



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
CENTRO DE CIÊNCIAS MATEMÁTICAS E DA NATUREZA
INSTITUTO DE GEOCIÊNCIAS
DEPARTAMENTO DE METEOROLOGIA

LUIZ ALEXANDRE FRAGOZO DE ALMEIDA

**PREVISÃO DE *WIND SHEAR* BASEADA EM APRENDIZADO DE MÁQUINA
PARA O AEROPORTO INTERNACIONAL DE SÃO PAULO/GUARULHOS**

RIO DE JANEIRO
2024

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
CENTRO DE CIÊNCIAS MATEMÁTICAS E DA NATUREZA
INSTITUTO DE GEOCIÊNCIAS
DEPARTAMENTO DE METEOROLOGIA

LUIZ ALEXANDRE FRAGOZO DE ALMEIDA

**PREVISÃO DE *WIND SHEAR* BASEADA EM APRENDIZADO DE MÁQUINA
PARA O AEROPORTO INTERNACIONAL DE SÃO PAULO/GUARULHOS**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Meteorologia (PPGM) da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do grau de Mestre em Meteorologia.

Orientadores: Gutemberg Borges França, *PhD*.
Francisco Leite de Albuquerque Neto, *DSc*.

RIO DE JANEIRO

2024

LUIZ ALEXANDRE FRAGOZO DE ALMEIDA

**PREVISÃO DE *WIND SHEAR* BASEADA EM APRENDIZADO DE MÁQUINA
PARA O AEROPORTO INTERNACIONAL DE SÃO PAULO/GUARULHOS**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Meteorologia (PPGM) da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do grau de Mestre em Meteorologia.

Aprovado por:



Documento assinado digitalmente
GUTEMBERG BORGES FRANCA
Data: 16/12/2024 08:22:22-0300
Verifique em <https://validar.itl.gov.br>

Gutemberg Borges França, *PhD.* (Orientador)



Documento assinado digitalmente
FRANCISCO LEITE DE ALBUQUERQUE NETO
Data: 17/12/2024 08:24:19-0300
Verifique em <https://validar.itl.gov.br>

Francisco Leite de Albuquerque Neto, *D Sc.* (Coorientador)



Documento assinado digitalmente
HAROLDO FRAGA DE CAMPOS VELHO
Data: 16/12/2024 12:16:54-0300
Verifique em <https://validar.itl.gov.br>

Haroldo Fraga Campos Velho, *D Sc.*



Documento assinado digitalmente
WANDERSON LUIZ SILVA
Data: 16/12/2024 13:40:26-0300
Verifique em <https://validar.itl.gov.br>

Wanderson Luiz Silva, *D Sc.*



Documento assinado digitalmente
FABRICIO POLIFKE DA SILVA
Data: 16/12/2024 14:01:05-0300
Verifique em <https://validar.itl.gov.br>

Fabricio Polifke da Silva, *D Sc.*

Rio de Janeiro

2024

AGRADECIMENTOS

A realização desta dissertação marca o fim de uma jornada desafiadora, repleta de aprendizado e crescimento pessoal e profissional. Para chegar até aqui, contei com o apoio e o incentivo de muitas pessoas que, direta ou indiretamente, contribuíram para a concretização deste trabalho. A todos, deixo meu mais sincero agradecimento.

Assim, em primeiro lugar, agradeço a Deus, por ter me concedido força, sabedoria e perseverança ao longo de todo o caminho. Certamente, sem Sua presença constante, as dificuldades que encontrei seriam impossíveis de serem superadas. A fé e a confiança no propósito maior foram essenciais para alcançar esta vitória.

Aos meus pais, cuja dedicação e amor incondicionais sempre foram a base sólida na qual construí a minha vida e os meus sonhos. Vocês me ensinaram o valor do esforço e da integridade, sendo desta forma este trabalho fruto do apoio de vocês.

À minha amada família, pelo carinho, pela paciência e por muitos momentos de compreensão e ausência. Vocês estiveram sempre ao meu lado nos momentos de alegria e também nos desafios, sempre me incentivando a seguir em frente. Sou eternamente grato pelos sacrifícios feitos em prol de mais uma conquista em minha vida.

Agradeço também às minhas chefias, tanto no DECEA quanto no CIMAER, que compreenderam as demandas deste processo acadêmico e me deram o suporte necessário para equilibrar os meus compromissos profissionais com os acadêmicos.

Aos meus amigos e colegas, por todas as palavras de encorajamento, pelas conversas que me ajudaram a encontrar novas perspectivas e pela paciência ao ouvir sobre cada etapa desta jornada.

Por fim, agradeço aos meus orientadores e professores da UFRJ, cujos ensinamentos e orientações foram fundamentais para a realização desta pesquisa. Suas contribuições, tanto técnicas quanto humanas, foram indispensáveis para o desenvolvimento deste trabalho.

Muito obrigado!

RESUMO

Este estudo apresenta a análise de desempenho de algoritmos de aprendizado de máquina para prever eventos de cisalhamento de vento (WS) forte e severo com 15 minutos de antecedência no Aeroporto Internacional de São Paulo/Guarulhos (SBGR) dentro da camada atmosférica entre 30 e 200 metros acima da superfície do solo. Dezesesseis algoritmos foram treinados e testados utilizando-se de dados oriundos de equipamento de Detecção e Alcance Sônicos de alta frequência (SODAR), de 2016 a 2019, usando validação cruzada de 10 amostras em conjuntos de dados segmentados sazonalmente. O algoritmo denominado Random Forest, NBTree e Logistic Model Trees (LMT) demonstraram os melhores desempenhos, com uma Probabilidade de Detecção (POD) atingindo 0,716 e uma F-Measure de 0,774 sob condições atmosféricas estáveis durante o período de junho a agosto, enquanto Naive Bayes e BayesNet alcançaram altas taxas de detecção, POD igual a 0,985 e 0,866, respectivamente, mas tenderam a superestimar as ocorrências do WS. Devido à falta de índices termodinâmicos, os modelos dependeram apenas de preditores específicos gerados a partir de dados oriundos unicamente do SODAR, limitando sua precisão para eventos de WS de origem convectiva, mas produzindo forte desempenho em condições atmosféricas predominantemente estáveis. Adicionar índices termodinâmicos essenciais para identificar o cisalhamento convectivo pode melhorar as previsões para os eventos convectivos. Este estudo ressalta a importância da previsão adaptativa para condições atmosféricas variáveis, aumentando a segurança no SBGR e oferecendo uma estrutura para aeroportos que enfrentam desafios meteorológicos semelhantes.

Palavras-chave: previsão de cisalhamento do vento, aprendizado de máquina, SODAR, análise de dados meteorológicos.

ABSTRACT

This study presents the performance analysis of machine learning algorithms to predict strong and severe wind shear (WS) events 15 minutes in advance at São Paulo/Guarulhos International Airport (SBGR) within the atmospheric layer between 30 and 200 meters above the ground surface. Sixteen algorithms were trained and tested using data from high-frequency Sonic Detection and Ranging (SODAR) equipment from 2016 to 2019, using 10-sample cross-validation on seasonally segmented datasets. The Random Forest, NBTree and Logistic Model Trees (LMT) algorithms demonstrated the best performances, with a Probability of Detection (POD) of 0.716 and an F-Measure of 0.774 under stable atmospheric conditions during the period from June to August, while Naive Bayes and BayesNet achieved high detection rates, POD equal to 0.985 and 0.866, respectively, but tended to overestimate WS occurrences. Due to the lack of thermodynamic indices, the models relied only on specific predictors generated from data coming exclusively from SODAR, limiting their accuracy for WS events of convective origin, but producing strong performance in predominantly stable atmospheric conditions. Adding thermodynamic indices essential to identify convective shear can improve forecasts for convective events. This study highlights the importance of adaptive forecasting for changing atmospheric conditions, increasing safety at SBGR and providing a framework for airports facing similar meteorological challenges.

Keywords: wind shear prediction, machine learning, SODAR data, meteorological data analysis.

LISTA DE FIGURAS

Figura 1 Geração de empregos, salários, produção e impostos no Brasil decorrentes da aviação	10
Figura 2 Acidentes relacionados e não-relacionadas às condições de tempo	13
Figura 3 Ocorrências aeronáuticas tendo como fator contribuinte as condições meteorológicas, no período entre 2014 e 2023	13
Figura 4 Wind Shear e suas outras nomenclaturas	14
Figura 5 Ranking de aeródromos – 2023	16
Figura 6 Relatórios de WS entre 1999 e 2012.....	16
Figura 7 Tipos de Wind Shear.....	22
Figura 8 Camada limite atmosférica. Estrutura e ciclo diário	27
Figura 9 Forças atuantes sobre uma aeronave	29
Figura 10 Funcionamento do SODAR	33
Figura 11 Paradigma da Programação.....	35
Figura 12 Tela da plataforma WEKA.....	38
Figura 13 Região do Aeroporto Internacional de Guarulhos/SP	40
Figura 14 Imagem de satélite do SBGR, com a localização do SODAR.....	41
Figura 15 Gráfico de Direção e Velocidade do vento no SBGR	42
Figura 16 Composição do APRun 1.51, apresentando na tela superior os dados de WS obtidos pelo SODAR na relação tempo (a cada 15 minutos) versus altura relativa do solo (a cada 10 metros) e na tela inferior os dados das barbelas com dados de SPD e DIR obtidos pelo SODAR na relação tempo versus altura.....	46
Figura 17 Mensagens de Observação à Superfície do tipo METAR, Carta Sinótica de Superfície e Imagem de Satélite (Goes 13). Fonte: O Autor.....	48
Figura 18 Fluxograma do método	50
Figura 19 Disponibilidade de dados de WS com a altura (2016 a 2019)	51
Figura 20 a) Distribuição do WS para intensidades fortes-severas, detectadas pelo sistema SODAR, para os períodos mensais de dezembro-fevereiro, março-maio, junho-agosto e setembro-novembro, de 1º de janeiro de 2016 a 31 de dezembro de 2019; b) Distribuição dos pousos e decolagens de aeronaves no SBGR, segundo o anuário estatístico de 2023 do DECEA	53

LISTA DE TABELAS

Tabela 1 Classificação de intensidade do Wind Shear	23
Tabela 2 Características dos Algoritmos	39
Tabela 3 Dados usados para o período de 1º de janeiro de 2016 a 31 de dezembro de 2019. Características de recursos (preditores de entrada) e saída (alvos) foram utilizadas para treinar e testar algoritmos de aprendizado de máquina para previsões de WS de 15 minutos. O WS significa que são variáveis derivadas de dados de wind shear, o SPD às variáveis derivadas de dados de velocidade do vento e o DIR às variáveis derivadas da direção do vento.....	45
Tabela 4 Matriz de contingência WEKA.....	49
Tabela 5 Métricas de validação cruzada com 10 amostras para 16 algoritmos de aprendizado de máquina, usando dados de todos os meses (janeiro a dezembro) durante o período de 2016 a 2019, capturando métricas para eventos de WS com gênese de atmosfera convectiva e estável.....	56
Tabela 6 Métricas de validação cruzada com 10 amostras para 16 algoritmos de aprendizado de máquina usando dados do JJA e JJA_Stable. O JJA inclui eventos de cisalhamento de vento forte-severo relacionados a atmosferas convectivas e estáveis, enquanto o JJA_Stable inclui apenas eventos de WS forte-severo em atmosfera estável com detecção contínua superior a 30 minutos, para o período de 2016–2019	59

LISTA DE ABREVIATURAS E SIGLAS

ABEAR	Associação Brasileira de Empresas Aéreas
AGL	Above Ground Level
CGNA	Centro de Gerenciamento de Navegação Aérea ⁷
CENIPA	Centro Investigação e Prevenção de Acidentes Aeronáuticos
CPTEC	Centro de Pesquisa e Tecnologia
DASA	Digital Airspace System Analysis
DECEA	Departamento de Controle do Espaço Aéreo
FAA	Federal Aviation Administration
ICEA	Instituto de Controle do Espaço Aéreo
ICAO	International Civil Aviation Organization
INPE	Instituto Nacional de Pesquisas Espaciais
IPCC	Intergovernmental Panel on Climate Change
LIDAR	Light Detection and Ranging
LLWAS	Low Level Wind Shear Alert System
METAR	Meteorological Aerodrome Report
NASA	National Aeronautics and Space Administration
RNA	Rede Neural Artificial
RWY	Runway
SBGR	Aeroporto Internacional de São Paulo/Guarulhos
SISCEAB	Sistema de Controle do Espaço Aéreo Brasileiro
SODAR	SONic Detection and Ranging
TDWR	Terminal Doppler Weather Radar
WEKA	Waikato Environment for Knowledge Analysis
WS	Wind Shear

SUMÁRIO

1. INTRODUÇÃO.....	10
1.1 Contexto do Estudo.....	10
1.2 Motivação.....	12
1.3 Objetivo.....	17
1.3 Revisão Bibliográfica.....	17
2. FUNDAMENTAÇÃO TEÓRICA.....	22
2.1. <i>Wind Shear</i>	22
2.2. Condições Meteorológicas Associadas a Eventos de WS.....	23
2.3. Camada Limite Atmosférica (CLA).....	26
2.4. Efeitos do WS no Desempenho das Aeronaves.....	29
2.5. Sistemas de Detecção de WS.....	30
2.5.1. SODAR.....	32
3. PROCESSAMENTO E ANÁLISE DE DADOS.....	34
3.1. Aprendizado de Máquina.....	34
3.2. WEKA.....	37
4. DADOS E MÉTODO.....	40
4.1. Área de Estudo.....	40
4.2. Dados.....	43
4.3. Método.....	47
4.4. Métricas de Avaliação.....	49
5. RESULTADOS E DISCUSSÕES.....	51
5.1. Análise da ocorrência de tURbulência e do WS nas proximidades do SBGR.....	51
5.2. Treinamento e teste de algoritmo de aprendizado de máquina.....	53
5.2.1. Avaliação dos Algoritmos.....	55
5.2.1.1. Análise de algoritmos de aprendizado de máquina para previsão de WS.....	55
5.2.1.2. Análise comparativa dos conjuntos de dados de JJA e JJA_Stable.....	57
6. CONCLUSÃO.....	60
7. REFERÊNCIAS BIBLIOGRÁFICAS.....	63

1. INTRODUÇÃO

1.1 CONTEXTO DO ESTUDO

O setor de transporte aéreo do Brasil é vital para a conectividade nacional e o crescimento econômico, suportando mais de 2,8 milhões de movimentos de pousos e decolagens anualmente (DECEA, 2023). Observa-se, através da figura 1, que em 2019, ano anterior aos efeitos globais da pandemia da COVID-19, o transporte aéreo contribuiu com aproximadamente 1,4% do PIB do Brasil, gerando 1,4 milhão de empregos, direta e indiretamente, incluindo empregos relacionados ao turismo. O setor arrecadou mais de R\$ 32 bilhões em tributos naquele ano para os cofres públicos (ABEAR, 2020).

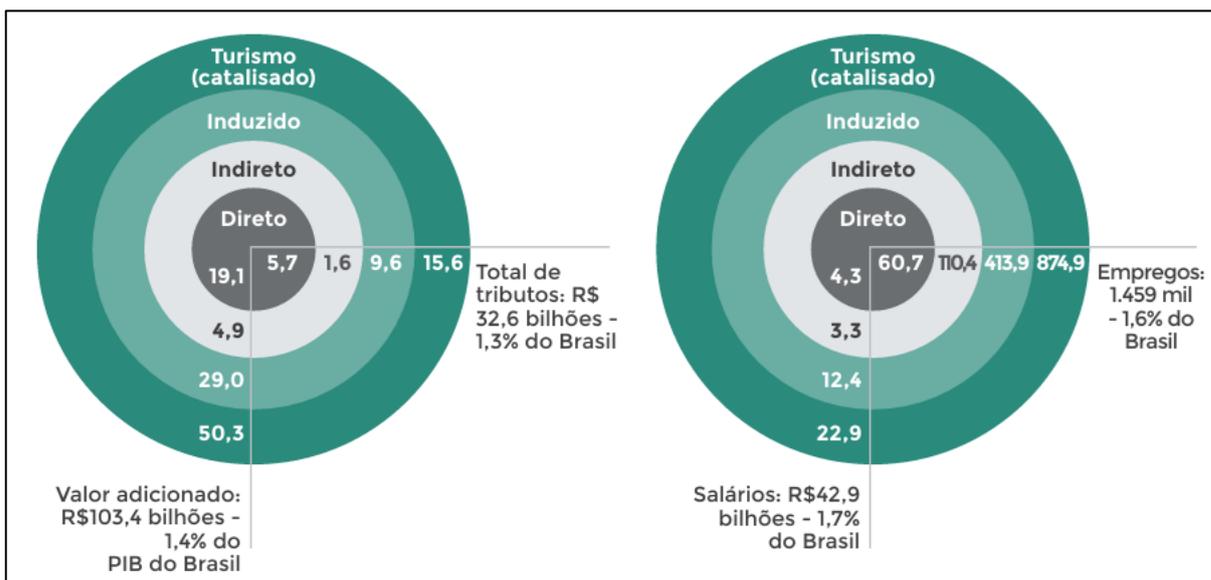


Figura 1 Geração de empregos, salários, produção e impostos no Brasil decorrentes da aviação
Fonte: ACSP, ANAC, IBGE, IPEA e MTur. Elaboração ABEAR 2020

Conforme a *International Civil Aviation Organization* (ICAO, 2016), em seu relatório “DOC 9750-NA/963”, o tráfego aéreo vem dobrando de tamanho a cada 15 anos, desde 1977, e continuará nessa tendência de taxa de crescimento. Assim, a aviação como parte de um setor em constante crescimento e com projeções futuras de aumento significativo na demanda pela utilização de seus serviços, necessita da maximização de sua eficiência, com conseqüente redução de custos operacionais,

otimização de utilização do espaço aéreo e implementação constante de medidas de prevenção e segurança.

A ICAO, criada a partir da decisão da Convenção de Aviação Civil Internacional (Chicago, 1944) é uma agência especializada das Nações Unidas. Seu objetivo é promover o desenvolvimento de princípios e técnicas para a navegação aérea internacional visando garantir a segurança, a eficiência, a economia e o desenvolvimento dos serviços aéreos

O DECEA, como órgão representante do Brasil junto à ICAO e responsável pelo provimento e controle do serviço de navegação aérea no Brasil, estabelece normas e regulamentações para que essa atividade seja guiada pelo comprometimento de todos na prestação desse serviço com total segurança e máxima eficiência.

