

Universidade de Brasília - UnB
Faculdade UnB Gama - FGA
Engenharia Aeroespacial

Modelagem do Problema da Alocação de Frota Aplicado à Aviação Regional A Partir da Redução da Complexidade Computacional

Autora: Julia Ribeiro Felipe
Orientador: Dr. Victor Rafael Rezende Celestino
Coorientador: Dr. Mateus Rodrigues Miranda

Brasília, DF
2022



Julia Ribeiro Felipe

**Modelagem do Problema da Alocação de Frota Aplicado
à Aviação Regional A Partir da Redução da
Complexidade Computacional**

Monografia submetida ao curso de graduação em Engenharia Aeroespacial da Universidade de Brasília, como requisito parcial para obtenção do Título de Bacharel em Engenharia Aeroespacial.

Universidade de Brasília - UnB

Faculdade UnB Gama - FGA

Orientador: Dr. Victor Rafael Rezende Celestino

Coorientador: Dr. Mateus Rodrigues Miranda

Brasília, DF

2022

Julia Ribeiro Felipe

Modelagem do Problema da Alocação de Frota Aplicado à Aviação Regional
A Partir da Redução da Complexidade Computacional/ Julia Ribeiro Felipe. –
Brasília, DF, 2022-

102 p. : il. (algumas color.) ; 30 cm.

Orientador: Dr. Victor Rafael Rezende Celestino

Trabalho de Conclusão de Curso – Universidade de Brasília - UnB
Faculdade UnB Gama - FGA , 2022.

1. Aviação Regional. 2. Problema da Alocação de Frota. 3. Pesquisa Operacional. 4. Python. 5. PuLP. 6. Intervalo de Otimalidade. I. Dr. Victor Rafael Resende Celestino, Dr..Matheus Rodrigues Miranda. II. Universidade de Brasília. III. Faculdade UnB Gama. IV. Modelagem do Problema da Alocação de Frota Aplicad à Aviação Regional A Partir da Redução da Complexidade Computacional.

CDU 02:141:005.6

Julia Ribeiro Felipe

Modelagem do Problema da Alocação de Frota Aplicado à Aviação Regional A Partir da Redução da Complexidade Computacional

Monografia submetida ao curso de graduação em Engenharia Aeroespacial da Universidade de Brasília, como requisito parcial para obtenção do Título de Bacharel em Engenharia Aeroespacial.

Trabalho aprovado. Brasília, DF, 30 de setembro de 2022:

Dr. Victor Rafael Rezende Celestino
Orientador

Dr. Mateus Rodrigues Miranda
Coorientador

**Dra. Poliana Candida Oliveira
Martins**
Convidado 1

**Dr. Manuel Nascimento Dias Barcelos
Júnior**
Convidado 2

Brasília, DF
2022

*Este trabalho é dedicado aos meus pais,
aos sonhos que realizamos
e aos que ainda vamos realizar.*

Agradecimentos

Agradeço, primeiramente, aos meus pais, Júlio César e Marta Helena, pelos ensinamentos ministrados e pelo investimento em minha formação pessoal e profissional. Gostaria de demonstrar minha gratidão a minha tia Solead por ter cedido sua casa durante estes anos graduação e por acompanhar de perto meus primeiros passos na vida adulta. Agradeço também aos tios Woxthon Luís e Elizabeth pelos investimentos em minha educação e infinitos conselhos de vida.

Estendo os agradecimentos, também, aos professores e orientadores Dr. Victor Celestino e Dr. Mateus Miranda pelos ensinamentos prestados e também pelas oportunidades de crescimento profissional ofertadas. Tenho os senhores como exemplos de engenheiros e de pessoas para minha vida.

Demonstro minha gratidão e respeito pelas instituições e profissionais da indústria de transporte aéreo consultados durante a confecção deste trabalho. Sou grata também a todos os professores da Universidade de Brasília (UnB) que contribuíram com a minha formação acadêmica, que contribuem com o desenvolvimento tecnológico do Brasil e driblam diariamente a falta de investimento na pesquisa brasileira.

