

INSTITUTO TECNOLÓGICO DE AERONÁUTICA



Carlos Alexandre Silva dos Santos

**MEDIDAS DE EFICIÊNCIA DO TRÁFEGO AÉREO:
UMA COMPARAÇÃO ENTRE TRAJETÓRIAS
PREVISTAS E REAIS**

Trabalho de Graduação
2021

Curso de Engenharia Civil-Aeronáutica

Carlos Alexandre Silva dos Santos

**MEDIDAS DE EFICIÊNCIA DO TRÁFEGO AÉREO:
UMA COMPARAÇÃO ENTRE TRAJETÓRIAS
PREVISTAS E REAIS**

Orientador

Prof. Dr. Marcelo Xavier Guterres (ITA)

ENGENHARIA CIVIL-AERONÁUTICA

**SÃO JOSÉ DOS CAMPOS
INSTITUTO TECNOLÓGICO DE AERONÁUTICA**

Dados Internacionais de Catalogação-na-Publicação (CIP)
Divisão de Informação e Documentação

Santos, Carlos Alexandre Silva dos
Medidas de eficiência do tráfego aéreo: uma comparação entre trajetórias previstas e reais /
Carlos Alexandre Silva dos Santos.
São José dos Campos, 2021.
47f.

Trabalho de Graduação – Curso de Engenharia Civil-Aeronáutica– Instituto Tecnológico de
Aeronáutica, 2021. Orientador: Prof. Dr. Marcelo Xavier Guterres.

1. Controle de tráfego aéreo. 2. Eficiência. 3. Monitoramento. 4. Aeroportos. 5. Aprendizagem
(inteligência artificial). 6. Transportes. I. Instituto Tecnológico de Aeronáutica. II. Título.

REFERÊNCIA BIBLIOGRÁFICA

SANTOS, Carlos Alexandre Silva dos. **Medidas de eficiência do tráfego aéreo: uma comparação entre trajetórias previstas e reais**. 2021. 47f. Trabalho de Conclusão de Curso (Graduação) – Instituto Tecnológico de Aeronáutica, São José dos Campos.

CESSÃO DE DIREITOS

NOME DO AUTOR: Carlos Alexandre Silva dos Santos

TÍTULO DO TRABALHO: Medidas de eficiência do tráfego aéreo: uma comparação entre trajetórias previstas e reais.

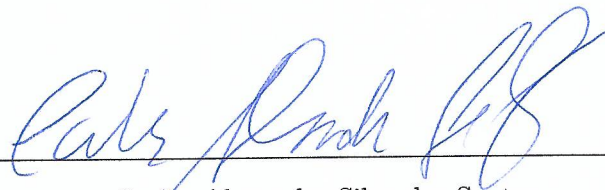
TIPO DO TRABALHO/ANO: Trabalho de Conclusão de Curso (Graduação) / 2021

É concedida ao Instituto Tecnológico de Aeronáutica permissão para reproduzir cópias deste trabalho de graduação e para emprestar ou vender cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte deste trabalho de graduação pode ser reproduzida sem a autorização do autor.

Carlos Alexandre Silva dos Santos
Rua. Figueiredo de Magalhães, 442
22.031-012 – Rio de Janeiro–RJ

MEDIDAS DE EFICIÊNCIA DO TRÁFEGO AÉREO: UMA COMPARAÇÃO ENTRE TRAJETÓRIAS PREVISTAS E REAIS

Essa publicação foi aceita como Relatório Final de Trabalho de Graduação

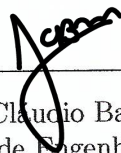


Carlos Alexandre Silva dos Santos

Autor

Marcelo Xavier Guterres (ITA)

Orientador



Prof. Dr. João Claudio Bassan de Moraes
Coordenador do Curso de Engenharia Civil-Aeronáutica

São José dos Campos, 17 de novembro de 2021.

Aos meus pais: Maria José e Francisco
Carlos.

Agradecimentos

Gostaria de agradecer primeiramente à Deus, fonte e fundamento de todas as coisas, aos meus pais, meus colegas de faculdade, meus professores, ao ITA por ter me oferecido educação gratuita e de qualidade.

Ao projeto Primeira Chance, em especial à dona Armênia, que ainda em Fortaleza me deu suporte e apoio para que eu ingressasse numa ótima escola do Ceará sem a qual nada disso estaria acontecendo.

Aos professores do ITA, em especial ao professor Armando um grande mentor em todos os âmbitos e durante toda a faculdade, ao meu orientador, professor Guterres, por toda sua paciência e ensinamentos durante esse longo trabalho, ao Wallace por ter me guiado nos primeiros passos sobre esse vasto assunto, sem o qual esse trabalho talvez não tivesse sido produzido.

*"The night is darkest just before the dawn.
And I promise you, the dawn is coming."*
— HARVEY DENT/TWO-FACE

Resumo

Com a crescente demanda por transporte aéreo, preocupações com eficiência de voo tem crescido igualmente. A adoção de sistemas como ADS-B permitiram o monitoramento contínuo da grande maioria dos voos comerciais realizados em nível mundial. A partir da base de dados fornecidos pelo ADS-B e através de sites de monitoramento como FlightRadar24 além da crescente aplicação de métodos de machine-learning em situações com densa base de dados vem aumentando continuamente em especial para cálculos de eficiência e tentativas em se minimizá-lo. Este trabalho avalia a eficiência de pouso de voos baseado em duas métricas distintas, a primeira através de rotas nominais obtidas a partir do algoritmo DBSCAN de clusterização e através da comparação das rotas reais com as aerovias, rotas teóricas em que os pilotos se baseiam em seus voos. O parâmetro de eficiência analisado foi o HTE, e uma análise estatística sobre as amostras em cada caso, cluster e aerovias, foram comparados via teste de Kolmogorov-Smirnov a fim de se avaliar se tais valores partiram de uma mesma distribuição. Por fim, nosso caso de estudo foi o aeroporto de São Paulo/Congonhas (SBSP) onde se analisou os voos de chegada nesse aeroporto entre setembro e dezembro de 2019, o resultado do teste K-S se mostrou significativo estatisticamente e podemos com alto grau de confiança rejeitar a hipótese de as métricas serem equivalentes.

Abstract

With the growing demand for air transport, concerns about flight efficiency have grown equally. The adoption of systems such as ADS-B allowed the continuous monitoring of the vast majority of commercial flights performed worldwide. From the database provided by ADS-B and through monitoring sites such as FlightRadar24 in addition to the increasing application of machine-learning methods in situations with the dense database has been continuously increasing in particular for calculations of efficiency and attempts to minimize. This work evaluates the efficiency of the landing of flights based on two distinct metrics, the first through nominal routes obtained from the DBSCAN clustering algorithm and by comparing the actual routes with the airways, theoretical routes on which pilots are based their flights. The efficiency parameter analyzed was the HTE, and statistical analysis of the samples in each case, cluster, and airways, were compared a Kolmogorov-Smirnov test in order to evaluate whether these values came from the same distribution. Finally, our study case was the airport of São Paulo/Congonhas (SBSB) where we analyzed the arrival flights at this airport between September and December 2019, the results of the K-S test proved statistically significant and we can with a high degree of confidence reject the hypothesis that the metrics are equivalent.

Lista de Figuras

FIGURA 2.1 – Perfil das aerovias. Fonte: (DECEA, 2015)	18
FIGURA 2.2 – Cartas de Rota (ENRC). Fonte: DECEA	19
FIGURA 2.3 – Carta de Rota ENRC-L2	20
FIGURA 3.1 – Fixos ao redor de 40nm de CGH	24
FIGURA 3.2 – Carta de aproximação: RNAV DOSPI 1A RWY 17R/17L	24
FIGURA 3.3 – Rotas obtidas	25
FIGURA 3.4 – Região do aeroporto de CGH com raio de 40nm.	27
FIGURA 3.5 – Representação visual dos conceitos principais do DBSCAN. Fonte: (MURÇA <i>et al.</i> , 2020)	30
FIGURA 3.6 – Esquema da identificação do voo para posterior clusterização. Fonte: (GARIEL <i>et al.</i> , 2011)	31
FIGURA 3.7 – Exemplo de discretização da trajetória. Fonte: (MCLEAN, 2021)	32
FIGURA 3.8 – Distribuição dos voos	32
FIGURA 3.9 – Esquema da interpolação linear. Fonte: (BRETT, 2016)	33
FIGURA 3.10 – Reamostragem da trajetória de cada voo. Fonte: (MURÇA, 2018)	33
FIGURA 3.11 – Plot kNNdist para escolha do eps	34
FIGURA 3.12 – Coeficiente de silhouette	35
FIGURA 3.13 – Histograma da clusterização	36
FIGURA 4.1 – Representação de cada voo e seu cluster	37
FIGURA 4.2 – Representação de cada cluster obtido através da média aritmética dos voos	38
FIGURA 4.3 – Boxplot da medida de eficiência HTE para cada cluster	39

FIGURA 4.4 – Boxplot da medida de eficiência HTE para cada aerovia	39
FIGURA 4.5 – Boxplot da medida de eficiência HTE para as aerovias e clusters . .	40
FIGURA 4.6 – Aerovia 1, cluster 3	40
FIGURA 4.7 – Aerovias 4 e 5, cluster 2	41
FIGURA 4.8 – Aerovia 6, cluster 1	41
FIGURA 4.9 – Aerovia 13, cluster 4	42
FIGURA 4.10 – Exemplo da avaliação no teste de Kolmogorov-Smirnov. Fonte: (AR- SENAULT, 2017)	43

Lista de Tabelas

TABELA 3.1 – Pontos fixos do espaço aéreo brasileiro	23
TABELA 3.3 – Primeiras linhas da base de dados	26
TABELA 4.1 – Resultado do teste K-S para cada par aerovia x cluster	43

Lista de Abreviaturas e Siglas

ADS-B	Automatic dependent surveillance - broadcast
DBSCAN	Densit-based spatial clustering of applications with noise
HTE	Horizontal Traffic Efficiency
SBSP	Código ICAO do aeroporto de São Paulo/Congonhas
K-S	Teste Kolmogorov-Smirnov
ANAC	Agência Nacional de Aviação Civil
VFR	Visual flight rules
IFR	Instrument flight rules
ENRC	Enroute charts
GPC	Global Positioning System
BIRCH	Balanced iterative reducing and clustering using hierarchies
CGH	Código IATA do aeroporto de São Paulo/Congonhas
GRU	Código IATA do aeroporto de Guarulhos
DECEA	Departamento de controle do espaço aéreo
STAR	Standard instrument arrival
IATA	International Air Transport Association
ICAO	International Civil Aviation Organization
UTC	Coordinated Universal Time
DAE	Denosing Autoencoder
GMM	Gaussian Mixture Modeling
HAC	Hierarchical Agglomerative Clustering

Sumário

1	INTRODUÇÃO	15
1.1	Apresentação do problema	15
1.2	Objetivo Geral	16
1.3	Justificativa do Trabalho	16
1.4	Escopo	16
2	REVISÃO DA LITERATURA	17
2.1	Transporte Aéreo	17
2.1.1	Aerovias	17
2.1.2	Cartas de voo	18
2.2	Machine Learning no setor de transporte aéreo	20
3	METODOLOGIA	23
3.1	Escolha das aerovias	23
3.2	Análise dos dados	25
3.3	As medidas de eficiência de voo	27
3.4	Método de Clusterização	28
3.4.1	Apresentação dos Métodos	28
3.5	Tratamento dos dados	30
3.5.1	Escolha dos Parâmetros	34
4	RESULTADOS E DISCUSSÕES	37
4.1	Análise qualitativa dos resultados	38
4.1.1	Comparação Final	39

SUMÁRIO	xiv
4.2 Análise quantitativa dos resultados	42
5 CONCLUSÃO	45
REFERÊNCIAS	46

1 Introdução

1.1 Apresentação do problema

O transporte aéreo é um importante fator na economia de qualquer país sobretudo no Brasil, com dimensões continentais. Só em 2019, 805 mil voos foram realizados no mercado doméstico, 95,3 milhões de passageiros pagos e 456 mil toneladas de carga paga e correio foram transportados (ANAC, 2019b). Além disso por volta de 70% do transporte interestadual no país é feito de modo aéreo (ANAC, 2019a). Dito isso a procura de soluções que melhorem a eficiência de voo nos mais diversos aspectos tem impactos não apenas nos faturamentos das empresas, como também na nossa capacidade aeroportuária de oferecer mais serviços de melhor qualidade e menor custo.

O desenvolvimento recente de técnicas de *machine learning* e *data science* promoveram grande impacto no estudo sobre eficiência na trajetória do voo (MURÇA *et al.*, 2018). Na prática, embora o voo se baseie nas rotas aéreas, questões meteorológicas além de fatores na própria aeronave fazem com que as rotas reais praticadas acabem se diferenciando das aerovias, devido ao grande número de voos em um determinado trecho, dependendo de quais aeroportos se escolham.

Uma possível abordagem que simplifique esse desvio, seria a definição de *rotas médias* que são aproximações de um conjunto de rotas reais que melhor caracteriza aquele trajeto. Em (MURÇA *et al.*, 2020) lança-se mão de algoritmos de clusterização onde se usam as rotas reais para criar uma rota média relativa (rota nominal) para cada *cluster* e a partir deste calcular a eficiência do voo.

Outra forma de se calcular a eficiência seria comparando as rotas reais com a alguma aerovia correspondente. Este trabalho pretende analisar as vantagens de cada abordagem, além de verificar se há equivalência estatística entre ambas.

1.2 Objetivo Geral

O objetivo geral é avaliar se há equivalência qualitativa e quantitativa entre a eficiência de voos medida de duas formas: através das aerovias de um trecho e uma rota nominal obtida a partir das rotas reais efetuadas naquele trecho através de métodos de machine learning, em específico técnicas de clusterização. Em não havendo tal equivalência, avaliar qual método representa melhor o nível de eficiência do voo.

1.3 Justificativa do Trabalho

Na prática, para cada método de machine-learning aplicado na análise dos voos, pode-se ter um resultado diferente para a rota nominal (ou rota média). De modo que diferentes algoritmos podem resultar em diferentes valores de eficiência para uma mesma base de dados, o que dificulta a comparação entre si dos resultados obtidos. A justificativa de analisar o cálculo através de aerovias é a possibilidade de se unificar as metodologias e assim obter valores que possam ser comparados entre si, de acordo com o método de clusterização usado.

