	31
5.8	Atividade do atleta que durou mais de 1 mês.	32
5.9	Atividade do atleta que teve <i>heartRates</i> muito baixo.	32
5.10	Provável residência do atleta pelo alto número de eventos.	33
5.11	Frequência dos períodos das trajetórias do atleta.	33
5.12	Redução das atividades do ciclista no ano.	34
5.13	Rotas com grande variação da dimensão <i>elevation</i>	34
5.14	Análise das duas dimensões: <i>elevation</i> e <i>heartRates</i>	35

Lista de Tabelas

3.1	Uma amostra do conjunto de dados dos furacões do Atlântico	16
3.2	Uma amostra do conjunto de dados das atividades dos ciclistas	16

Lista de Acrônimos

SVG *Scalable Vector Graphics*

D3.js *Data Driven Documents*

JSON *JavaScript Object Notation*

HTML *HyperText Markup Language*

Sumário

1	Introdução	10
1.1	Objetivos	10
1.2	Estrutura do trabalho	11
2	Trabalhos Relacionados	12
2.1	Ferramentas <i>Onlines</i>	12
2.2	Literatura	13
3	Análise	15
3.1	Pretensão da visualização	15
3.2	Conjuntos de dados	15
4	VTMV - Visualização de Trajetórias Multidimensionais	18
4.1	Padrão de entrada	18
4.1.1	Dados	18
4.1.2	Configuração	20
4.2	VTMV - Visualização de Trajetórias Multidimensionais	21
4.2.1	Mapa	22
4.2.2	Histograma	24
4.2.3	Gráfico de linhas	24
4.2.4	Gráfico de Gantt	25
5	Casos de Uso	27
5.1	Caso de uso: Dados dos furacões	27
5.1.1	Identificar inconsistências nos dados e <i>outliers</i>	27
5.1.2	Padrões e extração de conhecimentos das variáveis	29
5.2	Caso de uso: Dados dos ciclistas	31
5.2.1	Identificar inconsistências nos dados e <i>outliers</i>	31
5.2.2	Padrões e extração de conhecimentos das variáveis	33
6	Conclusão	36
6.1	Melhorias e Futuros Trabalhos	36
Referências		37

1

Introdução

A compreensão das circunstâncias e consequências de um movimento é um dos desafios da humanidade, um exemplo disso é o movimento de organismos vivos e matérias. Diante desse desafio, surgem as principais questões que incluem o porquê, como, quando e onde os organismos se movem e como esse processo está conectado a fatores externos. Exemplos disso podem ser vistos na análise de movimentos das correntes marítimas e dos ventos para adquirir informações ajudando a previsões meteorológicas, na observação do deslocamento de grupos de animais para descobrir padrões alimentares e migratórios como também na análise de comutações nos centros urbanos para auxiliar no seu planejamento (FERREIRA et al. (2013)). A redução dos preços e a popularização dos dispositivos móveis e sensores, assim como o forte avanço tecnológico do meio computacional, fez com que grandes quantidades de dados de trajetória fossem coletados (ZHU; ZHONG; XIONG (2009)).

Em particular, um grande número desse conjunto de dados é multidimensional, isto é, não inclui somente a posição de indivíduos, mas também medidas associadas a cada uma destas posições. Estes conjuntos de dados abrem grandes oportunidades para melhor entendimento e consequentemente melhor gerenciamento de aplicações nos diversos domínios acima citados.

Entretanto, a análise deste tipo de dados é um problema não trivial. De fato, o volume de dados bem como o caráter espaço-temporal dos mesmos traz grandes desafios, em especial no que diz respeito às técnicas de visualização focadas nestes dados (FERREIRA, 2015). Apesar de algumas técnicas e ferramentas terem sido criadas para transpor esses desafios, como por exemplo os estudos da população de pássaros (FERREIRA et al., 2011), das lambretas elétricas (KRÜGER et al., 2013) e dos movimento populacional (BOYANDIN et al., 2011), não existe uma ferramenta de visualização padrão para análise de dados de trajetórias multidimensionais.

1.1 Objetivos

Diante do panorama exposto na seção anterior, a visualização interativa de dados proporciona excelentes resultados na análise e identificação de padrões, porém atualmente ainda tem sua limitação. Isso por causa do déficit de uma ferramenta para abranger trajetórias multidimen-

sionais independente do âmbito de estudo. Essa falta ocorre por conta que os estudos existentes produziram ferramentas para uso específicos e com foco nas dimensões espaço e tempo. Assim a finalidade do trabalho foi desenvolver um sistema de visualização interativa para a análise de trajetórias multidimensionais, VTMV - Visualização de Trajetórias Multidimensionais, (descrita em detalhes no Capítulo 4). No qual utilizou o objetivo central da técnica de visualização para transmitir conhecimento de forma clara e eficiente através de gráficos estatísticos e mapas, e deste modo proporciona mais um recurso na exploração de padrões em dados complexos. Essas técnicas usam a capacidade de percepção do sistema visual humano e instrumentos gráficos para enriquecer a capacidade de exploração de dados complexos e auxiliar na elaboração e verificação de suposições de modo interativa. Tornando assim ferramentas essenciais quando se considera esse cenário.

Na elaboração da aplicação foi desenvolvido um estudo das metodologias utilizadas para este tipo de dados, selecionando as técnicas que mais se adequaram as ferramentas utilizadas no projeto e as que vêm apresentando melhores resultados na literatura atual. O sistema foi projetado com a tecnologia interativa na web por causa da disponibilidade de acesso em qualquer lugar, a sua facilidade de ser usada na mais ampla gama de dispositivos através de navegadores recentes, a dinamicidade para que os usuários explore suas consultas por si mesmo e por ser um meio simples para expor ao público a ferramenta e os resultados. Para o desenvolvimento dessa aplicação foi utilizado a *Data Driven Documents* (D3.js)¹, uma poderosa biblioteca em javascript para manipulação de dados e produção complexa de efeitos visuais dinâmicos e desenhos interativos. Para os experimentos, utilizamos alguns conjuntos de dados de trajetórias multidimensionais incluindo conjuntos de dados de furacões (HURDAT2: THE REVISED ATLANTIC HURRICANE DATABASE, 2017) e de atividades de 7 ciclistas (AL., 2016).

1.2 Estrutura do trabalho

A estrutura deste documento é descrita nesta secção. O capítulo a seguir, o 2, refere-se a base do conhecimento para este trabalho com trabalhos científicos pertencentes a mesma área de estudo na literatura e ferramentas onlines existentes para análise de trajetórias. O capítulo 3 aborda as pretensões de visualizações para serem realizadas e as bases de dados utilizadas para executar o estudo da ferramenta. O design do VTMV - Visualização de Trajetórias Multidimensionais é exposto no capítulo 4, no qual explica o padrão de entrada, as funcionalidades e as técnicas de visualização utilizadas. O capítulo 5 demonstra a ferramenta em execução com os casos de uso. Por último, no capítulo 6, as conclusões e considerações finais sobre o trabalho e quais serão os possíveis trabalhos futuros.

¹<https://d3js.org/>

2

Trabalhos Relacionados

Neste capítulo os diversos trabalhos relacionados ao tópico de análise visual de trajetórias multivariáveis são apresentados e discutidos. Nos últimos anos, muitas técnicas de geo-visualização vem sendo elaborada e aprofundada para análise de dados exploratórios, e essas técnicas tem sido empregada para inspirar o pensamento criativo e fornecer inéditas informações nas quais eram desconhecidas anteriormente no cenário dos dados originais. Porém as inúmeras ferramentas desenvolvidas são focadas apenas em um conjunto específico de dados com técnicas peculiares a aquele problema, dificultando a reutilização para dados de natureza diferente.

2.1 Ferramentas *Onlines*

De forma geral, é usual o desenvolvimento de ferramentas web tais como o TRAVIC (Transit Visualization Client)¹ e a Windy² (Figura 2.2) com intuito de exteriorizar os resultados da pesquisa para o público leigo e facilitar a exploração dos mesmos dados por outros pesquisadores.

Para isso apresentam algumas funcionalidades, por exemplo: filtros para a dimensão do tempo, camadas de cores que exibem a densidade ou o valor do parâmetro selecionado e as funcionalidade de navegação de zoom e pan no mapa. Além disso exibem visualizações animadas com marcadores presente no mapa destacando a direção do objeto em movimento.

Essas ferramentas apresentam serem eficazes no contexto que são empregadas, pois segundo o criador da plataforma Windy, a ferramenta foi a principal fonte de informações meteorológicas para governos, instituições e indivíduos nas áreas afetadas na época de furacões. Entretanto, elas não disponibilizam funcionalidades necessárias para uma análise mais detalhadas e também não permite o estudo de diferentes contextos. Além disso, não possibilita a comparação de diferentes períodos, consequentemente essas plataformas não sanam a carência abordada neste trabalho.

¹<http://tracker.geops.ch>

²<https://www.windy.com>

2.2 Literatura

A exploração visual dos dados espaço-temporal tem sido bastante investigado pela comunidade científica, pois é uma tarefa desafiadora, no qual existem milhares de rotas, com diferentes tamanhos e formas, mais também diversos destinos e origens. Além disso esse dados estão contendo outras dimensões no quais elevam a dificuldade do estudo.

Essa análise de modo geral ainda é um dos tópicos mais significativos em muitos campos de pesquisa, dentre eles a mineração de dados e a análise visual (ANDRIENKO et al., 2010). A maior parte dos trabalhos existentes se concentra principalmente na exploração dos aspectos espaciais e temporais das trajetórias, ou nas relações entre trajetórias e outros tipos de eventos. Porém, a quantidade de estudos focados nas variáveis desses dados não é tão expressiva.

Dentre as primeiras tentativas com foco nos atributos da trajetória, a que mais se destacou foi o famoso mapa de Charles Minard. No qual, representa a campanha de Napoleão na Rússia, exibindo tanto o tamanho da tropa, como a temperatura do ar vinculado com as dimensões temporal e espacial, como mostra a Figura 2.1.

