



ESCOLA DE APERFEIÇOAMENTO DE OFICIAIS DA AERONÁUTICA
CURSO DE APERFEIÇOAMENTO DE OFICIAIS DA AERONÁUTICA
2/2022

GUILHERME VIEIRA DA ROCHA, Cap Av

**Construção de modelo matemático para predição de ocorrências aeronáuticas
na Academia da Força Aérea**

Rio de Janeiro

2022

ESCOLA DE APERFEIÇOAMENTO DE OFICIAIS DA AERONÁUTICA
CURSO DE APERFEIÇOAMENTO DE OFICIAIS DA AERONÁUTICA
2/2022

GUILHERME VIEIRA DA ROCHA, Cap Av

**Construção de modelo matemático para predição de ocorrências aeronáuticas
na Academia da Força Aérea**

Trabalho de conclusão de curso apresentado no Curso de Aperfeiçoamento de Oficiais da Aeronáutica como requisito parcial para aprovação no Curso de Especialização *lato-sensu* em Liderança com Ênfase em Gestão no COMAER.

Linha de Pesquisa: Segurança de Voo

Orientador: Thiago Diorgilis Ribeiro **DANIEL**,
Maj Av

Rio de Janeiro

2022

GUILHERME VIEIRA DA ROCHA, Cap Av

**Construção de modelo matemático para predição de ocorrências aeronáuticas
na Academia da Força Aérea**

Trabalho de conclusão de curso apresentado
no Curso de Aperfeiçoamento de Oficiais da
Aeronáutica.

Aprovado por:

Thiago Diorgilis Ribeiro **Daniel**, Maj Av
EAOAR

Carlos **Eduardo** José da Silva, Maj Esp Av
EAOAR

Rio de Janeiro

2022

RESUMO

Sistemas sociotécnicos modernos como a aviação têm por natureza a alta complexidade de interação entre seus elementos, podendo esta interação resultar em acidentes aeronáuticos catastróficos. A fim de entender e prever estes comportamentos, autoridades aeronáuticas internacionais têm investido em modelagem matemática e *machine learning*. A Força Aérea Brasileira, entretanto, não possui direcionamento explícito em suas concepções estratégicas, transcendendo esta conduta às suas Organizações Militares, como a Academia da Força Aérea e o 1º Esquadrão de Instrução Aérea. Nesse contexto, este ensaio acadêmico propõe como tese a utilização de dados sobre o tipo de ocorrência aeronáutica e a quantidade de horas/experiência de voo de pilotos da aeronave T-27 com o intuito de construir modelo matemático para predição de ocorrências aeronáuticas e, conseqüentemente, prover nova interposição na política de prevenção de acidentes aeronáuticos. Como argumento, primeiramente, considera-se a experiência de voo como fator influente em uma ocorrência, havendo correlação entre estes. Outro ponto indica que é possível prever uma variável com base no valor de outra por meio de um modelo matemático, bem como ajustá-lo para atender à complexidade do cenário e aproximá-lo progressivamente da realidade. Em finalização, considera-se o benefício de aplicar conceitos econométricos para prever comportamentos e riscos em ascensão não somente no âmbito em tela, mas em outros Esquadrões Operacionais e áreas como a Segurança do Trabalho, baseando nesta conduta a adoção de medidas e políticas convergentes à prevenção contra acidentes aeronáuticos, resguardando vidas, material de alto valor agregado e, em última instância, o erário brasileiro.

Palavras-chave: Segurança de voo. Modelagem matemática. Predição de acidentes. Análise Econométrica.

1 INTRODUÇÃO

Conforme Leveson (2011), elementos componentes de sistemas sociotécnicos como a aviação interagem de forma complexa, com difícil previsibilidade, podendo resultar em acidentes aeronáuticos catastróficos, perda de material e pessoal de alto valor agregado e prejuízo ao erário (PERROW, 1986).

À despeito disso, com o intuito de entender e prever estes comportamentos, autoridades responsáveis pela prevenção de acidentes aeronáuticos têm utilizado métodos para criação de modelos matemáticos (MM), predição de cenários e gerenciamento da política de prevenção de acidentes. Como exemplo, a *Federal Aviation Administration* (FAA) indica a “otimização de informações para reduzir o risco”, focando na elaboração de MM para tomada de decisão (FAA, 2022, p. 8). Na mesma linha, a *European Aviation Safety Agency* (EASA) tem implementado uma estrutura para gerenciar riscos com *machine learning* (ML) para prever ocorrências aeronáuticas por meio de MM flexíveis (EASA, 2021).