Deixo também um agradecimento especial aos colegas Tiago Jesus e Gustavo Pereira que iniciaram a pesquisa, desenvolveram importantes trabalhos e mostraram-se completamente disponíveis para contribuir da melhor forma possível. Por fim, gostaria de celebrar a as grandes amizades que cultivei até o momento, fiz grandes amigos que estiveram comigo nas conquistas e dificuldades do ensino superior.

*"O sucesso canta mais doce
Na boca do fracassado.
A compreensão de tal néctar
Requer a dor do necessitado.[...];"*

Emily Dickinson (Trad. Alexei Gonçalves de Oliveira)

Resumo

O presente trabalho tem por objetivo a elaboração de uma malha aérea ótima aplicada a aviação regional brasileira a partir da resolução do Problema de Alocação de Frota (FAP) por meio dos recursos da Pesquisa Operacional (PO). Foram utilizados na pesquisa 6 tipos diferentes de aeronaves para 710 rotas que passam por 12 *hubs* no território brasileiro. O projeto foi estruturado a partir de um modelo matemático pre-existente desenvolvido por outros pesquisadores, tendo como principal abordagem a redução da complexidade computacional do algoritmo igualmente desenvolvido em trabalhos anteriores na linguagem Python com o auxílio das bibliotecas Numpy e Python e do *framework* Pulp. A estratégia utilizada para resolver as questões computacionais foi a inserção de um intervalo de otimalidade na solução do problema de Pesquisa Operacional, o que permitiu uma análise completa do modelo previamente elaborado. Além disso, o modelo foi aprimorado com alterações no indicador CASK (*Cost per Available Set Kilometer*) e inclusão de novas restrições no modelo de PO, a fim de aproximar a alocação de frota à realidade da indústria brasileira.

Palavras-chaves: Aviação Regional, Problema da Alocação de Frota, Pesquisa Operacional, Python, Pulp, Intervalo de Otimalidade.

Abstract

The paper's goal is to create an optimal air network applied to Brazilian regional aviation from the resolution of the Fleet Allocation Problem (FAP) using the resources of Operational Research (OR). Six different types of aircraft were used to be allocated in 710 routes that pass through 12 *hubs* in the Brazilian territory. A pre-existing mathematical model developed by other researchers was used in the project, with the main strategy being the reduction of computational complexity of the algorithm also developed in previous works in the Python language with the help of Numpy and Pandas libraries and the Pulp *framework*. To solve the computational issues the insertion of an optimality gap was used in the solution of the Operational Research problem, which allowed a complete analysis of the previously model. Besides, the model was enhanced by changes of CASK (*Cost per Available Set Kilometer*) indicator and inclusion of new restrictions in the PO problem, in order to bring the fleet allocation closer to the reality of the Brazilian industry.

Key-words: Regional Aviation, Fleet Assessment Problem, Operations Research, Python, Pulp, optimality gap.

Lista de ilustrações

Figura 1 – Representação figurativa de uma malha aérea do tipo <i>hub and spoke</i> . Fonte: adaptado de Lederer (1998)	32
Figura 2 – Análise do processo de Pesquisa Operacional (PO). Fonte: adaptado de Rodrigues (2013)	33
Figura 3 – Projeto de artefatos a serem desenvolvidos.	42
Figura 4 – Metodologia do <i>Design Science Research</i> aplicada ao projeto.	44
Figura 5 – Etapas do desenvolvimento dos artefatos.	45
Figura 6 – Mapa Mental usado para a Criação do Código. Fonte: (PEREIRA, 2021)	49
Figura 7 – Variação da curva de CASK de acordo com as diferentes aeronaves antes da alteração nos valores de CASK.	66
Figura 8 – Variação da curva de CASK de acordo com as diferentes aeronaves após a alteração nos valores de CASK.	67
Figura 9 – Variação de lucro de acordo com as diferentes aeronaves	71