A ICAO (2016) aponta que devido à complexidade do sistema aeronáutico a capacidade da atividade de operação aérea é significativamente impactada por eventos externos, como as condições de tempo. Assim, para manter altos níveis de segurança e eficiência, devem-se priorizar investimentos significativos em novas tecnologias em todas as áreas que venham a impactar na prestação do serviço aeronáutico.

O Brasil, através da atuação do DECEA comprometeu-se a manter um sistema de vigilância e divulgação de informações meteorológicas aeronáuticas aos usuários do espaço aéreo brasileiro visando a elevação contínua dos níveis de segurança e eficiência da atividade aeronáutica, e para isso responsabilizou-se em prover o serviço de Meteorologia Aeronáutica aos usuários do SISCEAB.

Através de uma rede complexa de estações meteorológicas de observação à superfície, estações de observação meteorológicas de altitude e centros operacionais de meteorologia aeronáutica, o DECEA fornece apoio aos usuários do SISCEAB com informações meteorológicas primordiais à segurança das operações aéreas.

Em vista disso, o DECEA, através da atuação de seus Subdepartamentos, está em constante busca pelo aprimoramento de equipamentos, sensores, ferramentas, serviços, sistemas e capacitação de seu pessoal para assegurar à sociedade

brasileira e internacional seu comprometimento na maximização da eficiência do transporte aéreo e na segurança de todos aqueles que, de alguma forma, estejam envolvidos com a atividade aeronáutica.

1.2 MOTIVAÇÃO

Conforme o Relatório de Situação 2008-2015 (volume 1) do Instituto de Física Atmosférica (Institute of Atmospheric Physics - IPA), pertencente ao Centro Aeroespacial Alemão (Deutsches Zentrum für Luft-und Raumfahrt - DLR), as condições atmosféricas adversas com fenômenos perigosos como turbulência, cisalhamento do vento, raios, granizo e gelo, assim como eventos raros como nuvens de cinzas vulcânicas, podem levar a interrupções significativas no tráfego aéreo.

Projeta-se também que poderá ocorrer o aumento de impactos na atividade aérea em decorrência do aumento de eventos meteorológicos extremos, principalmente àqueles relacionados aos impactos causados pelo aquecimento global. Nesse sentido, o Painel Intergovernamental sobre Mudanças Climáticas (IPCC) enfatiza a importância de uma abordagem colaborativa e global para enfrentar esse desafio, reconhecendo que o aquecimento global é um problema que afeta a todos e requer esforços conjuntos para ser resolvido. Desta forma, fatores relacionados a eficiência e a segurança do espaço aéreo justificam os custos destinados ao aprimoramento de sistemas que forneçam dados e informações mais confiáveis e disponíveis no menor espaço de tempo, assim como o aperfeiçoamento de métodos e ferramentas de previsões meteorológicas mais oportunas e precisas.

A agência subordinada ao Departamento de Transporte dos Estados Unidos que regulamenta a aviação civil e militar no país, a *Federal Aviation Administration* (FAA), realizou estudo relacionando os acidentes de aviação e as condições meteorológicas, entre os anos de 2003 e 2007 (NTSB, 2010), conforme pode-se observar na figura 2. Nele foram contabilizados 8.657 acidentes aéreos, sendo 1.740 deles (20,1%) tendo as condições meteorológicas do tempo como uma das causas ou como fator contribuinte.

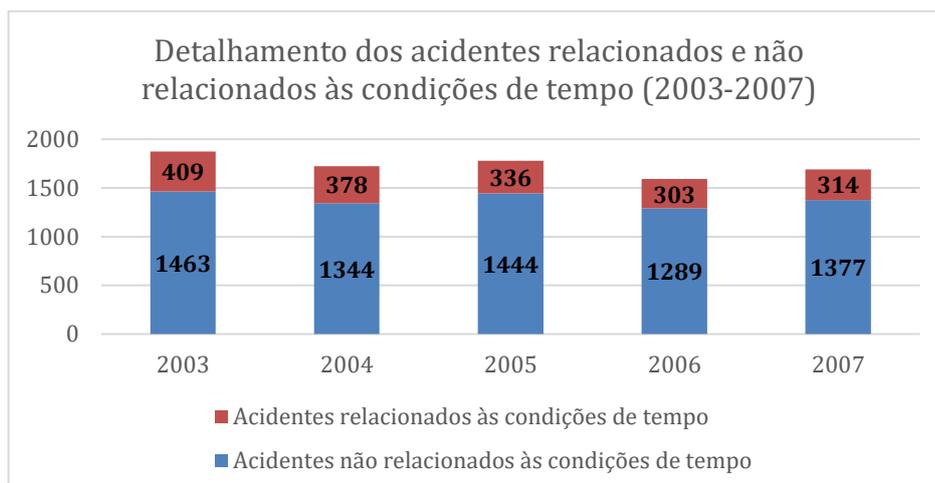


Figura 2 Acidentes relacionados e não-relacionadas às condições de tempo
Fonte: Adaptado de NTSB 2010

No Brasil, o CENIPA disponibiliza através do Painel do Sistema de Investigação e Prevenção de Acidentes Aeronáuticos (SIPAER) os dados relativos às ocorrências aeronáuticas ocorridas no Brasil nos últimos 10 anos. Assim, é possível observar pela figura 3 que, entre 2014 e 2023, as condições meteorológicas ainda foram um fator contribuinte em 94 ocorrências aeronáuticas no país.

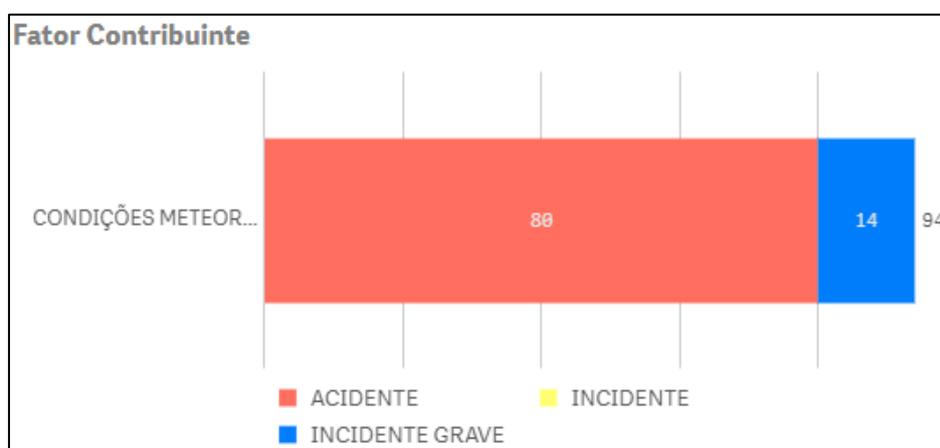


Figura 3 Ocorrências aeronáuticas tendo como fator contribuinte as condições meteorológicas, no período entre 2014 e 2023
Fonte: Adaptado de Painel SIPAER 2024

Nesse sentido, estudos e pesquisas buscando compreender e prever (ou prognosticar) a interferência das condições atmosféricas em acidentes aeronáuticos torna-se de fundamental importância para prevenir acidentes e garantir a segurança das operações aéreas, seja pelo incremento de conhecimentos acerca dos fenômenos meteorológicos ou pelo aperfeiçoamento de sistemas e ferramentas de observação, alerta e previsão do tempo.

Dentre os fenômenos meteorológicos que impactam as atividades aeronáuticas pode ser citado o *Wind Shear* (WS), também conhecido por vários outros nomes, como apresentado na figura 4, tais como: cisalhamento de vento, cortante de vento, gradiente de vento e tesoura de vento.

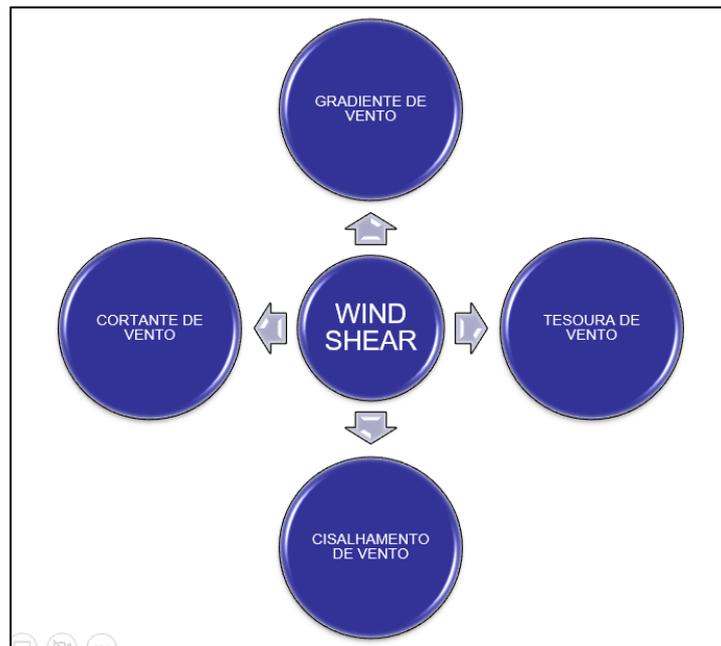


Figura 4 Wind Shear e suas outras nomenclaturas
Fonte: O autor

O cisalhamento do vento é uma mudança na velocidade e/ou direção do vento em uma curta distância, podendo ocorrer tanto horizontalmente quanto verticalmente. Ele é mais frequentemente associado a tempestades e a fortes inversões de temperatura ou gradientes de densidade e pode ocorrer tanto em altas quanto em baixas alturas (FAA, 2008). Nessa perspectiva, o estudo do WS torna-se de extrema importância para a aviação, pois esse fenômeno meteorológico pode representar um grande risco para a segurança de voo, principalmente quando de sua ocorrência em baixas alturas em relação ao solo.

Assim, no momento que uma aeronave adentra em uma região afetada pelo cisalhamento de vento ela pode sofrer uma perda ou ganho súbito de altitude e perda de sustentação devido a mudanças inesperadas na velocidade e direção do vento. E o fato mais relevante é que se a aeronave estiver em baixa altura, tais como durante os procedimentos de pouso ou decolagem, pode levar a ocorrência de um acidente aeronáutico. Portanto, a compreensão, a detecção e, se possível, a previsão do WS são essenciais para garantir a segurança das aeronaves.

Villela (2005) relata que o reconhecimento sobre o WS, como causa de acidentes nas fases de pouso e decolagem, teve início em 1975 com a análise do Professor Tetsuya Theodore Fujita sobre o acidente com o Boeing 727 da Eastern Airlines, em 27 de junho daquele ano, no aeroporto John F. Kennedy, em Nova Iorque. Entretanto, recorda que o acidente ocorrido em 24 de junho de 1956, em Kano, região norte da Nigéria foi o pioneiro a citar nas investigações o WS como causa efetiva de um acidente aeronáutico.

Contribuindo na relevância do tema, conforme ICAO (2005), de 1964 a 1983, o cisalhamento do vento em baixos níveis foi associado como fator contribuinte em, pelo menos, 28 grandes acidentes e incidentes aeronáuticos e resultando em mais de 500 mortes e 200 feridos.

Assim, as companhias aéreas e os órgãos responsáveis pela aviação civil têm investido cada vez mais em tecnologias para monitorar e detectar o WS, como radares meteorológicos, sistemas de alerta e prevenção de acidentes, instrumentos de medição de velocidade e direção do vento, entre outros. Além disso, os pilotos são estimulados a participar de inúmeros treinamentos para reconhecer o mais rápido possível o fenômeno e lidar com o WS, através de manobras e procedimentos específicos para minimizar os riscos desse fenômeno.

Contribuindo de forma relevante à importância do estudo do tema, pode-se observar que de acordo com o Anuário Estatístico de Tráfego Aéreo de 2023 (DECEA, 2023) o Aeroporto Internacional de São Paulo/Guarulhos (SBGR) é o de maior movimento no cenário nacional (figura 5), com 276.809 pousos e decolagens. Conforme Matschinske (2013), o aeroporto de Guarulhos também foi o aeródromo com o maior número de relatórios de ocorrência do fenômeno WS por aeronaves durante os procedimentos de pouso e decolagem, com 1857 ocorrências (figura 6).



Figura 5 Ranking de aeródromos – 2023

Fonte: Adaptado de Anuário Estatístico de Tráfego Aéreo 2023. CGNA, DECEA

Nesse sentido, devido ao alto risco que o fenômeno WS representa para as operações aéreas, a ICAO (2005) recomenda o investimento em equipamentos capazes de registrar a ocorrência do cisalhamento de vento e que haja esforços na realização de estudos e pesquisas voltadas para o desenvolvimento e aprimoramento de métodos capazes de prever o fenômeno.

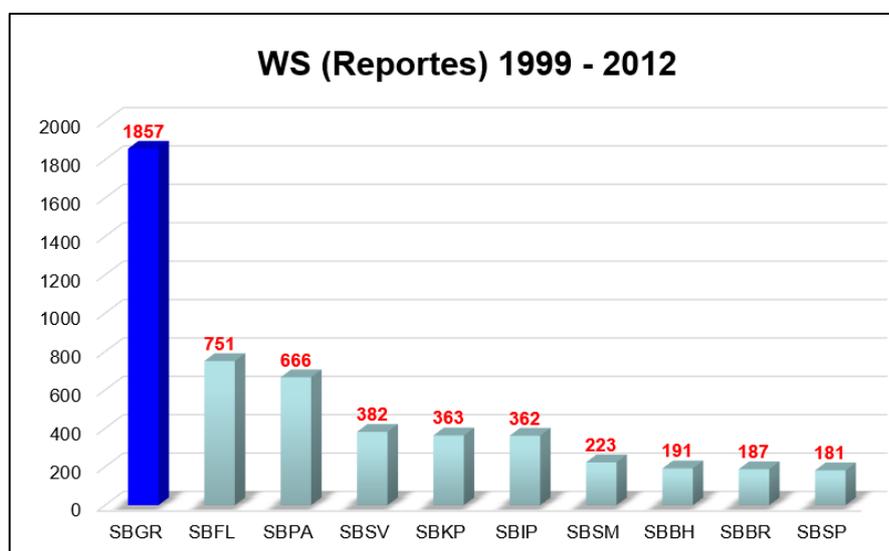


Figura 6 Relatórios de WS entre 1999 e 2012

Fonte: Adaptado de Matschinske (2013)

Sendo assim, apesar do cisalhamento do vento poder ocorrer tanto em altas quanto em baixas alturas (FAA, 2008), conforme SILVA et al (2016) os primeiros 500m da camada atmosférica mais próximos à superfície do solo, são os mais críticos devido as aeronaves estarem nos estágios iniciais e finais das operações de voo, possuem

velocidade relativamente mais baixas e com suas capacidades de manobra comprometidas.

1.3 OBJETIVO

Este estudo teve como objetivo desenvolver um método baseado em aprendizado de máquina para a previsão de ocorrência de WS com 15 minutos de antecedência, na camada atmosférica compreendida entre 30 e 200 metros acima da superfície do solo, no Aeroporto Internacional de São Paulo/Guarulhos, utilizando dados do SODAR.

Ao contribuir para o desenvolvimento de métodos de previsão de WS mais confiáveis, esta pesquisa propõe-se a melhorar a segurança nos procedimentos de aproximação, pouso e decolagem, particularmente no SBGR, onde o WS é um fenômeno frequente.

1.3 REVISÃO BIBLIOGRÁFICA

No início do século XX, Vihelm Bjerknes, físico, matemático e meteorologista norueguês, ousou descrever as condições futuras da atmosfera terrestre através da integração das equações diferenciais que governam o comportamento dos parâmetros meteorológicos. Entretanto, conforme Ynoue et al (2014) foi o cientista britânico Lewis Fry Richardson que realizou a primeira integração numérica compreensível dessas equações. Percebe-se que a busca incessante pelo conhecimento, motivou inúmeras pesquisas e estudos, que se concentravam no estudo físico dos fenômenos meteorológicos e na tentativa de classificá-los e reconhecer suas origens e seus efeitos decorrentes.

Ynoue et al (2014) citam que aproveitando-se do desenvolvimento dos recursos computacionais, e do interesse na predição do estado futuro das condições de tempo e de clima, em 1950, os cientistas Jule Charney, Ragnar Fjørtoft e J. Von Neuman realizaram a primeira previsão do tempo para um dia com auxílio de um computador. E, a partir de 1955, nos Estados Unidos, esse tipo de previsão utilizando-se de computadores passou a ser realizada de forma contínua.

A modelagem numérica do tempo aplicada a meteorologia evoluiu significativamente desde suas origens na década de 1950. Inicialmente, esses modelos eram baseados em equações diferenciais que descreviam a dinâmica e a termodinâmica da atmosfera, mas eram limitados pela capacidade computacional da época, entretanto com o avanço da computação, especialmente a partir dos anos 1980 e 1990, os modelos numéricos tornaram-se mais sofisticados, incorporando maiores resoluções espacial e temporal, além de processos físicos mais complexos.

No entanto, apesar dos avanços tecnológicos, a modelagem numérica do tempo ainda enfrenta alguns desafios significativos, tais como a representação de fenômenos de pequena escala, como por exemplo a previsão de áreas de tempestades convectivas e de regiões de formação de nevoeiro. Outro fator a ser considerado negativamente é a gama de incertezas nas previsões devido a inconsistências causadas por erros e inconsistências, assim como por limitações na obtenção de dados de entrada.

Nesse contexto, os algoritmos de aprendizado de máquina têm se mostrado promissores, pois podem analisar grandes volumes de dados meteorológicos e identificar padrões mais complexos que podem não ser tão facilmente capturados por modelos numéricos do tempo convencionais. Além disso, o aprendizado de máquina poderia ser aproveitado também como forma de otimizar a assimilação de dados e melhorar a qualidade das condições iniciais usadas pelos próprios modelos numéricos.

Assim, Aguiar e Silva (2022) citam que na década de 1950, iniciaram-se estudos na área de aprendizado de máquina e devido ao constante desenvolvimento tecnológico, o processo de previsão meteorológica utilizando-se das técnicas de *machine learning* passou a representar uma alternativa viável aos modelos físicos de previsão do tempo.

Hsieh et al. (2015) realizaram um estudo sobre a aplicação de modelos de redes neurais artificiais (RNA) para previsão e análise de dados em meteorologia e oceanografia. Nesse trabalho, inicialmente, os autores dividiram os métodos empíricos e estatísticos utilizados na meteorologia e oceanografia em 4 (quatro) estágios distintos: regressão linear (e correlação), análise de componentes principais,

análise de correlação canônica, e modelos de RNA. Nesse trabalho enfatizaram que apesar da grande popularidade dos modelos de redes neurais em muitos campos, existem alguns obstáculos para a adaptação do método de redes neurais à meteorologia-oceanografia, mas que pesquisas recentes mostram que esses obstáculos podem ser superados e o método de redes neurais está evoluindo para uma técnica versátil e poderosa, capaz de aumentar os métodos estatísticos lineares tradicionais na análise e previsão de dados, e que sua vinculação a modelos dinâmicos de assimilação de dados resulta em uma nova classe de modelos híbridos neurodinâmicos.