1.4 Escopo

No presente trabalho nos restringimos ao Aeroporto de São Paulo/Congonhas, em especial aos voos comerciais com destino a esse aeroporto no período entre setembro e dezembro de 2019.

2 Revisão da Literatura

2.1 Transporte Aéreo

2.1.1 Aerovias

Quando uma aeronave realiza um voo, o piloto não tem liberdade total sobre qual caminho seguir até seu destino. Para isso ele deve seguir aerovias, que são rotas aéreas pré-estabelecidas e que no Brasil são definidas pelo DECEA (Departamento de Controle do Espaço Aéreo) (CASSOL, 2020).

Segundo a Instrução do Comando da Aeronáutica 100-37, uma aerovia é toda área de controle, ou parte dela, disposta em forma de corredor. Dito de outro modo, uma aerovia é uma trajetória desenhada sobre coordenadas do espaço aéreo, dotada de informações específicas (identificação, posicionamento, rumo, altitude, etc.), destinada ao voo controlado de uma aeronave (DECEA, 2020).

As aerovias são divididas em dois grupos:

1. Aerovias Superiores — cujos voos ocorrem acima de 24,500 pés (7,468 metros)
2. Aerovias Inferiores — cujos voos ocorrem abaixo de 24,500 pés (7,468 metros)

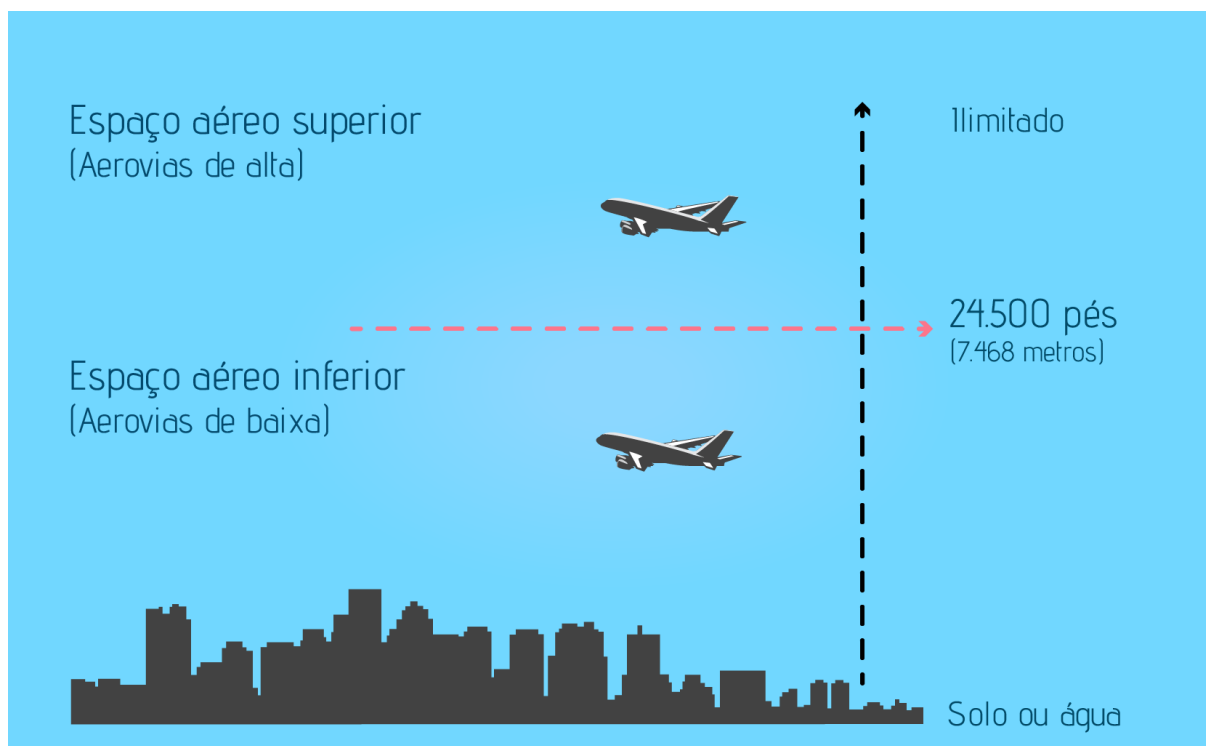


FIGURA 2.1 – Perfil das aerovias. Fonte: (DECEA, 2015)

As aerovias de alta (superiores) são usadas, em geral, por jatos que voam mais alto, companhias aéreas em cruzeiros. As aerovias de baixa (inferiores) são usadas por aviões de menor porte, turboélices e helicópteros. (MARINHO, 2016).

2.1.2 Cartas de voo

Para auxiliar na locomoção entre as aerovias, o piloto pode, em seu plano de voo, fazer uso de mapas de voo ou cartas de navegação. De acordo com seu tipo, essas cartas trazem informações de altitude, distância entre pontos, relevo da região, aerovias a seguir e inclusive a frequência com que o piloto deve se sintonizar com uma torre de controle. Essas cartas de navegação se dividem em três grupos: aeródromos e TMAS, cartas de rota e cartas visuais (BIANCH, 2019).

- **Aeródromos e TMAS** — abrangem os procedimentos de chegada e saída além de poderem variar de acordo com o tamanho do aeroporto. As cartas de saída (SID) trazem informações sobre decolagem e afastamento do aeroporto. As de chegada (STAR) são utilizadas em aeroportos de maior movimentação e trazem informações sobre a transição entre rotas e áreas próximas, além disso possuem maior detalhamento da área terminal (TMA).
- **Cartas de rota** — usadas durante o voo e indicam as rotas até o destino, seus dados incluem nomes dos fixos, direção, aeródromos próximos, entre outros.

- **Cartas visuais** — as cartas visuais (WAC) apresentam duas seções: uma superior que trás a rota de aproximação e pouco visual e a inferior que trás informações sobre o percurso. Há ainda a carta de aproximação (IAC) que auxilia na região de visualização da pista (BIANCH, 2019).

De forma geral, o voo pode se dar de duas formas: através do VFR (Visual Flight Rules - Regras de Voo Visual) quando as condições permitem ao piloto controlar visualmente a navegação da aeronave ou IFR (Instrument Flight Rules - Regras de Voo por Instrumentos) onde o piloto se orienta através de instrumentos de bordo, instrumentos no solo e em órbita, em vez de referências visuais exteriores a aeronave (DECEA, 2016).

As cartas de rotas (ENRC) recebem uma nomenclatura de acordo com sua região e tipo, recebendo os sufixos L1 a L9 e H1 a H9, donde H são para as aerovias de superiores (high) e L para as inferiores (low). O esquema é representado nas figuras abaixo:

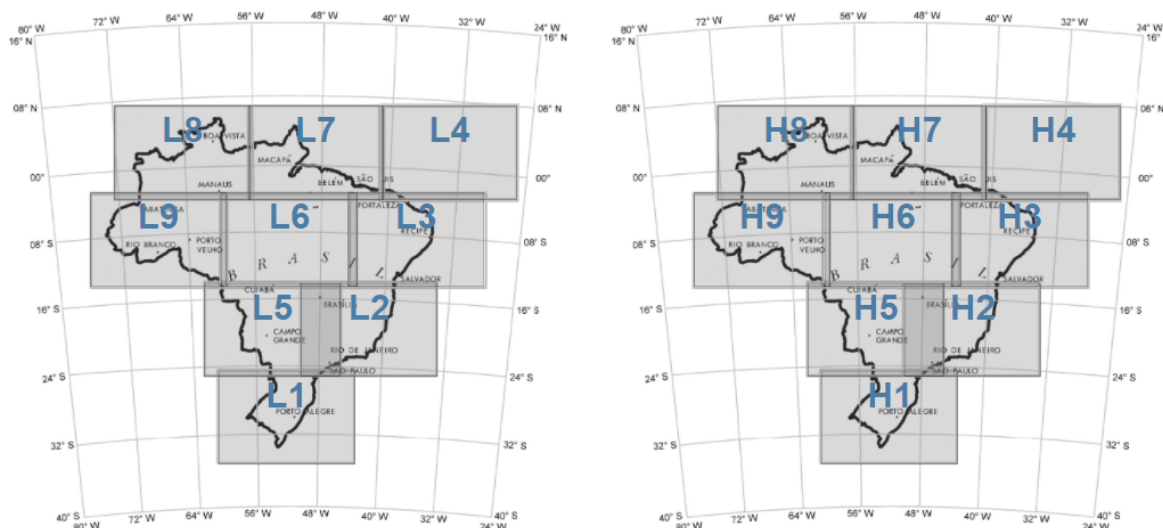


FIGURA 2.2 – Cartas de Rota (ENRC). Fonte: DECEA

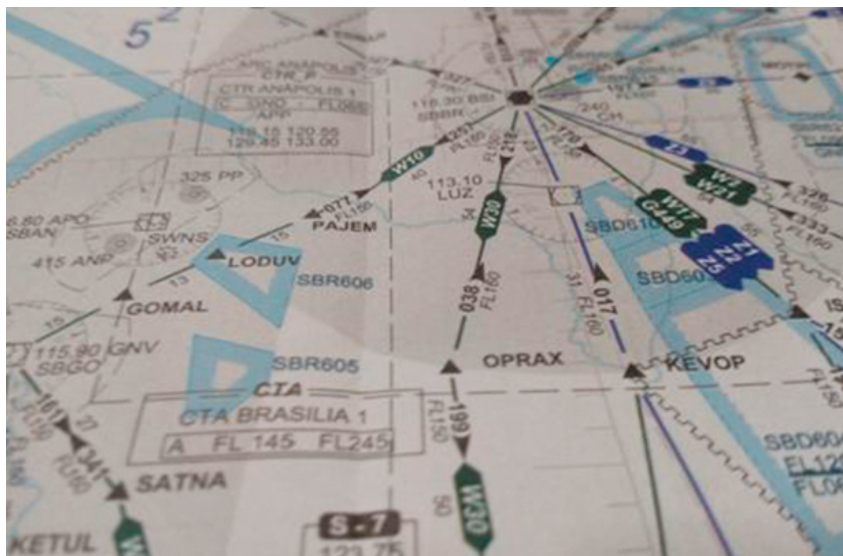


FIGURA 2.3 – Carta de Rota ENRC-L2

2.2 Machine Learning no setor de transporte aéreo

A capacidade das ferramentas de machine learning de trabalhar com dados em grande volume, pouco estruturados e dos mais diversos tipos são características que naturalmente se expandem para os mais diversos campos. Nos dias de hoje com a abundância de dados e informações de todo tipo, não seria de se estranhar a aplicação cada vez mais forte de tais ferramentas em problemas de transporte aéreo. A aparição de ferramentas como o ADS-B contribuiriam ainda mais para a disseminação desses métodos.

Da sua sigla, Automatic Dependent Surveillance - Broadcast ou Sistema de vigilância aérea automático, permite que aeronaves equipadas com essa tecnologia possam ser monitoradas por meio de satélites obtendo com isso informações sobre voo, latitude e longitude, velocidade, entre outro, em tempo real. Através dessas ferramentas, várias iniciativas, como FlightRadar24 (FLIGHTRADAR24, 2021), foram criadas a fim de prover acesso público a tais dados, tornando mais facilitado seu uso por pesquisadores.

Do ponto de vista geral, vários trabalhos já foram desenvolvidos a fim de se entender comportamentos e identificar padrões a partir de informações no espaço-tempo, em (GUDMUNDSSON *et al.*, 2007) os autores apresentam várias possibilidades tanto do ponto de vista de monitoramento autônomo, via GPS por exemplo, como da abordagem via mineração de dados de tais informações. Exemplo de aplicações em comportamento animal, locomoção humana, gestão de tráfego, segurança e vigilância, aplicações militares e até no campo esportivo são apresentados.

Para aplicações voltadas apenas para questões de tráfego aéreo a literatura também é extensa, deste modo optou-se por apresentar abaixo um panorama geral dos estudos

desenvolvidos nessa área que melhor direcional o contexto geral do presente trabalho.

Em se tratando de informações eletrônicas, sujeitas a interferências e ruídos nos dados o primeiro cuidado a se tomar é identificar e filtrar ruídos a fim de se extrair informações relevantes e confiáveis sobre os dados. Nesse caminho, o trabalho de (SUN *et al.*, 2016) apresenta alguns métodos e técnicas de machine learning, como DBSCAN e BIRCH, a fim de lidar com tais dificuldades.

Outro exemplo é o trabalho de (GARIEL *et al.*, 2011) que aplica clusterização e técnica de componentes principais para comparar rotas reais percorridas com rotas previamente identificadas ou típicas daquele percurso, além de introduzir uma medida de complexidade para medir a conformidade de um voo corrente com o padrão típico.

(BASORA *et al.*, 2017) propõe dois métodos de clusterização baseados numa melhoria do DBSCAN, o hierarquical density-based-spatial clustering of Applications with Noise (HDBSCAN) que é capaz de trabalhar com clusters de diferentes densidades com um único parâmetro como input. Os métodos consistem em usar duas métricas de distâncias distintas, a distância euclidiana e symmetrized segment-path distance (SSPD) (BESSE *et al.*, 2015), uma métrica de distância que leva em conta o formato da trajetória. O trabalho mostrou que a métrica SSPD fornece clusters mais precisos sobre áreas limitadas além de ser bem mais lento, para dados e áreas maiores a distância euclidiana se mostrou uma opção melhor.

Em (MURÇA *et al.*, 2018), novamente sobre uma base de dados obtidas via tecnologias como ADS-B, os autores utilizam de métodos de machine-learning para extrair padrões espaço-temporais e assim entender melhor o comportamento e performance do fluxo de tráfego nos aeroportos de Nova York, Hong Kong e São Paulo. Três métricas quantitativas de análise são aplicadas para cálculos de: eficiência, capacidade e previsibilidade. Através dessas, obtém-se uma comparação de performance nas operações de chegada de cada aeroporto. Além disso os fatores climáticos mostraram grande influência na performance de cada aeroporto.