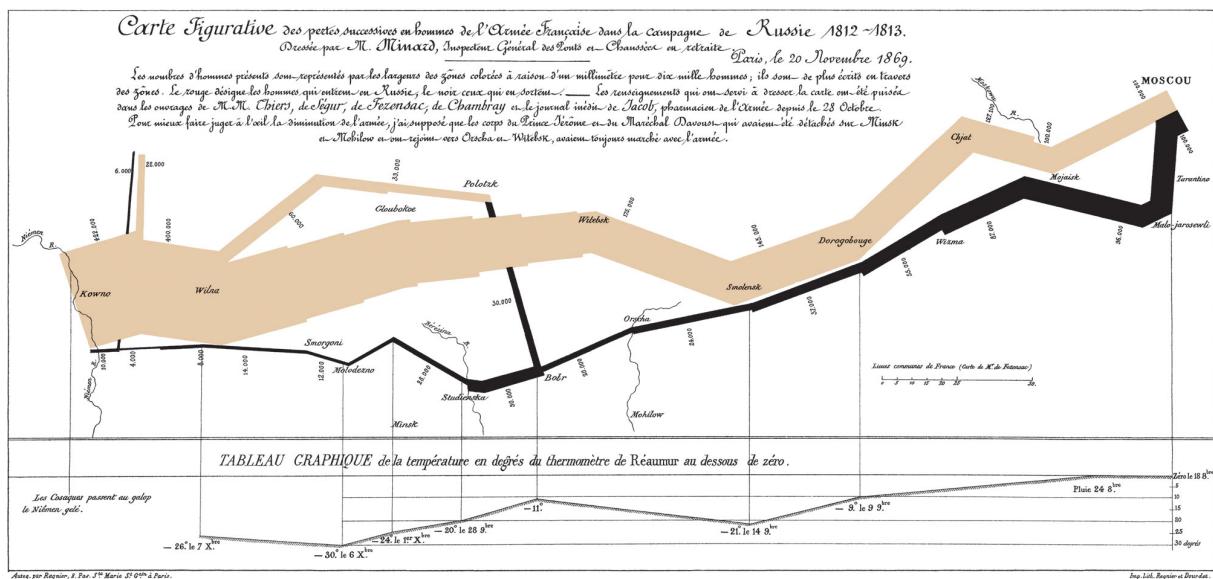
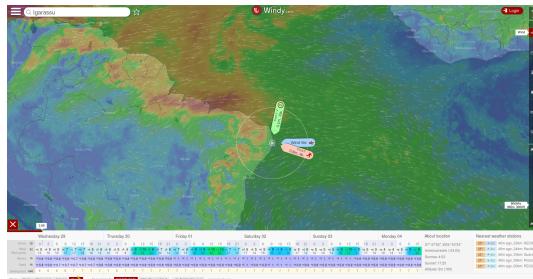


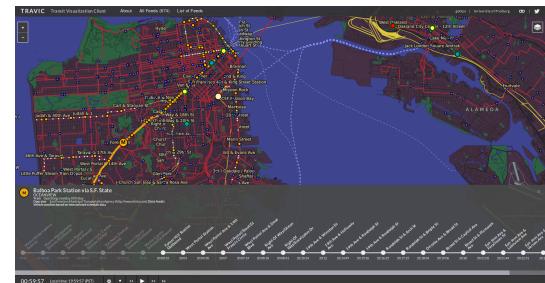
Figura 2.1: Mapa de Charles Minard - Campanha de Napoleão na Rússia

Em estudos contemporâneos, os dados espaço-temporais tem sido usado em diversos contextos, como por exemplo: população de pássaros (FERREIRA et al., 2011), lambretas elétricas (KRÜGER et al., 2013), corridas de taxi (FERREIRA et al., 2013), movimento populacional (BOYANDIN et al., 2011), respectivamente nas figuras 2.3 (a), (b), (c) e (d).

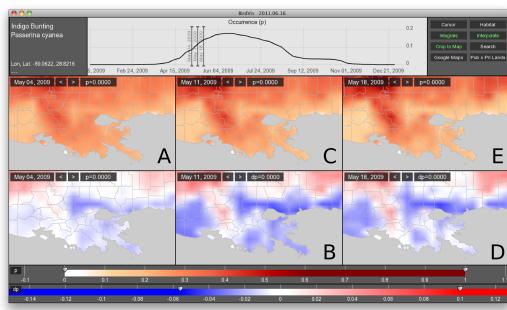
Seguindo a tendência dos trabalhos anteriores, este trabalho pretende explorar o carácter das diversas variáveis presente nos dados de natureza espaço-temporal, através dos, mapas, gráficos e filtros.



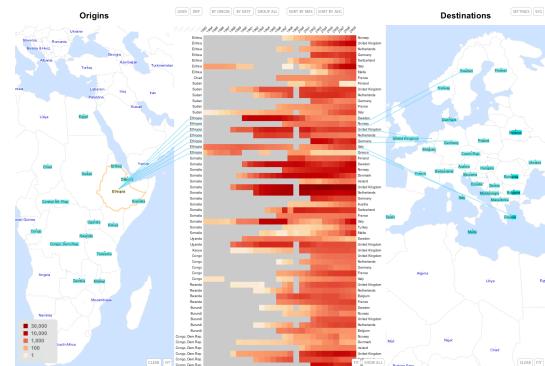
(a) WINDY Website



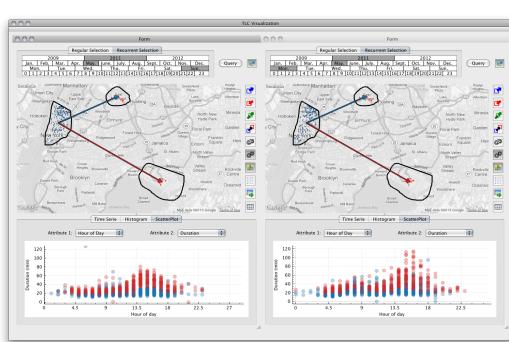
(b) TRAVIC Website

Figura 2.2: Sites das ferramentas Web de Trajetórias

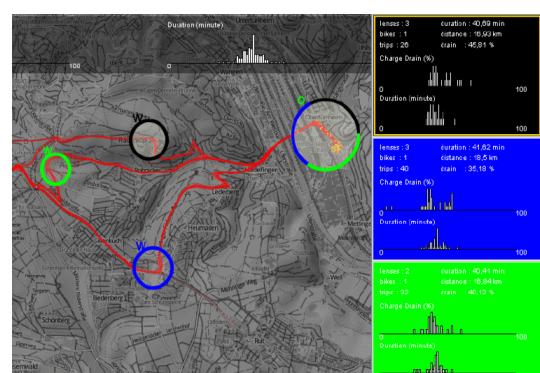
(a) BirdVis: Visualizing and Understanding Bird Populations



(b) Flowstrates: An Approach for Visual Exploration of Temporal Origin-Destination Data



(c) Visual Exploration of Big Spatio-Temporal Urban Data: A Study of New York City Taxi Trips



(d) TrajectoryLenses: A Set-based Filtering and Exploration Technique for Long-term Trajectory Data

Figura 2.3: Ferramentas de pesquisas feitas sobre trajetórias

3

Análise

Este capítulo refere-se a análise da questão levantada nesse trabalho e também do detalhamento de alguns padrões. Esses padrões foram determinados para possibilitar a utilização do VTMV - Visualização de Trajetórias Multidimensionais com diferentes bases de dados. Assim como o trabalho realizado em FERREIRA et al. (2013), a abordagem desse trabalho é considerada uma abordagem mais geral em relação a análise exploratória e visual de dados. Deste modo aproveita-se do poder visual humano para a detecção de padrões, podendo assim, logo em seguida restringir o campo de análise e explorar aquela área na qual deseja maior atenção.

3.1 Pretensão da visualização

O objetivo inicial desse projeto foi o desenvolvimento de uma ferramenta interativa de visualização de trajetórias multivariadas, na qual permitisse auxiliar cientistas de diversos domínios. A partir dessa premissa foi possível desenvolver o sistema VTMV - Visualização de Trajetórias Multidimensionais para suportar as seguintes tarefas analíticas:

- (T1) Encontrar valores extremos e discrepantes de uma variável.
- (T2) Caracterizar a distribuição dos valores, possibilitando a filtragem e uma representação visual da região com suas respectivas trajetórias e variáveis ilustradas.
- (T3) Encontrar anomalias e padrões presentes nos dados.
- (T4) Correlacionar os atributos com dimensões espaço temporal, viabilizando um filtro por tempo.

3.2 Conjuntos de dados

Os conjuntos de dados utilizados para analisar o poder da ferramenta apresentam domínios diferentes. Esses conjuntos são duas bases de dados que foram utilizadas neste trabalho

e foram obtidas através de base de dados públicas, uma sobre furações (Tabela 3.1) e a outra sobre a atividade de 7 ciclistas (Tabela 3.2).

Tabela 3.1: Uma amostra do conjunto de dados dos furacões do Atlântico

Hurricane AL162011: OPHELIA						
Date	Time	Status of system	Latitude	Longitude	Maximum sustained wind(in knots)	Minimum Pressure(in millibars)
20110920	0600	LO	11.6N	37.0W	25	1008
20110920	1200	LO	11.9N	37.7W	25	1007
20110920	1800	TD	12.2N	38.6W	30	1007
20110921	0000	TS	12.5N	39.7W	40	1005
20110921	0600	TS	12.7N	41.1W	45	1004

Essas bases além de conter os atributos da trajetória, continham outras dimensões adicionais que foram utilizadas no estudo como pressão atmosférica e velocidade. Dessa forma sendo possível explorar o poder da ferramenta construída para investigar uma quantidade diversificada de problemas e satisfazer as tarefas analíticas.

Tabela 3.2: Uma amostra do conjunto de dados das atividades dos ciclistas

coordTimes	heartRates	latitude	longitude	elevation
2014-04-25T00:46:14	124	-16.7482639	145.6643462	30.799999
2014-04-25T00:46:18	125	-16.7482284	145.6641309	31.399999
2014-04-25T00:46:19	125	-16.7482182	145.6640766	31.7999997
2014-04-25T00:46:20	125	-16.7482068	145.6640276	32.200000
2014-04-25T00:46:21	125	-16.7481975	145.6639791	32.599998

O banco de dados revisado dos furacões no Atlântico é mantido e atualizado pelo Centro Nacional de Furacões (National Hurricane Center - NHC)¹. Esse centro é a divisão do Centro de Previsão Tropical do Serviço Nacional de Meteorologia dos Estados Unidos, responsável pelo monitoramento e previsão de ciclones tropicais. Essa base de dados tem registro que regressam ao ano de 1851, porém ainda está longe de ser completa e precisa. Visto que nessa base ainda encontra-se valores ausentes em algumas colunas, principalmente nos dados mais antigos. Esses valores são datados com valores iguais a '-999'. Essa larga massa de dados têm 1830 furações e 49691 momentos datados no novo formato HURDAT2 (HURricane DATa 2nd generation), como podemos ver em HURDAT2: THE REVISED ATLANTIC HURRICANE DATABASE (2017). Essa extensa base de dados possibilita uma exploração precisa da ferramenta. Para isso, foi necessário a conversão para o formato padrão utilizado nesse projeto, no qual foi realizado por meio de códigos em *python*.

¹<https://www.nhc.noaa.gov/>

Já a segunda base de dados para o estudo consiste em dados de 7 ciclistas com perfis nas plataformas Strava e Garmin Connect (AL., 2016). Essa base de dados conta com 4874 trajetórias e mais de 4 *Gigabytes* de informações no formato *GPX*, que é basicamente um formato *XML*, que foram voluntariamente doados pelos ciclistas. Portanto a conversão desse formato para o padrão do projeto foi realizada através de um módulo em *JavaScript* chamado **toGeoJSON** para converter o *GPX* para o formato *GeoJSON*. E logo após utilizou-se um código em *python* para converter para o formato de entrada da ferramenta. Devido ao enorme tamanho dos dados, decidimos explorar uma área menor dessa massa e isso será descrito nos Casos de uso (no capítulo 5).

4

VTMV - Visualização de Trajetórias Multidimensionais

O design e a interface do sistema VTMV - Visualização de Trajetórias Multidimensionais serão descritos neste capítulo. Este desenvolvimento da ferramenta foi embasado nas tarefas analíticas presentes na Seção 3.1. Conseqüentemente de forma interativa foi possível definir e elaborar as interações usuário-sistema. Como também a delimitação da concepção física e funcional da ferramenta.

Elaborar visualizações que elucidam o usuário informações não é uma tarefa fácil. Pois anteriormente esta informação estava oculta pela quantidade de dados presente e dessa maneira dificulta a identificação das mesmas. Em vista disso, o método de tentativa e erro foi indispensável para o desenvolvimento desse sistema. Uma vez que uma implementação limitada não viabiliza um estudo adequado dos dados.

4.1 Padrão de entrada

Nessa seção será descrito o padrão para a entrada no sistema VTMV - Visualização de Trajetórias Multidimensionais e dessa maneira será possível a inserção de dados por meio de um arquivo. Como os dados de trajetórias têm diversos padrões e formatos de arquivos à depender do usuário, a definição de um padrão comum para entrada dos dados se tornou inevitável. Por essa motivo o sistema contará com dois arquivos presentes na pasta **Data**: o dados.json e configuracao.config.