Gultepe et al. (2019) citam em seu trabalho que os impactos dos processos atmosféricos na aviação é reconhecido desde 1900, e nesse sentido, realizaram um estudo revisando o alto impacto das condições do tempo sobre as atividades aeronáuticas, tais como nevoeiro, precipitação, gelo, turbulência, wind shear e tornados. O artigo agrupa o conhecimento atual sobre meteorologia aeronáutica, apresentando os recursos disponíveis e os problemas enfrentados pela aviação relacionados aos fatores meteorológicos, realizando uma série de avaliações quanto as melhorias necessárias nas medições, nas técnicas e modelos de previsão do tempo e os impactos decorrentes das mudanças climáticas na atividade aeronáutica.

Nessa perspectiva, percebe-se a necessidade de intensificar a parcela de pesquisas e estudos voltados especificamente à meteorologia aeronáutica. Assim, a implementação de novas técnicas e ferramentas são primordiais para a redução da subjetividade causada pelas percepções individuais de cada previsor meteorologista na elaboração de seus prognósticos e gerar o aperfeiçoamento do conhecimento acerca dos fenômenos meteorológicos que impactam as operações aéreas contribuindo na tarefa de apoio eficiente e seguro da atividade de tráfego aéreo.

Almeida (2009) utilizou dados de 56 anos da estação meteorológica de superfície do SBGR e técnicas de aprendizado de máquina, no desenvolvimento de um sistema de previsão de curtíssimo prazo (nowcasting), para prever o teto da camada de nuvens e a visibilidade horizontal no SBGR. Nesse trabalho, foram testadas três abordagens metodológicas: uma com Regressão Linear Múltipla (RLM) e duas com Redes Neurais Artificiais (RNA).

Pereira (2016) utilizando os dados meteorológicos oriundos da estação meteorológica de superfície e da estação meteorológica de altitude do Aeroporto Internacional Afonso Pena - Curitiba/PR - analisou o emprego de três redes neurais artificiais na previsão de nevoeiro naquele aeroporto. Em seu estudo, verificou que a técnica utilizando climatologia apresentou razão de falso alarme em torno de 96% e a técnica de persistência apresentou índices de desempenho muito próximos aos apresentados pelas redes neurais.

Aliado às técnicas utilizando as RNA, além da utilização de dados oriundos das estações meteorológicas de superfície e de altitude, dados de radar e satélites meteorológicos como fonte de dados, a manipulação e análise dos dados obtidos por equipamentos de medição instalados em aeronaves e em solo, igualmente estão sendo importantes no entendimento de fenômenos meteorológicos que impactam na aviação e na consolidação de técnicas que contribuem de forma mais ágil e eficiente nas previsões desses fenômenos, tais como o nevoeiro e o cisalhamento de vento.

Nesse sentido, Crescenti (1996) realizou um estudo compilando e discutindo os resultados de duas décadas de vários experimentos utilizando SODAR Doppler, participando que sua utilização possibilitou avanços significativos na compreensão da camada limite atmosférica e concluindo que os resultados de muitos desses estudos mostraram que o equipamento pode obter perfis confiáveis da velocidade média do vento e da direção do vento. Do mesmo modo, o SODAR foi utilizado para estudar vários outros fenômenos meteorológicos, tais como: circulações costeiras de terra/brisa marítima, a camada limite estável da Antártica, o fluxo de vento em terreno complexo, ondas de gravidade, jatos de baixo nível e passagens frontais.

Silva (2016) utilizou dados do SODAR do SBGR aliados as informações obtidas das mensagens de observação à superfície METAR/SPECI, de imagens de satélite e radar meteorológico para analisar as ocorrências de WS no SBGR. Verificou que nessas condições, eventos severos e fortes são influenciados pela topografia da Serra da Cantareira e pela intensificação da componente norte da direção do vento sobre a região do aeroporto. No entanto, casos de WS com intensidade forte e moderada também podem ocorrer em situações não necessariamente associadas a condições frontais ou pré-frontais.

Silva *et al.* (2016) qualificaram e quantificaram as vantagens do uso do SODAR com o objetivo de estabelecer um modelo conceitual para o estabelecimento da indicação de mudança de pista no SBGR. Pela análise de 234 casos de mudanças de pista foi verificado que ocorreu a mudança antecipada na direção do vento em altitude e posteriormente essa mudança passou a predominar nos ventos mais próximos à superfície. Desta forma o uso do SODAR permite ajudar os controladores de tráfego aéreo a visualizar antecipadamente essa mudança de direção do vento e redirecionar o fluxo de tráfego aéreo promovendo vantagens de viés econômico e de segurança com um tempo hábil para uma medida de segurança de cerca de 40 minutos.

Mantendo esse foco, França *et al.* (2018) criaram um modelo de *nowcasting* utilizando dados do SODAR e utilizaram-se de uma rede neural como ferramenta para prever os perfis verticais de vento a baixos níveis em até 45 minutos para o SBGR. Os estudos demonstraram que as previsões obtiveram melhor desempenho em níveis superiores (300m), declinando em níveis mais próximos à superfície (10m), assim como quanto ao período da previsão, onde o modelo apresentou diminuição da precisão à medida que o período da previsão aumentava.

Dessa forma, verifica-se que os eventos meteorológicos impactam as atividades aeronáuticas e que o estudo desses fenômenos é de extrema relevância para utilização eficiente e segura do espaço aéreo. Nesse sentido, vários estudos e pesquisas visam contribuir com o aprofundamento dos conhecimentos correlacionados aos impactos da meteorologia na aviação e que o advento tecnológico promoveu a busca e o aperfeiçoamento de ferramentas que possam colaborar com a tarefa de previsão do tempo. Assim, através do intercâmbio de ferramentas e técnicas desenvolve-se o aprimoramento dos métodos de pesquisa e na elevação da precisão almejada desde os primórdios da evolução humana.

2. FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, apresenta-se uma revisão sobre os conceitos envolvendo o fenômeno *Wind Shear*, os seus efeitos nas aeronaves e os principais sistemas e ferramentas utilizadas atualmente para sua detecção e previsão.

2.1. WIND SHEAR

Conforme definido no Manual sobre Cisalhamento do Vento de Baixo Nível (ICAO, 2005), o *Wind Shear* é uma mudança na velocidade e/ou direção do vento no espaço (figura 7), incluindo correntes ascendentes e descendentes. Dessa forma, pode-se associar a esse fenômeno meteorológico, a depender de sua intensidade, a alguns impactos significativos para a aviação, principalmente se ocorrerem nas camadas mais próximas à superfície.

Portanto, a ocorrência do *Wind Shear* até 500m (1600 pés), torna-se fator preocupante em termos de segurança, pois seus efeitos sobre a aeronave impactam na sustentação nos momentos mais críticos de um voo (procedimentos de pouso e decolagem) e podem contribuir para a ocorrência de incidentes ou acidentes aeronáuticos, pois nessas fases do voo os pilotos têm pouco tempo para identificar a ocorrência do fenômeno e efetivar as medidas de controle necessárias para a recuperação das condições de segurança.

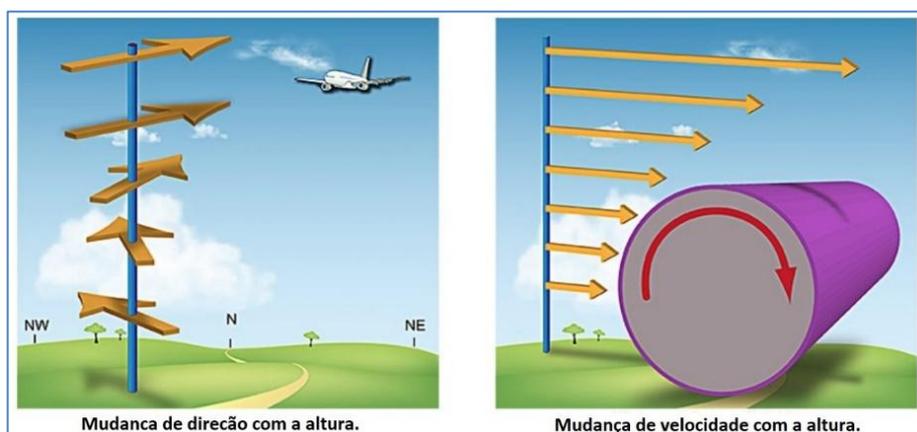


Figura 7 Tipos de Wind Shear

Fonte: Adaptado <https://slcc.pressbooks.pub/physicalgeography/chapter/9-4/>

Nesse sentido, conforme FAA (1988), o piloto possui apenas de 5 a 15 segundos para reconhecer e reagir a um encontro com o WS durante os procedimentos de aproximação e pouso da aeronave.

A intensidade do WS é calculada dividindo-se a magnitude da diferença vetorial entre dois pontos pela distância entre eles. Nesse sentido, sua intensidade é calculada em unidades de velocidade do vento por unidade de distância (tanto horizontal quanto vertical), sendo que se tornou conveniente no meio aeronáutico a utilização do cálculo do WS ocorrido pelas diferenças de intensidade do vento na vertical, utilizado-se as unidades Km/h por 30 m, m/s por 30 m ou kt (nós) por 100 pés.

Assim, são adotados pelo DECEA, os critérios para classificação da intensidade de cisalhamento de vento, recomendados na 5ª Conferência de Navegação Aérea ocorrida em Montreal, no ano de 1967 (tabela 1).

INTENSIDADE	VARIAÇÃO
LEVE	0 A 4 kt / 30 m (100 ft)
MODERADA	5 A 8 kt / 30 m (100 ft)
FORTE	9 A 12 kt / 30 m (100 ft)
SEVERA	> 12 kt / 30 m (100 ft)

Tabela 1 Classificação de intensidade do Wind Shear
Fonte: Adaptado de ICAO

Entretanto, cabe ressaltar que nesses cálculos, são consideradas as medições em uma atmosfera livre e a intensidade do WS para fins aeronáuticos torna-se mais subjetiva, pois neste setor está sujeita a inúmeros fatores, tais como: o tipo de aeronave, a velocidade de passagem sobre a área afetada, a extensão da área de ocorrência do cisalhamento de vento, o tempo de exposição da aeronave ao fenômeno e a relação entre as direções do vetor vento e da trajetória de voo da aeronave no decorrer dos procedimentos de pouso e decolagem.

2.2. CONDIÇÕES METEOROLÓGICAS ASSOCIADAS A EVENTOS DE WS

Apesar de várias condições meteorológicas estarem associadas à geração de áreas com a presença de WS, de acordo com FAA (2008) as quatro fontes mais comuns que originam o WS em baixos níveis são: os sistemas frontais, as tempestades, as inversões de temperatura e os obstáculos à superfície.

Conforme FAA (2008), nem todas os sistemas frontais estarão associados a regiões com WS. Entretanto, sistemas frontais, como frentes frias, quentes ou oclusas, podem causar WS significativo e perigoso para as aeronaves. Massas de ar com diferentes características se encontram, ocorrendo uma mudança abrupta de temperatura e umidade, que pode criar cisalhamento vertical do vento. Quando uma frente fria se aproxima, por exemplo, pode haver um aumento súbito da velocidade do vento e uma mudança na sua direção à medida que a frente se desloca. Essa mudança abrupta na velocidade e direção do vento pode causar cisalhamento de vento horizontal e vertical. Além disso, a presença de nuvens cumulonimbus associadas a sistemas frontais pode indicar a presença de turbulência e outras condições meteorológicas adversas que podem afetar a segurança das aeronaves.

Por sua vez, tempestades podem estar associadas a diferentes condições atmosféricas, não exclusivamente associadas a sistemas frontais. Áreas de baixa pressão, calor excessivo e umidade elevada também são fatores desencadeadores de geração de tempestades. Essas condições podem desencadear, de forma extremamente rápida, a formação local de nuvens cumulonimbus, que são as principais responsáveis pela ocorrência de tempestades com raios e trovões. Assim, à medida que a tempestade se desenvolve, as correntes ascendentes e descendentes de ar quente e frio dentro da nuvem podem criar o cisalhamento de vento, que pode afetar as aeronaves em diferentes altitudes.

O WS pode estar associado também a inversões térmicas em baixas alturas. As inversões térmicas ocorrem quando o ar frio fica preso abaixo do ar quente. Uma inversão térmica ocorre quando a temperatura aumenta com a altitude, ao invés de diminuir, que é o que normalmente ocorre na atmosfera. Essa condição pode levar à formação de uma camada de ar estável, onde o vento pode mudar de direção e de velocidade abruptamente em relação às camadas atmosféricas adjacentes. Esse tipo de WS também tem potencial para afetar a aviação de forma semelhante ao WS associado a tempestades, entretanto sua presença é mais difícil de ser percebida visualmente pelos pilotos.

Obstáculos à superfície, como edifícios, montanhas, colinas e árvores, podem gerar WS devido ao efeito que têm na circulação do vento em seu entorno. Quando o vento encontra um obstáculo, ele é desviado e sua velocidade pode diminuir na área de "sombra" do obstáculo, criando um gradiente de velocidade vertical e horizontal.

Isso pode levar à formação de turbulência e cisalhamento de vento na região imediata do obstáculo, onde o fluxo de ar é direcionado após passar pelo obstáculo. O tamanho e a forma do obstáculo, bem como a velocidade e direção do vento, são fatores importantes que influenciam a magnitude e a localização do cisalhamento de vento gerado. Assim, regiões com terrenos montanhosos, vales e desfiladeiros podem gerar cisalhamento de vento em diferentes alturas, especialmente quando o vento é obrigado a passar por essas áreas e é forçado a mudar de direção.

A turbulência em atmosfera convectiva também é fortemente influenciada pelo efeito de ilha de calor urbana e em SBGR pelo influxo de ar quente e úmido oriundo da costa atlântica próxima a SBGR, o que intensifica a atividade convectiva diurna. Nesse sentido, conforme discutido por Oke (1982), as áreas urbanas absorvem e retêm calor, aumentando as temperaturas do ar local e promovendo correntes convectivas.

Atkinson (2003) explica como as estruturas urbanas e o aquecimento da superfície amplificam ainda mais a turbulência, particularmente em áreas densamente construídas. Essa turbulência convectiva é mais intensa durante a tarde, especialmente no verão, com uma persistência diurna que se alinha com o pico de aquecimento solar. Essas condições podem levar a turbulência moderada a severa, apresentando desafios operacionais significativos para a aviação, especialmente quando o aquecimento diurno atinge seu pico.

Por outro lado, o WS em Guarulhos também é moldado por características topográficas como as montanhas da Serra da Cantareira e da Serra do Mar, que criam gradientes de vento acentuados e influenciam os padrões de fluxo de ar (Banta, 2007). Esse WS de baixa altitude, que normalmente persiste durante o início da manhã ou em condições calmas, como discutido mais tarde aqui, é considerado frequentemente mais previsível em uma atmosfera estável.

Além disso, a presença de jatos de baixo nível (JBNs) — correntes de ar em movimento rápido na atmosfera inferior — intensificam ainda mais o cisalhamento, causando mudanças abruptas na velocidade e direção do vento, especialmente à noite ou no início da manhã, quando a atmosfera é mais estável. Essas interações de cisalhamento e jato de baixo nível são particularmente perigosas para aeronaves

durante a decolagem e o pouso, exigindo respostas rápidas do piloto para conter mudanças repentinas na sustentação.

Tanto a turbulência convectiva quanto o cisalhamento do vento de baixa altitude, incluindo a influência de JBNs, exigem respostas distintas dos pilotos e são afetadas pelas características urbanas, geográficas e meteorológicas únicas da área. Essa combinação ressalta a necessidade de estratégias operacionais personalizadas para abordar esses desafios específicos de forma eficaz.

Essas são apenas algumas das condições meteorológicas que podem estar associadas ao cisalhamento de vento. Deste modo, é importante que pilotos, controladores de tráfego aéreo e especialistas em meteorologia estejam cientes dessas condições e monitorem continuamente as mudanças no vento e nas condições meteorológicas para garantir a segurança da aviação.

Assim, torna-se importante estudar as diferentes condições meteorológicas geradoras de cisalhamento de vento, sejam associadas a condições de instabilidade atmosférica quanto em regiões relacionadas a regiões de estabilidade.

2.3. CAMADA LIMITE ATMOSFÉRICA (CLA)

Segundo ICAO (2005), para a detecção e medição do cisalhamento do vento em aeródromos, o ideal é a medição contínua das três componentes do vento até cerca de 500m (1.600 pés) AGL ao longo das trajetórias de aproximação e subida. Portanto, torna-se importante ter algum conhecimento sobre a camada atmosférica mais próxima ao solo, a Camada Limite Atmosférica (CLA).

Stull (1988) define a CLA como sendo constituída pelos poucos quilômetros mais baixos da troposfera e que são diretamente influenciados pela presença da superfície terrestre, respondendo às forçantes da superfície com uma escala de tempo de cerca de uma hora ou menos.

A altura da CLA varia dependendo das condições meteorológicas e das características da superfície terrestre, tais como topografia, tipo de cobertura do solo e taxas de aquecimento e resfriamento da superfície. Assim, sua espessura varia no

decorrer do dia, sendo que em condições atmosféricas de calmaria, há o papel preponderante dos efeitos radiativos de emissão de ondas longas no desenvolvimento de uma camada de inversão (Stull, 1988).

A compreensão da CLA é importante em muitas áreas, não apenas para a meteorologia e climatologia, mas também para a engenharia ambiental, para a construção civil, para a dispersão de poluentes, para estudos da eficiência de turbinas eólicas e muitos outros aspectos da vida humana e do meio ambiente.

Garratt (1994) reforça que a CLA desempenha um papel muito importante em muitos campos, incluindo poluição do ar, agricultura, hidrologia, meteorologia aeronáutica, entre outras áreas. Dentro do campo da meteorologia aeronáutica as principais preocupações são a compreensão dos processos de formação, manutenção e dissipação de nebulosidade baixa e dos jatos de baixo nível, assim como captar e prever o cisalhamento do vento mais intenso que pode afetar a segurança das aeronaves nos procedimentos de pouso e decolagem.

Conforme a figura 8, podemos observar a estrutura e o ciclo diurno da CLA, de acordo com Stull (1988), onde os três principais componentes dessa estrutura são a Camada de Mistura (Camada Limite Convectiva), a Camada Residual e a Camada Limite Estável (Camada Limite Noturna).