(MURÇA *et al.*, 2020) um perfil de eficiência é traçado por meio da comparação entre as trajetórias reais e trajetórias nominais obtidas, via método de clusterização. Além disso, para os aeroportos de Congonhas (CGH) e Guarulhos (GRU), um modelo de regressão linear é desenvolvido a fim de investigar os fatores por trás das performances observadas. O trabalho mostrou, com aceitável significância estatística, que fatores como demanda no trecho, eventos climáticos e baixa visibilidade reduzem a eficiência de voo.

Por fim, no recente trabalho (ZENG *et al.*, 2021), propõe-se um modelo de denoising autoencoder (DAE) para redução da dimensionalidade da base de dados e extração dos fatores que melhor refletem as informações da trajetória, na base reduzida é aplicado um modelo de mistura gaussiana (GMM) para obter clusters para as trajetórias. A metodo-

logia se mostro mais rápida e eficiente que o uso apenas do método de clusterização.

3 Metodologia

3.1 Escolha das aerovias

Dada a escolha do Aeroporto de São-Paulo/Congonhas (CGH), e através dos pontos fixos do espaço aéreo brasileiro disponível em (DECEA, 2021b), que disponibiliza o arquivo no formato .csv com latitude, longitude e a identificação do ponto fixo, um total de 8638 pontos.

TABELA 3.1 – Pontos fixos do espaço aéreo brasileiro

ident	latitude	longitude
USAMO	-22.69263515	-45.82582149
MORPI	-21.30555556	-49.89027778
QV012	-14.91666667	-40.78666667
OBDIK	-22.8155	-47.22516667
MUDKA	-21.57233333	-43.57516667
ARUKU	-22.14696944	-51.63998333
OGPIR	-27.86423333	-48.59089722
IBDOD	-27.60110278	-48.67177778
ENVUD	-18.62755	-54.69020556
ASOXU	-13.16724444	-38.33548056
KP103	-22.95670556	-47.20527778

Descrição dos dados:

<i>ident</i>	Identificador do fixo
<i>latitude</i>	Latitude do ponto fixo, em graus.
<i>longitude</i>	Longitude do ponto fixo, em graus.

De mãos desses dados e do aeroporto de estudo, filtramos todos os pontos a uma

distância de no máximo 40nm do centro do aeroporto CGH, obtendo assim 484 pontos.

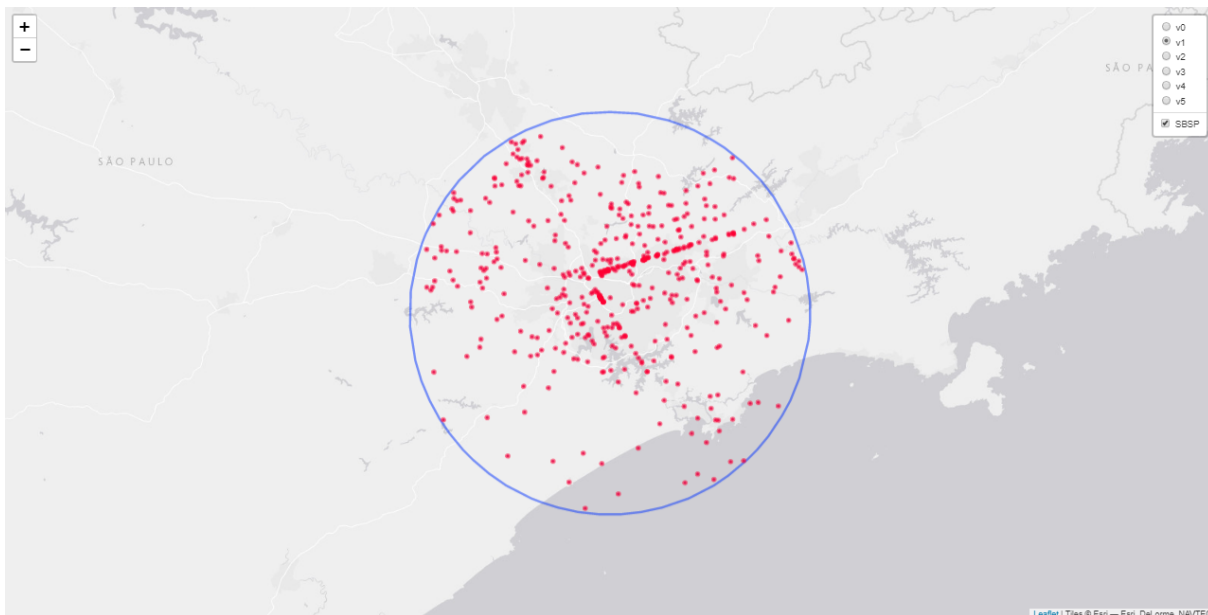


FIGURA 3.1 – Fixos ao redor de 40nm de CGH

Observe que mesmo após o filtro, muitos pontos foram capturados, desse modo, a fim de se ter uma análise mais precisa fez-se a correspondência do *ident* dos pontos com os presentes nas cartas de chegada (STAR) do aeroporto obtidas em (DECEA, 2021a)

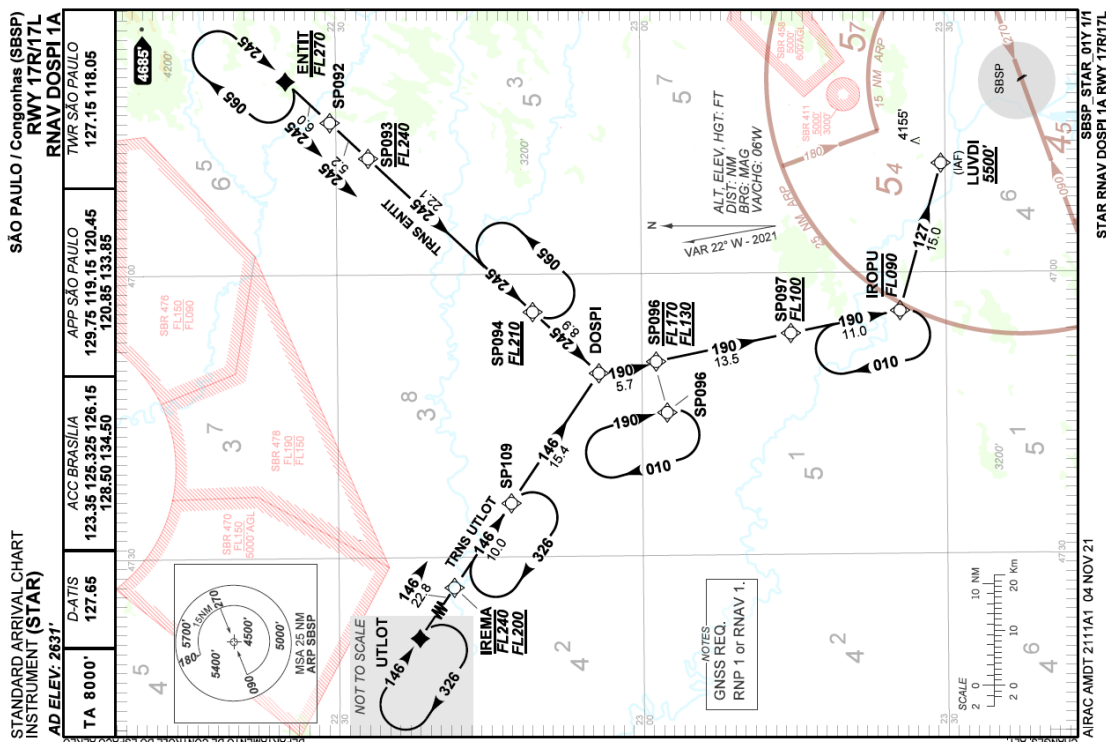


FIGURA 3.2 – Carta de aproximação: RNAV DOSPI 1A RWY 17R/17L

Desse modo coletamos todas as cartas disponíveis em (DECEA, 2021a) e coletamos visualmente cada ponto fixo, obtendo no final a configuração abaixo.

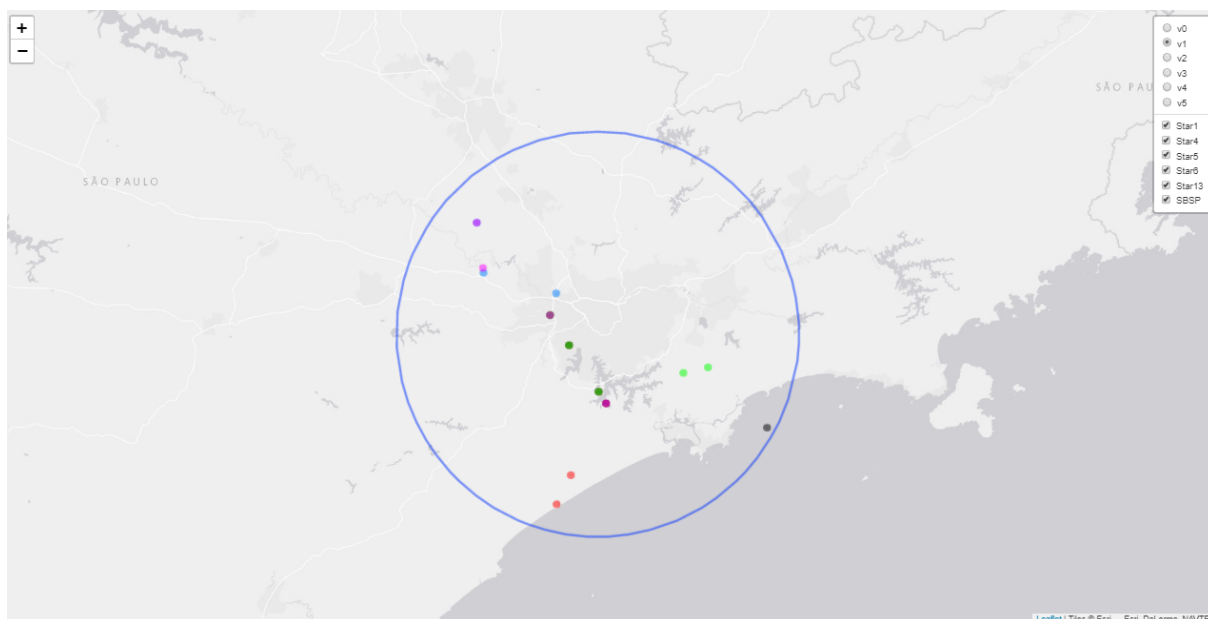


FIGURA 3.3 – Rotas obtidas

3.2 Análise dos dados

A base de dados, obtidos através do serviço global de rastreamento FlightRadar24 (FLIGHTRADAR24, 2021), correspondem ao perfil aéreo da região do Aeroporto de São Paulo/Congonhas (IATA: CGH, ICAO: SBSP), no período entre setembro e dezembro de 2019, para voos com destino em Congonhas. A base de dados contém as informações minuto a minuto de ID do voo, latitude, longitude, altitude, tipo de aeronave, origem, destino, entre outras informações. Ao todo, um total de 188872 linhas de dados foram obtidos, desses tivemos 11703 voos e 41 origens distintas. Para fins de análise, a região foi modelada está centrada no aeroporto de destino (CGH) e corresponde a um círculo de raio 74 km (40 nm).

TABELA 3.3 – Primeiras linhas da base de dados

	<i>indicat</i>	<i>time</i>	<i>unix</i>	<i>lat</i>	<i>lon</i>	<i>alt</i>	<i>orig</i>	<i>dest</i>	<i>raio</i>
1	3	2019-09-05 07:04:00	1567677897	-23.15750	-47.10720	14.450000	PLU	CGH	69.589512
2	3	2019-09-05 07:05:00	1567677958	-23.23270	-47.09610	13.125000	PLU	CGH	62.730162
3	3	2019-09-05 07:06:00	1567678015	-23.30840	-47.08510	10.275000	PLU	CGH	56.278161
4	3	2019-09-05 07:07:00	1567678076	-23.38650	-47.07350	8.100000	PLU	CGH	50.244119
5	3	2019-09-05 07:08:00	1567678136	-23.44550	-47.04480	7.575000	PLU	CGH	44.447971
6	3	2019-09-05 07:09:00	1567678196	-23.48140	-47.00300	7.125000	PLU	CGH	38.865585
7	3	2019-09-05 07:10:00	1567678256	-23.48550	-46.93820	6.500000	PLU	CGH	32.740945
8	3	2019-09-05 07:11:00	1567678318	-23.48900	-46.88030	5.975000	PLU	CGH	27.477423
9	3	2019-09-05 07:12:00	1567678378	-23.49170	-46.83530	5.575000	PLU	CGH	23.604423
10	3	2019-09-05 07:13:00	1567678434	-23.49500	-46.79850	4.950000	PLU	CGH	20.574205
11	3	2019-09-05 07:14:00	1567678497	-23.50520	-46.75980	4.400000	PLU	CGH	17.102674
12	3	2019-09-05 07:15:00	1567678555	-23.52150	-46.73040	4.400000	PLU	CGH	13.879256

Descrição dos dados tabelados

indicat Identificador do voo

time Data e horário da observação

unix Número de segundos entre o tempo da medição e o momento 00:00:00 UTC de 01-01-1970

lat Latitude do voo, em graus, no instante descrito

lon Longitude do voo, em graus, no instante descrito

alt Altitude do voo, em mil pés, no instante descrito

orig Código aeroportuário IATA do aeroporto de origem do voo

dest Código aeroportuário IATA do aeroporto de destino do voo

raio Distância ao aeroporto de destino, em quilômetros, no instante descrito

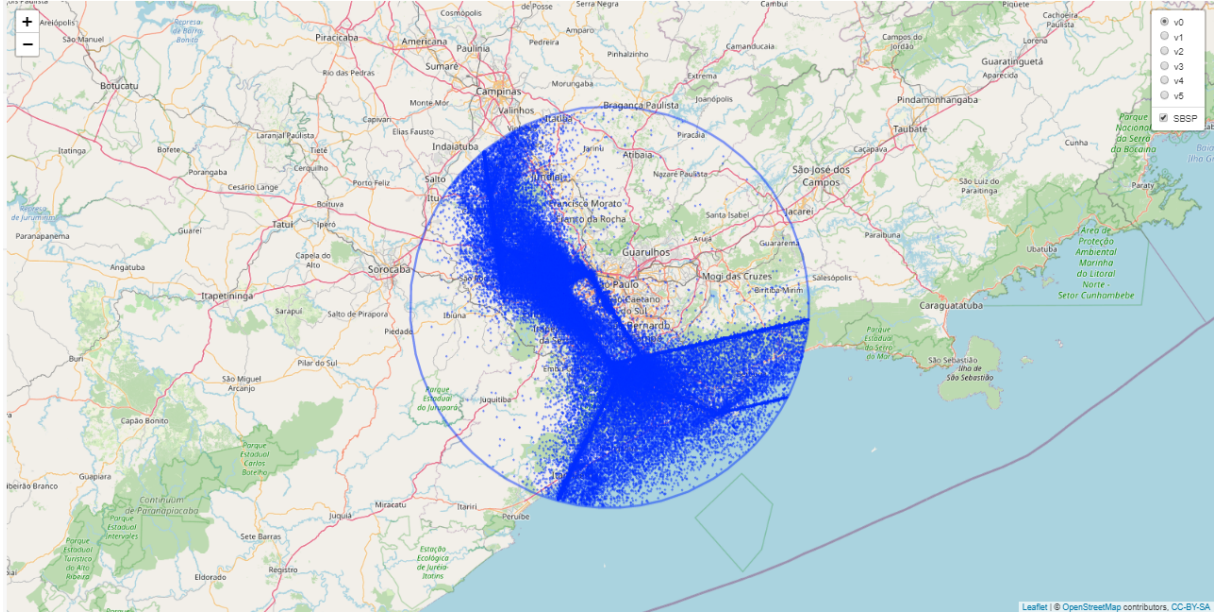


FIGURA 3.4 – Região do aeroporto de CGH com raio de 40nm.