Lista de tabelas

Tabela 1 – Comparativo entre pesquisas	40
Tabela 2 – Demanda e Oferta do transporte aéreo. Adaptado de ANAC, 2021	41
Tabela 3 – Correspondências de símbolos utilizados no modelo	47
Tabela 4 – Especificações da máquina virtual utilizada para simulações	50
Tabela 5 – Detalhamento do cenário reduzido.	55
Tabela 6 – Variação da tarifa de acordo com a distância entre um <i>hub</i> h e um <i>spoke</i> j	55
Tabela 7 – Variação de <i>CASK</i> de acordo com a rota e a aeronave alocada.	56
Tabela 8 – Atribuição de <i>Load factor</i> de acordo com cada rota.	56
Tabela 9 – Variação de lucro de acordo com a mudança no número de vezes que uma determinada aeronave foi alocada.	62
Tabela 10 – Resultado do cenário teste utilizado o código em Python.	62
Tabela 11 – Resultados do cenário teste, $gap = 0.01$	63
Tabela 12 – Resultados da variação do valor de intervalo de otimalidade.	64
Tabela 13 – Variação de aeronaves alocadas.	64
Tabela 14 – Resumo dos resultados obtidos a partir das simulações definitivas.	69
Tabela 15 – Participação de cada aeronave na alocação de passageiros e nos lucros.	69

Lista de abreviaturas e siglas

ANAC	Agência Nacional de Aviação Civil
ASK	Available Seat Kilometers
BB	<i>Branch-and-bound</i>
eVTOL	<i>Electric Vertical Take-off and Landing</i>
CASK	<i>Cost per Available Seat Kilometer</i>
FAP	<i>Flight Assessment Problem</i>
LF	<i>Load Factor</i>
LP	<i>Linear Programming</i>
Matriz OD	Matriz Origem-Destino
MILP	Mixed Integer Linear Programming
PI	Programação Inteira
PIB	Programação Inteira Binária
PL	Programação Linear
PO	Pesquisa Operacional
RPK	Revenue Passenger Kilometers
TCC 1	Trabalho de Conclusão de Curso 1
TCC 2	Trabalho de Conclusão de Curso 2

Lista de símbolos

α	Coeficiente linear
β	Coeficiente angular
v	Diferença entre valores sugeridos
Y	Variável dependente quantitativa
X	Variável explícita

Sumário

1	INTRODUÇÃO	25
1.1	Contextualização	25
1.2	Problemática	26
1.3	Hipótese	26
1.4	Objetivos	26
1.4.0.1	Objetivo Geral	26
1.4.1	Objetivos Específicos	27
1.5	Justificativa	27
1.6	Organização do trabalho	27
2	REFERENCIAL TEÓRICO	29
2.1	Aviação regional	29
2.2	Indicadores da Aviação	30
2.3	Malhas aéreas	31
2.4	Aeronaves na aviação regional	32
2.5	Pesquisa Operacional	32
2.6	Programação inteira	34
2.7	Complexidade Computacional na Programação Inteira	35
2.8	Problema da Alocação de Frota - FAP	36
2.9	Métodos de regressão linear	37
3	METODOLOGIA	39
3.1	<i>Design Science Research</i>	39
3.1.1	Etapas da metodologia proposta	40
3.2	Desenvolvimento do Modelo	43
3.3	Modelo Matemático	46
3.4	Algoritmo pré existente	48
3.5	Alterações no algoritmo pré-existente	48
3.5.1	O Parâmetro K e a Complexidade Computacional	48
3.5.2	Adaptações para o cálculo de CASK	51
3.5.3	Inserção do fator de ocupação (<i>load factor</i>)	51
3.5.4	Algoritmo atualizado	52
4	RESULTADOS E DISCUSSÃO	55
4.1	Resultados preliminares e verificação do modelo	55
4.2	Impactos do intervalo de otimalidade	63