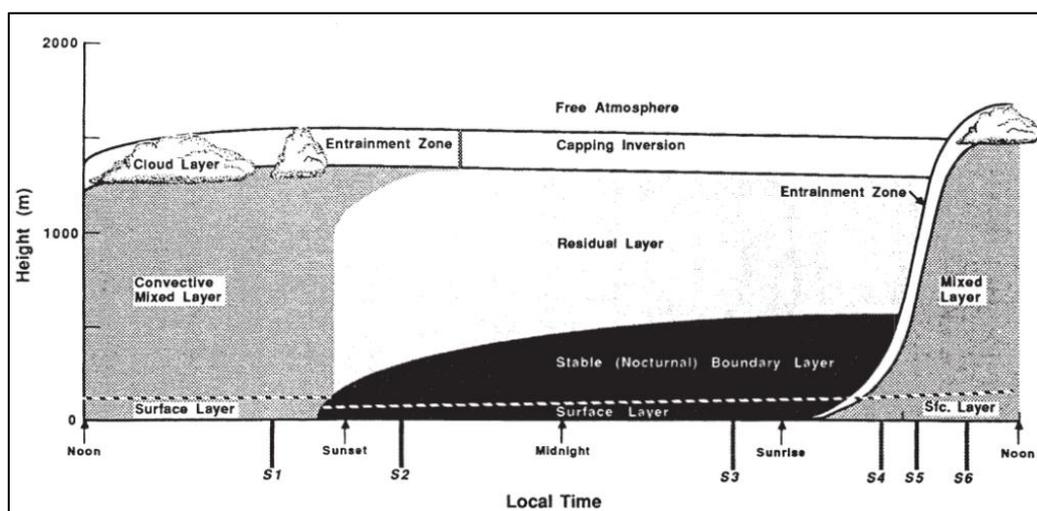


Figura 8 Camada limite atmosférica. Estrutura e ciclo diário

Fonte: STULL, 1988

A Camada de Mistura, se origina no decorrer do dia estando associada aos movimentos verticais do fluxo de calor decorrentes do aquecimento solar da superfície terrestre. Em dias sem nuvens, a Camada de Mistura começa a se aprofundar

verticalmente meia hora após o nascer do sol. Quando em presença de nebulosidade ela é subdividida em uma Camada de Nuvem e uma Camada de Sub-Nuvem.

No decorrer da noite, forma-se a Camada Limite Estável decorrente da diminuição dos fluxos turbulentos associados a convecção térmica diurna decorrentes do resfriamento da superfície terrestre. Essa camada caracteriza-se por ser estaticamente estável e por uma inversão de temperatura próxima ao solo.

A Camada Residual forma-se também no período noturno, após o pôr do sol, acima da Camada Limite Estável, preservando algumas características físicas da Camada de Mistura, tais como a concentração de poluentes. Assim, além da Camada Residual não estabelecer contato direto com o solo, o seu limite inferior vai subindo à medida que a Camada Limite Estável vai aumentando gradualmente de espessura.

Stull (1988) esclarece também que embora os ventos próximos à superfície possuam baixas velocidades, nas proximidades da interseção entre a Camada Limite Estável e a Camada Residual, os ventos normalmente sofrem aceleração significativa devido aos ventos geostróficos e podem dar origem aos Jatos de Baixos Níveis ou Jatos Noturnos.

Além dessas camadas principais, Stull (1988) faz menção à Camada Superficial, que representa em média os primeiros 10% da atmosfera mais próximas do solo, independentemente de fazer parte da Camada de Mistura ou da Camada Limite Estável. Assim, por possuir maior interação com a superfície terrestre, é mais suscetível a formação de turbulência e variações locais de temperatura, umidade e concentração de poluentes. De mesma sorte, o vento na camada superficial é mais influenciado pela rugosidade da superfície do terreno que pode ser alterada pela presença de obstáculos como árvores, construções e pela topografia da região.

A principal correlação entre a CLA e o cisalhamento do vento é que a CLA é a região da atmosfera que fica em contato direto com a superfície da Terra e é influenciada pelas características da superfície de seu entorno, como a rugosidade do terreno, a presença de edifícios e construções, vegetação e outros obstáculos. Assim, essas características da superfície geram variações no escoamento no interior da CLA, que podem ocasionar mudanças bruscas na velocidade e na direção do vento,

especialmente em alturas próximas à superfície gerando regiões sob influência de cisalhamento de vento mais intensos e colocando em risco as operações aéreas.

2.4. EFEITOS DO WS NO DESEMPENHO DAS AERONAVES

Para que uma aeronave possa voar, basicamente, temos que observar a atuação de quatro forças: sustentação, peso, tração e arrasto. (figura 9)



Figura 9 Forças atuantes sobre uma aeronave
Fonte: Adaptado de <https://www.boeing.com/>

Assim, de acordo com Anderson (2010), o voo de um avião é possível graças a interação dessas quatro forças principais. A sustentação que é a componente de força aerodinâmica que é perpendicular à direção da corrente livre, sendo gerada pelas asas do avião, que são projetadas para criar uma diferença de pressão entre a parte superior e inferior, permitindo que o avião suba. O peso como a força da gravidade que puxa o avião para baixo. A tração (ou empuxo, no caso de aviões a jato) sendo produzida pelos motores, que empurram o avião para frente, enquanto a resistência do ar é a força que se opõe ao movimento do avião. O equilíbrio entre essas forças é fundamental para que o avião possa decolar, voar e pousar de maneira segura.

Basicamente, deve-se entender que a sustentação é uma força aerodinâmica mecânica vetorial que se opõe diretamente ao peso e mantém uma aeronave no ar através, principalmente, de suas asas. Entretanto, para que essa força de sustentação seja gerada é necessário que haja a interação e o contato, assim como uma diferença de velocidade, entre a aeronave e o meio fluido que a envolve, ou seja, o ar.

Dessa forma, para que um avião voe é necessário, inicialmente, que a aeronave possua um motor com potência capaz de gerar tração (ou empuxo) e movimentar-se através desse fluido de modo que a aeronave possa vencer a força de arrasto, que é a força de resistência ao seu avanço e atinja a velocidade necessária para que o diferencial de pressão criado nas superfícies de sustentação da aeronave seja superior ao seu peso.

Conforme o “*Pilot Wind Shear Guide*” (FAA, 1988) o encontro com uma área de WS pode ocasionar alguns efeitos sobre as aeronaves e seus equipamentos, tais como: aumento ou diminuição da velocidade indicada, flutuações no ângulo de ataque, turbulência, queda no desempenho, perda da estabilidade e distorções nas indicações altimétricas.

Assim, as indicações incorretas nos diversos equipamentos da aeronave geradas pelo WS tem como principal efeito negativo a possibilidade da perda de sustentação e conseqüente eliminação de controle dos pilotos de manter a aeronave em condições seguras de voo.

Assim, devido a várias interferências que o WS pode ocasionar nas aeronaves em um voo, existem recomendações da FAA para como os pilotos devem lidar em áreas com indicação do fenômeno. Algumas dessas recomendações incluem a preocupação constante na obtenção de informações atualizadas sobre as condições meteorológicas em rota e nos aeródromos, manter atenção nas condições de vento durante a decolagem e aterrissagem, especialmente em áreas conhecidas por apresentar ventos fortes ou instáveis, monitoramento constante do radar meteorológico a bordo da aeronave visando auxílio na detecção de possíveis áreas com WS, e familiarização com manuais e treinamentos específicos, onde os pilotos devem estar familiarizados com as recomendações e procedimentos para lidar com condições de WS.

2.5. SISTEMAS DE DETECÇÃO DE WS

Gultepe et al. (2019) cita que os sistemas e ferramentas de observação são essenciais para o discernimento e monitoramento dos parâmetros meteorológicos necessários para operações aéreas seguras e eficientes. Desta forma, torna-se indispensável a garantia de confiabilidade nos dados obtidos por esses sistemas e

ferramentas. Assim, os métodos de obtenção de dados devem passar constantemente por avaliações, aferições e melhorias, além da combinação e interação entre os mesmos e a inclusão de novas tecnologias que visem aperfeiçoar a identificação e a previsão de fenômenos que ofereçam perigo as atividades aeronáuticas.

Globalmente, a indústria da aviação já fez avanços significativos na detecção e previsão do cisalhamento do vento. Tecnologias como os radares meteorológicos Doppler, SODAR (Sonic Detection and Ranging) e LIDAR (Light Detection and Ranging) foram amplamente adotadas para identificar e monitorar o cisalhamento do vento em aeroportos. Esses sistemas fornecem dados de velocidade e direção do vento de alta frequência, permitindo que os controladores de tráfego aéreo antecipem condições perigosas de vento e possam emitir avisos aos pilotos.

Nos Estados Unidos, a Federal Aviation Administration (FAA) introduziu o sistema Terminal Doppler Weather Radar (TDWR) na década de 1980, que continua sendo um componente essencial da detecção de cisalhamento do vento hoje (FAA, 2008). O TDWR fornece um aviso antecipado de cisalhamento do vento convectivo na presença de nuvens ou chuva e sob condições de *microburst*, permitindo que os pilotos possam ajustar suas trajetórias de voo adequadamente.

No Japão, o Aeroporto Internacional de Kansai implementou com sucesso sistemas LIDAR para monitorar o cisalhamento do vento. O LIDAR usa feixes de laser para detectar partículas atmosféricas, fornecendo perfis de vento detalhados mesmo em ar limpo, onde os sistemas de radar tradicionais podem ter dificuldades. Estudos mostraram que o LIDAR é particularmente eficaz na detecção de cisalhamento do vento em áreas complexas, como a área da baía perto de Kansai, onde as interações entre brisas marítimas e ventos terrestres frequentemente geram condições de vento perigosas (Shun e Chan, 2008).

Na Europa, por exemplo, o Aeroporto de Heathrow no Reino Unido adotou uma combinação de tecnologias LIDAR e SODAR para monitorar as condições do vento durante fases críticas do voo, como decolagem e pouso. Esses sistemas reduziram significativamente o número de incidentes relacionados ao cisalhamento do vento em um dos aeroportos mais movimentados do mundo (Crosier et al., 2002).

Pode-se dizer que existem, atualmente, vários tipos de equipamentos que podem ser usados para estimar a direção e velocidade do vento na baixa troposfera e indicar a presença do WS, como LLWAS, TDWR, LIDAR, Doppler Radar e SODAR.

Entretanto, conforme ICAO (2005), alguns desses equipamentos são sofisticados e caros em termos de aquisição e manutenção, apontando que a maioria dos aeroportos estariam restritos a utilização de métodos de observação indireto para a detecção da presença e da indicação da intensidade do WS.

Em relação a essas ferramentas, a ICAO (2005) relata que o sistema SODAR se utiliza de ondas sonoras para estimar a velocidade e direção do vento em diferentes níveis da baixa atmosfera. Nesse sentido, Little (1969) destaca que as ondas sonoras interagem mais fortemente com a atmosfera inferior do que as ondas eletromagnéticas. Portanto, a utilização do SODAR é mais sensível do que o RADAR para observar a estrutura fina e a turbulência na camada limite atmosférica.

2.5.1. SODAR

O SODAR é um dispositivo cujo princípio de funcionamento (figura 10) é baseado no conceito de eco, semelhante ao radar, porém utilizando ondas sonoras para medir as características da atmosfera, como perfil vertical de velocidade e direção do vento sendo útil em estudos para aplicação em várias áreas, tais como a identificação de potencial eólico de uma região e no apoio em atividades aeronáuticas. Algumas vantagens de sua utilização são seu baixo custo, a medição em 3D, não ser invasivo (não interfere em outros equipamentos de comunicação e navegação) e eficaz em condições de ar seco.

Silva et al (2016) acrescentam a vantagem que o SODAR é um equipamento de baixo custo e fácil instalação que utiliza a reflexão de pulsos acústicos nas inhomogeneidades atmosféricas para estimar a direção e a velocidade do vento em três dimensões (3D), em intervalos de tempo predeterminados.

Nos estudos de Silva et al. (2016) e França et al. (2018) demonstrou-se a eficácia dos sistemas Doppler SODAR na detecção de WS de baixa altitude no SBGR. Dados do SODAR em tempo real permitem observações de um perfil de vento preciso, auxiliando os controladores de tráfego aéreo a prever mudanças bruscas de vento.

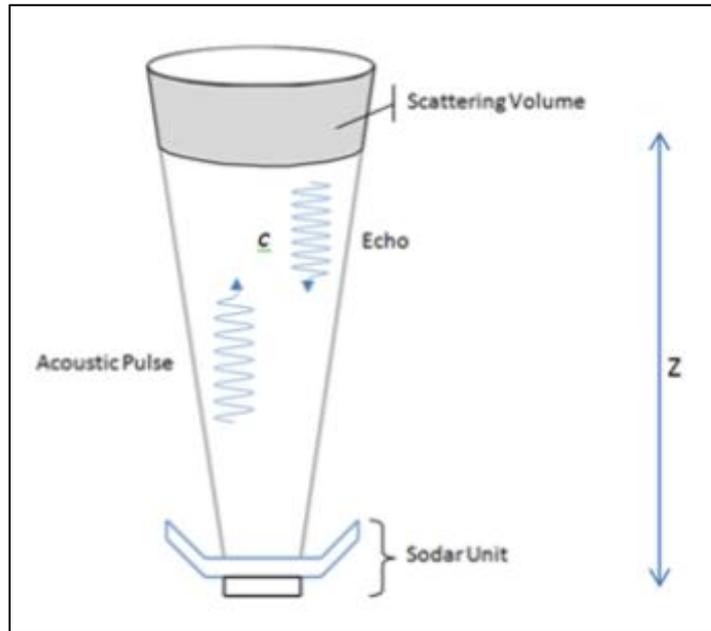


Figura 10 Funcionamento do SODAR

Fonte: <https://isaacbrana.wordpress.com/2012/02/18/sodar-operation-limitations-12/>

França et al. (2018) desenvolveram ainda mais um modelo de *nowcasting* usando dados do SODAR e redes neurais, alcançando alta precisão na previsão de perfis de vento verticais em alturas de 10, 100, 200 e 300 metros no SBGR. Esses avanços ressaltam a importância de integrar o aprendizado de máquina com sensores meteorológicos avançados para aprimorar a detecção e a previsão de WS.

Assim, dando continuidade a outros estudos e pesquisas realizados pela equipe do Laboratório de Meteorologia Aplicada da Universidade Federal do Rio de Janeiro (LMA – UFRJ), utilizando os dados obtidos pelo SODAR, que no Brasil apenas os Aeroportos de Guarulhos/SP (SBGR) e Santos Dumont/RJ (SBRJ) contam com esse tipo de equipamento instalado, essa pesquisa busca integrar os dados obtidos pelo SODAR a técnicas de aprendizado de máquina para desenvolver método de previsão de eventos de WS e auxiliar na manutenção da segurança do tráfego aéreo no aeroporto de maior movimento do país.

3. PROCESSAMENTO E ANÁLISE DE DADOS

3.1. APRENDIZADO DE MÁQUINA

No decorrer das últimas décadas, várias empresas e diversas organizações começaram a gerar uma grande quantidade de dados em suas operações diárias, como dados de vendas, dados de clientes, dados de produção, dados financeiros, entre outros. Além disso, os avanços na tecnologia de sensores e dispositivos permitiram que uma grande quantidade de dados fosse coletada a partir de sensores em tempo real em diferentes tipos de ambientes, como em residências, em fábricas e na natureza.

Assim, com a crescente disponibilidade de dados, surgiu um novo campo de pesquisa conhecido como ciência de dados, que impulsionada, principalmente, pelo avanço da tecnologia da informação e dos sistemas de comunicação permitiu a coleta, o armazenamento e o processamento de grandes quantidades de dados de forma mais eficiente e econômica.

Uma das formas de tratamento e processamento dessa enorme quantidade de dados foi concebida pelo desenvolvimento de técnicas de Inteligência Artificial (IA), tais como o aprendizado de máquina. Nesse sentido, Murphy (2012) definiu o aprendizado de máquina como um conjunto de métodos que podem detectar padrões automaticamente nos dados e, em seguida, usar esses padrões descobertos para a previsão de dados futuros ou para realizar outros tipos de tomada de decisão sob incerteza, levando-se em consideração métodos estatísticos aliados aos fundamentos da teoria da probabilidade.

Chollet (2017), aponta que o aprendizado de máquina é uma área de pesquisa que surgiu da pergunta se um computador poderia aprender sozinho como executar uma tarefa, em vez de ser programado manualmente por humanos. Ou seja, em vez de criar-se regras de processamento de dados manualmente, o aprendizado de máquina permite que o computador aprenda automaticamente essas regras observando os dados de entrada. Isso possibilitou um novo paradigma na ciência de programação (figura 11) em que os humanos fornecem os dados e as respostas esperadas, e as regras são geradas de forma automatizada pelo computador. Assim,

essas regras poderiam ser aplicadas a novos dados para produzir as respostas originais.

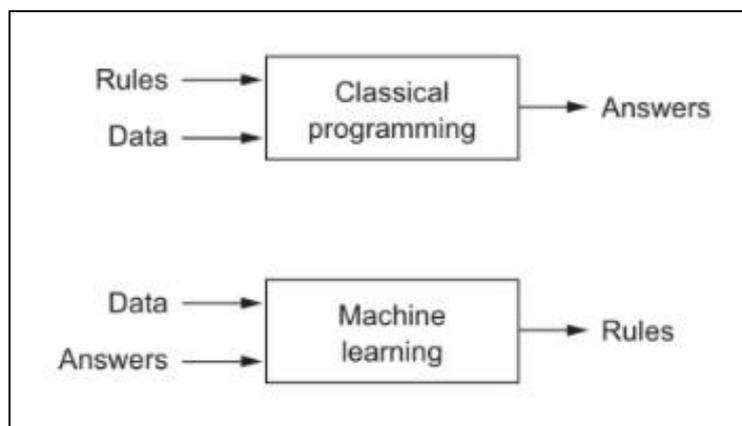


Figura 11 Paradigma da Programação
Fonte: Chollet, Francois. 2017. Deep Learning with Python

Da mesma forma, Allende-Cid (2019) corrobora que o aprendizado de máquina é uma subárea da Inteligência Artificial e que tem apresentado um crescimento enorme nas últimas décadas, tratando a forma como algoritmos matemáticos, estatísticos e computacionais são capazes de realizar um processo de inferência por meio de aprendizado baseado em exemplos.

De forma geral, o processo de aprendizado de máquina geralmente envolve três etapas principais: a coleta de dados; o treinamento do modelo, onde um algoritmo de aprendizado de máquina é alimentado com esses dados para que possa aprender com eles e seja ajustado para encontrar padrões e relações nos dados; e o teste do modelo para avaliar sua precisão e eficácia na resolução dos problemas.