4.1.1 Dados

A padronização dos dados tem que ser feita para o perfeito funcionamento da ferramenta. Para esse fim o usuário terá que transformar os seus dados para o formato estabelecido nesta subseção e ter o arquivo de configuração regulado. O formato do arquivo tem que ser *JavaScript Object Notation* (JSON) e esse arquivo é contém um *array* de objetos com as propriedade de ID e trajetórias. As trajetórias são uma propriedade com nome trajetória, que é um *array* de

objetos, no qual tem as informações obrigatórias(longitude,latitude e datahora). Além disso ela pode conter as variáveis opcionais que tem que tá especificadas no arquivo de configuração.

A **datahora** é uma propriedade que pode atender a dois padrões, o *timestamp* ou uma palavra que é formada pela concatenação da data e hora. Uma dessas opções tem que ser informada no arquivo de configuração.

O padrão seguido para o formato de concatenação da propriedade **datahora** é o seguinte: "YYYYMMDD-HHMM". No qual o ano (YYYY) é representado primeiro com quatro dígitos, depois o mês (MM) com dois dígitos e o dia (DD) com dois numerais. Logo após o hífen a hora (HH) representada com o formato de 24 horas e os minutos (MM). Podemos ver esse exemplo no trecho de código 4.1.

O **timestamp** é o registro de data e hora do unix, uma maneira de contabilizar o tempo como um total de milissegundos em execução. Esse tempo se inicia desde de um tempo fixado, chamado de *Unix Epoch*, no qual é 1 de janeiro de 1970 na UTC (do inglês *Universal Time Coordinated*). Assim o timestamp é a quantidade de milissegundos passado depois da *Unix Epoch* e o registro desejado. Este ponto no tempo tecnicamente não se altera, não importando onde esteja localizado a aferição no globo. Esse padrão trás grande benefícios para sistemas computacionais na ordenação e no rastreamento de informações datadas por sistemas dinâmicos e distribuídos.

Quanto a **latitude** atribui-se sinal positivo para as latitudes norte e negativo para as latitudes sul. Esse mesmo raciocínio é seguindo o para a **longitude**, atribui-se valores negativo para oeste e positivo para leste.

Trecho de Código 4.1: "Exemplo de padrão para a entrada dos dados".

```
[  
  {  
    "idObj": "obj1",  
    "trajetoria": [  
      {  
        "longitude": "-94.8",  
        "datahora": "18510625-0001",  
        "latitude": "28.0",  
      },  
      {  
        "longitude": "-95.4",  
        "datahora": "18510625-0600",  
        "latitude": "28.0",  
      }  
    ]  
  },
```

```
{
    "idObj": "obj2",
    "trajetoria": [
        {
            "longitude": "-48.0",
            "datahora": "18510816-0100",
            "latitude": "13.4"
        },
        {
            "longitude": "-49.5",
            "datahora": "18510816-0600",
            "latitude": "13.7"
        },
        {
            "longitude": "-51.0",
            "datahora": "18510816-1300",
            "latitude": "14.0"
        }
    ]
}
```

4.1.2 Configuração

O arquivo de configuração contém informações indispensáveis para o funcionamento do VTMV - Visualização de Trajetórias Multidimensionais. Esse arquivo configuracao.config também tem a formatação JSON e contém as seguintes propriedades: **att_extras**, **nomes** , **datahora** e **mapColor**.

O **att_extras** é responsável pelo número de atributos extras que será analisado, as variáveis opcionais e que tem papel importante nesse trabalho. Já a propriedade **nomes** é um array com os títulos de cada variável numérica a mais presente nos dados. Assim essas variáveis são responsáveis pelas visualizações geradas e potencializa o material de estudo.

O atributo **datahora** é encarregado pela definição da formatação da variável tempo presente nas trajetórias. Essa formatação pode ter duas formas, a *concatenação do tempo* e o *timestamp*, selecionando as respectivamente com os valores default ou timeParse.

Já a propriedade **mapColor** é responsável por ajustar o tipo de exibição do mapa. Assim alterando o *Graphical Map Tiles* e possibilitando ser o modelo padrão ou o Preto e branco. Para isso deve-se configurar com valores default e blackAndWhite respectivamente. Essa opção foi adicionada para facilitar a visualização de casos que as rotas se confundem com as cores do

mapa, exemplo disso podemos ver na Figura 4.1.

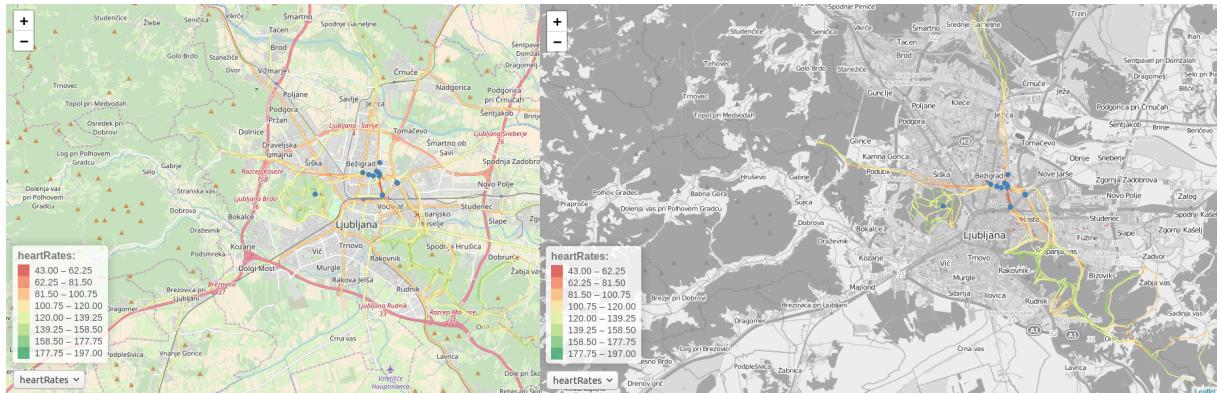


Figura 4.1: Comparando os diferentes *Graphical Map Tiles*.

Um exemplo ao utilizar os dados do furacões e decidir pela análise das variáveis *wind* e *pressure*. Logo o arquivo de configuração deve ser como exemplo abaixo.

Trecho de Código 4.2: "Exemplo para o arquivo de configuração".

```
{
    "att_extras" : 2 ,
    "nomes" : [ "wind", "pressure" ],
    "datahora": "default",
    "mapColor": "blackAndWhite"
}
```

4.2 VTMV - Visualização de Trajetórias Multidimensionais

O sistema implementado com auxílio de duas bibliotecas em *JavaScript* bastante utilizadas em aplicações interativas, a D3.js e a *Leaflet*. Essa tecnologia interativa na web permitiu a disponibilidade de acesso em qualquer lugar, por causa da facilidade de ser usada na mais ampla gama de dispositivos através de navegadores recentes. Bem como facilitou com que os usuários explore suas consultas por si mesmo.

A interface do sistema pode ser vista na figura 4.2, no qual foi desenvolvida para viabilizar execução das pretensões de visualizações listadas anteriormente (Seção 3.1). Essa interface é integrada por quatro funcionalidades principais interligadas: o aspecto geográfico realizado pelo mapa, a visão da distribuição de frequências no histograma, a perspectiva de série temporal com gráfico de múltiplas linhas e o diagrama de Gantt com enfoque no tempo dos eventos e suas durações. Todos estes componentes são cuidadosamente detalhadas nas subseções deste capítulo.

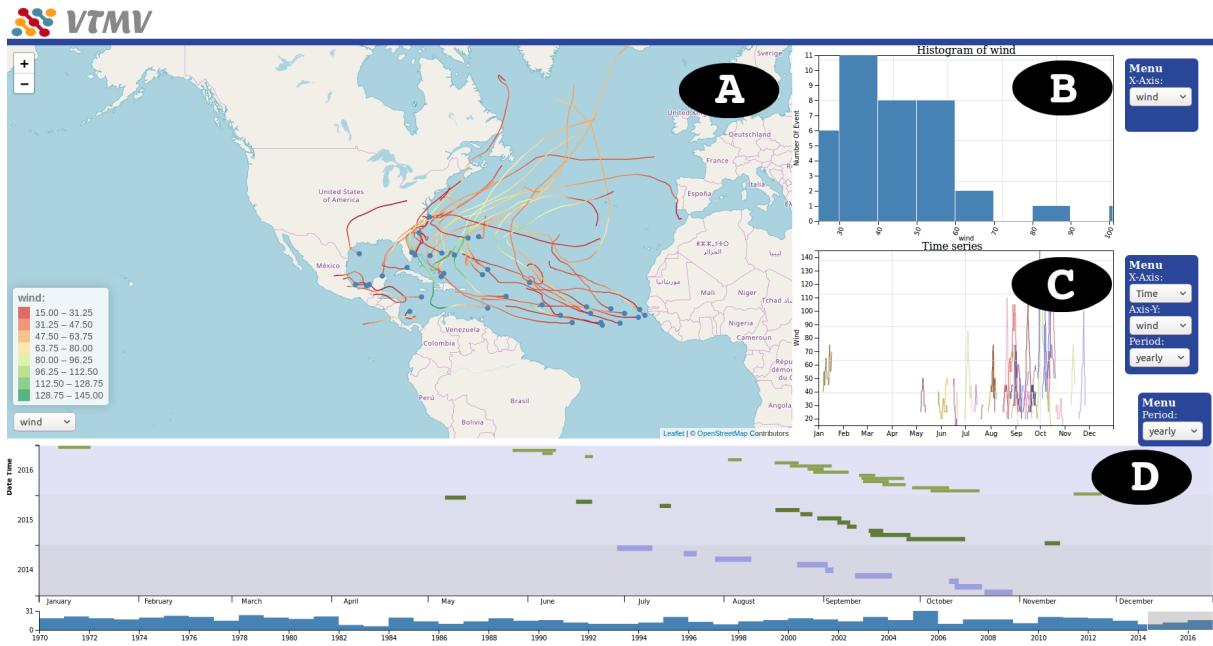


Figura 4.2: Interface do VTMV - Visualização de Trajetórias Multidimensionais é composta por 4 componentes: (A) o mapa, (B) o histograma, (C) o gráfico de múltiplas linhas e (D) o diagrama de Gantt.

4.2.1 Mapa

A funcionalidade do aspecto geográfico se concedeu através da ilustração de um mapa e das trajetórias presentes nos dados. Essa visão espacial auxilia para que a tarefa analítica (T2) seja cumprida. Uma vez que além de ter a existência do evento espacial ilustrado representado pela trajetória do objeto, também existe o valor presente daquela variável estudada. Essas linhas são formadas por segmentos, nos quais são concebidos pela média dos atributos dos pontos consecutivos de amostragem da trajetória. Em outras palavras, podemos explicar que os pontos tem seus atributos e os segmentos são formados por dois pontos consecutivos, logo os atributos do segmento são gerados através da média dos atributos dos pontos.

As trajetórias poderiam ser representadas apenas pelo aspecto espacial, no qual seria ilustrado exclusivamente a localização exata do evento pelas coordenadas geográficas. Assim esses eventos conectados formariam o movimento do objeto estudado. Porém, levando em consideração as propriedades numéricas, essa perspectiva bidimensional apresentada anteriormente se torna uma tridimensional. Dado que essas propriedades são variáveis contínuas nas quais representam valores quantitativos do objeto naquele instante estudado. A nova dimensão foi incluída através da codificação por cores dos valores dos atributos, em razão de ser uma técnica amplamente respeitada pelos pesquisadores. Visto que essa técnica apresenta propriedades seletiva e associativa (TOMINSKI et al. (2012)).