4.3	Alterações no modelo matemático	65
4.4	Resultados Definitivos	68
5	CONCLUSÃO	73
	REFERÊNCIAS	75
	APÊNDICE A – CÓDIGO DO MODELO MATEMÁTICO ATUA- LIZADO	77
	ANEXOS	85
	ANEXO A – CÁLCULO DE DISTÂNCIAS <i>HUB - HUB</i>	87
	ANEXO B – CÓDIGO DO MODELO MATEMÁTICO PRE-EXISTENTE	89

1 Introdução

1.1 Contextualização

A existência de um país com as características geográficas e socioeconômicas como o Brasil, implicam em uma desigualdade considerável entre as cinco regiões brasileiras e tais diferenças podem ser refletidas em vários setores da indústria, inclusive na indústria de transporte aéreo. Sendo assim, faz-se necessário entender melhor como a aviação civil pode ser um fator de impacto para o cenário econômico e cultural de uma determinada região ao mesmo tempo que pode ser igualmente impactada pela própria economia brasileira.

Trazendo essa ideia para o campo da aviação civil, é esperado que haja uma concentração de rotas operadas por poucas companhias nos grandes centros urbanos. Entretanto, encontra-se na aviação regional a possibilidade de fomentação de atividade de turismo e negócios em regiões interioranas, podendo representar uma contribuição no desenvolvimento do país (BETTINI, 2007).

A fim de um esclarecimento da proposta, define-se como aviação regional atividades regulares em rotas com uma menor densidade de tráfego e que façam o uso de aeronaves que suportem até cem passageiros (BETTINI, 2007). Ao mesmo tempo que é sabido os possíveis impactos positivos do investimento da aviação regional no país, este ainda é um setor pouco explorado por fatores como a falta de infraestrutura de aeroportos e a elevação no preço dos combustíveis. Além disso, também existe um oligopólio na indústria de transporte aéreo, visto que apenas três companhias aéreas ocupam 99,7% de toda a faixa de mercado brasileiro (ANAC, 2022a).

Um dos fatores bastante influentes para a condição atual da aviação no país é o impacto da pandemia causada pelo vírus da COVID-19. Desde o ano de 2020 o mercado de transporte aéreo tem passado por uma crise devido ao cancelamento de operações aéreas, interrupção do turismo e até mesmo a mudança nos hábitos de trabalho, como a popularização forçada do *home office*, por exemplo. De acordo com a ANAC (2022a), a indústria ainda não conseguiu recuperar-se completamente dos impactos da pandemia visto que, comparado com os valores registrados no ano de 2019, o número total de passageiros de transporte aéreo ao longo de 2021 apresenta um recuo de 43,5%.

Acontecimento mais recente, porém também importante a ser pontuado para uma análise mais completa do panorama do setor da aviação brasileiro é a condição geopolítica global. Entende-se que um dos desdobramentos da guerra na Ucrânia seja o aumento considerável nos preços do barril de petróleo, fato que contribuiu para a redução de voos por partes de empresas da indústria de transporte aéreo brasileira (AEROIN, 2022).

Por fim, acredita-se que o investimento na aviação regional possa ser uma alternativa para a recuperação do setor aeronáutico como um todo, bem como um auxílio no desenvolvimento de regiões interioranas brasileiras. Entretanto, para que o resultado seja de fato satisfatório é imprescindível que o planejamento estratégico de malha aérea seja feito de forma otimizada a fim de reduzir despesas e maximizar lucros.

1.2 Problemática

A partir das análises registradas na seção 1.1, percebe-se um potencial de expansão do setor por meio do investimento na aviação regional do país. Seguindo essa ideia, verifica-se que apesar da existência de trabalhos voltados para o desenvolvimento de malhas aéreas regionais ótimas, o modelo em questão ainda não foi completamente selecionado, principalmente devido a questões computacionais. Partindo da ideia de necessidade de aprimoramento dos modelos já trabalhados, indaga-se: como a redução da complexidade computacional pode influenciar na modelagem de uma malha aérea exclusiva para rotas regionais brasileiras, considerando a alocação de diferentes rotas e aeronaves de forma simultânea?