No Brasil, o Comando da Aeronáutica (COMAER) não possui direcionamento explícito para a análise preditiva de dados com vistas à prevenção de acidentes em termos de DIPLAN (Diretriz de Planejamento Institucional) e PEMAER (Planejamento Estratégico Militar da Aeronáutica).

Por sua vez, o Centro de Investigação e Prevenção de Acidentes Aeronáuticos (CENIPA) não possui iniciativas nesta linha e, sendo assim, esta conduta transcende para outras Organizações Militares (OM) da Força Aérea Brasileira (FAB) conectadas sistemicamente, como a Academia da Força Aérea (AFA), mais especificamente o 1º Esquadrão de Instrução Aérea (1º EIA), conforme Programa de Prevenção a Acidentes Aeronáuticos da organização (BRASIL, 2021).

Perante o submetido acima e como tese deste ensaio, defende-se que a utilização dos registros administrativos dos Relatórios de Ação Inicial (RAI) da AFA (tipo de ocorrência e quantidade de horas de voo de pilotos) possibilita a elaboração de um MM preditivo de ocorrências aeronáuticas.

Para fundamentar o exposto, dois argumentos serão desenvolvidos. O primeiro argumento considera que há correlação estatística entre uma ocorrência aeronáutica e a quantidade de horas de voo.

O segundo argumento tratará que é possível a previsão de uma variável com base no valor de outra por meio de um MM, bem como ajustá-lo para atender à complexidade do cenário e aproximá-lo progressivamente da realidade.

Ressalte-se que a tomada de dados para previsão de um comportamento não necessariamente precisa ser feita a partir de ML, havendo métodos matemático-estatísticos suficientemente aplicáveis para este fim. Ademais, ML não leva em consideração a relação entre variáveis, questão preponderante na análise de causalidade de ocorrências aeronáuticas (CHARPENTIER; FLACHAIRE; LY, 2019).

2 ECONOMETRIA E MODELAGEM ECONOMÉTRICA

Econometria é a “ciência e a arte de usar a teoria econômica e técnicas estatísticas para analisar dados econômicos” (STOCK; WATSON, 2015, p. 1, tradução nossa). Seu objetivo maior é a formulação de modelos para estudar e entender a realidade bem como auxiliar a tomada de decisão e analisar políticas públicas (MADDALA, 1992). Estes modelos, por sua vez, são uma representação da realidade em escala reduzida, porém, por impossibilidade de delinear as infinitas variáveis da aviação em uma proposição matemática, não irão exprimir a exata realidade, sendo possível listar e observar muitos outros fatores que podem desencadear uma ocorrência e outros que não podemos listar por impossibilidade de observar, mas que também são influentes. Todos esses, de algum modo, devem ser considerados no modelo como um fator de erro (MCCULLAGH; NELDER, 1983).

Tomando por base a taxonomia proposta por Stock e Watson (2015) e o recorte de dados sugerido, o enfoque deve ser na análise de dados observacionais¹, sendo estes classificados como dados de corte transversal².

2.1 Influência das horas de voo de um piloto em uma ocorrência aeronáutica

Na obra “*The killing zone*” (utilizando dados do *National Transportation Safety Board*), Craig (2001) discorre sobre a “zona da morte”, faixa de horas de voo que determina a maior recorrência de acidentes aeronáuticos³, sendo a maior propensão

¹ Que não são acumulados por experimentos controlados

² Os quais não são influenciados por diferenças temporais

³ Destruição total do equipamento e/ou no mínimo uma fatalidade

de um acidente quando o piloto se encontra entre 51 e 350h de voo. Essa informação advém da hipótese de que nesta faixa os pilotos se expõem a um risco maior, devido ao excesso de confiança⁴, falta de experiência e habilidades para lidar com eventos raros e desafiadores.