Ao transformar os valores dos atributos para cores é necessário um mapeamento apropriado, pois a informação visual deverá ser interpretável. Para isso, utilizamos o esquema de cores divergentes presente na biblioteca D3.js, conhecido como *RdYlGn*, no qual possibilita a

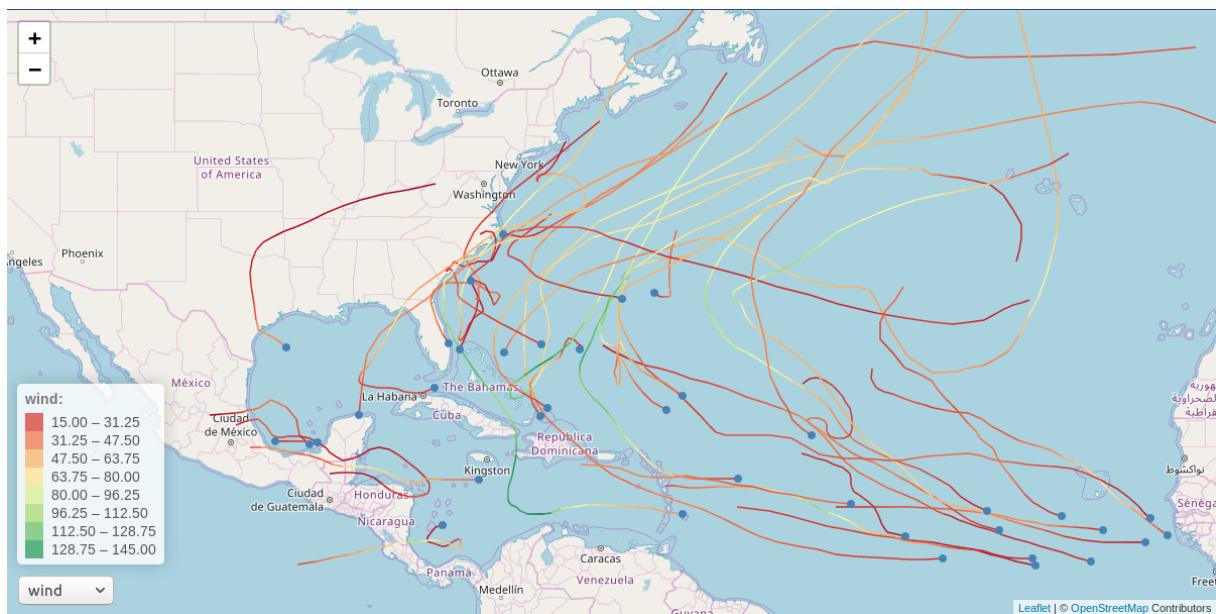


Figura 4.3: A visualização do mapa com a análise da variável wind.

interpolação dos valores contínuos e o esquema de cores sequenciais. Esse esquema *RdYlGn* foi escolhido com a finalidade de enfatizar os outliers e as mudanças de padrões presente nos dados, enquanto o sequencial é apropriado para valores numéricos (HARROWER; BREWER (2003)). Com isso se tem um conjunto de 8 cores exibidas na legenda, no quais os valores menores são mapeados para as cores avermelhada, já os montantes médios para as cores de tendência amarela e os maiores montantes têm a cor norteada para o verde. Esse valores mínimos e máximos presente na legenda pode variar de acordo com a escolha feita pelo usuário. Então os conjuntos e intervalos entre as cores alteram a depender da seleção feita, mas as cores se mantém a mesma, como podemos ver no canto esquerdo inferior do mapa. Também é possível interagir com a linhas através *combobox* localizado no canto inferior esquerdo da figura 4.3 e com isso modifica a variável observada no mapa. Outra maneira de interagir é ao clicar na legenda, com isso altera o esquema da legenda. Além das linhas, utilizamos outra maneira de mapear informações através dos círculos azuis presente no mapa. Eles servem para indicar o início do movimento da trajetória selecionadas.

Essa visualização só foi possível com o auxílio da biblioteca em *javascript Leaflet.js*, no qual é vastamente utilizada para utilitários na web com mapas profundamente interativos. Essa biblioteca possui farta quantidade de funcionalidades, um exemplo disso é o de *zoom and pan* (YOU et al., 2007) no qual foi empregue neste componente. Além disso, propicia o uso de camadas com os recursos de elementos gráficos tanto em *Scalable Vector Graphics (SVG)* como *CANVAS*. Além do mais possibilita a inserção de legendas no formato *HyperText Markup Language (HTML)* como foi utilizado, entre outras funcionalidades.

4.2.2 Histograma

Esse componente oferece suporte às tarefas analíticas(T1 e T2), através do formato gráfico que retrata a distribuição de frequências de um atributo (Figura 4.4).

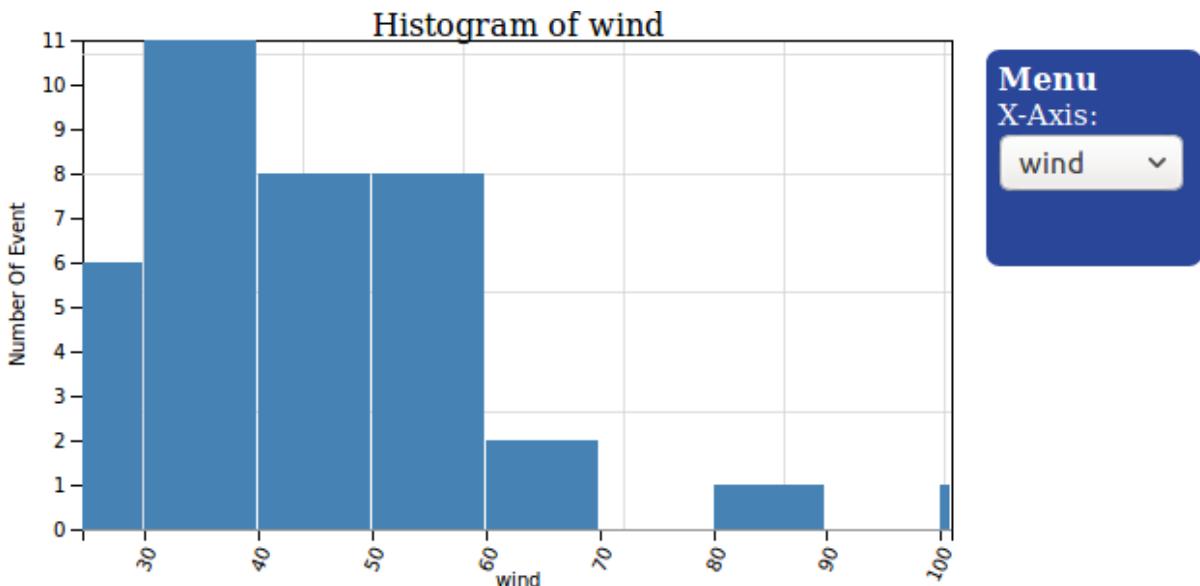


Figura 4.4: A visualização do histograma com a análise da variável wind.

Essa representação tem cada barra simbolizando a frequência de cada classe que foi previamente dividida. O histograma, além de dar uma noção da distribuição de probabilidade, fornece uma visão da concentração dos dados. Assim como quais os seus extremos e a presença de valores incomuns ou ausências de dados em determinado intervalo. Muitos estudos de visualização utilizam histograma para focar no ponto de vista estatístico das propriedades numéricas (CHEN; GUO; WANG, 2015). Esse gráfico reflete as informações dos dados filtrados pelo usuário no gráfico Gannt e pode ser utilizado para refinar mais a filtragem para a visualização geográfica e no gráfico de linhas. Essa filtragem pode ser vista através de dois eventos do mouse: o primeiro é ao passar o mouse por cima das barras,e por consequência, destacando as trajetórias que pertencem e o outro por meio do click no qual mostra apenas aqueles movimentos selecionadas. Adicionalmente existe um *combobox* na parte superior direita na qual é possível analisar a distribuição de frequências de outra variável.

4.2.3 Gráfico de linhas

A visualização do gráfico de linhas é comumente aplicado por pesquisadores para mostrar uma tendência nos dados em intervalos de tempo e assim pode se analisar a continuidade do evento no tempo ou comparar eventos no mesmo intervalo (KHAN; SHAH, 2011).

Esse gráfico facilita verificar as relações de séries temporais, pois ele possibilita a visualização das tendências, aclives, declives e oscilações. Em razão disso, esse componente foi utilizado para cooperar a solucionar as tarefas analíticas(T3 e T4), pois é útil para confrontar

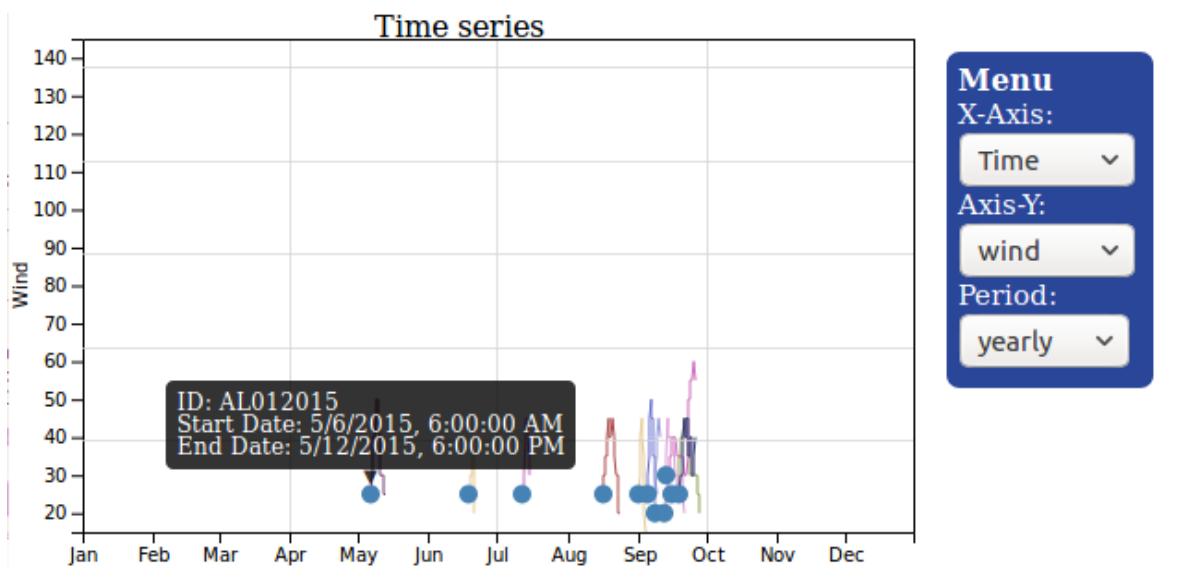


Figura 4.5: A visualização do gráfico de linhas com a análise da variável wind.

diferenças e semelhanças entre vários movimentos no mesmo período de tempo. Isso fez com que decidimos por escolher a igualdade em alguma unidade de tempo e assim pudéssemos ver as linhas empilhadas no mesmo intervalo de tempo, por exemplo no ano. Para isso existe o filtro com 3 opções(anual, mensal e diário) para o usuário selecionar qual a unidade de tempo que deseja. Assim, é possível comparar alguns eventos que divergem bastante na escala no tempo como se tivesse ocorrendo no mesmo instante.