3.3 As medidas de eficiência de voo

Para medida de eficiência de voo usaremos a mesma métrica usada em (MURÇA *et al.*, 2020), a Eficiência de Tráfego Horizontal (HTE - Horizontal Traffic Efficiency) que relaciona a trajetória real de um dado espaço aéreo durante um dado período de tempo com as rotas nominais (obtidas da clusterização).

O HTE durante um dado período de tempo T é definido como:

$$HTE_T = \frac{\sum_{i;t(i) \in T} S_i}{\sum_{i;t(i) \in T} S_i} \quad (3.1)$$

onde S_i é a distância horizontal real com respeito a trajetória i e s_i é o mínimo entre a distância horizontal de voo real e a sua rota nominal de referência, ou seja: $s_i = \min\{S_i, n_S(i)\}$, donde $n_S(i)$ é a distância horizontal da rota nominal referente a trajetória i .

É fácil observar que $HTE_T \in [0, 1]$.

3.4 Método de Clusterização

3.4.1 Apresentação dos Métodos

Para o nosso caso, a ideia é se utilizar de métodos não-supervisionados de machine learning. Além disso pretendemos identificar padrões semelhantes de rotas (clusters) e tratá-los de modo similar. Com isso os métodos de clusterização parecem ser os mais adequados.

Entre os principais métodos de clusterização estão: k-means, Gaussian Mixture Models (GMM), Hierarchical Agglomerative Clustering (HAC) e Density-Based Spatial Clustering of Application with Noise (DBSCAN). Cada um desses métodos podem ser enquadrados em um dos quatro itens abaixo (DOBILAS, 2021):

1. **Centroid-based** — usa alguma métrica de distância para associar cada ponto do conjunto ao seu centro mais próximo. Exemplo: k-means
2. **Connectivity-based** — assume que pontos próximos são mais correlacionados que pontos distantes entre si. Exemplo: Aglomeração Hierárquica (HAC).
3. **Density-based** — os clusters são definidos como regiões densas separados por regiões de baixa densidade. Exemplo: DBSCAN.
4. **Distribution-based** — assume a existência de um número específico de distribuições, cada uma com sua média (μ) e variância (σ^2)/covariância(Cov). Exemplo: Modelo de mistura de Gaussianas (GMM).

Entre os itens citados acima, o que mais se adéqua ao nosso propósito é o 3. Dentre os método baseados em densidade talvez o mais conhecido seja o DBSCAN, como pode ser visto em (ESTER *et al.*, 1996) esse algoritmo tem 3 características fundamentais para o nosso problema, são elas: ”base de dados espaciais de grande porte, requer o mínimo de conhecimento a priori das características ”

O método DBSCAN foi proposto inicialmente em 1996 (ESTER *et al.*, 1996). O algoritmo busca identificar regiões densas no espaço separado por regiões menos densas. Para que o algoritmo entenda concretamente o que seria essa região densa, alguns parâmetros precisam ser definidos, são eles:

1. **Epsilon (ϵ)** — o raio que define a vizinhança de um ponto, ou seja, qualquer ponto a uma distância menor ou igual a ϵ faz parte da vizinhança
2. **Min Points (MinPts)** — o número mínimo de ponto requeridos para uma vizinhança ser ”densa”.

A partir desses dois parâmetros, o algoritmo segue os seguintes passos:

Definição 1 (ε -vizinhança de um ponto). *A ε -vizinhança de um ponto p , denotado por $N_\varepsilon(p)$, é definido por $N_\varepsilon(p) = \{q \in D \mid \text{dist}(p, q) \leq \varepsilon\}$.*

Definição 2 (Diretamente alcançável). *Um ponto p é diretamente alcançável por um ponto q com respeito a $Eps, MinPts$ se:*

- 1) $p \in N_{Eps}(q)$
- 2) $|N_\varepsilon(q)| \geq MinPts$ (condição de núcleo).

Definição 3 (Alcançável por densidade). *Um ponto p é alcançável por densidade por um ponto q com respeito a ε e $MinPts$ se há uma cadeia de pontos $p_1, \dots, p_n, p_1 = q, p_n = p$ tal que p_{i+1} é diretamente alcançável por p_i .*

Definição 4 (Conectado por densidade). *Um ponto p é conectado por densidade a um ponto q com respeito a ε e $MinPts$ se existe um ponto o tal que ambos, p e q são conectados por densidade com o com respeito a ε e $MinPts$.*

Definição 5 (Cluster). *Seja D um database de pontos. Um cluster C com respeito a $\varepsilon, MinPts$ é um subconjunto não vazio de D que satisfaz as seguintes condições:*

- 1) $\forall p, q$ se $p \in C$ e q é alcançável por densidade por p com respeito a $\varepsilon, MinPts$, então $q \in C$. (Maximalidade)
- 2) $\forall p, q \in C$: p e q são conectados por densidade com respeito a $\varepsilon, MinPts$. (Conectividade)

Definição 6 (Noise). *Sejam C_1, \dots, C_k clusters do database D com respeito aos parâmetros ε_i e $MinPts_i$, $i = 1, \dots, k$. Então nós definimos noise como o conjunto de pontos no database D que não pertence a nenhum cluster C_i , i.e. $\text{noise} = \{p \in D \mid \forall i : p \notin C_i\}$.*

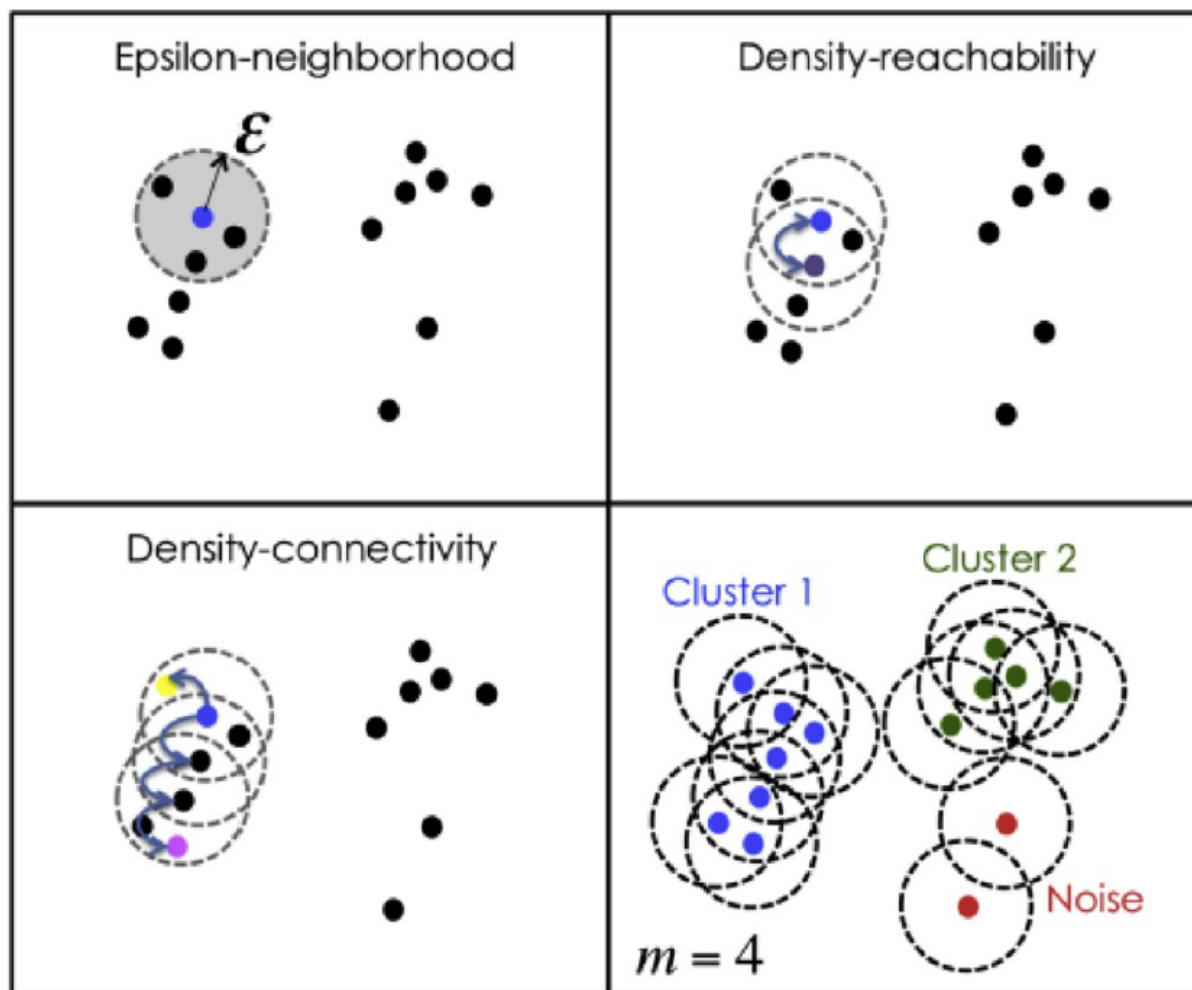


FIGURA 3.5 – Representação visual dos conceitos principais do DBSCAN. Fonte: (MURÇA *et al.*, 2020)

3.5 Tratamento dos dados

Visto que os dados provavelmente apresentaram discrepâncias de má distribuição seja temporal seja espacial, um prévio tratamento se faz necessário a fim de que o algoritmo obtenha resultados satisfatórios.

Alguns cuidados a se levar em conta é que cada dado que seja entregue ao modelo deve ter, a priori, igual contribuição. Deste modo algum tipo de normalização ou padronização deve ser adotado nessa etapa a fim de que cada observação contribua igualmente para o algoritmo de clusterização.

Pela figura 3.4 pode-se perceber algumas regiões de maior densidade de pontos, de modo que faz-se necessário uma identificação de cada rota com seu voo correspondente a fim de que voos não sejam indexados ao mesmo cluster apenas por possuírem o mesmo ponto de chegada em comum. Para isso transforma-se cada voo em um vetor de mesma dimensão, desse modo cada voo será um ponto para a nossa clusterização e resolve-se o

problema apresentado acima. Essa técnica se mostra adequada através de diversos autores como (GARIEL *et al.*, 2011) e (MURÇA, 2018).

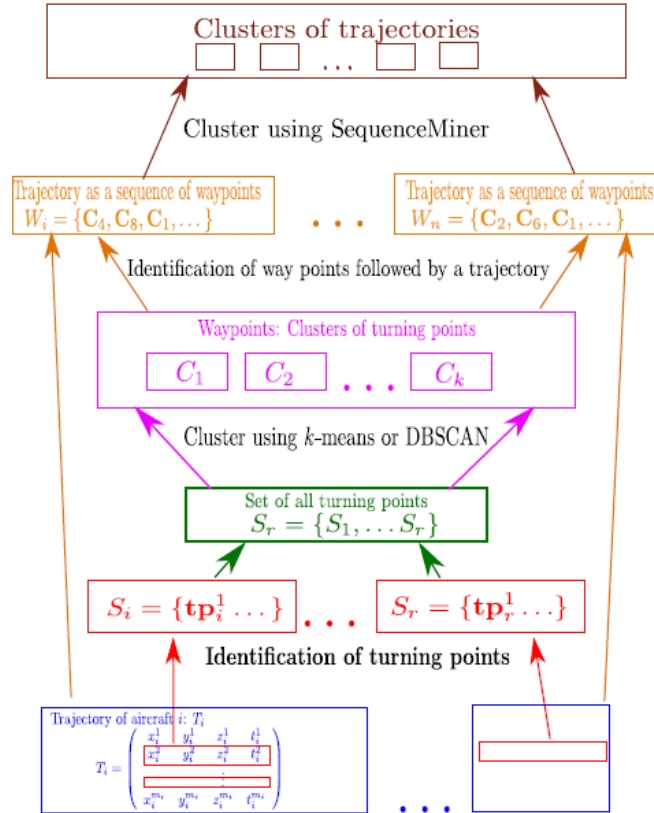


FIGURA 3.6 – Esquema da identificação do voo para posterior clusterização. Fonte: (GARIEL *et al.*, 2011)

A literatura oferece diversas técnicas para lidar com esse problema, um deles é fixar a extensão de cada trajetória usando métodos de reamostragem, a fraqueza desse método pode aparecer quando se tem trajetórias que variam muito em suas dimensões, fazendo assim que algumas trajetórias percam suas características. Outro método é usar diferentes métricas de distância que absorvam características particulares de curvas semelhantes, o que também acabem aumentando o custo computacional do processo.