Knecht (2015) indica, em complemento e por meio de MM, que esta faixa de horas de voo pode ser diferente do analisado por Craig, tendo apresentado medianas de 250,5h e 823,5h para pilotos sem qualificação de voo por instrumentos e com qualificação de voo por instrumentos, respectivamente, podendo a zona da morte ter maior amplitude quando considerando a taxa de acidentes por horas de voo, apresentando alta correlação entre estas variáveis (coeficiente de determinação⁵ de 0,654 e 0,775, respectivamente).

No entanto, essa zona pode ser muito mais ampla do que se imaginava anteriormente. O risco relativo de acidente pode se estender muito além da marca de 2 mil horas antes de ter sua tendência reduzida para uma linha de base mais nivelada. (KNECHT, 2015, p. 8, tradução nossa).

Considerando-se o ambiente aeronáutico da Academia da Força Aérea, os registros administrativos extraídos dos Relatórios de Atividades Anuais (RAA) da AFA indicam uma média de horas de voo de 137:28h anuais por piloto nos últimos 5 anos bem como média de horas totais por piloto de 1361:30h. Desta forma, tomando por base o apresentado acima considera-se que a composição desta experiência dos pilotos da AFA ano a ano pode ter influência na incidência de uma ocorrência aeronáutica (BRASIL, 2021).

2.2 Construção de modelo matemático e previsão de valor de uma variável com base em outra

Um MM é um processo formal de estimativa de coeficientes, ou seja, relaciona-se uma variável resposta a uma ou mais variáveis explicativas (MADDALA, 1992). No caso em tela, as ocorrências aeronáuticas (por tipo) serão explicadas pela quantidade de horas de voo dos pilotos da aeronave T-27.

Sabe-se, porém, que nunca há relação exata entre duas variáveis, mas deve-se considerar uma relação funcional entre as mesmas, podendo-se descrevê-las matematicamente por meio de parâmetros, que indicarão o comportamento do

⁴ Em termos de domínio da aeronave

⁵ Coeficiente que varia entre 0 e 1 e expressa a proporção da variação de uma medida (variável resposta) que é explicada pela variação de outra (variável explicativa)

modelo, ou seja, indicarão em qual proporcionalidade uma mudança nas explicações afetam as respostas (MADDALA, 1992). Há que se levar em consideração, ainda, todos os fatores que afetam as respostas, mas que não estão nas explicações, chamados de erros (fatores de erro), pois são todos os fatores que não são possíveis de serem explicados a partir do modelo (STOCK; WATSON, 2015).

Os parâmetros devem ser estimados pelo Método dos Mínimos Quadrados Ordinários (MQO) à amostra de dados com o objetivo de minimizar a soma dos quadrados dos erros⁶, permitindo que os parâmetros sejam os que melhor descrevem a relação entre as variáveis (MCCULLAGH; NELDER, 1983). Ressalte-se que já existem softwares aptos ao cálculo da curva de tendência dos dados aplicando os MQO, como MATLAB, Python, Microsoft Excel, entre outros, de modo a facilitar a dispersão e análise de dados e, quando estipulados e descritos os parâmetros em quantidades numéricas, “devemos considerar as teorias (ou valores teóricos) como geradores de padrões numéricos, os quais, em alguns casos, podem substituir os dados.” (MCCULLAGH; NELDER, 1983, p. 4, tradução nossa).

Em suma, os dados observados permitirão desenhar um MM, o qual admitirá a formulação de valores teóricos, que permitem que a amostra coletada seja extrapolada e os padrões de comportamento sejam projetados fora do intervalo observado, caracterizando uma predição de comportamento. Porém, a extrapolação deve ser utilizada com cautela, uma vez que não se sabe o exato comportamento dos dados fora do intervalo especificado.

A composição de variáveis explicativas pode ser aumentada, contabilizando outros fatores, a fim de se obter mais fatores influentes na ocorrência. Este nível de complexidade de interação de variáveis pode ser amainado por meio de uma regressão múltipla de dados.

A regressão múltipla permite maior flexibilidade por ser mais receptiva ao princípio *ceteris paribus*⁷ e possibilita o controle claro de muitos outros fatores que afetam as variáveis dependentes ao mesmo tempo (WOOLDRIDGE, 2016). Sendo assim “o parâmetro de cada variável explicativa medirá a modificação média desta variável quando todas as outras variáveis explicativas permanecerem fixas.” (FREUND; WILSON; SA, 2006, p. 74, tradução nossa).