1.3 Hipótese

A hipótese de pesquisa está pautada no objetivo de aprimoramento do modelo matemático desenvolvido por Pereira (2021). Aponta-se que voltando-se para a redução da complexidade computacional do modelo desenvolvido anteriormente seria possível encontrar soluções ótimas para diferentes cenários para alocação de recursos utilizando a Pesquisa Operacional aplicados ao Problema de Alocação de Frota, do inglês *Fleet Assignment Problem* (FAP), além da possibilidade de verificação de adequações no modelo matemático para aproximar o cenário elaborado da realidade do setor do transporte aéreo brasileiro.

1.4 Objetivos

1.4.0.1 Objetivo Geral

Caracteriza-se como objetivo do trabalho modelar uma malha aérea ótima em termos de escolha e alocação de diferentes aeronaves para a realidade da aviação regional brasileira, tendo um foco na resolução da problemática a partir da redução da complexidade computacional de modelos já existentes.

1.4.1 Objetivos Específicos

Visando o avanço da pesquisa, objetiva-se:

- Revisar trabalhos anteriores e propor alterações na configuração da malha aérea;
- Identificar pontos a serem aprimorados no modelo matemático apontado como solução da problemática em questão por pesquisas anteriores;
- Definir a operação de uma malha aérea otimizada aplicando pesquisa operacional, levando em consideração as melhorias apontadas e limitações computacionais existentes;

1.5 Justificativa

A opção pela realização do trabalho justifica-se a partir da importância do impulsionamento da aviação regional para o desenvolvimento do setor aéreo no país. Acredita-se na aviação como uma forma de fomentar atividades econômicas fora dos grandes centros urbanos brasileiros, onde há a maior concentração de renda do país (BETTINI, 2007). O desenvolvimento de um projeto voltado para as análises de possibilidades de rotas aéreas em áreas pouco exploradas pode contribuir para o esclarecimento de medidas a serem tomadas por parte das iniciativas públicas - principalmente - e privadas.

Por fim, optou-se pela utilização do Problema da Alocação de Frota (FAP), pois aliado aos dados das bases de órgãos públicos brasileiros, acredita-se na possibilidade de desenvolvimento de uma malha aérea condizente com a realidade brasileira. A partir de trabalhos desenvolvidos na temática anteriormente por Jesus (2019) e por Pereira (2021), enxerga-se lacunas a serem endereçadas no modelo matemático utilizado até o momento. Assim sendo, a confecção do presente trabalho está atrelada a necessidade de atualização do modelo considerando as limitações computacionais encontradas além de considerar as mudanças no setor que ocorreram desde o início das pesquisas no tema.

1.6 Organização do trabalho

No primeiro capítulo do trabalho encontra-se a introdução do projeto, estruturada em seis tópicos, sendo eles: contextualização do cenário envolvendo a temática a ser trabalhada, problemática a ser estudada, hipótese levantada, objetivos do trabalho, justificativa das escolhas de tema e abordagens e por fim a organização do texto.

Em seguida, o segundo capítulo é dedicado ao referencial teórico, em que são feitas breves explicações conceituais acadêmicas que serviram de base para o desenvolvimento do modelo proposto. São apresentados os tópicos: aviação regional, indicadores da aviação,

malhas aéreas, pesquisa operacional, programação inteira, complexidade computacional, Problema da Alocação de Frota (FAP) e métodos de regressão linear.

O capítulo três é dedicado para a descrição da metodologia *Design Science Research* e como ela se aplica à hipótese proposta. Ao longo do texto, além da explicação conceitual da abordagem e como cada etapa da metodologia foi desenvolvida. Também são descritos os passos para seleção de aeronaves e o método para aprimoramento do modelo.