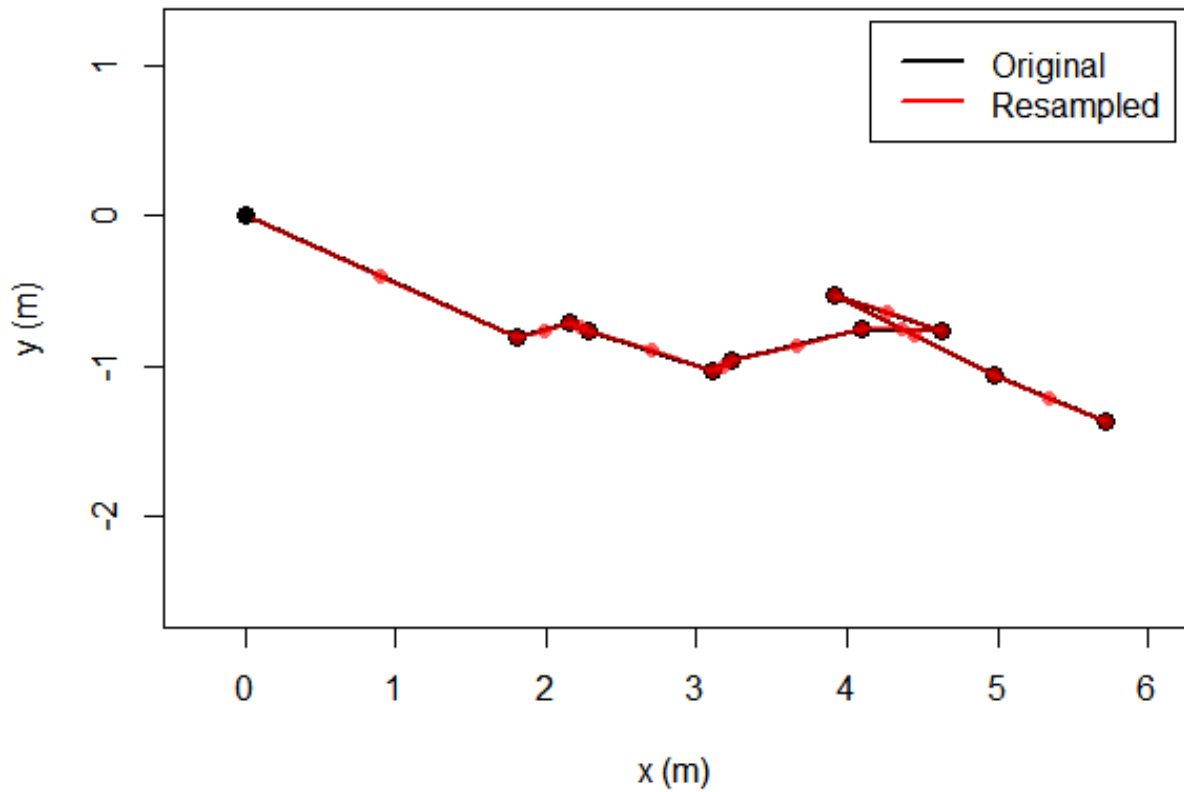


FIGURA 3.7 – Exemplo de discretização da trajetória. Fonte: (MCLEAN, 2021)

Para o objetivo deste trabalho, a ferramenta que mais se adequou foi uma normalização no tempo de voo da aeronave na região pré-estabelecida. Para a escolha da melhor discretização para a trajetória, foi plotado um histograma da distribuição de pontos em cada voo, como pode ser observado na figura 3.8.

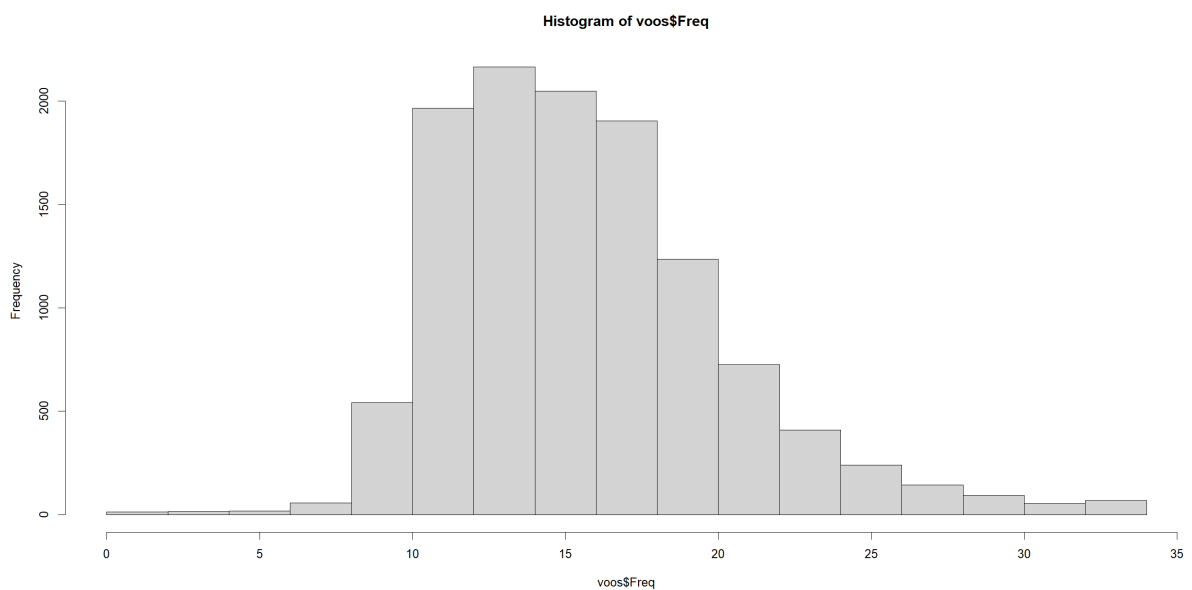


FIGURA 3.8 – Distribuição dos voos

Pela figura pode-se observar que a maior parte dos voos apresenta pelo menos 14 pontos, desse modo escolheu-se remover da base todos os voos com menos que 14 pontos e os demais voos tiveram sua trajetória rediscritizada a fim de que todos apresentassem 15 pontos em sua trajetória. Para os pontos da discretização que ficam entre dois pontos da representação pré-discretizada é aplicada uma interpolação linear para determinar os valores de latitude e longitude correspondentes como esquematizado na figura 3.9.

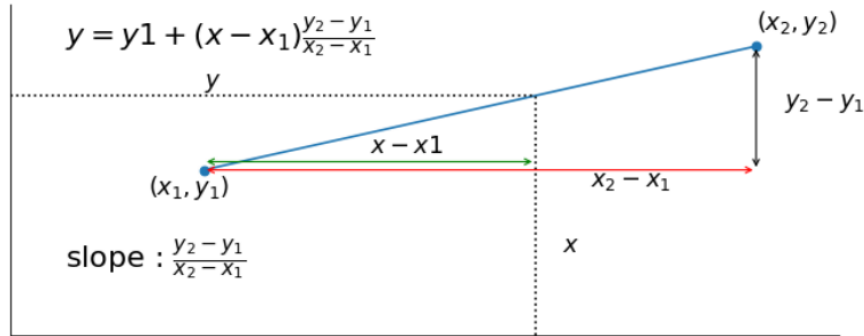


FIGURA 3.9 – Esquema da interpolação linear. Fonte: (BRETT, 2016)

Desse modo, cada aeronave analisada na região é observada em intervalos de tempo de aproximadamente 60 segundos entre uma observação e outra, o que foi feito nesse trabalho foi normalizar essas observações para que cada voo apresentasse o mesmo número de pontos observados e com mesmo espaçamento temporal entre eles.

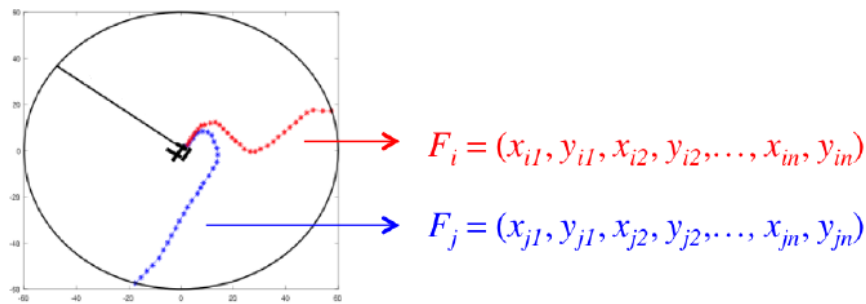


FIGURA 3.10 – Reamostragem da trajetória de cada voo. Fonte: (MURÇA, 2018)

Para intervalos da reamostragem que caíssem entre duas observações consecutivas foi feito uma interpolação linear entre os respectivos parâmetros de latitude e longitude. Resumindo, cada voo ficou caracterizado como uma matriz $4 \times n$ no seguinte formato:

$$R_i = (lon_{i1}, lat_{i1}, lon_{i2}, lat_{i2}, \dots, lon_{in}, lat_{in}) \tag{3.2}$$

Desde que a dimensão t_i foi padronizada ela pode ser omitida do vetor.

Por fim, cada vetor foi normalizado a fim de todos terem média 0 e desvio padrão 1.

$$x_{ik}^* = \frac{x_{ik} - \bar{x}_k}{\sigma_k} \quad (3.3)$$

Donde:

$$\bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_{ik} \quad (3.4)$$

$$\sigma_k = \sqrt{\frac{\sum_{i=1}^n (x_{ik} - \bar{x}_k)^2}{n}} \quad (3.5)$$

3.5.1 Escolha dos Parâmetros

Como vimos na seção anterior, o método DBSCAN requer dois parâmetros: *eps*, e *MinPts*. Para a escolha do *eps* usamos o método *kNNdist* que calcula, para cada ponto da nossa base de dados, qual a distância ao ponto mais próximo. Ordenando esses pontos de modo crescente e plotando num gráfico podemos ter uma ideia de qual *eps* mais adequado para o nosso método.

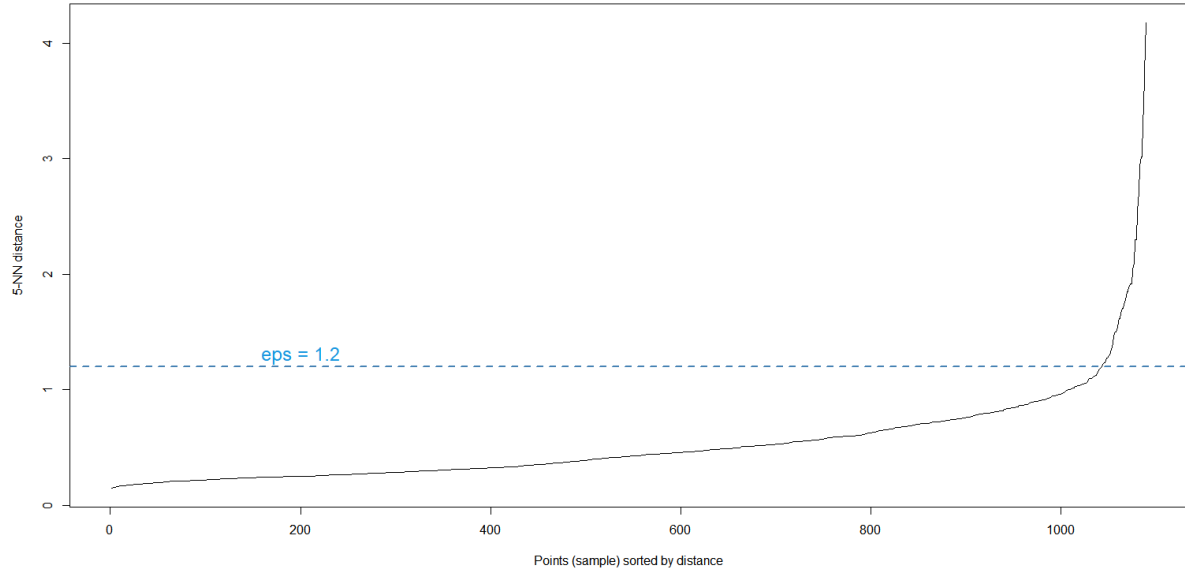


FIGURA 3.11 – Plot kNNdist para escolha do eps

Baseado na figura acima, o parametro escolhido foi $eps = 1,2$.

A qualidade da clusterização foi avaliada pelo coeficiente de Silhouette, que descreveremos abaixo:

O coeficiente Silhouette é definido de modo que quanto maior seu valor relacionado a um modelo, melhor definida estará aquela clusterização. O coeficiente é definido para

cada amostra e é composto por dois scores:

- **a:** É a distância média entre a amostra e todos os outros pontos na mesma classe.
- **b:** É a distância média entre a amostra e todos os outros pontos da classe mais próxima.

De posse desses valores o coeficiente de Silhouette (s) é então definido como:

$$s = \frac{b - a}{\max(a, b)} \quad (3.6)$$

Por fim, o coeficiente de um conjunto é a média dos coeficientes de cada amostra.