⁶ Diferença entre os valores da população e os valores estimados pelo modelo

⁷ Quando tudo além da variável considerada como preditora ou explicativa deve ser considerado constante

Um modelo de regressão múltipla pode ser descrito como a relação entre variáveis explicativas que resultam em uma resposta. Todas as variáveis explicativas que não forem observáveis, devem compor o modelo como um termo de erro. (MADDALA, 1992).

A estimação dos parâmetros deve ser feita também pelo MQO com o intuito de melhor descrever a relação entre as variáveis. A seleção cuidadosa de variáveis a se inserir na análise é primordial para manter a simplicidade do modelo, possibilitar aproximação gradual da realidade e permitir melhores previsões. Para tanto, deve-se atingir o conceito de parcimônia máxima, balizada pelo princípio da navalha de Occam (MCCULLAGH; NELDER, 1983).

Na parcimônia máxima, uma variável explicativa que não tenha efeito detectável na variável resposta deve ser excluída como fator preditor. Se a adição de uma variável explicativa diminui o erro e torna o modelo mais ajustado à amostra de dados, deve-se adicioná-la. Caso contrário, pelo princípio da navalha de Occam⁸, a parcimônia máxima é atingida e deve-se excluir esta variável da análise. O princípio permite melhor assertividade, pois impede a composição de parâmetros extras que sejam desnecessários ou que resultam em anomalias, como a colinearidade de dados⁹ entre outros. (MCCULLAGH; NELDER, 1983).

Compondo a tomada de variáveis em tela, poder-se-ia acrescentar à formulação dados relacionados com quantidade de treinamento teórico (carga-horária anual com instruções técnicas e relacionadas com segurança de voo), quantidade de instrutores do quadro de tripulantes no ano em tela, tipo de especialização operacional, entre outros, a fim de aproximar o modelo cada vez mais da realidade por meio do incremento de variáveis explicativas que aumentem o seu valor agregado como preditor de ocorrências.

3 CONSIDERAÇÕES FINAIS

Neste trabalho, comentou-se que em um sistema sociotécnico complexo como a aviação os elementos interagem de forma complexa, podendo resultar em acidentes aeronáuticos catastróficos. Ressaltou-se as iniciativas de órgãos internacionais no ramo da análise de dados e modelamento matemático para

⁸ Exigência de reconhecer, para cada objeto analisado, apenas uma explicação suficiente

⁹ Correlação entre duas variáveis explicativas que não computa influência real na variável resposta

predição de cenários e consequente interposição na política de prevenção, no entanto, tomando por base o PEMAER e a DIPLAN, o COMAER carece de iniciativas nesta área, transcendendo este comportamento às suas OM, como a AFA, sendo considerado para este ensaio mais especificamente o 1º EIA.

O trabalho apresentou como tese que os registros administrativos de experiência de voo e tipo de ocorrência aeronáutica contidos nos RAI do 1º EIA possibilitam a construção de MM preditor de ocorrências aeronáuticas e esses foram distinguidos como dados observacionais e de corte transversal.

Como argumento inicial, considerou-se que a experiência de voo tem correlação com ocorrências aeronáuticas, conforme atestado pelos estudos de Craig (2001) e Knecht (2015). Nestes estudos, apresenta-se uma faixa de horas de voo na qual pilotos estariam mais propensos a um acidente aeronáutico. A condição é reforçada pela quantidade média de horas de voo anuais e média de horas totais por piloto na AFA, as quais indicam proximidade da “zona da morte”.

O segundo argumento tratou que, havendo correlação entre as variáveis explicativa e resposta consideradas, seria possível a previsão de uma variável com base no valor de outra por meio de um MM, utilizando com cautela o recurso da extrapolação, podendo computar mais variáveis explicativas por meio de regressão múltipla, aproximando-se gradualmente da realidade, atentando à navalha de Occam e parcimônia máxima, com o fito de não compor o modelo com informações demasiadas e irrelevantes que ocasionariam anomalias.