O quarto capítulo será dedicado aos resultados finais obtidos, ou seja, a metodologia é detalhada juntamente com o processo todo do projeto e os resultados obtidos na primeira parte do trabalho de conclusão. Por fim, o quinto e último capítulo é dedicado para as considerações finais do trabalho. Nesse tópico é feita uma revisão dos resultados em comparativo com os objetivos iniciais, bem como uma sugestão de próximos pontos que podem ser aprimorados na pesquisa.

2 Referencial Teórico

Ao decorrer do capítulo, encontram-se registrados os resultados da revisão narrativa da literatura que foi realizada para a confecção da pesquisa. O objetivo principal é buscar entendimento sobre as ferramentas que foram utilizadas para a confecção do modelo anterior bem como foram considerados temas relevantes: aviação regional, indicadores da aviação, malhas aéreas, aeronaves na aviação regional, pesquisa operacional, programação inteira, complexidade computacional na programação inteira, problema da alocação da frota (FAP) e métodos de regressão linear.

2.1 Aviação regional

Assim como explicitado no capítulo anterior, a aviação regional busca a conexão de cidades de pequeno e médio porte com grandes centros urbanos do país. Entende-se que existem diversos fatores que podem influenciar as diretrizes de companhias aéreas nas tomadas de decisões relacionadas à alocação de recursos para operação. Sendo assim, até mesmo a definição da aviação regional depende bastante da funcionalidade do projeto, uma vez que não há um conceito fechado sobre a modalidade (BETTINI, 2007).

Apesar de não existir uma única definição da modalidade, em sua maioria, as rotas regionais são representadas por voos com distâncias e passageiros reduzidos. Partindo das características citadas, entende-se, em um primeiro momento, que essa modalidade pode ser lida como pouco competitiva quando comparada aos demais modelos de aviação, principalmente devido aos custos operacionais, desvantagens essas que podem ser verificadas na participação da aviação regional na malha aérea geral do Brasil.

Entretanto, o investimento na conexão entre cidades que não fazem parte dos principais centros urbanos já está no planejamento do governo brasileiro. De acordo com [Infraestrutura \(2022\)](#), o governo brasileiro tem a intenção de intensificar a presença do modal aéreo no interior do país. O objetivo é que até o ano de 2025, exista estrutura aeroportuária em 200 cidades do Brasil, um aumento de 37% comparado ao ano de 2019.

No entanto, outro fator a ser considerado para a construção de uma malha aérea regional é a demanda de transporte entre determinadas regiões. Ou seja, a partir de um planejamento estratégico correto que considere as melhores rotas e aeronaves apropriadas, a aviação regional deixa de ser uma desvantagem e passa a ser uma opção viável economicamente.

Além disso, também há de ser considerado o impacto social e econômico oriundo da implantação de rotas aéreas regionais em cidades de pequeno e médio porte em setores

como o turismo e até mesmo em modalidades de trabalho que requerem deslocamentos esporádicos de colaboradores (BETTINI, 2007). Sendo assim, o investimento na aviação regional é sustentável, desde que haja demanda para suprir os custos operacionais, podendo ser algo que reflete positivamente nas regiões abordadas.

2.2 Indicadores da Aviação

Objetivando o desenvolvimento voltado para a alocação de frotas no setor aeronáutico, foi realizada uma seleção de indicadores da aviação que nortearam o trabalho. Essas métricas estão, em sua maioria, ligadas aos custos operacionais de uma companhia aérea e, por isso, estão bastante relacionadas ao planejamento estratégico e operacional de qualquer empresa aérea. No caso do trabalho em questão, os indicadores apontados foram essenciais para o entendimento do modelo pré-existente e para o apontamento de possíveis mudanças necessárias.