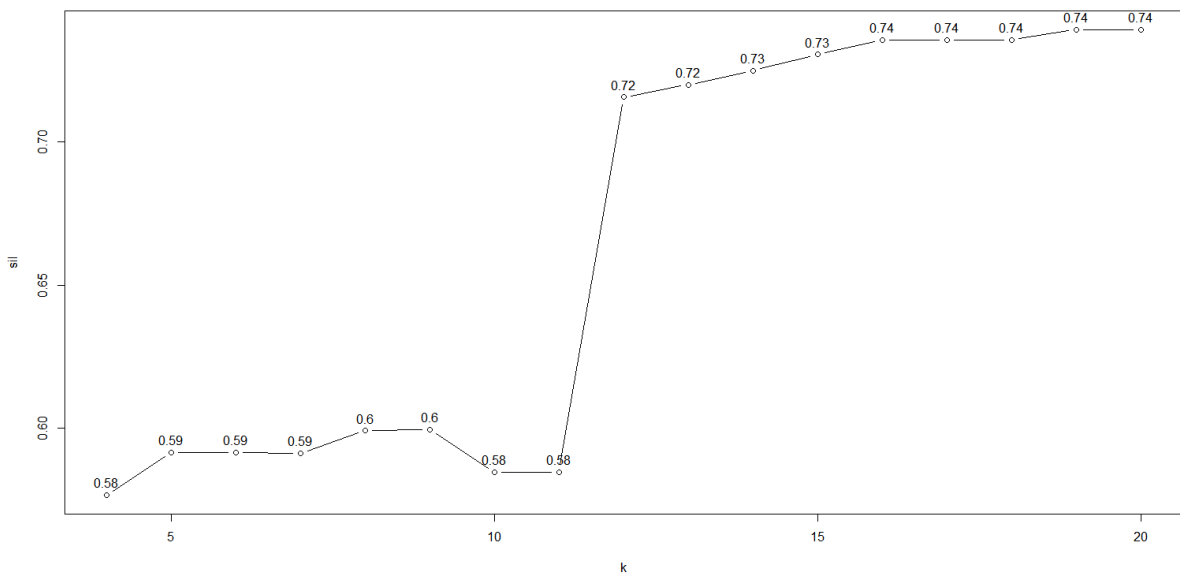


FIGURA 3.12 – Coeficiente de silhouette

O coeficiente foi escolhido de modo que fosse obtido 6 clusters e o melhor coeficiente. Deste modo pelo gráfico da figura 3.12, o valor escolhido foi $MinPts = 16$. O resultado da clusterização com esses parâmetros, foi representada abaixo pelo gráfico da distribuição de voos em cada cluster:

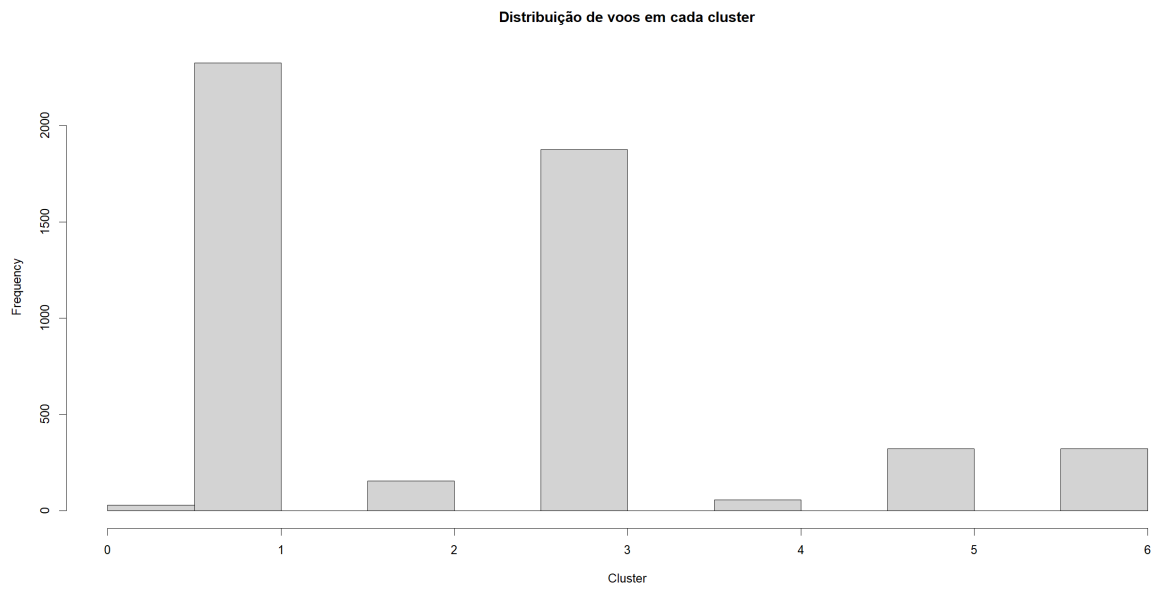


FIGURA 3.13 – Histograma da clusterização

4 Resultados e Discussões

O resultado da clusterização pode ser visto abaixo, onde cada cluster é representado por uma cor distinta e voos no mesmo cluster recebem a mesma cor.

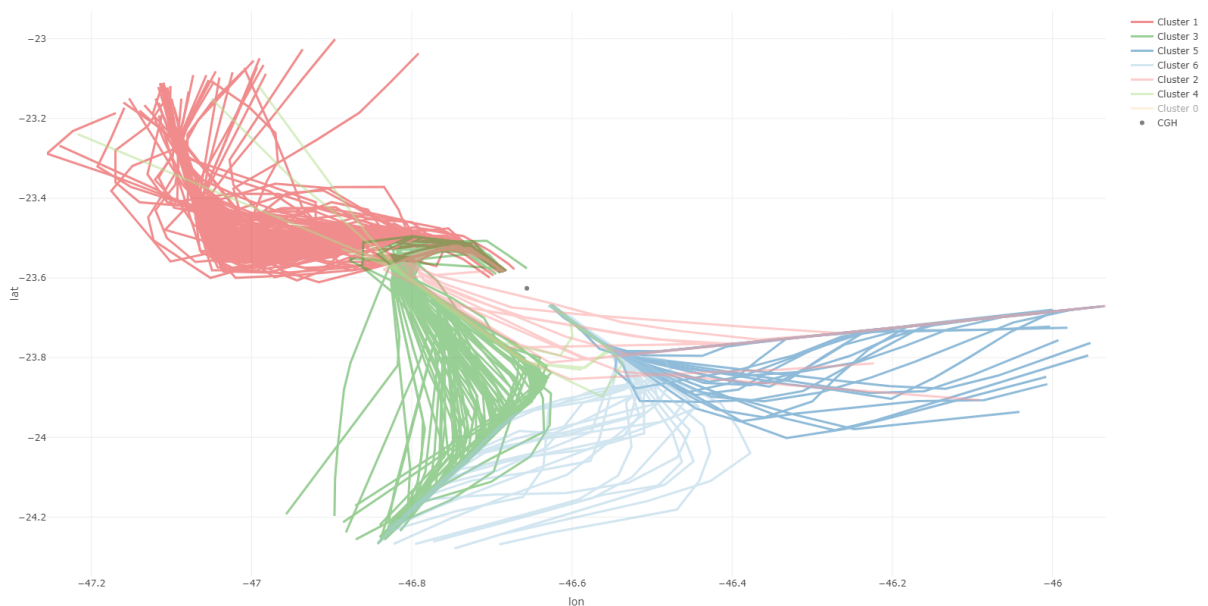


FIGURA 4.1 – Representação de cada voo e seu cluster

Como cada voo foi representado por um vetor de mesma dimensão, pode-se calcular uma rota que melhor representa cada cluster através de uma média aritmética de cada vetor trajetória por cluster. O resultado dessa metodologia é mostrado na figura abaixo.

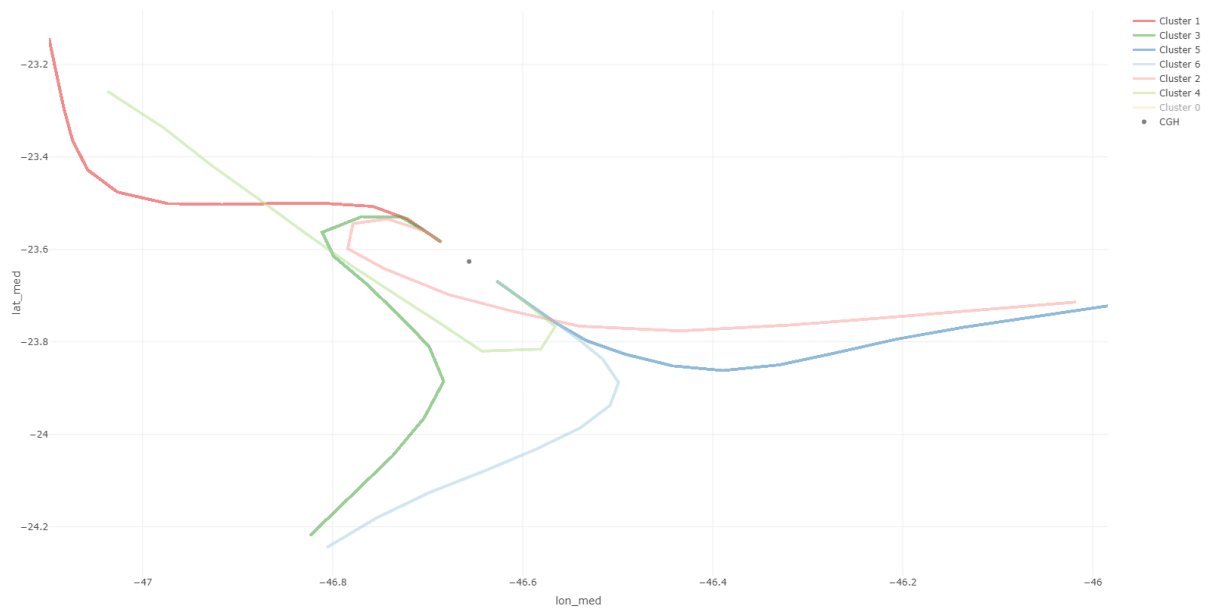


FIGURA 4.2 – Representação de cada cluster obtido através da média aritmética dos voos

4.1 Análise qualitativa dos resultados

De posse das rotas nominais obtidas na clusterização, podemos calcular a eficiência horizontal (HTE) pela equação 3.1. Vale ressaltar que os pontos fixos obtidos para cada aerovia também foram rediscritizados a fim de que cada aerovia tivesse 15 pontos correspondentes, tornando assim possível a comparação direta entre as métricas calculadas.

O resultado da eficiência HTE para os voos de cada cluster, assim como para as aerovias, foram plotados num gráfico boxplot que exibe o valor da mediana assim como o primeiro e terceiro quartil. Segue abaixo esse gráfico para os clusters obtidos.

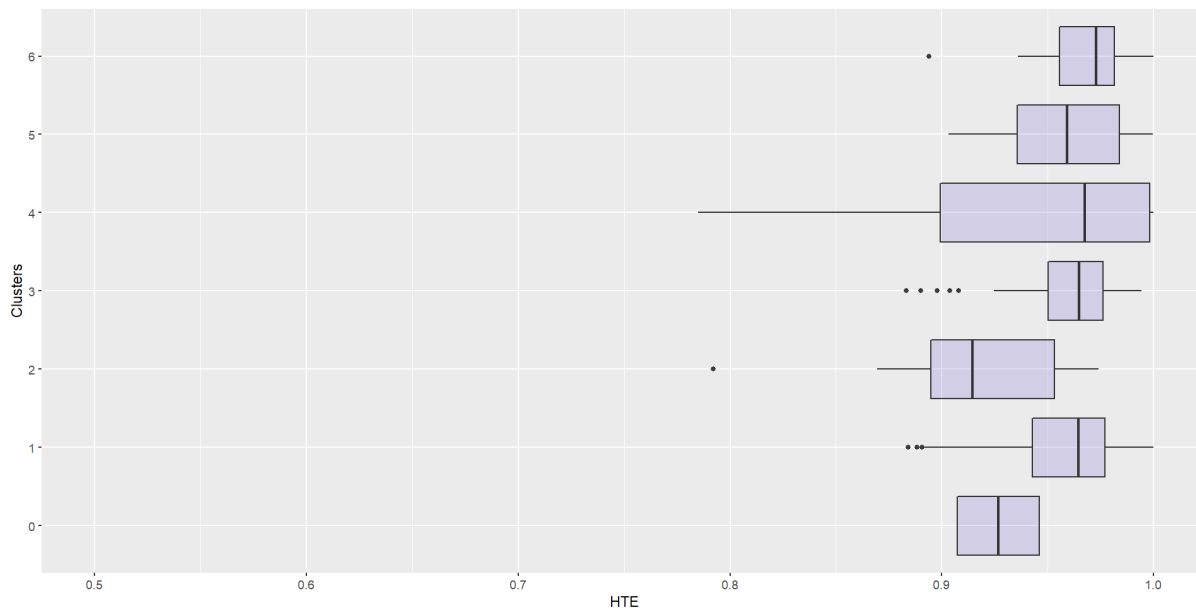


FIGURA 4.3 – Boxplot da medida de eficiência HTE para cada cluster

Abaixo plotou-se o mesmo gráfico, agora para as aerovias.

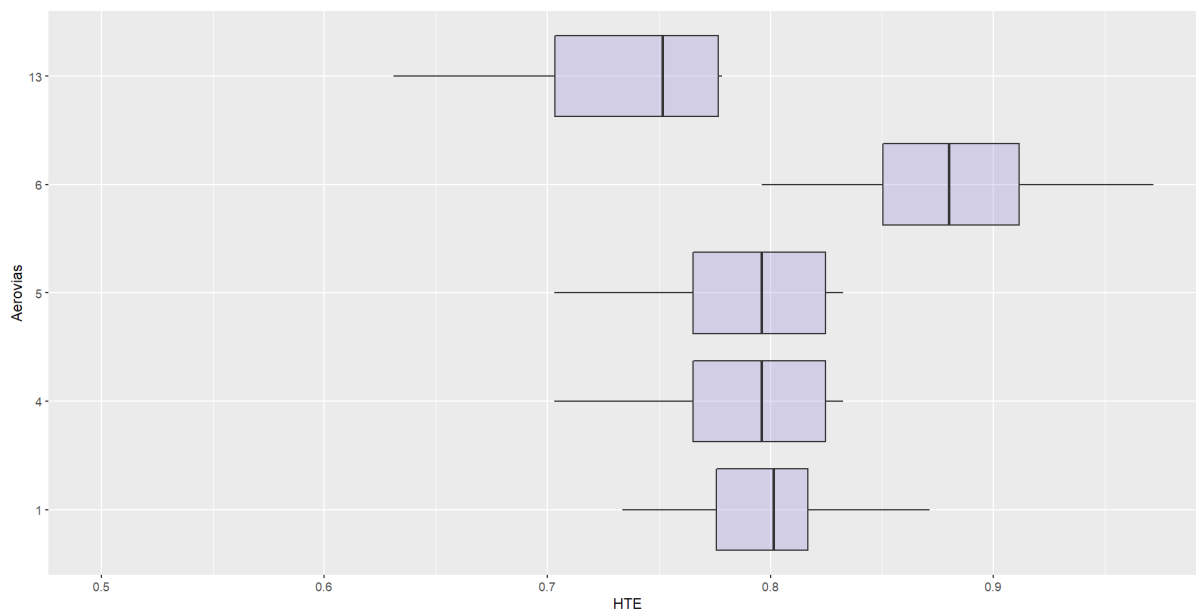


FIGURA 4.4 – Boxplot da medida de eficiência HTE para cada aerovia

4.1.1 Comparação Final

A fim de se ter uma representação visual da discrepância das duas métricas, os gráficos plotados acima foram dispostos lado a lado e ajustada a escala do eixo horizontal.