Além disso, existe a possibilidade de comparar as multi-variáveis estudadas, pois é possível mudar os eixos estudados. Por exemplo, como padrão está sendo usado o tempo no eixo X, porém pode se utilizar outra dimensão para ser comparada (Figura 4.5). Essa visualização retrata também das filtragens realizadas nos gráficos Gannt e no histograma. Uma das interações presente nesse gráfico é quando as duas seleções forem realizadas, tanto em relação ao tempo(Gannt) como no intervalo de valores da variável(histograma). Em consequência disso, aparece círculos azuis no começo dos eventos, nos quais quando o mouse passa por cima exibe um *tooltip* com as principais informações sobre a trajetória e realça nos outros componentes.

4.2.4 Gráfico de Gantt

O diagrama de Gannt foi utilizado como base para atender a pretensão de visualização T4, no qual viabiliza o filtro pelo tempo. Esse gráfico, descrito em GANTT (1919), é um gráfico de barra no qual é utilizado para ilustrar o cronograma de um projeto. Seu principal propósito é mostrar de modo visual como se desenvolve o cronograma de um projeto e assim a compreensão ocorra sem dificuldade pelo leitor. Isso é possível através das barras que simbolizam uma linha do tempo e desse modo permitindo uma visão comprehensível do tempo.

Desse modo, utilizamos ele como base para ter a percepção da duração, do inicio, do fim e periodicidade dos eventos. Para isso, dividimos o componente em duas partes conectadas:

superior e inferior. Na parte superior (Parte A da Figura 4.6) exibimos as durações dos eventos em barras, o eixo X representa a unidade de tempo que foi uniformizada para o estudo e no eixo Y a unidade de tempo exato que ocorreu a trajetória. Essa unidade de tempo pode ser alterada com o filtro presente no canto superior direito, existindo 3 opções: anual, mensal e diário. Esse filtro altera tanto a parte superior como a parte inferior do diagrama.

Essa parte superior é alterada de acordo com a seleção feita na parte inferior. O tamanho dos retângulos é dado pelo período de duração do movimento e a altura do quadrilátero leva em consideração quantos eventos existe naquela unidade de tempo padronizada. Por exemplo, se a unidade de tempo padronizada o ano e a seleção tiver período de 3 anos, o eixo Y será dividido em 3 partes e a altura do retângulo dependerá de quantos eventos tem naquele ano.

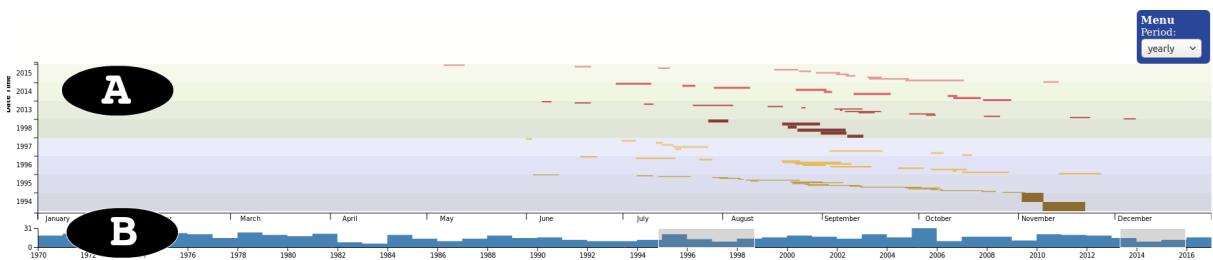


Figura 4.6: A visualização do gráfico de Gantt com múltiplos filtros. (A) Parte superior e (B) Parte Inferior

Já na parte inferior (Parte B da Figura 4.6) ficou responsável pelo filtro, no qual exibe todo o domínio de tempo presente nos dados. Como também exibe uma gráfico de frequência dos eventos por unidade de tempo, por exemplo ano. Dessa forma o usuário pode comparar eventos que ocorrem próximos na escala de tempo, bem como comparar eventos que estão afastados. Pois pode-se aplicar múltiplas seleções na escala de tempo (Figura 4.6). A manipulação do filtro é feita em cima do gráfico de frequência, no qual a seleção dos filtros é representada pelos retângulos cinzas. Esses retângulos podem ser alterados os seus intervalos através do movimento de clicar e arrastar, tanto nas bordas para diminuir ou aumentar na escala do tempo, como no centro para deslocar o período. Para se adicionar uma nova seleção é necessário apenas clicar e arrastar em uma área que não existe o retângulo.

5

Casos de Uso

A finalidade deste capítulo é evidenciar a utilidade do VTMV - Visualização de Trajetórias Multidimensionais através de conjuntos de dados reais já que se trata de uma ferramenta de análise exploratória. Para isso utilizamos dois conjuntos de dados citados anteriormente na Seção 3.2 para comprovar o poder do sistema na detecção de padrões gerais e tendências presentes em dados de diferentes domínios. Além disto, certificar a capacidade da ferramenta para identificar outliers, informações ausentes, errôneas ou inconsistentes nos dados.

5.1 Caso de uso: Dados dos furacões

Com o intuito de facilitar a análise e tirar um melhor proveito da ferramenta, reduzimos a quantidade de dados utilizado nesse estudo para os eventos que ocorrem após 1970. Pois os eventos que antecedem a essa data apresentavam problemas que serão mostrado nessa seção.

5.1.1 Identificar inconsistências nos dados e *outliers*

Ao utilizar a ferramenta e filtrar os dados para intervalos que antecederam o ano de 1987 é possível notar a presença de valores negativos para o atributo *Wind* (Figura 5.1). Enquanto para o atributo *Pressure* os valores negativos aparecem para as trajetórias que precedem o ano de 1991 (Figura 5.2).

Ao deslizar para intervalos mais antigos é possível constatar que a variável *Wind* foi melhor datada, pois apresenta um número muito inferior de casos errôneos ou inconsistentes comparado a variável *Pressure*. Esse valores podem ser visto facilmente no gráfico de linha, como também podem ser reparados pelo histograma e no mapa com a ajuda da legenda. Esses valores muitos discrepantes terminam criando um domínio muito extenso e dificultam outras análises naquele intervalo. O usuário pode minimizar essa adversidade através dos filtros do histograma.

Outro conhecimento notável no atributo *Wind* é que quando a média da trajetória é maior que 30 nós não ocorrem mais inconsistência ou ausência. Enquanto que para *Pressure*

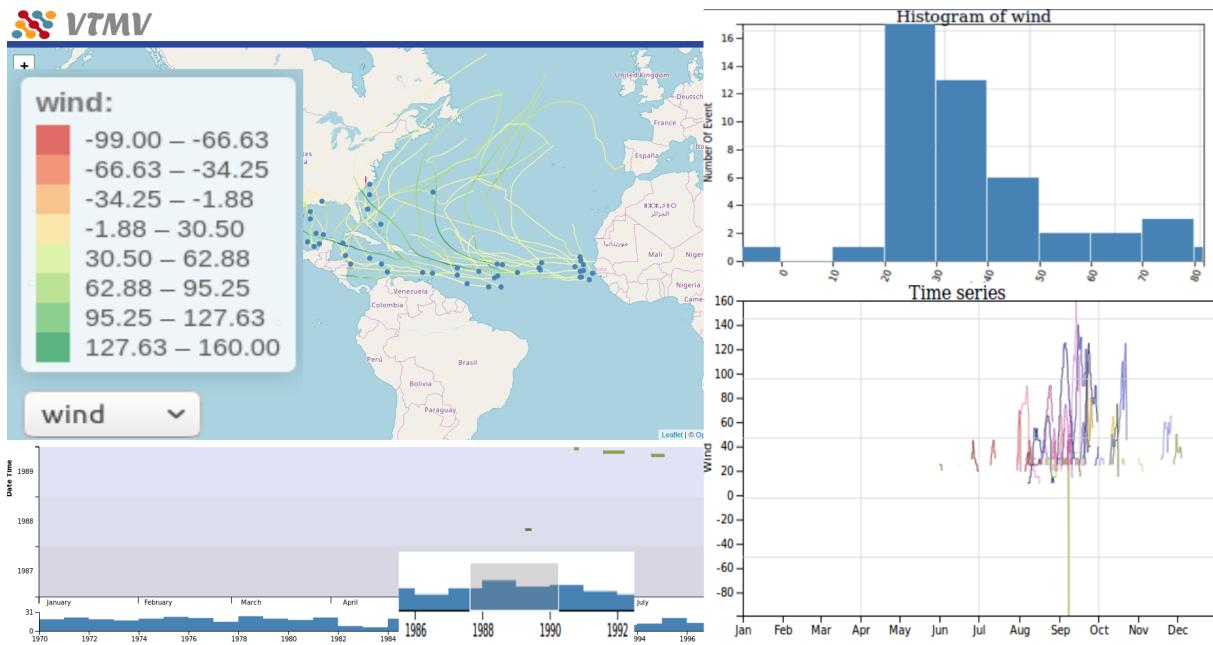


Figura 5.1: Intervalo que as incoerências no valores do atributo *Wind*

não aparece essa distinção para algum intervalo, entretanto, todos as trajetórias que tem a variável *Wind* com esse problema de inconsistência tem valor médio da *Pressure* abaixo de -800. Além disso, os atributos para localização não apresentaram nenhuma insensatez com relação ao posicionamento geográfico e o rumo das suas trajetórias.

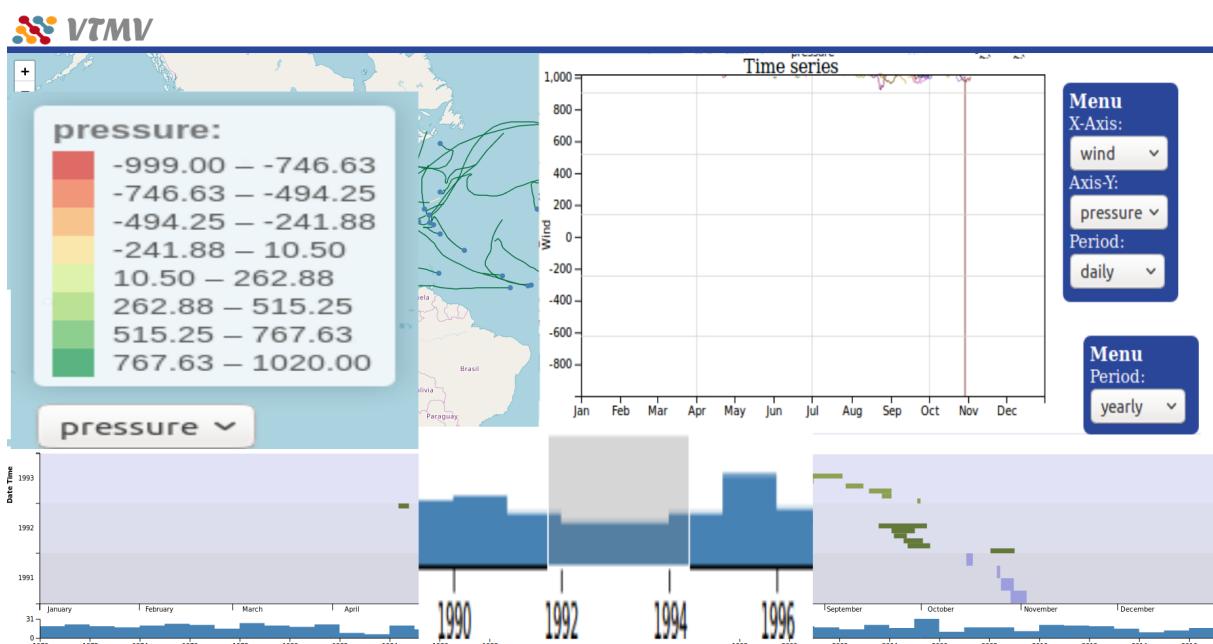


Figura 5.2: Intervalo que as incoerências no valores do atributo *Pressure*

5.1.2 Padrões e extração de conhecimentos das variáveis

Uma informação que foi possível extrair foi a que os furacões praticamente não ocorrem quando é inverno no hemisfério norte. Isso foi possível através da observação gráfico de Gantt (Figura 5.3), no qual exibiu apenas 3 ocorrências nesse período. Esse intervalo começa aproximadamente no dia 21 de dezembro e termina no dia 21 de março. Esses fenômenos meteorológicos começam a aparecer na primavera, porém com um número bem discreto comparado com as outras estações.