Nesse sentido, o aprendizado de máquina, como uma técnica de Inteligência Artificial, pode ser aplicado em muitas áreas, tais como medicina, finanças, marketing, educação, logística, segurança e no aperfeiçoamento da previsão meteorológica.

Murphy (2012) nos diz que o aprendizado de máquina pode ser dividido em dois tipos principais. O primeiro é o aprendizado preditivo ou supervisionado, que é a forma de aprendizado de máquina mais utilizada na prática, onde o objetivo é aprender um mapeamento das entradas para as saídas, com base em um conjunto de treinamento rotulado. Nesse tipo de aprendizado, as entradas são representadas por vetores de atributos, como altura e peso de uma pessoa, e as saídas podem ser categóricas (classificação) ou valores reais (regressão).

O segundo tipo é o aprendizado descritivo ou não supervisionado, onde recebemos apenas as entradas e o objetivo é encontrar padrões de interesse nos dados. Nesse tipo de aprendizado, não há rótulos ou métricas de erro predefinidas, e o foco está na descoberta de conhecimento.

Entretanto, além desses dois tipos principais, há ainda o aprendizado por reforço, que é menos comumente utilizado. Nesse tipo de aprendizado, a máquina aprende a tomar decisões e a agir de maneira a maximizar recompensas ou minimizar punições com base nos sinais de reforço recebidos do ambiente.

Assim, uma das principais vantagens do uso de aprendizado de máquina na previsão meteorológica é a sua capacidade de lidar com grandes volumes de dados e encontrar padrões que seriam difíceis ou impossíveis de detectar por meio de métodos tradicionais. Além disso, o aprendizado de máquina pode ser usado para melhorar a precisão da previsão, fornecendo previsões mais confiáveis e com maior antecedência.

No entanto, a aplicação do aprendizado de máquina na previsão meteorológica apresenta desafios, pois exige um grande volume de dados de alta qualidade e a seleção adequada dos algoritmos de aprendizado de máquina. Além disso, a previsão meteorológica é influenciada por muitos fatores, como mudanças climáticas, mudanças sazonais e variações geográficas, o que torna difícil a criação de um modelo de aprendizado de máquina único que possa prever com precisão as condições meteorológicas em todas as regiões e em todos os momentos.

Entretanto, apesar desses desafios, o uso do aprendizado de máquina na previsão meteorológica tem se mostrado promissor e tem sido cada vez mais utilizado nos últimos anos. À medida que os modelos de aprendizado de máquina se tornam mais sofisticados e o volume de dados disponíveis aumenta, espera-se que a precisão das previsões meteorológicas também continue a evoluir.

Nesse sentido, também já existem pesquisas científicas que aplicam técnicas de aprendizado de máquina para prever o comportamento de fenômenos associados ao parâmetro vento, tais como cisalhamento de vento, turbulência e padrão de campo de vento. Esses estudos e pesquisas mostram que as técnicas de aprendizado de máquina podem ser uma ferramenta útil para prever o cisalhamento de vento e

melhorar a confiabilidade das previsões meteorológicas desses fenômenos que impactam a segurança das aeronaves.

Williams (2014) realizou uma pesquisa tratando da utilização do algoritmo de aprendizado de máquina *Random Forest* para diagnosticar a turbulência na aviação. O autor coletou dados de diversas fontes, incluindo sensores em aeronaves e dados meteorológicos, para desenvolver um conjunto de dados de treinamento para a previsão de turbulência. Em seguida, ele utilizou esse conjunto de dados para treinar um modelo de *Random Forest* que pudesse identificar a presença ou a ausência de turbulência.

Gao et al. (2022) abordaram a aplicação de técnicas de aprendizado profundo (*deep learning*) para a previsão do campo de vento em tempo real. O artigo apresentou um novo método de previsão de curto prazo do campo de vento, que utiliza uma rede neural convolucional profunda (*Deep Convolutional Neural Network* - DCNN) com uma camada de atenção especializada para eventos altamente variáveis.

Os resultados dessa pesquisa mostraram que o método proposto pode prever de forma eficiente e precisa as mudanças no campo de vento em tempo real, mesmo durante eventos altamente variáveis, como tempestades. O trabalho destacou a importância da previsão precisa do campo de vento para aplicações em áreas como aviação, geração de energia eólica e prevenção de desastres naturais.

Assim, a pesquisa de Gao et al. (2022) colaborou para o avanço da previsão do tempo utilizando técnicas de aprendizado profundo e destacando a importância da precisão na previsão do campo de vento para diversas aplicações práticas.

3.2. WEKA

Nesse contexto, foi utilizado neste trabalho o software WEKA – versão 3.8.6 (Waikato Environment for Knowledge Analysis - <https://ml.cms.waikato.ac.nz/weka/>) que é um software de código aberto projetado para mineração de dados e aprendizado de máquina, desenvolvido pela Universidade de Waikato, na Nova Zelândia, sendo amplamente utilizado tanto em ambientes acadêmicos quanto em ambientes industriais, devido à sua interface gráfica intuitiva e ao suporte a uma ampla gama de algoritmos de aprendizado de máquina.

Além disso, o WEKA é uma ferramenta versátil e compatível com vários tipos e formatos de dados, podendo ser executado em qualquer sistema operacional que suporte Java. A plataforma oferece ainda diversas funcionalidades para o processamento de dados, incluindo as fases de pré-processamento, de classificação, de seleção de atributos, aplicação dos algoritmos e a própria avaliação dos modelos (figura 12). Assim, o WEKA torna-se uma ferramenta popular de aprendizado de máquina e mineração de dados que fornece um ambiente integrado para desenvolvimento e aplicação de algoritmos de previsão. Esses algoritmos de previsão, nomeados como "*classifiers*" no WEKA, são organizados em várias categorias, onde as principais delas são: *Bayes*, *Functions*, *Lazy*, *Rules* e *Trees*.

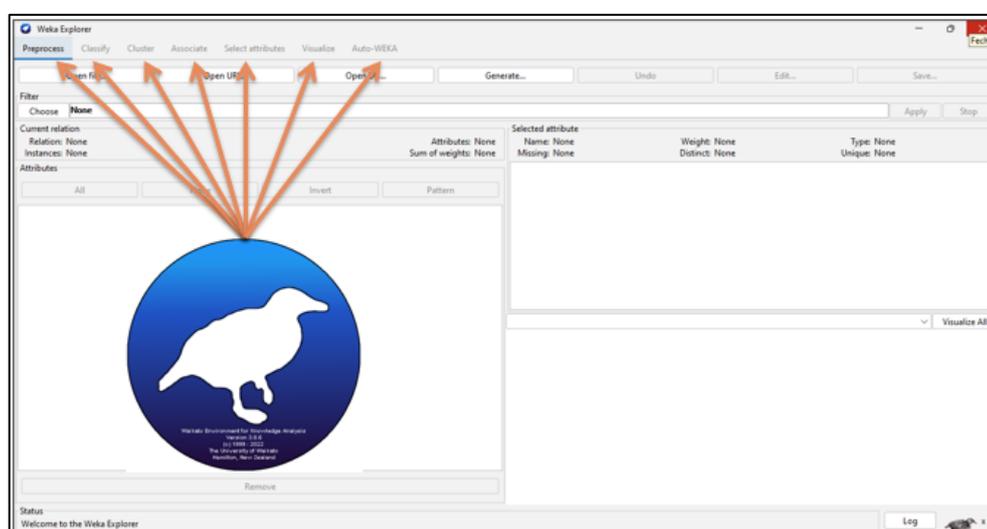


Figura 12 Tela da plataforma WEKA

Fonte: Adaptado pelo autor

O WEKA, por possuir uma ampla e extensa biblioteca de algoritmos de aprendizado de máquina pré-implementados, oferece também uma variedade de opções para explorar e experimentar diferentes abordagens, permitindo comparar e avaliar o desempenho dos seus diferentes algoritmos em termos de precisão, sensibilidade, F-Measure e outras métricas de avaliação. Na tabela 2 são apresentadas as principais características dos algoritmos utilizados na pesquisa.

Assim, o WEKA além de apreciar o desempenho de seus modelos através de diversas métricas de avaliação, oferece ainda uma visualização rápida da matriz de confusão, e viabiliza o uso de técnicas utilizando-se tanto da divisão dos dados em conjuntos de treino e teste, como utilizando-se de técnicas de validação cruzada.

A validação cruzada (“*cross-validation*”) trata-se de uma técnica amplamente utilizada em estatística e machine learning para avaliar o desempenho de modelos preditivos. O objetivo principal é garantir que o modelo possa generalizar bem para dados não vistos, evitando problemas como o sobreajuste (“*overfitting*”). Assim, o método envolve a divisão do conjunto de dados em vários subconjuntos, amostras ou “*folds*”.

Em cada iteração, um desses subconjuntos é utilizado para teste, enquanto os outros são usados para treinar o modelo. Esse processo é repetido diversas vezes, permitindo que cada subconjunto seja utilizado como conjunto de teste uma vez. A performance do modelo é então avaliada com base na média das acurácias obtidas em todas as iterações (Kohavi, 1995).

Algoritmos	Características
RandomForest	Método de aprendizado de conjunto; combina várias árvores de decisão; reduz overfitting; alta precisão e robustez em conjuntos de dados complexos (Breiman, 2001).
IBk	Aprendizado baseado em instância (k-vizinhos mais próximos com k=5); armazena instâncias; faz previsões com base na similaridade; sensível à escala de recursos (Aha, 1991).
NBTree	Híbrido de Naive Bayes e árvores de decisão; usa probabilidade condicional; adequado para dados categóricos e numéricos; saída interpretável (Kohavi, 1996).
J48	Implementação da árvore de decisão C4.5; divide dados com base no ganho de informações; poda árvores para evitar overfitting; interpretável (Quinlan, 1993).
LMT	Combina regressão logística e árvores de decisão; usa regressão linear em folhas; bom equilíbrio entre interpretabilidade e poder preditivo (Landwehr, 2005).
Multilayer Perceptron	Modelo de rede neural; várias camadas com ativação não linear; adequado para padrões complexos; requer tempo de treinamento e ajuste significativos (Rumelhart et al., 1986).
RandomTree	Árvore de decisão sem poda; seleção aleatória de recursos; rápido, mas propenso a overfitting; usado principalmente para comparação (Breiman, 2001).
REPTree	Árvore de poda de erro reduzido; árvore de decisão rápida; poda de erro para simplicidade; pode lidar com dados contínuos e categóricos (Witten & Frank, 2005).
DecisionTable	Tabela de decisão simples (modelo baseado em regras); fácil de interpretar; realiza seleção de atributos; limitado a padrões simples (Witten & Frank, 2005).
PART	Árvores de decisão parciais; híbrido de árvore de decisão e aprendiz baseado em regras; gera regras com árvores parciais; interpretável (Frank & Witten, 1998).
OneR	Algoritmo de uma regra; gera uma regra por recurso; altamente interpretável, mas simplista; útil como um modelo de linha de base (Holte, 1993).
Logistic	Modelo de regressão linear para resultados binários; útil para classificação binária; coeficientes interpretáveis; manipula dados categóricos com codificação (Cox, 1958).
HoeffdingTree	Árvore de decisão de aprendizagem on-line; eficiente para streaming de dados; usa limite de Hoeffding para decisões; atualização incremental do modelo (Domingos & Hulten, 2000).
SimpleLogistic	Classificador linear; combina regressão logística e seleção de recursos; interpretável; adequado para problemas binários e multiclasse (Landwehr, 2005).
BayesNet	Rede Bayesiana; modelo gráfico probabilístico; aprende dependências condicionais; interpretável, mas complexo; manipula dados ausentes (Pearl, 1988).
NaiveBayes	Classificador probabilístico baseado no teorema de Bayes; você assume independência de recursos; eficiente com dados pequenos; bom para classificação de texto (Duda & Hart, 1973).

Tabela 2 Características dos Algoritmos

4. DADOS E MÉTODO

Neste Capítulo, serão apresentados e descritos os dados utilizados na análise de ocorrência dos eventos de WS e o método proposto para estabelecer um modelo de previsão desse fenômeno meteorológico no SBGR.

4.1. ÁREA DE ESTUDO

A área de estudo é representada pelo Aeroporto Internacional de São Paulo/Guarulhos - Governador André Franco Montoro, sob as coordenadas geográficas 23° 26' 08" S e 046° 28' 23" W e a uma altitude de 750 metros (2461 pés). Sob o indicador de localidade da ICAO: SBGR, o aeroporto é o mais movimentado do Brasil, conforme o Anuário Estatístico de Tráfego Aéreo 2023 do DECEA.

A topografia do município de Guarulhos (figura 13) é predominantemente plana, constituída pela bacia do Rio Baquirivu-Guaçu, afluente do Rio Tietê e limitando-se ao norte-noroeste pela Serra da Cantareira, com densa vegetação de Mata Atlântica e ao sul por morrotes de relevo suave com predomínio de vegetação rasteira e pastos de atividades rurais. (Estudo de Impacto Ambiental – EIA; 2004)

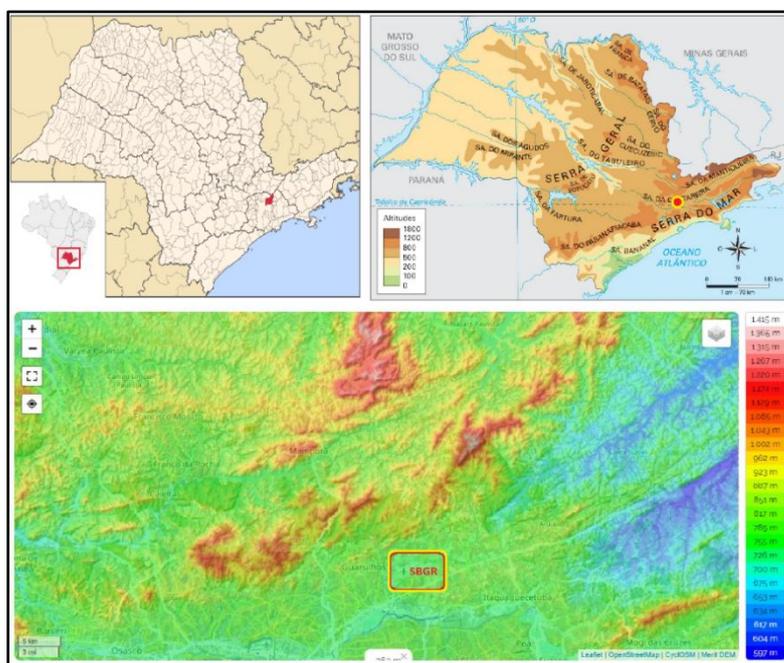


Figura 13 Região do Aeroporto Internacional de Guarulhos/SP
Fonte: Autor, adaptado de Ministério do Meio Ambiente (2002)

O Aeroporto Internacional de Guarulhos possui duas pistas paralelas: 10 (280° magnético) e 28 (10° magnético), localizadas a uma altitude de 750 metros. A pista 10R/28L mede 3000 metros, enquanto a 10L/28R se estende por 3700 metros, ambas com uma largura de 45 metros. A Figura 14 fornece uma visão geral do aeroporto e da localização do SODAR, marcada pelo círculo vermelho.



Figura 14 Imagem de satélite do SBGR, com a localização do SODAR
Fonte: Adaptado de DASA

O Gráfico de Velocidade e Direção do Vento (figura 15), com dados do período compreendido entre 2001 e 2020, obtido através do Banco de Dados Climatológicos do Instituto de Controle do Espaço Aéreo (ICEA), através de consulta ao Sistema de Geração e Disponibilização de Informações Climatológicas, aponta como característica na região do SBGR que o vento predominante à superfície possui direção de 070° (leste/nordeste) com 05 nós de intensidade.

Neste gráfico, percebe-se ainda que o intervalo na direção predominante do vento à superfície fica entre 60° e 70° (leste/nordeste), quando a intensidade varia de 1 a 10 nós. Entretanto, quando essa intensidade varia de 11 a 20 nós, ou seja, com a presença de ventos mais intensos, sua direção predominante passa a predominar entre 290° a 320° (noroeste).

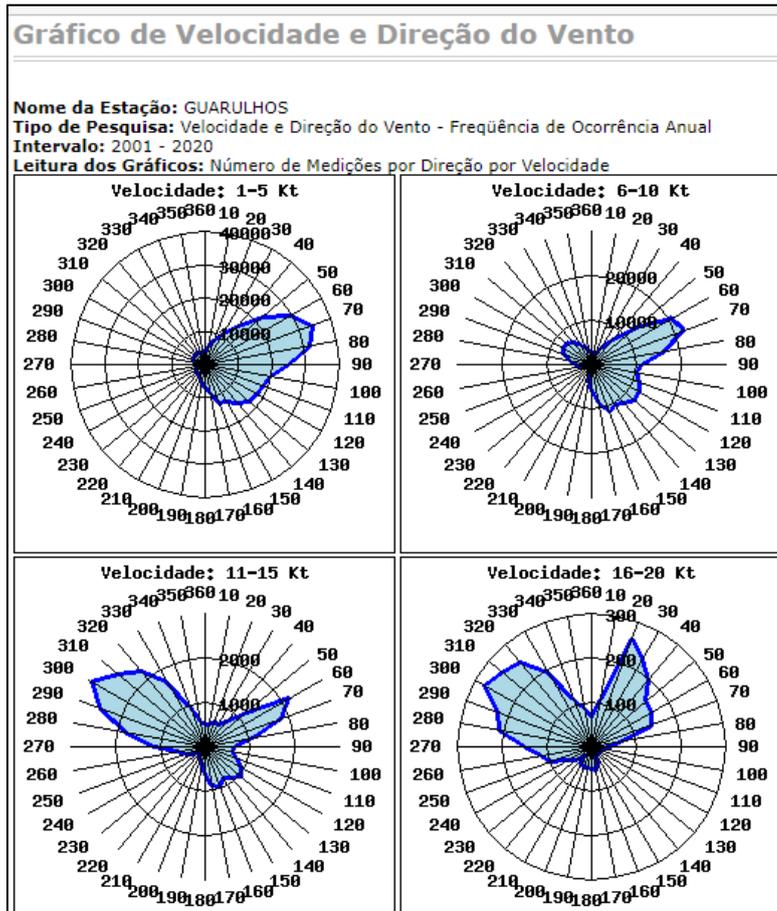


Figura 15 Gráfico de Direção e Velocidade do vento no SBGR
 Fonte: <http://clima.icea.decea.mil.br/clima/>

Costa (2008) verificou em seu trabalho de pesquisa que os ventos de superfície predominantes no SBGR vem de leste (30,9%), entretanto velocidades de vento mais intensas, superiores a 10KT, ocorrem quando o vento sopra de noroeste, ou seja, de 300° a 330°. De acordo com a pesquisa, esses ventos mais intensos de noroeste, que sopram perpendicularmente à Serra da Cantareira, favorecem a ocorrência do fenômeno WS de origem orográfica em Guarulhos devido ao posicionamento das pistas do aeródromo em relação a essa formação orográfica.