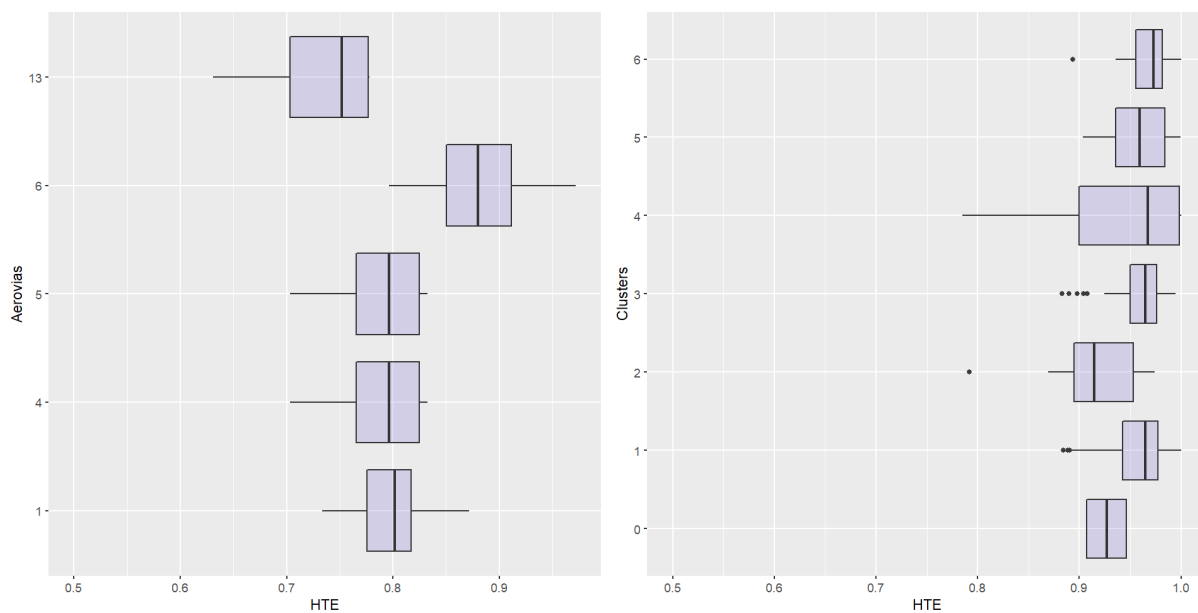


FIGURA 4.5 – Boxplot da medida de eficiência HTE para as aerovias e clusters

Por fim, faz-se o comparativo de cada cluster com sua aerovia mais próxima.

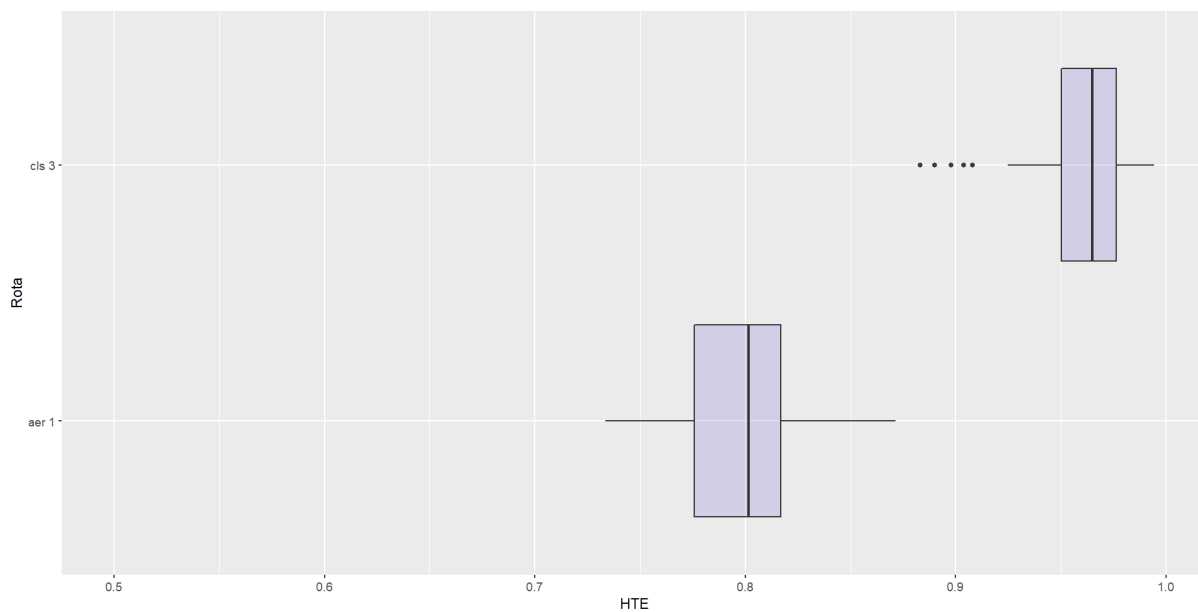


FIGURA 4.6 – Aerovia 1, cluster 3

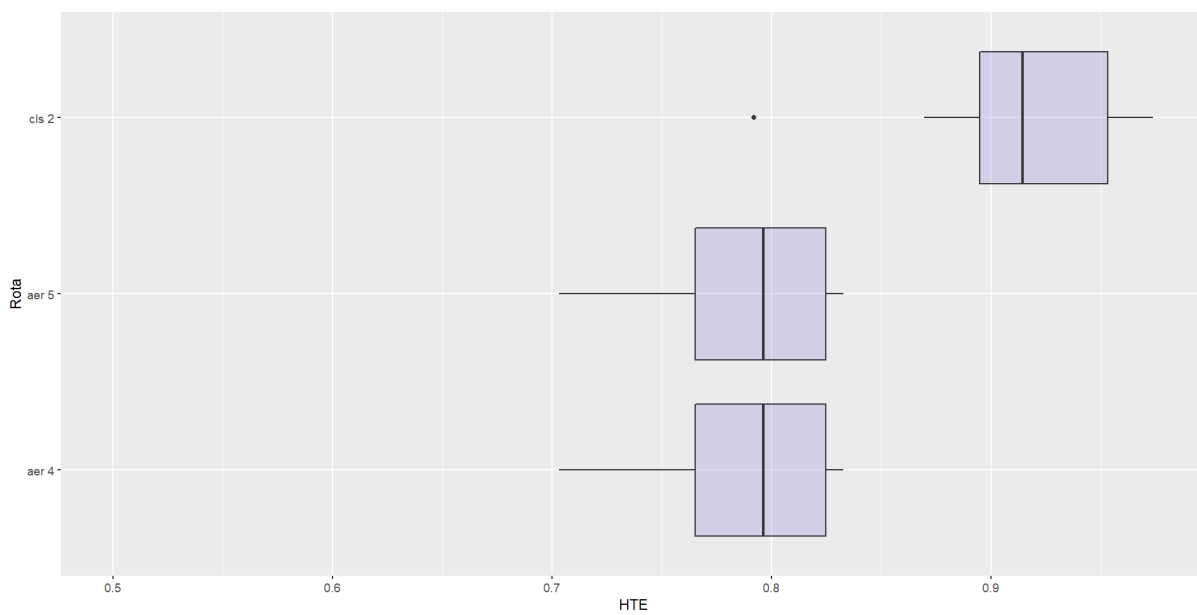


FIGURA 4.7 – Aerovias 4 e 5, cluster 2

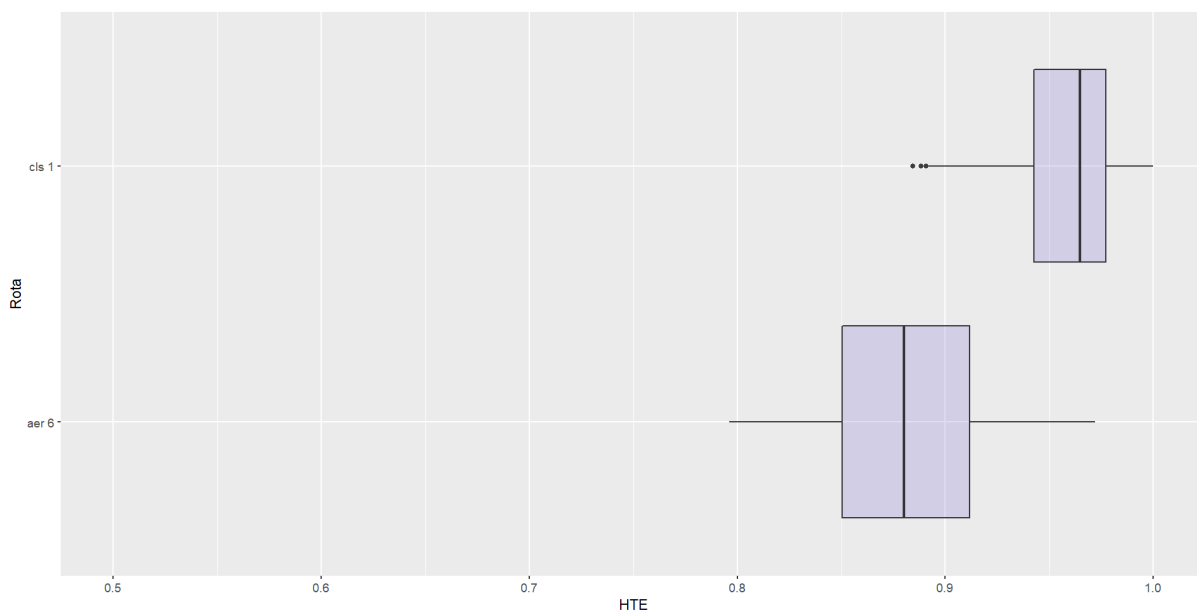


FIGURA 4.8 – Aerovia 6, cluster 1

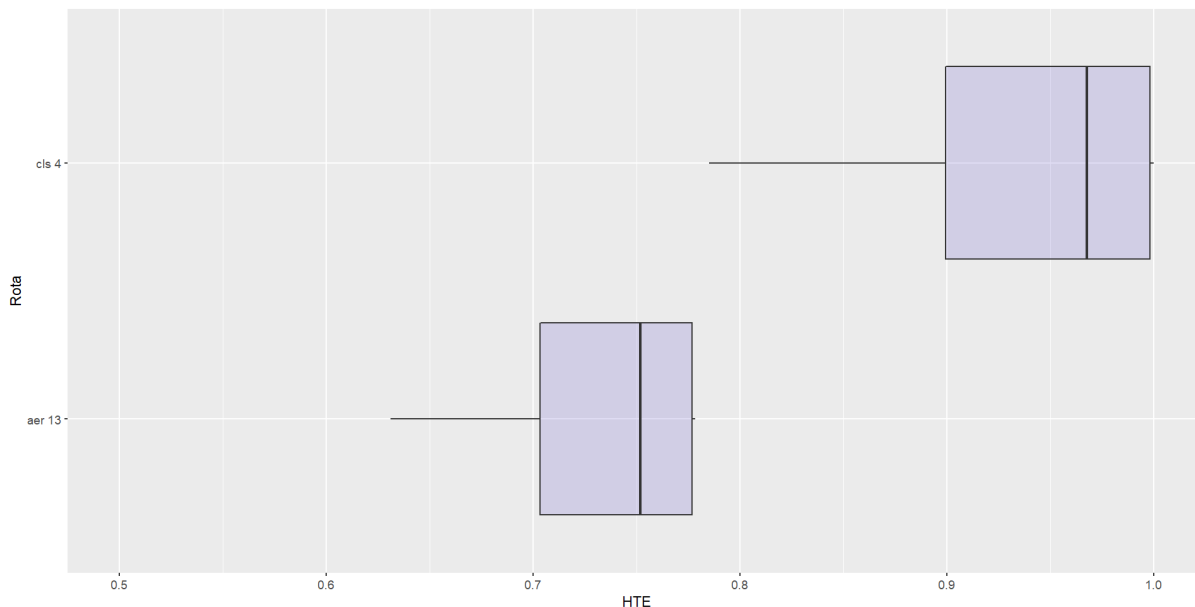


FIGURA 4.9 – Aerovia 13, cluster 4

Qualitativamente, os resultados não parecem partir de uma mesma distribuição, defendendo a tese de que os diferentes métodos produzem resultados estatisticamente distintos. Outra observação que pode-se notar é que os valores não apresentam uma distribuição normal, em vez disso apresentam um desvio maior para a região de menor eficiência em relação a média.

4.2 Análise quantitativa dos resultados

A fim de se analisar, quantitativamente, sob algum grau de confiança se sobre a igualdade ou não das distribuições, o mais comum a se fazer é realizar um teste de hipóteses que tenha como hipótese nula a aceitação de que as amostras partem de uma mesma distribuição ou, caso haja evidência estatística para rejeitar a hipótese nula, a hipótese alternativa será aceita. Há diversos métodos capazes de resolver tal problema os mais comuns são o teste de Shapiro-Wilk e o teste de Kolmogorov-Smirnov. A diferença entre eles geralmente se dá por sua maior ou menor eficácia sobre amostras pequenas. O presente caso de estudo não apresenta tal deficiência de modo que podemos escolher entre ambos. A decisão sobre o teste de Kolmogorov-Smirnov se deu por simplicidade e facilidade no estendimento da teoria por trás do método.

Para aplicar o teste faremos a seguinte correspondência:

- X — distribuição da eficiência de cada voo para o cluster.
- Y — distribuição da eficiência de cada voo para a aerovia correspondente ao cluster de X .

Teste de hipóteses

- H_0 : X e Y têm a mesma distribuição $\rightarrow p > 0,05$
- H_1 : X e Y não têm a mesma distribuição $\rightarrow p \leq 0,05$

O teste consiste em avaliar para cada número real r , qual a porcentagem de valores em cada amostra que são inferiores ou iguais a r . Se o valor absoluto da maior diferença for pequeno o suficiente, então a hipótese H_0 é aceita.