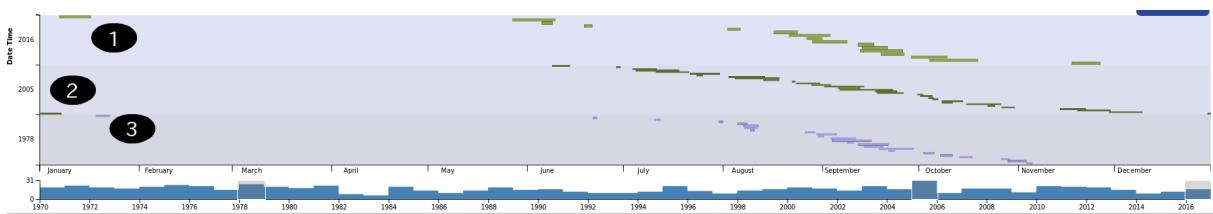


Figura 5.3: Os períodos nos quais apresentam um baixo número de ocorrência no ano.

Analizando a relação entre as variáveis por meio do gráfico de linhas e seus filtros foi possível constatar uma relação das grandezas da *Pressure* e *Wind*. Assim quando o valor da *Pressure* aumenta o valor do *Wind* diminui e vice-versa (Figura 5.4). Porém não apresenta um comportamento proporcional, pois apresenta algumas variações de comportamento. Apenas de ser possível ver a dependência entre as duas variáveis.

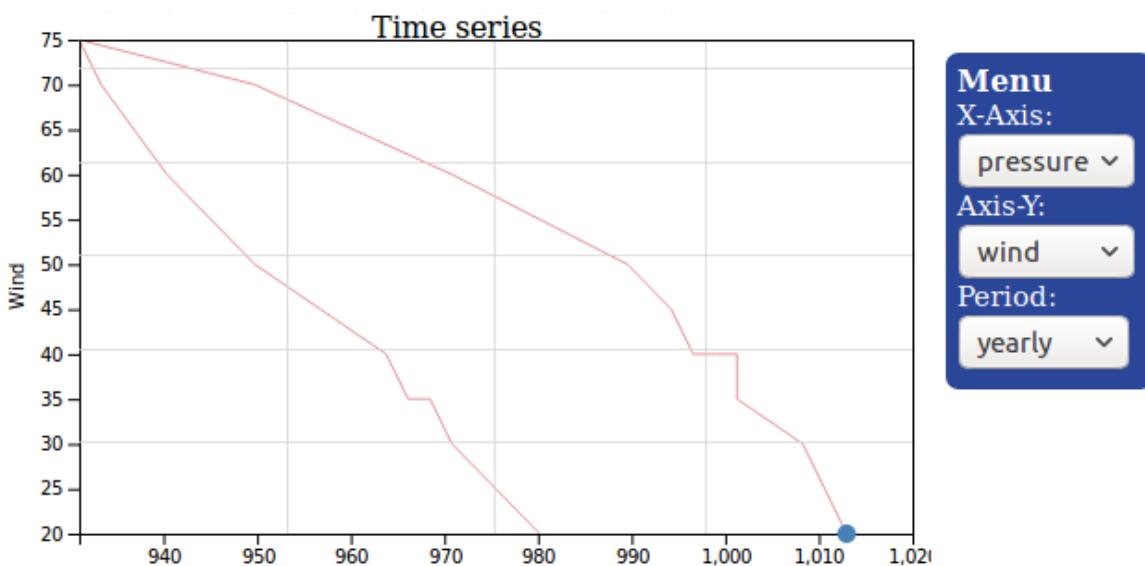


Figura 5.4: Relação de dependência entre as grandezas *Pressure* e *Wind*.

Pode-se notar também que esses fenômenos meteorológicos tendem a começar em um determinado intervalo para as duas variáveis presente no estudo. Foi possível perceber isso utilizando o mesmo padrão do gráfico anterior, no qual o início desses fenômenos (Círculo azul) se aglomeraram na mesma região (Figura 5.5).

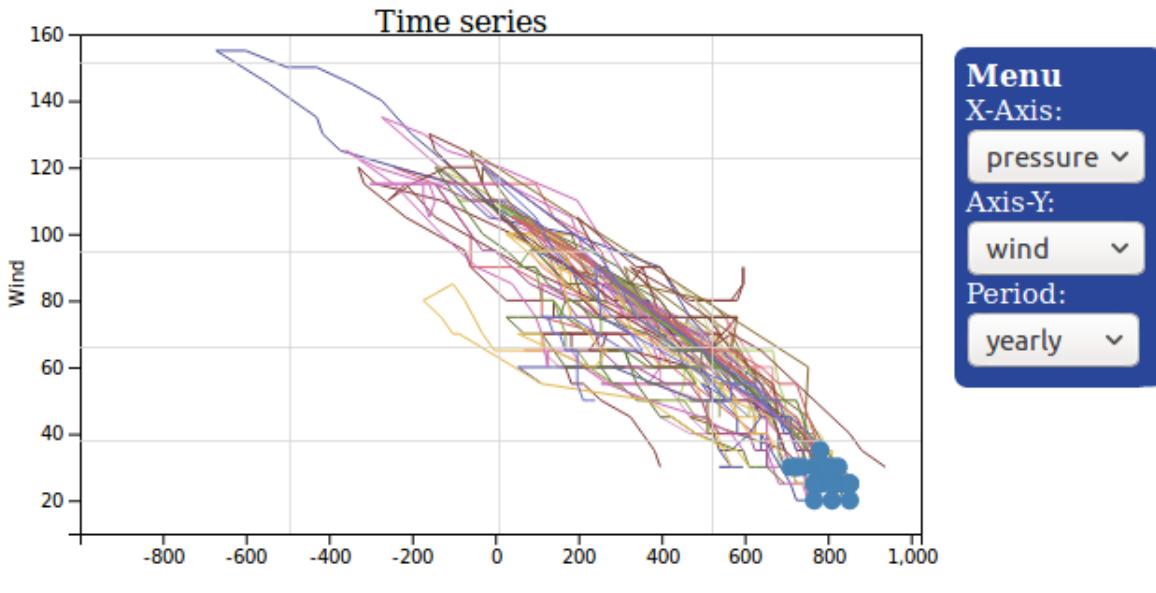


Figura 5.5: As trajetórias desses dados tendem a iniciar no mesmo região do gráficos.

Outro conhecimento que foi possível extrair através desse gráfico de linhas é o comportamento de picos e vales presente no gráficos tanto da *Pressure* quanto do *Wind*. Mostrando a instabilidade presente nessas variáveis e a quantidade de variações total na tendência da grandeza ao longo do tempo (Figura 5.6).

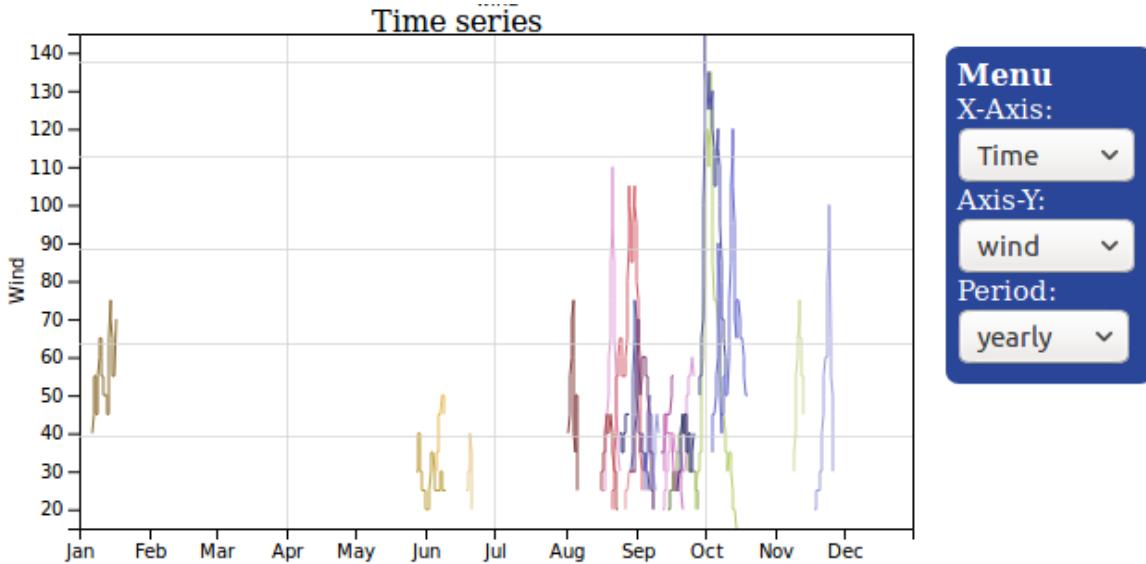


Figura 5.6: Picos e vales presente na grandeza *Wind* ao longo do tempo.

Ao explorar os furacões que tiveram média para a variável *Wind* acima de 50 nós, podemos constar que quase a totalidade de eventos ocorrem entre os períodos entre agosto e setembro (Figura 5.7). Além disso, é capaz de notar nessas trajetórias que apresentam essa condição uma tendência. Essa tendência é o início do evento que ocorre próxima a uma mesma região, com a direção ao subcontinente da América do Norte e valores da grandeza *Wind* aproximadamente entre 20 e 40 nós. Bem como, ao se aproximar do continente podemos ver que ocorre um au-

mento dessa grandeza e a mudança da direção da trajetória, indo de encontro a Europa e logo após diminuindo o valor dessa grandeza.

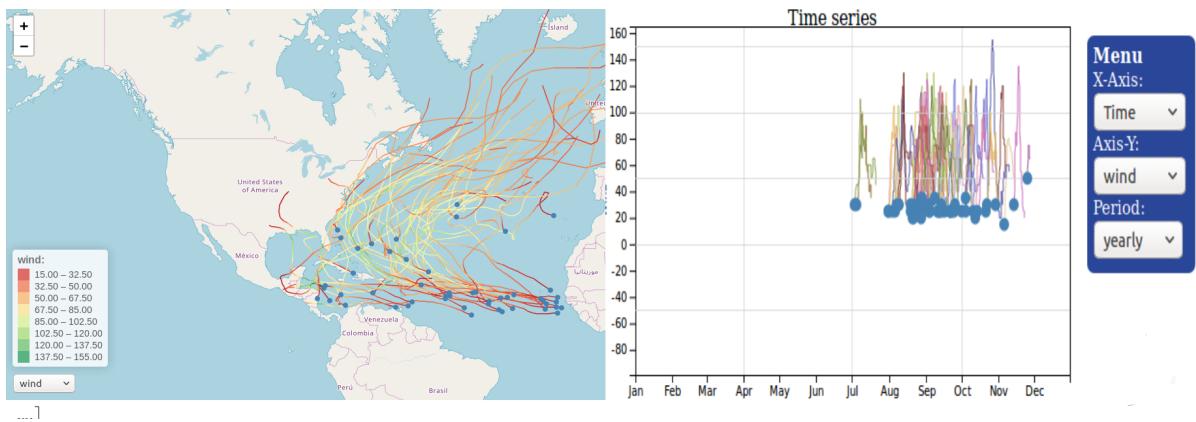


Figura 5.7: Os furacões que apresentam comportamento similares quando tem média a média do *Wind* acima de 50 nós.