Nesse sentido, Matschinske (2013), também corrobora que a principal causa da ocorrência do fenômeno de cisalhamento de vento em Guarulhos é a onda de montanha, que se forma sobre o alinhamento de picos existentes ao norte/noroeste do aeroporto, na região próxima a serra da Cantareira.

4.2. DADOS

Os dados utilizados para definir as características dos preditores (entradas) e alvos (saídas) foram coletados pelo SODAR, instalado nas proximidades das cabeceiras 28R e 28L do SBGR. A versão do SODAR utilizada é capaz de coletar o perfil dos componentes vetoriais $u(z_i, t_i)$, $v(z_i, t_i)$ e $w(z_i, t_i)$, onde Z_i representa o nível de altura inicial de 30 m (z_{30} representa a medição do perfil do vento inicial) até 520m (Z_{final} é variável e depende da instabilidade atmosférica) em um dado tempo t_i .

Dessa forma, é possível determinar para cada nível atmosférico a velocidade e direção do vento $Vel(z_i, t_i)$ que representa a variação da velocidade do vento em relação à altura, ou seja,

$$ws(topo - base, t_i) = \frac{\Delta vel(topo-base, t_i)}{\Delta alt(topo-base)}, \quad (1)$$

que pode ser classificado como: leve, $WS \leq 4,0$ KT/30 m; moderado, $5,0 \leq WS \leq 8,0$ KT/30 m; forte, $9,0 \leq WS \leq 12,0$ KT/30 m; e severo, $WS > 12,0$ KT/30 m.

A saída a ser prevista para os próximos 15 minutos foi definida como $WS(30-200, t_i+15)$, representando o WS na camada atmosférica entre 30 e 200 metros acima da superfície do solo. Essa faixa de altura é considerada crítica para aproximação e pouso e serve como limite para os pilotos decidirem se devem ou não realizar uma manobra de arremetida.

Ressalta-se que o tempo de previsão de 15 minutos foi selecionado pois é o intervalo de tempo entre as coletas de dados do SODAR e considerando-se que uma aeronave leva, em média, de 3 a 5 minutos entre a sua aproximação final e o pouso.

As oito variáveis preditoras (entradas) foram selecionadas de um conjunto de vinte e sete variáveis derivadas de WS, SPD e DIR geradas através da interpolação de dados coletados em diversas alturas pelo SODAR (30–200 m). Essa seleção foi feita por tentativa e erro, com base no desempenho dos algoritmos de aprendizado de máquina na previsão de WS (30–200, $t+15$), onde buscou-se manter variáveis interdependentes de WS, SPD e DIR. As características preditoras (entradas) selecionadas para a pesquisa foram:

- $WS(30-40, t-60)$: Soma de $WS(30) + WS(40)$ nos últimos 60 minutos,
- $WS(30-40, t-15)$: Soma de $WS(30) + WS(40)$ nos últimos 15 minutos,

- WSMAX1(110-200, t-60): Maior WS entre 110 m e 200 m nos últimos 60 minutos,
- WS3MAX(110-200, t-60): Soma dos três maiores WS entre 110 m e 200 m nos últimos 60 minutos,
- SPDMAX1(110-200, t-60): Maior velocidade do vento entre 110 m e 200 m nos últimos 60 minutos,
- SPD3MAX(110-200, t-60): Soma das três maiores velocidades do vento entre 110 m e 200 m nos últimos 60 minutos,
- SPDMAX60(110-200, t-60): Maior velocidade do vento entre 110 m e 200 m com uma diferença de 30 metros a 60 metros entre os níveis nos últimos 60 minutos, e
- DIRMAXDIF(30-60, t-60): Maior diferença na direção do vento entre 110 m e 200 m com uma diferença de 30 metros a 60 metros entre os níveis nos últimos 60 minutos.

A Tabela 3 fornece um resumo detalhado das entradas e saídas definidas.

Além disso, para examinar e compreender os sistemas meteorológicos ativos na região SBGR durante os períodos de ocorrência de WS detectados pelo SODAR, dados auxiliares, conforme mostrado na Tabela 3, foram utilizados. Isso inclui mensagens do tipo METAR, que fornecem as condições de tempo observadas atuais padronizadas para aeroportos (como direção e velocidade do vento, visibilidade horizontal e fenômenos meteorológicos significativos). Essas mensagens foram combinadas com cartas sinóticas de superfície, oferecendo uma visão geral detalhada dos padrões atmosféricos regionais e sistemas de pressão, permitindo uma análise sinótica mais ampla. Imagens dos satélites meteorológicos GOES-13 e GOES-16 também foram incorporadas para avaliar sistemas meteorológicos transitórios que afetam a região. Essas fontes auxiliares podem ajudar a identificar se a possível origem do WS possui características convectiva ou não convectiva, fornecendo uma visão mais abrangente das condições atmosféricas predominantes nos momentos de ocorrência do WS.

Data	Freq.	Entradas	Saídas	Fonte
SODAR	15 min	1. WS(30-40, t-60): ws(30)+ws(40) nos últimos 60 minutos.	1. WS forte a severo detectado 15 minutos à frente na camada atmosférica de 30-200 m durante todo o período de dados. 2. WS forte a severo detectado 15 minutos à frente na camada atmosférica de 30-200 m durante o período de junho a agosto. 3. WS forte a severo detectado 15 minutos à frente na camada atmosférica de 30-200 m com condições atmosféricas estáveis durante o período de junho a agosto.	Banco de dados da Universidade Federal do Rio de Janeiro (UFRJ)
		2. WS(30-40, t-15): ws(30)+ws(40) nos últimos 15 minutos.		
		3. WSMAX1(110-200, t-60): maior WS entre 110m e 200m nos últimos 60 minutos.		
		4. WS3MAX(110-200, t-60): soma dos três maiores WS entre 110m e 200m nos últimos 60 minutos.		
		5. SPDMAX1(110-200, t-60): maior SPD entre 110m e 200m nos últimos 60 minutos.		
		6. SPD3MAX(110-200, t-60): soma dos três maiores SPD entre 110m e 200m nos últimos 60 minutos		
		7. SPDMAX60(110-200, t-60): maior velocidade do vento entre 110 m e 200 m com diferença de 30 metros a 60 metros entre os níveis nos últimos 60 minutos.		
		8. DIRMAXDIF(30-60, t-60): maior diferença na direção do vento entre 110 m e 200 m com uma diferença de 30 metros a 60 metros entre os níveis nos últimos 60 minutos.		
Data	Freq.	Descrição (Informações auxiliares)	Fontes	
METAR	1 h	Mensagens que fornecem as condições meteorológicas observadas atuais nos aeroportos	www.redemet.aer.mil.br/	
Satélite (GOES-13-16)	15 min	Informações sobre as condições atmosféricas dos aeroportos	satellite.cptec.inpe.br/home/	
Carta Sinótica de Superfície	24 h	Fornece uma visão geral das condições atmosféricas em uma região	satellite.cptec.inpe.br/home/	

Tabela 3 Dados usados para o período de 1º de janeiro de 2016 a 31 de dezembro de 2019. Características de recursos (preditores de entrada) e saída (alvos) foram utilizadas para treinar e testar algoritmos de aprendizado de máquina para previsões de WS de 15 minutos. O WS significa que são variáveis derivadas de dados de wind shear, o SPD às variáveis derivadas de dados de velocidade do vento e o DIR às variáveis derivadas da direção do vento.

Fonte: O autor

Cabe ressaltar que o software APRun 1.51 (Scintec AG) é a ferramenta que permite a operação do SODAR MFAS, e possibilita a geração de diversos tipos de

arquivos de dados, e que os dados de saída gerados são armazenados em arquivos de texto com a extensão MND. Na figura 16, é possível observar os dados registrados de WS, SPD e DIR obtidos pelo SODAR em sua relação tempo (a cada 15 minutos) versus altura (intervalos de 10 metros, com base inferior a partir de 30 metros).

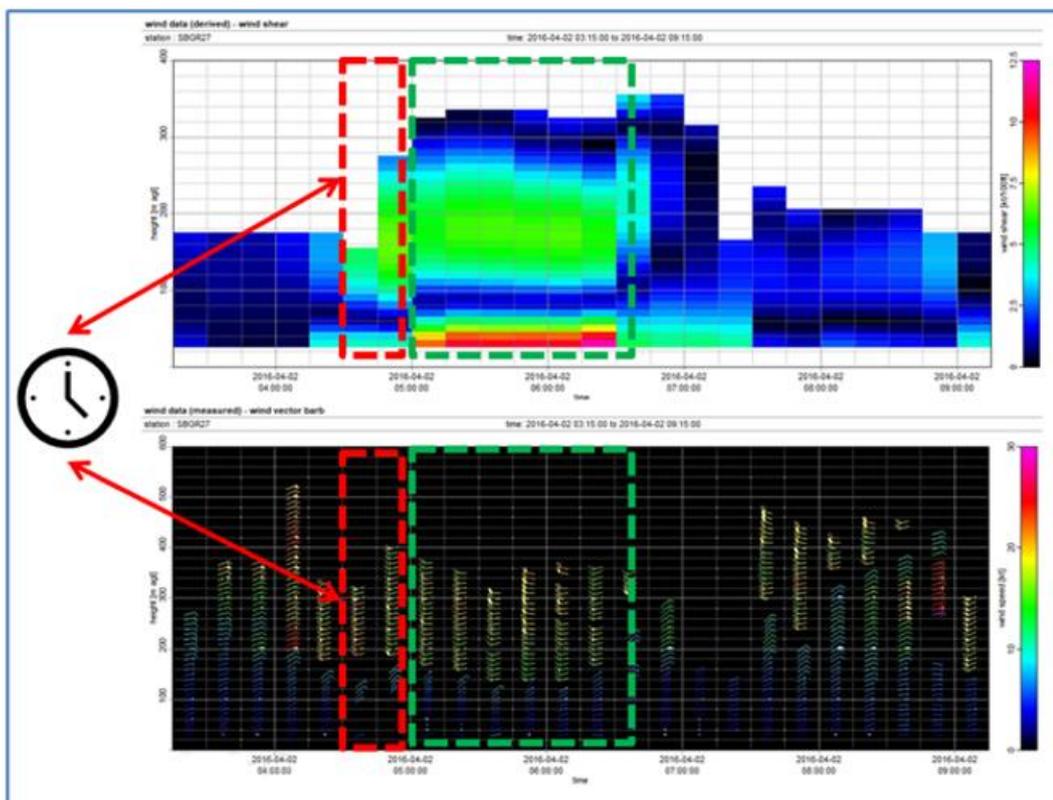


Figura 16 Composição do APRun 1.51, apresentando na tela superior os dados de WS obtidos pelo SODAR na relação tempo (a cada 15 minutos) versus altura relativa do solo (a cada 10 metros) e na tela inferior os dados das barbelas com dados de SPD e DIR obtidos pelo SODAR na relação tempo versus altura.

Fonte: Adaptado pelo Autor de APRun 1.51

Destaca-se ainda que para a pesquisa, foi fundamental garantir a padronização dos dados, especialmente em relação ao registro de tempo para evitar inconsistências e facilitar a comparação entre diferentes fontes. Neste contexto, todos os dados utilizados nesse estudo foram em horário Zulu (também conhecido como Tempo Universal Coordenado - UTC, ou Greenwich Mean Time - GMT). O horário Zulu é amplamente utilizado em estudos meteorológicos, aviação e outras áreas científicas por ser uma referência global, evitando assim confusões decorrentes de fusos horários locais ou mudanças sazonais, como por exemplo, o horário de verão.

4.3. MÉTODO

Este estudo teve como objetivo desenvolver um método baseado em aprendizado de máquina para prever a ocorrência do WS com 15 minutos de antecedência, entre 30 e 200 metros no SBGR. Para atingir esse objetivo, as seguintes etapas, descritas na Figura 18, foram projetadas:

- I. **Processamento de dados:** Coletar e processar os dados de 15 minutos do SODAR, verificando o período de dados disponível para evitar perda de continuidade e definindo o intervalo temporal para a pesquisa. Além disso, períodos com dados faltantes, inconsistentes ou imprecisos foram suprimidos para evitar comprometer a eficiência e a precisão das análises.

Assim, foi necessária uma análise da distribuição das ocorrências de WS indicadas pelo SODAR. Verificou-se a disponibilidade e a continuidade dos dados e realizou-se a remoção dos dados espúrios, ou seja, os períodos com dados faltantes, inconsistentes ou imprecisos, que poderiam comprometer a qualidade dos resultados da pesquisa. Essa etapa de limpeza inicial dos dados assegurou que apenas informações válidas pudessem ser utilizadas nas análises, aumentando a confiabilidade das conclusões.

- II. **Entradas e Saídas:** As entradas foram definidas conforme discutido anteriormente. Um conjunto de dados foi criado vinculando os preditores de entrada para t_i (8 preditores, conforme descrito na Tabela 3) a cada saída correspondente para $t_i+15\text{min}$ (saída, conforme detalhado na Tabela 3, coluna 4).

Para superar a limitação, devido à falta de sensores para medir os parâmetros de estabilidade da camada limite, como a razão de Bowen, número de Richardson ou comprimento de Obukhov, os registros de ocorrência de WS originados por processos convectivos foram distinguidos daqueles que ocorrem em condições atmosféricas estáveis por meio das seguintes análises:

1. Mensagens METAR: condições meteorológicas observacionais fornecidas para ajudar a diferenciar situações convectivas (por exemplo, tempestades) de condições estáveis (por exemplo, neblina e névoa)
2. Cartas sinóticas de superfície: limites frontais destacados para sugerir potencial convectivo nas proximidades do SBGR.
3. Imagens de satélite: usadas para identificar os tipos de nebulosidade, como nuvens cumuliformes indicando convecção e nuvens estratiformes sugerindo estabilidade.



Figura 17 Mensagens de Observação à Superfície do tipo METAR, Carta Sinótica de Superfície e Imagem de Satélite (Goes 13). **Fonte:** O Autor

III. Treinamento e Teste: O treinamento e o teste dos algoritmos de aprendizado de máquina foram conduzidos usando o pacote WEKA (Witten et al., 2011), que oferece uma ampla gama de algoritmos de classificação. Os algoritmos de melhor desempenho no pacote foram treinados e testados usando a técnica de validação cruzada com 10 amostras utilizando o software WEKA com a definição dos Hiperparâmetros estabelecidos pelos valores “default” da plataforma. Este método envolve dividir todo o conjunto de dados em k subconjuntos mutuamente exclusivos de tamanho igual, onde um subconjunto é usado para teste, e os k-1 restantes são usados para estimativa de parâmetros e avaliação da precisão do algoritmo (Hastie et al., 2009).

IV. Avaliação: Para a avaliação dos algoritmos, foi utilizada uma tabela de contingência (ou matriz de confusão), onde temos que:

- (a) representa o número de casos previstos e observados;

(b) representa os casos em que o evento foi previsto, mas não observado;

(c) representa o número casos não previstos, mas observados; e

(d) indica os casos não previstos e não observados.

PREVISTO			
Não-Classe	Classe		
(d) correta rejeição	(b) falso alarme	Não-Classe	OBSERVADO
(c) não detecção	(a) acerto	Classe	

Tabela 4 Matriz de contingência WEKA
Fonte: Adaptado de Wilks

4.4. MÉTRICAS DE AVALIAÇÃO

Com base nessa tabela de contingência (tabela 4), foram calculadas as seguintes métricas: Probabilidade de Detecção (POD), Taxa de Alarme Falso (FAR), F-Measure e BIAS.

POD, também conhecida como sensibilidade ou *recall*, mede a fração de eventos positivos observados que foram corretamente previstos, é calculada como:

$$\mathbf{POD} = \frac{a}{a+c} \quad (2)$$

e tem valores ideais próximos de 1, indicando detecção perfeita.

FAR ou Probabilidade de Falsa Detecção (POFD) indica a razão de alarmes falsos para o número total de não ocorrências do evento. Essa métrica é importante para avaliar a quantidade de previsões incorretas do modelo e é dado por:

$$\mathbf{FAR} = \frac{b}{b+d} \quad (3)$$

com um valor ideal de 0, representando nenhum alarme falso.

F-Measure, é a média harmônica entre a precisão e a sensibilidade. Essa métrica é útil para equilibrar a relação entre precisão e a capacidade de detecção do modelo, sendo amplamente utilizada em problemas de classificação onde há uma compensação entre a detecção de eventos verdadeiros e a minimização de alarmes falsos, ou seja, ela equilibra o *trade-off* entre previsões corretas e alarmes falsos, sendo calculada como:

$$\mathbf{F - MEASURE} = \frac{2 \times \text{sensibilidade} \times \text{precisão}}{\text{sensibilidade} + \text{precisão}} = \frac{2 * \left(\frac{a}{a+c} * \frac{a}{a+b} \right)}{\left(\frac{a}{a+c} + \frac{a}{a+b} \right)} \quad (4)$$

O valor ideal da F-Measure é 1, indicando um equilíbrio perfeito entre verdadeiros positivos e a minimização de falsas previsões.

BIAS, ou “viés”, mede a tendência do modelo em subestimar ou superestimar a ocorrência de eventos. Ele é a razão entre o número total de previsões do evento (acertos e falsos alarmes) e o número total de eventos reais observados (acertos e não detecção), e é calculado por:

$$\mathbf{BIAS} = \frac{a+b}{a+c} \quad (5)$$

com valor ideal de 1, indicando que o modelo não tem tendência a superestimar ou subestimar eventos.

Essas métricas foram essenciais para identificar os pontos fortes e fracos dos algoritmos, permitindo ajustes para melhorar o desempenho durante a fase de treinamento e teste dos algoritmos.