O primeiro indicador importante revisado está relacionado à demanda de passageiros pagos transportados, RPK (do inglês *Revenue Passenger Kilometers*). De acordo com Bazargan (2010), este cálculo se dá por meio da multiplicação entre o número de quilômetros voados e o número de passageiros pagantes. Outro parâmetro muito importante na aviação é a oferta de assentos disponíveis para serem utilizados pela empresa, ASK (do inglês *Available Seat Kilometer*) (ANAC, 2022a). Ambos indicadores podem ser utilizados tanto em cenários empresariais quanto no cenário do setor aéreo brasileiro no geral, ou seja, existe tanto o ASK aplicado a uma companhia aérea específica e também é possível calcular a quantidade de assentos disponíveis no mercado aeronáutico brasileiro.

Outros indicadores que foram cruciais para o entendimento do modelo são, o CASK (do inglês *Cost per Available Seat Kilometer*), que é o custo médio de se voar um assento em um quilômetro. De acordo com Fregnani, Ferreira e Griebeler (2009), o cálculo de custos operacionais de uma companhia aérea pode ser realizado utilizando-se de uma função do tipo Cobb-Douglas, que torna possível a relação entre custos dos insumos necessários com o custos de fatores diversos que podem ou não serem considerados. Esta ferramenta é utilizada em cenário microeconômico frequentemente e é capaz de quantificar a relação entre entradas diversas e saídas previamente selecionadas (FREGNANI et al., 2019).

No trabalho em questão optou-se por utilizar-se da formulação desenvolvida por Pereira (2021), visto que esta é a determinação do modelo pré-existente de malha aérea. Nesta formulação, a definição do CASK partiu de uma função Cobb-Douglas, Equação 2.1 e para a definição de todos os coeficientes, utilizou-se de regressões lineares de dados disponíveis pela base de dados Hórus, na matriz O&D (Horus, 2022), um conjunto de demandas de transporte (origem e destino) coletadas a partir de dados de telefonia móvel

(SILVA et al., 2021). Assim, foi possível alcançar a Equação 2.2 para calcular o CASK.

$$C = K_c(S^{\beta_1} \cdot D^{\beta_2}) \quad (2.1)$$

$$\exp^C = \exp^{1,72891238314985-0,34670847909125\ln(S)-0,404534157\ln(D)} \quad (2.2)$$

Em que k_c , β_1 e β_2 são coeficientes determinados por meio da matriz O&D (Horus, 2022), S é a quantidade de assentos em cada aeronave, D representa a distância voada e C é o CASK.

Por fim, existe o indicador de *Yield*, igualmente importante para o modelo desenvolvido, representa a receita da companhia por assento vendido a cada quilômetro voado, ou seja, é um parâmetro que recebe influência da quantidade de assentos disponíveis, da distância voada e do valor arrecadado. A definição do *Yield* foi realizado a partir do conceito de razão entre a tarifa média e a distância aérea do trecho (PEREIRA, 2021).

2.3 Malhas aéreas

As configurações de operação aérea de cada companhia variam bastante, logo, existem vários tipos de malhas aéreas que se adaptam de acordo com as características e estratégias de cada empresa, aeroporto e/ou da própria localidade. Dentre os tipos existentes, encontra-se a malha direta ou *point-to-point*, nessa organização as cidades são conectadas diretamente entre si, ou seja, não há nenhuma interceptação intermediária entre origem e destino de passageiros. Essa forma de serviço pode ser bastante positiva para companhias, em casos de rotas com menores distâncias (LEDERER, 1998).

Existem também as malhas aéreas baseadas em circuitos, em que cada aeronave é responsável por voar todas as rotas sequencialmente. Essa modalidade pode ser utilizada em conjunto com princípios de outros tipos de malha, havendo possibilidades de escalas intermediárias para saída de passageiros.

Por fim, existem também as malhas aéreas do tipo *hub-and-spoke*. Nessa modalidade há a determinação de um determinado local de referência, os chamados *hubs*. Esses centros de conexão são usados como ponto base da operação de uma determinada companhia aérea, assim sendo, os passageiros viajam de variados pontos de origem para algum *hub* em comum e, então, deste mesmo ponto, são deslocados novamente, desta vez para seus locais de destino.