5.2 Caso de uso: Dados dos ciclistas

Essa seção será destinada a análise da base de dados dos ciclistas na ferramenta. Como essa base apresenta um número muito grande de dados, decidimos reduzir o número de pontos de uma trajetória para um melhor comportamento da ferramenta. Para isso ignoramos alguns pontos que foram registrados na rota, aumentando assim o intervalo entre os pontos registrados. Consequentemente podendo explorar todas as trajetórias de um atleta com agilidade e viabilidade no sistema.

Arbitrariamente selecionamos o atleta número 2 para fazer o estudo e optamos pelas trajetórias que continham o atributo *heartRates*. Com isso os dados que anteriormente tinham 844 trajetórias e agora contêm 642. Porém ainda apresenta-se bastante volumoso, com 730.6 *Megabytes* e mais de 3 milhões de pontos por causa da frequência da coleta de informação utilizada. Essa coleta em certas ocasiões chega a ser de 1Hz. A fim de reduzir essa quantidade, retiramos 99 pontos entre dois instantes que seriam utilizadas. Assim uma trajetória de 4000 pontos dados, agora passa a ter aproximadamente 40 momentos utilizados no estudo. Diminuindo assim o tamanho para cerca de 32 mil pontos e 7.4 *Megabytes*.

5.2.1 Identificar inconsistências nos dados e *outliers*

O VTMV - Visualização de Trajetórias Multidimensionais possibilitou a visualização de alguns eventos que apresentaram comportamentos suspeitos quanto a sua veracidade. Um desses eventos é uma trajetória que durou mais de um mês, poucos pontos coletados e a dimensão *elevation* negativa, apontando ser algum erro na captura da atividade do ciclista. Pois essa

trajetória apresenta uma linha reta muito longa, destoando dos outros dados presentes no estudo como pode ser visto na Figura 5.8.

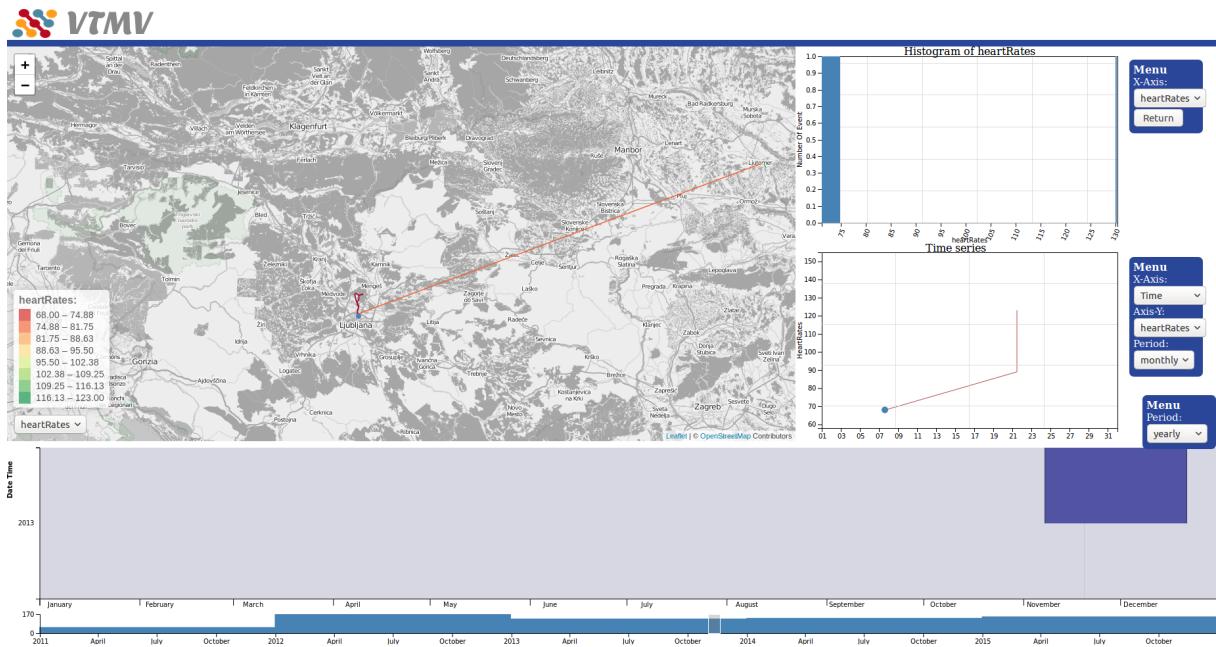


Figura 5.8: Atividade do atleta que durou mais de 1 mês.

Outro fato que gerou dúvida foi os valores da dimensão *heartRates* abaixo de 60 bpm. Pois esses valores são chamados de bradicardia e ocorrem quando o indivíduo está em repouso (FILIPE EUFRASIO FERREIRA, 2010). Diferentemente dos dados desse caso de uso, no qual apresenta trajetórias no qual o indivíduo está praticando atividade física e exibem valores abaixo de 40bpm (Figura 5.9).

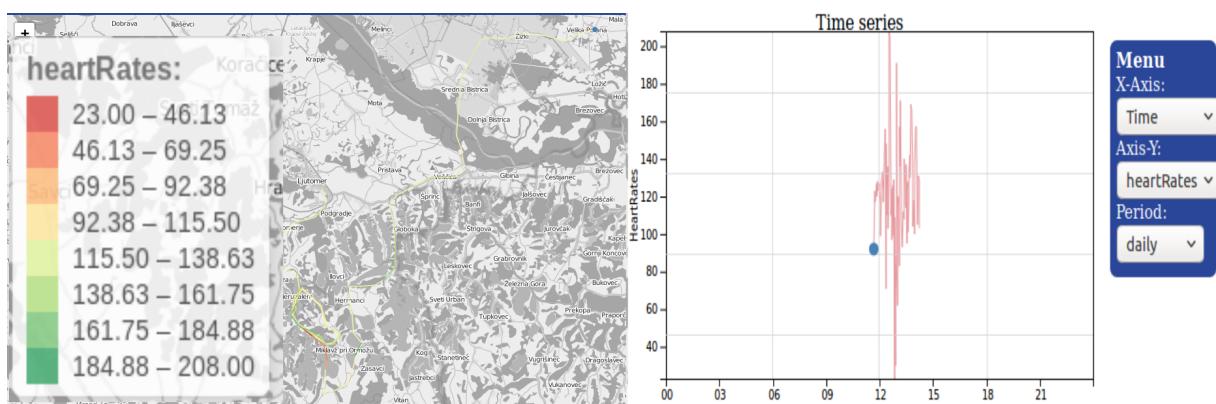


Figura 5.9: Atividade do atleta que teve *heartRates* muito baixo.

Além disso não foi possível reparar nenhuma inconsistência a mais presente nos dados, acreditamos que os dados cedidos pelos atletas apresentem uma filtragem de qualidade ou o método utilizado para melhorar o estudo tenha filtrado as inconsistências.

5.2.2 Padrões e extração de conhecimentos das variáveis

Uma das informações extraída através da visualização é a da possível residência desse atleta, no qual apresenta um grande número de eventos próximo a cidade Liublana e ao município Ljutomer na Eslovénia (Figura 5.10). Pois como atleta, ele não deve se deslocar amplas distâncias para treinar, dado que os treinos ocorrem com alta frequência.

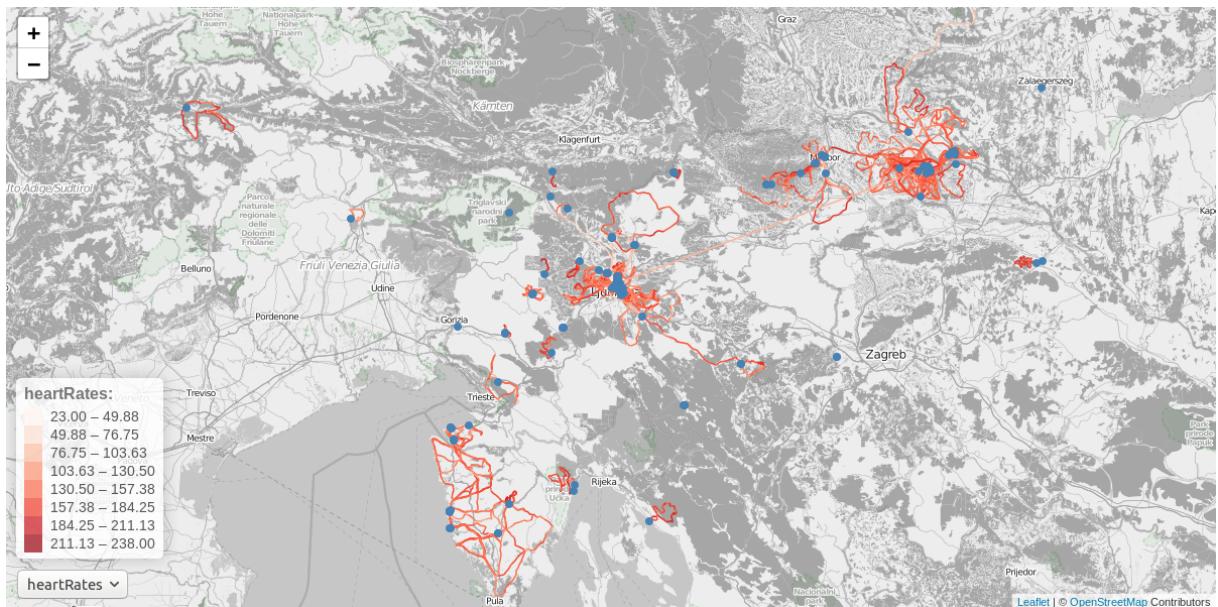


Figura 5.10: Provável residência do atleta pelo alto número de eventos.

Outro padrão que foi possível reparar foram os horários das atividades praticadas pelo atleta que na sua grande maioria aconteceram entre os intervalos de 06:00 e 18:00. Além disso, podemos ver que essas atividades têm uma duração maior que 2 horas (Figura 5.11).

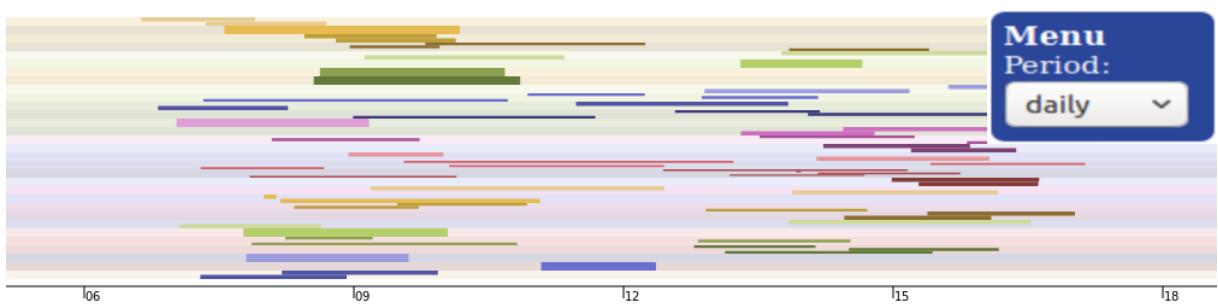


Figura 5.11: Frequência dos períodos das trajetórias do atleta.