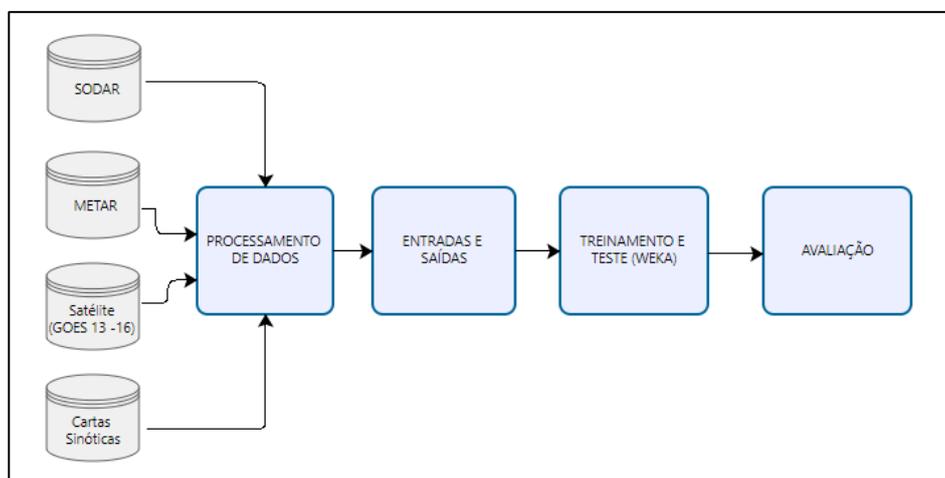


Figura 18 Fluxograma do método

Fonte: O autor

5. RESULTADOS E DISCUSSÕES

5.1. ANÁLISE DA OCORRÊNCIA DE TURBULÊNCIA E DO WS NAS PROXIMIDADES DO SBGR

A previsão de WS forte e severo, conforme definido nos objetivos deste estudo, envolve a identificação de preditores que capturam a gênese de tais eventos, seja devido a processos convectivos ou à influência das cadeias de montanhas da Serra da Cantareira e da Serra do Mar no fluxo de ar em uma atmosfera estável, ou ainda devido a condições de inversão térmica.

No estudo, foi realizada uma análise detalhada da disponibilidade dos dados de WS, DIR e SPD em relação à altura relativa ao solo, definindo o limite vertical para a pesquisa, de 30m até 200m, objetivando minimizar os prejuízos decorrentes da ausência de dados em níveis mais elevados (acima de 200 m). Esse limite permitiu focar em uma faixa de altura onde os dados disponíveis eram mais completos, otimizando a precisão das análises, buscando não comprometer a representatividade dos resultados.

Assim, foi possível observar, conforme a figura 19, que até 100m de altura existiam disponíveis 95,02% de dados e entre 110m e 200m existiam 82,10% de dados disponíveis de WS do SODAR. Entre 200m e 300m a redução na disponibilidade de dados foi bastante significativa, torna inviável sua utilização. Razão pela qual estabeleceu-se o limite superior em 200m de altura em relação ao solo.

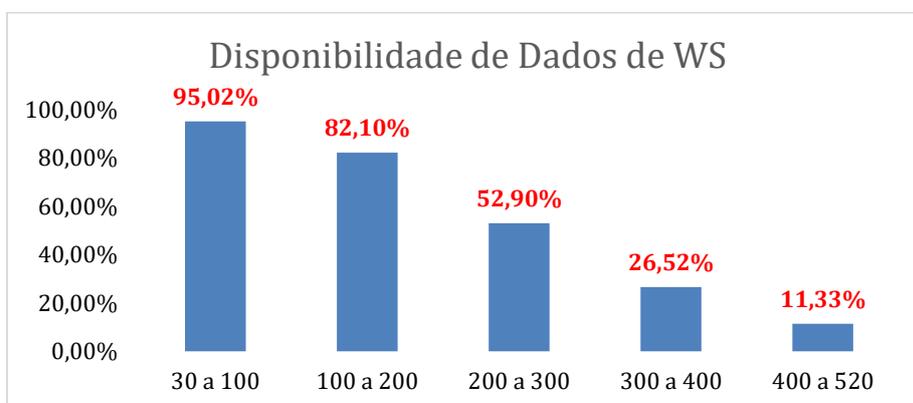


Figura 19 Disponibilidade de dados de WS com a altura (2016 a 2019)

Fonte: O autor

Essa diminuição de disponibilidade de dados do SODAR em níveis mais elevados de altura entende-se também como decorrência da influência das condições atmosféricas como umidade, temperatura e outros ruídos que atenuam de forma significativa as ondas sonoras emitidas pelo SODAR, causando assim uma variação natural da cobertura de dados disponíveis.

Posteriormente, a análise focou nas distribuições de ocorrências do WS, classificadas em três intensidades: moderada, forte e severa. Essa categorização permitiu entender a frequência e a intensidade do fenômeno, fornecendo uma visão mais detalhada sobre o padrão de comportamento do vento ao longo do período estudado, permitindo identificar como essas intensidades se distribuíram e qual sua relevância para o estudo.

A Figura 20 mostra as distribuições trimestrais, ou seja, DEZ-JAN-FEV, MAR-ABR-MAI, JUN-JUL-AGO e SET-OUT-NOV, nos registros de ocorrências de WS (Fig. 20a) e pousos e decolagens de aeronaves (Fig. 20b), de 1º de janeiro de 2016 a 31 de dezembro de 2019, no SBGR. A maior frequência de ocorrência de WS ocorreu durante os meses de inverno, coincidindo com a maior quantidade de pousos e decolagens em SBGR, totalizando 71.712 operações de acordo com o anuário estatístico de tráfego aéreo de 2023 do DECEA, representando 26% do tráfego anual no aeroporto.

Os registros de WS, excluindo as ocorrências de intensidade leves, totalizaram 22.938, com 22.390 classificados como moderados, representando 98% dos registros de ocorrência do WS. Registros fortes e severos totalizaram 527 e 21, respectivamente, compreendendo os 2% restantes dos registros analisados.

Acrescenta-se que o anuário estatístico de tráfego aéreo de 2023 do DECEA indica que aproximadamente 14% das operações de decolagem e pouso no SBGR, ocorrem entre meia-noite e 6h, um período de pico para casos de ocorrência de JBNS, aumentando o risco de encontros de WS associados a esse fenômeno, pois estudos observacionais na Região Metropolitana de São Paulo confirmam que JBNS são frequentes no inverno, especialmente à noite e perto da superfície, alinhando-se com o aumento de registros de ocorrência de WS de junho a agosto (Sánchez, 2022), aumentando os riscos para a aviação.

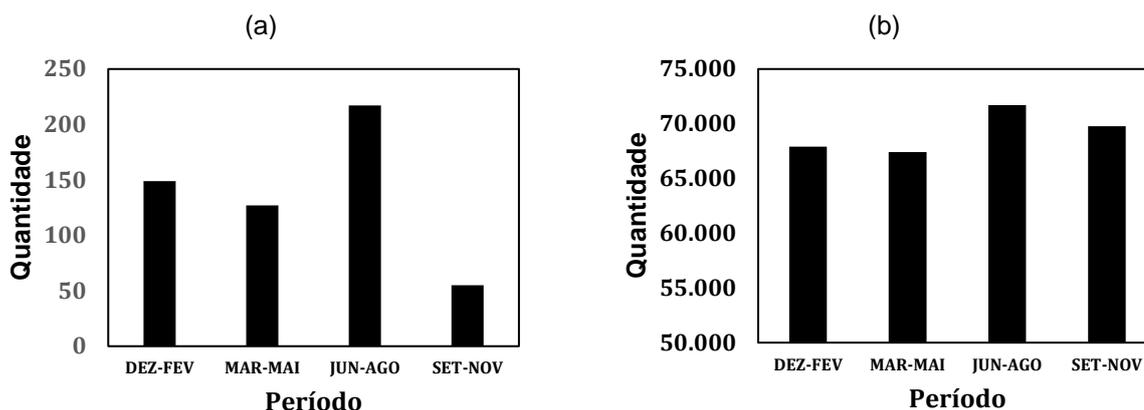


Figura 20 a) Distribuição do WS para intensidades fortes-severas, detectadas pelo sistema SODAR, para os períodos mensais de dezembro-fevereiro, março-maio, junho-agosto e setembro-novembro, de 1º de janeiro de 2016 a 31 de dezembro de 2019; b) Distribuição dos pousos e decolagens de aeronaves no SBGR, segundo o anuário estatístico de 2023 do DECEA
Fonte: O Autor

5.2. TREINAMENTO E TESTE DE ALGORITMO DE APRENDIZADO DE MÁQUINA

Esta seção descreve os procedimentos de treinamento e teste para os algoritmos de aprendizado de máquina utilizados para prever os eventos de WS no SBGR. Dada a interação complexa de fatores atmosféricos convectivos e estáveis, preditores derivados de perfis SODAR de 15 minutos, conforme mostrado na Tabela 2, formam o núcleo do processo de treinamento do modelo. Esses preditores capturam o comportamento do vento dentro da camada atmosférica entre 30 e 200 metros acima do solo, permitindo que o modelo se concentre nas principais características do vento essenciais para identificar padrões relacionados ao cisalhamento.

Segundo a metodologia passo a passo (Etapas I e II), os dados foram processados e, por tentativa e erro, observando-se a contribuição das variáveis na melhoria ou piora de desempenho dos algoritmos, oito características preditivas foram selecionadas (ver Tabela 3, coluna 3), que se acredita representarem a gênese de eventos de WS forte e severo.

Os dezesseis algoritmos de aprendizado de máquina foram então treinados e testados usando validação cruzada com 10 amostras, utilizando o software WEKA com a definição dos Hiperparâmetros estabelecidos pelos valores “default” da plataforma. Este método garante uma avaliação robusta e minimiza o *overfitting*, fornecendo uma avaliação abrangente da capacidade de cada modelo de generalizar em condições de vento variadas no SBGR.

O desempenho dos algoritmos foi avaliado usando quatro métricas específicas em três conjuntos de dados distintos que representam diferentes condições temporais:

(1) um conjunto de dados abrangendo todos os meses (**ALL**);

(2) um conjunto de dados para junho-julho-agosto (**JJA**) contendo registros de ocorrência de WS de várias origens (convectivas e não convectivas); e

(3) um conjunto de dados para junho-julho-agosto contendo apenas registros de ocorrência de WS decorrentes de condições atmosféricas estáveis (**JJA_stable**).

Embora 548 dos registros de ocorrência de WS fortes e severos tenham sido registrados durante o período do estudo (1º de janeiro de 2016 a 31 de dezembro de 2019), apenas 164 registros de ocorrência estavam disponíveis para o treinamento e teste do modelo devido a restrições, especificamente pela falta de dados de entrada (preditores) para todos os intervalos necessários dentro dos 60 minutos anteriores a ocorrência do WS.

A distribuição sazonal das ocorrências de WS foi a seguinte: 15 registros de ocorrência em dezembro-janeiro-fevereiro (DJF), 49 registros de ocorrência em março-abril-maio (MAM), 91 registros de ocorrência em junho-julho-agosto (JJA) e 9 registros de ocorrência em setembro-outubro-novembro (SON), representando 9,1%, 29,9%, 55,5% e 5,5% do total de ocorrências nesses trimestres, respectivamente.

Essa abordagem estruturada de treinamento e teste, combinada com dados SODAR de alta frequência, dá suporte à capacidade do modelo de prever eventos de WS no SBGR com precisão sob diversas condições atmosféricas.

5.2.1. Avaliação dos Algoritmos

5.2.1.1. Análise de algoritmos de aprendizado de máquina para previsão de WS

A Tabela 5 apresenta as métricas de avaliação para os 16 algoritmos de aprendizado de máquina usando o período de dados completo (ALL), que inclui todos os tipos de eventos de gênese de WS. Os resultados indicam desafios substanciais na geração de previsões precisas de 15 minutos. A POD teve uma média de $0,45 \pm 0,19$, revelando um problema comum com baixo desempenho na identificação de eventos relevantes. Enquanto a FAR é baixa em $0,01 \pm 0,00$, indicando que os modelos minimizam efetivamente os alarmes falsos, a baixa POD sugere potencial falta de sensibilidade na captura de ocorrências significativas.

Em particular, o NaiveBayes atinge uma precisão de 0,945 com uma FAR de 0,044, mostrando alguma variabilidade de previsão, enquanto o BayesNet tem uma precisão de 0,841 e uma FAR de 0,016, tornando esses modelos potencialmente adequados para tarefas priorizadas pela eficiência. No entanto, ambos exibem um alto nível de BIAS (30,756 e 11,872, respectivamente), o que pode impactar a confiabilidade geral.

A F-Measure, com média de $0,43 \pm 0,15$, ressalta ainda mais as dificuldades em equilibrar precisão e sensibilidade. Uma métrica BIAS de $3,17 \pm 7,63$ indica uma tendência à superestimação, onde as previsões excedem os eventos reais. Essa combinação de baixas POD e superestimação destacam a necessidade de mais ajustes de modelo para melhorar a precisão e a confiabilidade da detecção. A variabilidade observada em POD e BIAS sugere o potencial de incorporar preditores adicionais para aprimorar o desempenho do modelo para *nowcasting* de 15-minutos.

Algoritmo	POD	FAR	F-Measure	BIAS
RandomForest	0.494	0.000	0.600	0.646
IBk5	0.433	0.000	0.553	0.567
NBTree	0.457	0.000	0.545	0.677
J48	0.463	0.000	0.543	0.707
LMT	0.427	0.000	0.524	0.628
MultilayerPerceptron	0.415	0.000	0.515	0.610
RandomTree	0.506	0.001	0.500	1.024
REPTree	0.390	0.000	0.498	0.567
DecisionTable	0.360	0.000	0.478	0.506
PART	0.341	0.000	0.455	0.500
OneR	0.335	0.000	0.453	0.482
Logistic	0.280	0.000	0.385	0.457
HoeffdingTree	0.238	0.000	0.344	0.384
SimpleLogistic	0.207	0.000	0.319	0.299
BayesNet	0.841	0.016	0.131	11.872
NaiveBayes	0.945	0.044	0.060	30.756

Tabela 5 Métricas de validação cruzada com 10 amostras para 16 algoritmos de aprendizado de máquina, usando dados de todos os meses (janeiro a dezembro) durante o período de 2016 a 2019, capturando métricas para eventos de WS com gênese de atmosfera convectiva e estável

Fonte: O autor

Neste estudo, os preditores foram limitados e derivados apenas de perfis do SODAR de 15-minutos (conforme mostrado na Tabela 3), que capturam o comportamento do vento, incluindo cisalhamento, dentro da camada atmosférica de 30 a 200 metros nos últimos 60 minutos. No entanto, indicadores convectivos críticos — como índices termodinâmicos de radiômetros de micro-ondas de perfil (que podem estimar continuamente a temperatura do ar, umidade e estrutura do líquido atmosférico), radiossondas de baixa frequência ou perfis atmosféricos previstos de modelos meteorológicos numéricos — não foram incluídos.

Essa omissão pode restringir a capacidade do modelo de detectar o WS de origem convectiva, contribuindo para limitações observadas na precisão preditiva, conforme discutido acima e refletido nas métricas de avaliação dos algoritmos mostrados na Tabela 5.

O WS em condições estáveis também apresenta riscos significativos durante o pouso e a decolagem devido à ausência de indicadores visíveis, ao contrário das condições convectivas, onde nuvens cumuliformes sinalizam a possibilidade de encontro com regiões de turbulência localizada.

Ademais, uma atmosfera estável tende a favorecer a geração de áreas com a presença de WS devido à estratificação das camadas de ar, onde o ar mais frio e denso permanece abaixo do ar mais quente, criando barreiras que resultam em variações bruscas na velocidade e direção do vento entre essas camadas. Essa estabilidade inibe a convecção vertical, levando a uma concentração de áreas de ventos mais intensos em altitudes superiores, enquanto as camadas inferiores permanecem mais calmas.

Conseqüentemente, registros de ocorrência de WS forte e severo associados a atmosferas instáveis foram excluídos da análise seguinte, resultando em uma distribuição trimestral específica para casos de atmosfera estável, como segue: DJF - 0, MAM - 4, JJA - 67 e SON - 0. Esse conjunto de dados refinado permitiu uma avaliação focada de algoritmos de aprendizado de máquina para WS forte-severo em condições estáveis, onde apenas os casos contínuos e superiores a 30 minutos de ocorrência detectados pelo SODAR foram incluídos.

5.2.1.2. Análise comparativa dos conjuntos de dados de JJA e JJA_Stable

A Tabela 6 apresenta os resultados métricos para os 16 algoritmos de aprendizado de máquina definidos anteriormente no conjunto de dados (2) para JJA, que contém registros de ocorrência de WS forte-severo de várias origens (convectivas e não convectivas), e no conjunto de dados (3) para JJA, que inclui apenas registros de ocorrência de WS forte-severo decorrentes de condições atmosféricas estáveis e de continuidade superiores a 30 minutos de ocorrência (JJA_stable).

Essa distinção facilita a avaliação do desempenho do algoritmo sob condições variáveis de gênese de WS, oferecendo *insights* sobre a adaptação do modelo a condições estáveis versus condições convectivas-estáveis mistas durante o período de JJA.

Comparando o desempenho entre esses conjuntos de dados, surgem diferenças notáveis na eficácia de cada algoritmo:

- **RandomForest** se destaca no conjunto de dados JJA_Stable, alcançando um POD alto de 0,716 e uma F-Measure de 0,774, com FAR mínimo (0,000) e BIAS de 0,851, indicando detecção robusta e previsões balanceadas em atmosferas estáveis.

- **NBTree** também tem bom desempenho em condições estáveis, com um POD de 0,701 e uma F-Measure de 0,734, mantendo um FAR de 0,001 e um BIAS de 0,910, sugerindo superestimação controlada.

- **LMT** mostra resultados balanceados, particularmente em condições estáveis, com um POD de 0,687, F-Measure de 0,767 e BIAS de 0,791, refletindo detecção confiável sem superestimação excessiva.

Em contraste, NaiveBayes e BayesNet, respectivamente, alcançam 0,985 e 0,866 de POD, mas sofrem com pontuações extremas de BIAS de 17,881 e 2,746, indicando superestimação substancial, o que limita sua aplicação prática apesar das altas taxas de detecção. Modelos de desempenho inferior, como Logistic, SimpleLogistic e HoeffdingTree, exibem baixos valores de POD e F-Measure, refletindo eficácia limitada na detecção de eventos de WS em ambos os conjuntos de dados.

Esta análise ressalta a importância de selecionar algoritmos que podem se adaptar a condições atmosféricas específicas, com RandomForest, NBTree e LMT demonstrando o melhor desempenho geral em condições atmosféricas estáveis no período de JJA.