Em finalização, este estudo estimula a consideração do benefício de se aplicar os conceitos em tela, com o fito de, por meio do modelamento matemático, poder prever comportamentos e riscos em ascensão no 1º EIA. Ademais, o desenvolvimento deste MM permitirá à FAB desenvolver linhas de prevenção de acidentes aeronáuticos não somente no âmbito da AFA, mas extrapolando as fronteiras desta instituição para os Esquadrões Operacionais. Assim, poder-se-á projetar as características mais proeminentes de uma ocorrência aeronáutica e determinar propensões em cenários específicos, bem como, baseando-se nesta análise de dados, adotar medidas e políticas convergentes à prevenção contra acidentes aeronáuticos, resguardando o erário brasileiro. Ademais, este é o primeiro passo para uma modelagem mais complexa e precisa, podendo, inclusive, o método ser utilizado em outras áreas como Segurança do Trabalho.

REFERÊNCIAS

BRASIL. Ministério da Defesa. Comando da Aeronáutica. Portaria AFA nº 31/SIPAA, de 19 de fevereiro de 2021. Aprova a edição do Programa que estabelece as atividades e subprogramas que deverão ser cumpridos pelos elos SIPAER na gestão de segurança de voo da AFA. **Boletim Ostensivo da AFA**, Pirassununga, n. 1, p. 5-37, 19 fev. 2021.

BRASIL. Ministério da Defesa. Comando da Aeronáutica. Portaria GABAER nº 1.707/GC3, de 26 de setembro de 2019. Aprova a reedição da Diretriz de Planejamento Institucional (DCA 11-118). **Boletim do Comando da Aeronáutica**, Rio de Janeiro, n. 222, p. 189-232, 20 dez. 2018. Disponível em: <https://www.fab.mil.br/Download/arquivos/pemaer.pdf>. Acesso em: 20 jun. 2022.

BRASIL. Ministério da Defesa. Comando da Aeronáutica. Portaria GABAER nº 2.102/GC3, de 18 de dezembro de 2018. Aprova a reedição do Plano Estratégico Militar da Aeronáutica (PCA 11-47). **Boletim do Comando da Aeronáutica**, Rio de Janeiro, n. 175, p. 17, 30 set. 2019. Disponível em: https://www.fab.mil.br/Download/arquivos/prestacaodecontas/DCA_11_118_Set2019_DIPLAN.pdf. Acesso em: 20 jun. 2022.

CHARPENTIER, A.; FLACHAIRE, E.; LY, A. Econometrics and machine learning. **Economie et Statistique**, Rennes, v. 505, n. 1, p. 147-169, maio 2018.

CRAIG, P.A. **The killing zone. How and why pilots die**. New York: McGraw-Hill, 2001.

ESTADOS UNIDOS. Federal Aviation Administration. FY 2022 AVS business plan. United States. **Department of Transportation**, Washington, 17 fev. 2022. Disponível em: https://www.faa.gov/sites/faa.gov/files/2022-02/AVS_FY22_Business_Planv2.pdf. Acesso em: 20 jun. 2022.

UNIÃO EUROPÉIA. European Union Aviation Safety Agency. Concept Paper: First usable guidance for level 1 machine learning applications. **European Union**, Köln, 20 dez. 2021. Disponível em: <https://www.easa.europa.eu/downloads/134357/en>. Acesso em: 20 jun. 2022.

FREUND, R. N.; WILSON, W. J.; SA, P. **Regression analysis, statistical modeling of a response variable**. Burlington: Elsevier, 2006.

KNECHT, W. R. et al. Predicting accident rates from general aviation pilot total flight hours. United States. **Department of Transportation**. Federal Aviation Administration. Office of Aviation. Civil Aerospace Medical Institute, Washington, 2015.

LEVESON, N. G. **Engineering a safer world: systems thinking applied to safety**, Boston, Massachusetts: The MIT Press, 2011.

MADDALA, G. S. **Introduction to econometrics**. 2. ed. New York: MacMillan Publishing Company, 1992.

MCCULLAGH, P.; NELDER, J. A. **Generalized Linear Models**. 2. ed. London: Chapman and Hall, 1983.

PERROW, C. **Complex Organizations**, 3 ed. New York: Newbery Award Records Inc., 1986.

STOCK, J. H.; WATSON, M. W. **Introduction to econometrics**. 3 ed. New Jersey: Pearson Education, 2015.

WOOLDRIDGE, J. M. **Econometric analysis of Cross Section and Panel Data**, Boston, Massachusetts: The MIT Press, 2001.