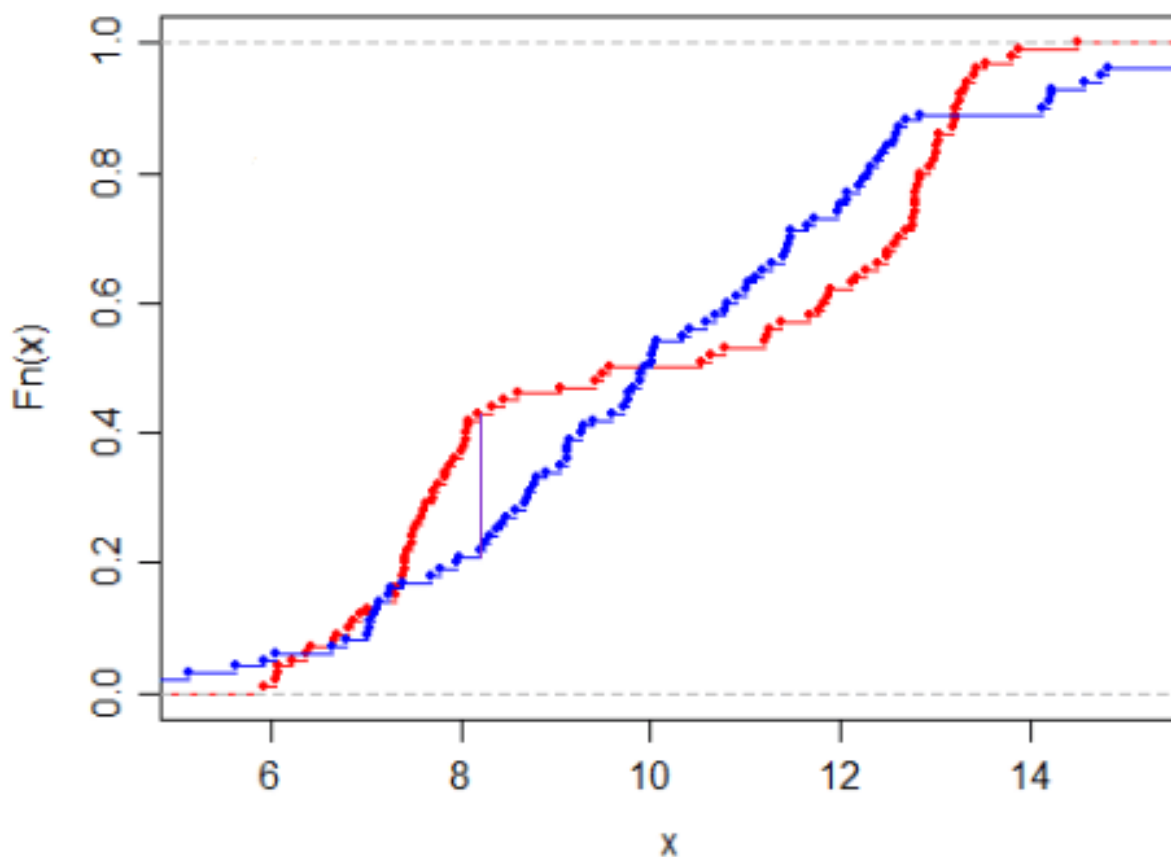


FIGURA 4.10 – Exemplo da avaliação no teste de Kolmogorov-Smirnov. Fonte: (ARSENAULT, 2017)

Desse modo, aplicando o teste temos os resultados apresentados na tabela abaixo:

TABELA 4.1 – Resultado do teste K-S para cada par aerovia x cluster

Par X, Y	p-value
aerovia 1 x cluster 3	$< 2,22e-16$
aerovia 4 x cluster 2	$< 2,22e-16$
aerovia 5 x cluster 2	$< 2,22e-16$
aerovia 6 x cluster 1	$< 2,22e-16$
aerovia 13 x cluster 4	$< 2,22e-16$

O valores baixo de p-value deixam claro que as distribuições são distintas confirmando a análise qualitativa feita na seção anterior.

5 Conclusão

Neste trabalho objetivou-se analisar a equivalência estatística entre a eficiência de voo calculada a partir de rotas nominais obtidas por meio de clusterização e por aerovias. Para o cálculo das rotas nominais usou-se o método DBSCAN de clusterização devido suas vantagens em relação a nossa base de dados e em função da sua vasta aplicação na literatura. As rotas teóricas foram obtidas através das cartas de rota em (DECEA, 2021a) e filtradas para nosso aeroporto de análise.

O método de clusterização mostrou bastante eficiência em encontrar clusteres coerentes além de apresentar relativamente um baixo custo computacional. A dificuldade principal se deu em obter os pontos fixos das aerovias do aeroporto que fizessem sentido para o nosso caso de estudo, além disso, por se tratar apenas de pontos estáticos sem referências temporais, não podemos avaliar métricas de eficiência distintas como por exemplo as analisadas em (MURÇA *et al.*, 2020).

Para comparar os dois métodos, o teste de Kolmogorov-Smirnov foi escolhido para avaliar a relação entre as métricas obtidas por cada modo, a escolha do método foi baseada no seu amplo uso na literatura além disso nosso caso especial não apresentava dificuldades recorrentes observadas nesses testes como tamanho pequeno de amostras, de modo que qualquer método padrão apresentaria resultados confiáveis de testes. No nosso exemplo em particular as métricas se mostraram bastante distintas, tanto pela análise qualitativa quanto pela análise quantitativa que ficou evidente pelo valor da estatística e pelos p-values obtidos. Um possível responsável por tal desvio pode se dever a fatores que afetam a performance em si do voo. Como já avaliado em (MURÇA *et al.*, 2020) fatores como clima e visibilidade afetam negativamente a performance da aeronave.

Para estudos futuros uma possível abordagem seria levar em conta os fatores que já se sabe afetarem negativamente a performance do voo. Além disso, uma análise comparativa de outros aeroportos assim como uma avaliação da partida e de voos em rota. Outro caminho seria a adição de mais variáveis, como a altitude, podem ser um complemento ao presente trabalho.

Referências

ANAC. **Dados Abertos**. 2019. Acessado em: 9 nov. 2021. Available at: <https://www.anac.gov.br/aceso-a-informacao/dados-abertos>.

ANAC. **DADOS ESTATÍSTICOS DO TRANSPORTE AÉREO**. 2019. Acessado em: 9 nov. 2021. Available at: <https://www.anac.gov.br/aceso-a-informacao/dados-abertos/areas-de-atuacao/voos-e-operacoes-aereas/dados-estatisticos-do-transporte-aereo>.

ARSENAULT, M.-O. 2017. Available at: <https://towardsdatascience.com/kolmogorov-smirnov-test-84c92fb4158d>, Acessado em: 20 nov. 2021.

BASORA, L.; MORIO, J.; MAILHOT, C. A trajectory clustering framework to analyse air traffic flows. *In: . Proceedings [...]. [S.l.: s.n.]*, 2017.

BESSE, P.; GUILLOUET, B.; LOUBES, J.-M.; ROYER, F. Review perspective for distance based trajectory clustering. 08 2015.

BIANCH. **Cartas de voo: por que são importantes e quais os tipos?** 2019. Available at: <https://blog.bianch.com.br/cartas-de-voo/>, Acessado em: 21 nov. 2021.

BRETT, M. **Linear interpolation**. 2016. Available at: https://matthew-brett.github.io/teaching/linear_interpolation.html, Acessado em: 21 nov. 2021.

CASSOL, L. **Aerovias: quais são e como funcionam as rotas utilizadas pelos aviões**. 2020. Acessado em: 9 nov. 2021. Available at: <https://www.melhoresdestinos.com.br/aerovias-rotas-avioes.html>.

DECEA. **Espaço Aéreo**. 2015. Available at: <https://www.fab.mil.br/noticias/mostra/22979/>, Acessado em: 21 nov. 2021.

DECEA. **Congonhas (SBSP)**. 2021. Available at: <https://aisweb.decea.mil.br/?i=aerodromos&codigo=SBSP>, Acessado em: 9 nov. 2021.

DECEA. **Fixos**. 2021. Available at: <https://aisweb.decea.mil.br/?i=espaco-aereo&p=fixos>, Acessado em: 9 nov. 2021.

- DECEA, M. d. D. Comando da A. **ICA 100-12**. 2016. Available at: <https://publicacoes.decea.mil.br/publicacao/ica-100-12>, Acessado em: 20 nov. 2021.
- DECEA, M. d. D. Comando da A. **ICA 100-37**. 2020. Available at: <https://publicacoes.decea.mil.br/publicacao/ica-100-37>, Acessado em: 20 nov. 2021.
- DOBILAS, S. 2021. Acessado em: 9 nov. 2021. Available at: <https://towardsdatascience.com/dbscan-clustering-algorithm-how-to-build-powerful-density-based-models-21d9961c4cec>.
- ESTER, M.; KRIEGEL, H.-P.; SANDER, J.; XU, X. *et al.* A density-based algorithm for discovering clusters in large spatial databases with noise. *In: kdd. Proceedings [...]*. [S.l.: s.n.], 1996. v. 96, n. 34, p. 226–231.
- FLIGHTRADAR24. 2021. Acessado em: 9 nov. 2021. Available at: <https://www.flightradar24.com/>.
- GARIEL, M.; SRIVASTAVA, A. N.; FERON, E. Trajectory clustering and an application to airspace monitoring. **IEEE Transactions on Intelligent Transportation Systems**, v. 12, n. 4, p. 1511–1524, 2011.
- GUDMUNDSSON, J.; LAUBE, P.; WOLLE, T. Movement patterns in spatio-temporal data. 01 2007.
- MARINHO, D. **HIGHWAYS DO CÉU – POR DENTRO DO FANTÁSTICO MUNDO DAS AEROVIAS**. 2016. Acessado em: 9 nov. 2021. Available at: <http://blog.hangar33.com.br/highways-do-ceu-por-dentro-do-fantastico-mundo-das-aerovias/>.
- MCLEAN, J. 2021. Available at: <https://cran.rstudio.com/web/packages/trajr/vignettes/trajr-vignette.html>, Acessado em: 18 nov. 2021.
- MURÇA, M. C. R.; GUTERRES, M. X.; OLIVEIRA, M. de; SZENCZUK, J. B. T.; SOUZA, W. S. S. Characterizing the brazilian airspace structure and air traffic performance via trajectory data analytics. **Journal of Air Transport Management**, Elsevier, v. 85, p. 101798, 2020.
- MURÇA, M. C. R.; HANSMAN, R. J.; LI, L.; REN, P. Flight trajectory data analytics for characterization of air traffic flows: A comparative analysis of terminal area operations between new york, hong kong and sao paulo. **Transportation Research Part C: Emerging Technologies**, Elsevier, v. 97, p. 324–347, 2018.
- MURÇA, M. **Data-driven modeling of air traffic flows for advanced Air Traffic Management**. Thesis (Doutorado), 01 2018.
- SUN, J.; ELLERBROEK, J.; HOEKSTRA, J. Large-scale flight phase identification from ads-b data using machine learning methods. *In: . Proceedings [...]*. [S.l.: s.n.], 2016.
- ZENG, W.; XU, Z.; CAI, Z.; CHU, X.; LU, X. Aircraft trajectory clustering in terminal airspace based on deep autoencoder and gaussian mixture model. **Aerospace**, v. 8, p. 266, 09 2021.

FOLHA DE REGISTRO DO DOCUMENTO

¹ CLASSIFICAÇÃO/TIPO <p style="text-align: center;">TC</p>	² DATA <p style="text-align: center;">25 de novembro de 2021</p>	³ REGISTRO N° <p style="text-align: center;">DCTA/ITA/TC-138/2021</p>	⁴ N° DE PÁGINAS <p style="text-align: center;">47</p>
⁵ TÍTULO E SUBTÍTULO: <p>Medidas de eficiência do tráfego aéreo: uma comparação entre trajetórias previstas e reais</p>			
⁶ AUTOR(ES): <p>Carlos Alexandre Silva dos Santos</p>			
⁷ INSTITUIÇÃO(ÕES)/ÓRGÃO(S) INTERNO(S)/DIVISÃO(ÕES): <p>Instituto Tecnológico de Aeronáutica - ITA</p>			
⁸ PALAVRAS-CHAVE SUGERIDAS PELO AUTOR: <p>Machine-Learning; Aerovias; DBSCAN</p>			
⁹ PALAVRAS-CHAVE RESULTANTES DE INDEXAÇÃO: <p>Controle de tráfego aéreo; Eficiência; Monitoramento; Aeroportos; Aprendizagem (inteligencia artificial); Transportes.</p>			
¹⁰ APRESENTAÇÃO: (X) Nacional () Internacional <p>ITA, São José dos Campos. Curso de Graduação em Engenharia Civil-Aeronáutica. Orientador: Prof. Dr. Marcelo Xavier Guterres. Publicado em 2021</p>			
¹¹ RESUMO: <p>Com a crescente demanda por transporte aéreo, preocupações com eficiência de voo tem crescido igualmente. A adoção de sistemas como ADS-B permitiram o monitoramento contínuo da grande maioria dos voos comerciais performados em nível mundial. A partir da base de dados fornecidos pelo ADS-B e através de sites de monitoramento como FlightRadar24 além da crescente aplicação de métodos de machine-learning em situações com densa base de dados vem aumentando continuamente em especial para calculos de eficiência e tentativas em se minimizalo. Este trabalho avalia a eficiência de pouso de voos baseado em duas métricas distintas, a primeira através de rotas nominais obtidas a partir do algoritmo DBSCAN de clusterização e através da comparação das rotas reais com as aerovias, rotas teoricas em que os pilotos se baseiam em seus voos. O parâmetro de eficiência analisado foi o HTE, e uma análise estatística sobre as amostras em cada caso, cluster e aerovias, foram comparados via teste de Kolmogorov-Smirnov a fim de se avaliar se tais valores partiram de uma mesma distribuição. Por fim, nosso caso de estudo foi o aeroporto de São Paulo/Congonhas (SBSP) onde se analisou os voos de chegada nesse aeroporto entre setembro e dezembro de 2019, o resultado do teste K-S se mostrou significativo estatisticamente e podemos com alto grau de confiança rejeitar a hipótese de as métricas serem equivalentes.</p>			
¹² GRAU DE SIGILO: <p style="text-align: center;">(X) OSTENSIVO () RESERVADO () SECRETO</p>			