No que se refere a frequência das atividades do ciclista, podemos notar da redução dos números de atividades registradas quando está no período do final do ano, entre outubro e dezembro (Figura 5.12). Isso ocorre em todos os 5 anos analisados nesse estudo, isso supostamente ocorre por se tratar do período de alta precipitação e temperaturas baixas na região da trajetórias.

Outra tendência extraída foi a da grande variação da dimensão *elevation*, o que sugere que o atleta nos seus treinamentos utiliza grande quantidade de aclives e declives (Figura 5.13).



Figura 5.12: Redução das atividades do ciclista no ano.

Esses treinamentos coincidem também na localização, pois acontecem em áreas montanhosas e de matas. Ao que tudo indica, esse tipo de treinamento pode determinar que a modalidade do ciclista seja o *Mountain Bike*.

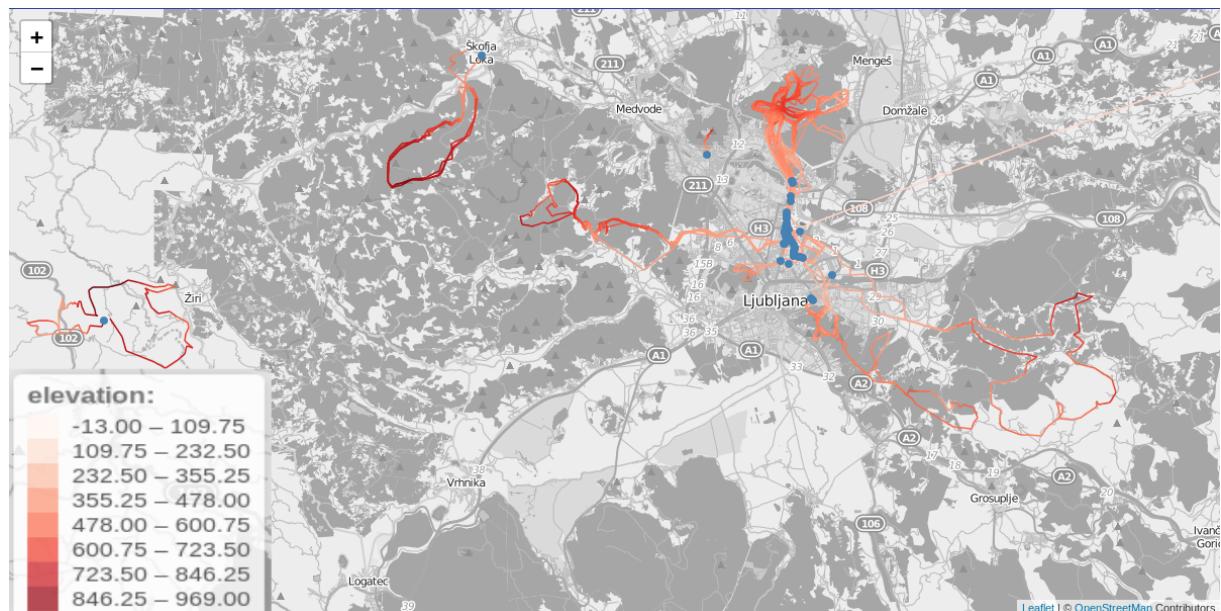


Figura 5.13: Rotas com grande variação da dimensão *elevation*.

E por último, foi possível notar que as trajetórias iniciam em baixo valores do *heartRates* e *elevation*, e com o tempo ocorre primeiramente o aumento do *heartRates* e logo após o da *elevation* (Figura 5.14). Isso aparenta ser um período de aquecimento para logo começar o treino. Logo em seguida, a *elevation* se mantém constante por um momento e ocorre a redução do *heartRates*. Por suposição, essa mudança aparenta como se o corpo se adapta com a atividade realizada naquela altura.

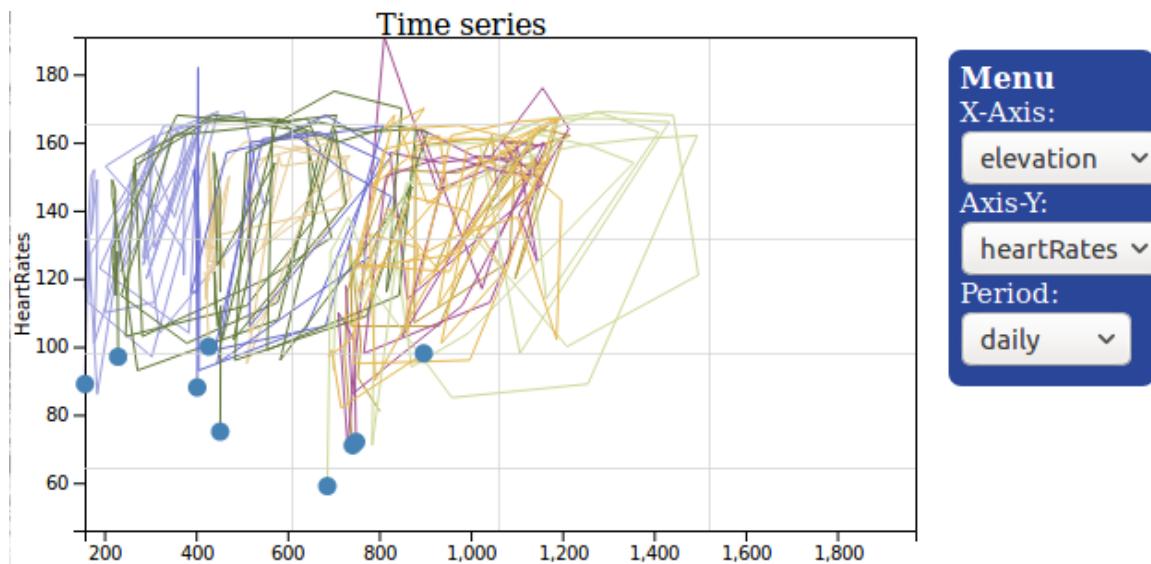


Figura 5.14: Análise das duas dimensões: *elevation* e *heartRates*.

6

Conclusão

Esse trabalho propôs a elaboração de uma ferramenta web de visualização interativa de dados denominada VTMV - Visualização de Trajetórias Multidimensionais, na qual se centraliza na exploração visual de dados de trajetórias multidimensionais. Essa ferramenta utiliza diversas técnicas de visualização de dados para analisar facilmente dados complexos e com isso contribuir com o usuário na extração de conhecimentos em tempo real. O VTMV - Visualização de Trajetórias Multidimensionais dispõe um design simples, porém cumpre excelentemente com o objetivo proposto pois realizou todas as tarefas analíticas (Seção 4.1.2) com dados de natureza diversa como descrito em todo capítulo 5. Assim com esse meio de exploração, o usuário adquire um novo recurso para lhe dar suporte nas suas pesquisas sobre trajetórias multidimensionais.

6.1 Melhorias e Futuros Trabalhos

Embora tenha alcançado os seus propósitos, o VTMV - Visualização de Trajetórias Multidimensionais ainda tem um vasto leque de melhorias. Essas melhorias seriam desde a existência de novas visualizações em novas camadas, como por exemplo boxplot e coordenadas paralelas, como a adição de novas formas de filtros como a possibilidade de filtrar por região no mapa ou por diversos valores das dimensões. Além disso possibilitar a leitura dos dados através de um banco de dados, pois a ferramenta apresenta limitação quanto ao tamanho do arquivo de dados, visto que as trajetórias da base do ciclista acaba sobrecarregando o sistema. Como também, aumentar o poder de exploração do gráfico de linha com a interação de *Pan & Zoom* e filtragens. Incorporar novos filtros para a parte superior do gráfico de Gannt. Adicionar na ferramenta a capacidade de trabalhar com variáveis qualitativas por meio das configurações. Por fim, incorporar um *dashboard* na parte superior da tela, resumindo algumas informações do estudo, por exemplo, a quantidade total de dados, o número filtrado e seu percentual, o período exibido, entre outras informações.

Referências

- AL., S. R. et. A collection of sport activity files for data analysis and data mining 2016a. **Technical report 0101, University of Ljubljana and University of Maribor 2016a**, [S.l.], 2016.
- ANDRIENKO, G. et al. Space, time and visual analytics. **International Journal of Geographical Information Science**, [S.l.], v.24, n.10, p.1577–1600, 2010.
- BOYANDIN, I. et al. Flowstrates: an approach for visual exploration of temporal origin-destination data. In: COMPUTER GRAPHICS FORUM. **Anais...** [S.l.: s.n.], 2011. v.30, n.3, p.971–980.
- CHEN, W.; GUO, F.; WANG, F. Y. A Survey of Traffic Data Visualization. **IEEE Transactions on Intelligent Transportation Systems**, [S.l.], v.16, n.6, p.2970–2984, Dec 2015.
- FERREIRA, N. **Visual analytics techniques for exploration of spatiotemporal data**. 2015. Tese (Doutorado em Ciência da Computação) — POLYTECHNIC INSTITUTE OF NEW YORK UNIVERSITY.
- FERREIRA, N. et al. Birdvis: visualizing and understanding bird populations. **IEEE Transactions on Visualization and Computer Graphics**, [S.l.], v.17, n.12, p.2374–2383, 2011.
- FERREIRA, N. et al. Vector Field K-means: clustering trajectories by fitting multiple vector fields. In: EUROGRAPHICS CONFERENCE ON VISUALIZATION, 15., Chichester, UK. **Proceedings...** The Eurographs Association &; John Wiley & Sons: Ltd., 2013. p.201–210. (EuroVis '13).
- FERREIRA, N. et al. Visual exploration of big spatio-temporal urban data: a study of new york city taxi trips. **IEEE Transactions on Visualization and Computer Graphics**, [S.l.], v.19, n.12, p.2149–2158, 2013.
- FILIPE EUFRASIO FERREIRA, E. Coracao de atleta. , [S.l.], 01 2010.
- GANTT, H. L. **Organizing for work**. [S.l.]: Harcourt, Brace and Howe, 1919.
- HARROWER, M.; BREWER, C. A. ColorBrewer.org: an online tool for selecting colour schemes for maps. **The Cartographic Journal**, [S.l.], v.40, n.1, p.27–37, 2003.
- HURDAT2: the revised atlantic hurricane database. 2017.
- KHAN, M.; SHAH, S. Data and Information Visualization Methods, and Interactive Mechanisms: a survey. , [S.l.], v.34, p.1–14, 12 2011.
- KRÜGER, R. et al. TrajectoryLenses A Set-based Filtering and Exploration Technique for Long-term Trajectory Data. **Computer Graphics Forum**, [S.l.], v.32, n.3pt4, p.451–460, 2013.
- STACKING-BASED VISUALIZATION OF TRAJECTORY ATTRIBUTE DATA. **Anais...** [S.l.: s.n.], 2012. v.18(12), p.2565–2574.

YOU, M. et al. A usability evaluation of web map zoom and pan functions. **International Journal of Design**, [S.l.], v.1, n.1, 2007.

ZHU, Y.; ZHONG, N.; XIONG, Y. Data Explosion, Data Nature and Dataology. In: INTERNATIONAL CONFERENCE ON BRAIN INFORMATICS, 2009., Berlin, Heidelberg. **Proceedings...** Springer-Verlag, 2009. p.147–158. (BI'09).