Algoritmo	JJA				JJA_Stable			
	POD	FAR	F-MEASURE	BIAS	POD	FAR	F-MEASURE	BIAS
RandomForest	0.615	0.001	0.683	0.802	0.716	0.000	0.774	0.851
NBTree	0.560	0.001	0.638	0.758	0.701	0.001	0.734	0.910
LMT	0.571	0.001	0.619	0.846	0.687	0.000	0.767	0.791
RandomTree	0.626	0.002	0.591	1.121	0.716	0.001	0.711	1.015
J48	0.571	0.001	0.654	0.747	0.612	0.000	0.707	0.731
PART	0.462	0.000	0.600	0.538	0.552	0.000	0.661	0.672
IBk5	0.462	0.001	0.568	0.626	0.567	0.001	0.639	0.776
DecisionTable	0.462	0.001	0.564	0.637	0.478	0.000	0.604	0.582
REPTree	0.451	0.001	0.522	0.725	0.507	0.000	0.602	0.687
Multilayer Perceptron	0.451	0.001	0.532	0.692	0.448	0.000	0.566	0.582
OneR	0.374	0.000	0.504	0.484	0.448	0.000	0.566	0.582
BayesNet	0.835	0.012	0.306	4.462	0.866	0.005	0.462	2.746
Logistic	0.275	0.001	0.376	0.462	0.343	0.000	0.447	0.537
SimpleLogistic	0.231	0.000	0.336	0.374	0.224	0.000	0.326	0.373
HoeffdingTree	0.220	0.001	0.305	0.440	0.134	0.000	0.220	0.224
NaiveBayes	0.945	0.045	0.123	14.407	0.985	0.041	0.104	17.881

Tabela 6 Métricas de validação cruzada com 10 amostras para 16 algoritmos de aprendizado de máquina usando dados do JJA e JJA_Stable. O JJA inclui eventos de cisalhamento de vento forte-severo relacionados a atmosferas convectivas e estáveis, enquanto o JJA_Stable inclui apenas eventos de WS forte-severo em atmosfera estável com detecção contínua superior a 30 minutos, para o período de 2016–2019

Fonte: O Autor

6. CONCLUSÃO

Motivado pela necessidade de aumentar a segurança e a eficiência da atividade aeronáutica, o objetivo deste estudo foi desenvolver uma abordagem baseada em aprendizado de máquina para prever eventos de WS forte-severo no SBGR. Usando perfis SODAR de 15 minutos capturando o comportamento do vento na camada atmosférica de 30 a 200 metros. O estudo analisou 16 algoritmos de aprendizado de máquina para avaliar seu desempenho na previsão de eventos de WS para uma previsão com 15 minutos de antecedência. Ao separar os registros de ocorrência de WS com base em sua gênese (convectiva *versus* estável), o modelo teve como objetivo fornecer *insights* personalizados para a segurança da aviação no ambiente meteorológico complexo do SBGR.

A previsão do fenômeno de WS em condições de atmosfera estável tem uma importância crítica para a segurança operacional em procedimentos de aproximação, pouso e decolagem. Em situações de estabilidade atmosférica, como ocorre frequentemente nos meses de inverno, os eventos de WS podem passar despercebidos, uma vez que os pilotos tendem a esperar uma menor atividade turbulenta nessas condições.

No entanto, esta pesquisa comprovou que mesmo em atmosferas predominantemente estáveis, o WS pode ocorrer com intensidade moderada a severa, representando um risco substancial. Assim, a antecipação desse fenômeno, especialmente em condições de estabilidade, é essencial para alertar equipes operacionais com antecedência suficiente, possibilitando uma tomada de decisão mais precisa e mitigando os riscos associados a essas ocorrências.

Para investigar esse fenômeno, foram utilizadas ferramentas de aprendizado de máquina, por meio do software WEKA, aplicado a dados de WS coletados pelo SODAR instalado no SBGR entre os anos de 2016 e 2019.

Outro fator importante da pesquisa foi a aplicação de validação cruzada com 10 amostras, pois essa técnica, amplamente reconhecida por sua capacidade de gerar modelos preditivos confiáveis, permitiu uma avaliação mais rigorosa do desempenho dos algoritmos de aprendizado de máquina, possibilitando a redução da variabilidade nos resultados e evitar problemas de *overfitting*, garantindo que os modelos fossem capazes de generalizar bem para novos dados.

As métricas de avaliação adotadas na pesquisa, incluindo POD, FAR, F-Measure, Bias, Precisão e Recall, permitiram uma análise criteriosa do desempenho dos modelos e ajudaram a identificar os algoritmos mais eficazes para a previsão de WS.

Os resultados revelaram variabilidade significativa entre os algoritmos, tanto que RandomForest, NBTree e LMT obtiveram melhores desempenhos, particularmente na previsão de eventos de WS forte-severo em condições atmosféricas estáveis, no período de JJA e para eventos de continuidade superior a 30 minutos de ocorrência. Esses algoritmos apresentaram altas pontuações de POD e F-Measure, juntamente com taxas mínimas de alarme falso, destacando sua eficácia na captura de padrões de WS em atmosfera estável.

A distinção entre os registros de ocorrência de WS convectivos e os de origem estável permitiu que esses modelos se adaptassem mais especificamente à dinâmica atmosférica que precede o WS em SBGR. No entanto, NaiveBayes e BayesNet exibiram POD alto, mas sofreram de BIAS excessivo, indicando uma tendência à superestimação. Isso reflete a necessidade de mais refinamento do modelo para evitar superestimar a frequência do evento, o que pode reduzir a confiabilidade da previsão.

Em resumo, os algoritmos geralmente tiveram melhor desempenho em condições atmosféricas estáveis, provavelmente devido à natureza estruturada e previsível dos padrões de cisalhamento estáveis. Em contraste, taxas de detecção mais baixas para eventos de WS convectivo podem ser atribuídas a limitações em variáveis preditoras, particularmente a ausência de índices termodinâmicos essenciais para identificar o cisalhamento convectivo.

O estudo não incluiu preditores convectivos, como índices termodinâmicos de radiômetros de micro-ondas de perfil (que medem continuamente a temperatura do ar, umidade e estrutura do líquido), dados de radiossonda de baixa resolução temporal ou perfis atmosféricos previstos de modelos meteorológicos. Essa omissão provavelmente restringiu a capacidade do modelo de detectar com precisão o WS de origem convectiva.

Assim, dada a eficácia dos modelos preditivos desenvolvidos, uma recomendação importante para futuras pesquisas é a incorporação dessas variáveis

adicionais para aprimorar as capacidades preditivas do modelo em cenários convectivos, onde o WS também representa um alto risco operacional.

Da mesma forma, recomenda-se o desenvolvimento de modelos híbridos, que combinem técnicas de aprendizado de máquina com modelagem numérica de tempo. Essa abordagem poderia melhorar ainda mais a precisão das previsões de WS, integrando dados de modelagem atmosférica de alta resolução com os padrões aprendidos pelos algoritmos de aprendizado de máquina.

Outra área promissora de investigação seria o uso combinado de sensores remotos, como Radars Doppler e LIDAR, podendo fornecer uma visão mais detalhada dos microclimas locais que favorecem a formação de WS, potencializando o uso operacional desses modelos em aeroportos de grande porte, como o SBGR.

Conclui-se, portanto, que esta pesquisa demonstrou a viabilidade e a importância de ferramentas preditivas baseadas em aprendizado de máquina para a detecção de eventos de WS. Os cenários analisados, especialmente os que focaram em atmosferas estáveis, indicaram que tais ferramentas têm o potencial de serem aplicadas em centros de monitoramento e vigilância meteorológicas, assim como em torres de controle de tráfego aéreo, aumentando a segurança e a eficiência das operações aéreas.

7. REFERÊNCIAS BIBLIOGRÁFICAS

ABEAR (2020) *Relatório de Impacto Econômico da Aviação no Brasil – 2019*. Associação Brasileira de Empresas Aéreas. Available at : <https://www.abear.com.br>

Aguiar GO, SILVA VR (2022) Sistema de previsão climática baseado em Machine Learning, Brasília, DF, 2022- Trabalho de Conclusão de Curso – Universidade de Brasília – UnB Faculdade UnB Gama – FGA , 2022 https://bdm.unb.br/bitstream/10483/34525/1/2022_GuilhermeAguiar_VictorSilva_tcc.pdf

Aha DW, Kibler D, and Albert MK (1991) Instance-based learning algorithms. *Machine Learning*, 6(1), 37-66. <https://doi.org/10.1007/BF00153759>

Allende-Cid H (2019) Machine Learning: catalisador da ciência. *Computação Brasil - Revista da Sociedade Brasileira de Computação*; nº39 ed.1/2019

Almeida MV (2009) Aplicação de técnicas de redes neurais artificiais na previsão de curtíssimo prazo da visibilidade e teto para o aeroporto de Guarulhos – SP. 2009. 206 f. Tese (Doutorado) - Curso de Engenharia Civil, Coppe, UFRJ, Rio de Janeiro, 2009.

Anderson JD (2010) *Fundamentals of Aerodynamics*. New York: Mcgraw-Hill Education. ISBN 0073398101

Atkinson B (2003) Numerical Modelling of Urban Heat-Island Intensity. *Boundary-Layer Meteorology*. <https://doi.org/10.1023/A:1025820326672>

Banta, Robert and Mahrt, Larry and Vickers, Dean and Sun, Jielun and Balsley, Ben and Pichugina, Yelena and Williams E (2007) The Very Stable Boundary Layer on Nights with Weak Low-Level Jets. *Journal of The Atmospheric Sciences - J ATMOS SCI*. <https://doi.org/10.1175/JAS4002.1>

Breiman L (2001) Random Forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>

Chollet F (2017) *Deep Learning with Python*. New York, NY: Manning Publications. [http://silverio.net.br/heitor/disciplinas/eeica/papers/Livros/\[Chollet\]-Deep_Learning_with_Python.pdf](http://silverio.net.br/heitor/disciplinas/eeica/papers/Livros/[Chollet]-Deep_Learning_with_Python.pdf)

Costa MMGA (2008) Meteorologia Aeronáutica no Aeroporto de Guarulhos. In: SITRAER VII – SIMPÓSIO DE TRANSPORTE AÉREO. Anais, Rio de Janeiro RJ

Cox DR (1958) The regression analysis of binary sequences. *Journal of the Royal Statistical Society: Series B (Methodological)*, 20(2), 215-232

Crescenti GA (1996) Look Back on Two Decades of Doppler Sodar Comparison Studies. *Bulletin of the American Meteorological Society*. 1997. 78. 10.1175/1520-0477(1997)078<0651:ALBOTD>2.0.CO;2

Crosier J et al (2002) *LIDAR and SODAR Systems at Heathrow Airport. Atmospheric Measurement Techniques*

DECEA (2023) Anuário Estatístico de Tráfego Aéreo 2023. Departamento de Controle do Espaço Aéreo. http://portal.cgna.decea.mil.br/files/uploads/anuario_estatistico/anuario_estatistico_2023.pdf

Domingos P, Hulten G (2000). Mining high-speed data streams. In *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 71-80). <https://doi.org/10.1145/347090.347107>

Duda RO, Hart PE (1973). *Pattern Classification and Scene Analysis*

Estudo de Impacto Ambiental – EIA. Aeroporto Internacional de São Paulo / Guarulhos. 2004. https://cetesb.sp.gov.br/eiarima/eia/EIA_13536_2002.pdf

FAA (1988) Pilot Wind Shear Guide. Federal Aviation Administration. https://www.faa.gov/documentLibrary/media/Advisory_Circular/AC00-54.pdf

FAA (2008) *Wind Shear*. Federal Aviation Administration. <https://www.faasafety.gov/files/gslac/library/documents/2011/aug/56407/faa%20p-8740-40%20windshear%5Bhi-res%5D%20branded.pdf>

França GB, de Almeida MV, Bonnet SM, Albuquerque Neto FL (2018). Nowcasting model of low wind profile based on neural network using SODAR data at Guarulhos Airport, Brazil. *International Journal of Remote Sensing*, 39(8), 2506–2517. <https://doi.org/10.1080/01431161.2018.1425562>

Frank E, Witten IH (1998) Generating accurate rule sets without global optimization. In *Proceedings of the Fifteenth International Conference on Machine Learning* (pp. 144-151)

Gao H, Shen C, Zhou Y, Xuesong W, Chan PW, Hon K, Li J (2022) A Deep Learning-Based Wind Field Nowcasting Method With Extra Attention on Highly Variable Events. *IEEE Geoscience and Remote Sensing Letters*. 2022.PP. 1-1. 10.1109/LGRS.2022.3212904

Garratt JR (1994) Review: The atmospheric boundary layer. *Earth-Science Reviews*, Volume 37, Issues 1–2, 1994, Pages 89-134, ISSN 0012-8252, [https://doi.org/10.1016/0012-8252\(94\)90026-4](https://doi.org/10.1016/0012-8252(94)90026-4). <<https://www.sciencedirect.com/science/article/pii/0012825294900264>>

Gultepe I, Sharman R, Williams PD, Zhou B, Ellrod G, Minnis P, Trier S, Griffin S, Yum SS, Gharabaghi B, et al. (2019) A review of high impact weather for aviation meteorology. *Pure Appl. Geophys.* 2018, 176, 1869–1921.

Hastie T, Tibshirani R, Friedman J (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2. ed. New York: Springer

Hsieh W, Tang B (1998) Applying Neural Network Models to Prediction and Data Analysis in Meteorology and Oceanography. *Bulletin of The American Meteorological Society - BULL AMER METEOROL SOC.* 1998.79. 1855-1870. DOI:10.1175/1520-0477(1998)079<1855:ANNMTP>2.0.CO;2

Holte RC (1993) Very simple classification rules perform well on most commonly used datasets. *Machine Learning*, 11(1), 63-91. <https://doi.org/10.1023/A:1022631118932>

ICAO (2005) Manual On Low-Level Wind Shear. DOC9817 AN/449. International Civil Aviation Organization

ICAO (2016) *Global Air Transport Outlook to 2030 and Trends to 2040* (DOC 9750-NA/963). International Civil Aviation Organization

Institute of Atmospheric Physics. Status Report 2008 – 2015, Volume 1. Deutsches Zentrum für Luft- und Raumfahrt (German Aerospace Center).

<https://www.dlr.de/pa/en/Portaldata/33/Resources/dokumente/publikationen/Status-Report-2015.pdf>

Kodama Y (1992) *Large-scale common features of subtropical precipitation zones (the Baiu frontal zone, the SPCZ, and the SACZ): Characteristics and circulation*. Journal of the Meteorological Society of Japan, 70(4), 813-836

Little CG (1969) Acoustic methods for the remote probing of the lower atmosphere. Proceedings of the IEEE, v. 57, n. 4, p. 571–578, 1 jan. 1969.

<https://ntrs.nasa.gov/api/citations/19720017734/downloads/19720017734.pdf>

Kohavi R (1996) Scaling up the accuracy of Naive-Bayes classifiers: A decision-tree hybrid. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining* (pp. 202-207)

Landwehr N, Hall M, Frank E (2005) Logistic model trees. *Machine Learning*, 59(1-2), 161-2. <https://doi.org/10.1007/s10994-005-0466-3>

Marth H, Banta RM, Pichugina YL, Newsom RK (1998) *The impact of low-level jets on turbulence and wind shear*. Journal of Atmospheric and Oceanic Technology, 15(5), 1076–1090

Matschinske MR (2013) *Wind Shear Events Reported at Guarulhos International Airport Between 1999 and 2012*. *Brazilian Journal of Aviation*, 2013

Murphy KP (2012) *Machine Learning: A Probabilistic Perspective*. The MIT Press. ISBN 978-0-262-01802-9

NTSB (2010) *Weather-related aviation accident study 2003–2007*. National Transportation Safety Board. <https://www.asias.faa.gov/i/studies/2003-2007weatherrelatedaviationaccidentstudy.pdf>

Oke TR (1982) “The Energetic Basis of the Urban Heat Island”, *Quarterly J. Royal Meteorol. Soc.*, 108(455): 1-24. <https://doi.org/10.1002/qj.49710845502>

Oke TR (1987) *Boundary Layer Climates* (2nd ed.). Routledge. ISBN 9780415043199

Pearl J (1988) *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann. <https://doi.org/10.1016/C2009-0-27609-4>

Pereira MCM (2016) *Emprego ne redes neurais artificiais RBF na previsão de nevoeiro no Aeroporto Internacional Afonso Pena*. 109 p. Dissertação (Mestrado em Meteorologia), Universidade Federal do Paraná, Paraná, 2016

Quinlan JR (1993) *C4.5: Programs for Machine Learning*. Morgan Kaufmann

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. *Nature*, 323(6088), 533-536

Sánchez, M. P. (2022) *Investigação dos Jatos de Baixos Níveis na Região Metropolitana de São Paulo*. Tese de Doutorado, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, São Paulo. doi:10.11606/T.14.2022.tde-02082022-140250. Recuperado em 2024-10-18, de www.teses.usp.br

Shun CM, Chan PW (2008) *Applications of an Infrared Doppler Lidar in Detection of Windshear*. *Journal of Atmospheric and Oceanic Technolog*. <https://doi.org/10.1175/2007JTECHA1057.1>

Silva HAJ (2016) *Análise das Ocorrências de Cisalhamento de Vento no Aeroporto de Guarulhos (SP) para a Prevenção de Acidentes Aeronáuticos*. 98 p. Dissertação (Mestrado em Meteorologia), Universidade de São Paulo, São Paulo, 2016.

Silva WL, Albuquerque Neto FL, França GB, Matschinske MR (2016) Conceptual model for runway change procedure in Guarulhos International Airport based on SODAR data. 2016. *The Aeronautical Journal*. 120. 725-734. 10.1017/aer.2016.33.

Stull RB (1988) *An introduction to boundary layer meteorology*. Dordrecht: Kluwer Academic Publishers, 1988. 670 p https://clouds.eos.ubc.ca/~phil/courses/atsc500/docs/Stull.1988_Boundary-Layer-Meteorology.pdf

Villela T (2005) *The Development of Wind Shear Safety Protocols in Aviation*. *Journal of Aviation Safety*

Wakimoto RM (1985) *The thunderstorm microburst: Dynamics of a wind shear hazard*. *Annual Review of Fluid Mechanics*, 17(1), 283–319

Wilks SD (2006) *Statistical Methods in the Atmospheric Sciences*. 2^a ed., New York, USA, Academic Press, 2006. <https://sunandclimate.wordpress.com/wp-content/uploads/2009/05/statistical-methods-in-the-atmospheric-sciences-0127519661.pdf>

Williams JK (2014) *Using random forests to diagnose aviation turbulence*. 2014. <https://doi.org/10.1007/s10994-013-5346-7>

Witten IH, Frank E (2005) *Data Mining: Practical Machine Learning Tools and Techniques*. 2nd Edition. Morgan Kaufmann

Ynoue RY et al (2014) *Meteorologia*. São Paulo: USP/Univesp/Edusp, 2014 https://midia.atp.usp.br/plc/plc0009/impressos/plc0009_12.pdf