Além de funcionar como um ponto comum para as conexões de voo, os *hubs* também podem ser utilizados como base geral da companhia, incluindo armazenamento de aeronaves ou demais instalações relacionadas à operação. A escolha por uma malha aérea

do tipo *hub and spoke* pode trazer algumas vantagens econômicas e operacionais, quando aplicada em um sistema que funcione de forma otimizada.

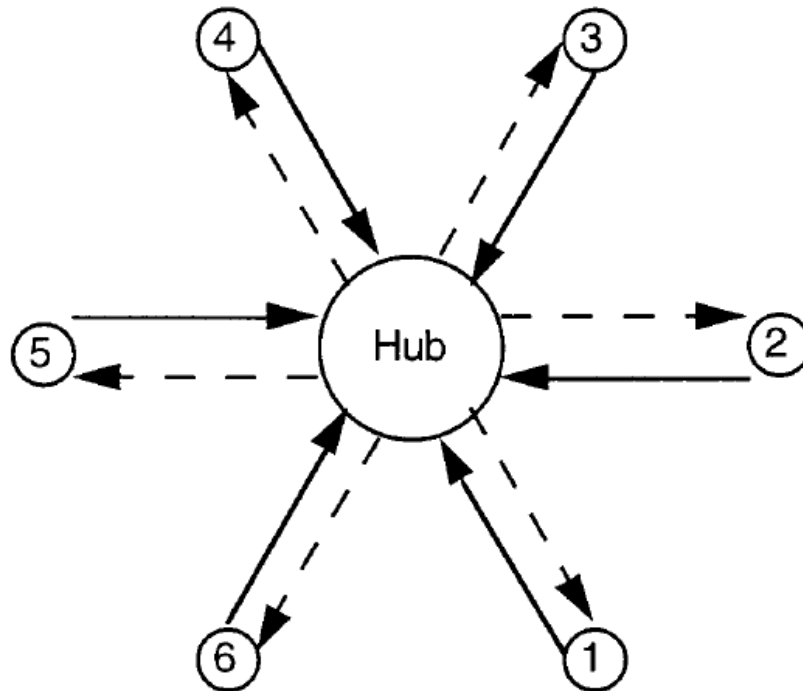


Figura 1 – Representação figurativa de uma malha aérea do tipo *hub and spoke*. Fonte: adaptado de Lederer (1998)

2.4 Aeronaves na aviação regional

Assim como citado na seção 2.1, não existe um consenso universal para o conceito de aviação regional. Portanto, para o desenvolvimento do trabalho, optou-se por seguir a definição de rotas regionais de Pereira (2021):

- Distância entre origem e destino a partir de 250 km, representando deslocamentos de pelo menos 2 horas em modais rodoviários;
- Rotas com demanda semanal de, pelo menos, 27 passageiros, que representa três vezes o número máximo de passageiros em uma aeronave Cessna 208 Caravan.

2.5 Pesquisa Operacional

A Pesquisa Operacional é pautada no uso da modelagem matemática, análise de recursos humanos, estatísticas e métodos computacionais variados com a intenção de aprimoramento na alocação de recursos de uma determinada produção. Uma análise de PO é baseada no levantamento de possibilidades, detecção de restrições do problema e

definição dos objetivos principais da tratativa. Ou seja, esta ciência está intrinsecamente ligada à tomada de decisões para a construção de um modelo produtivo ótimo, ou o mais próximo disto o possível. Podem ser consideradas etapas da modelagem de processos: identificação do problema, formulação de objetivos e análises de limitações e alternativas (RODRIGUES, 2013).

Assim, cabe aos profissionais encarregados pela análise de PO, inicialmente, a identificação da problemática a ser abordada e a identificação dos objetivos a serem alcançados. Após esta definição, faz-se necessário a elaboração de um modelo capaz de representar a realidade do sistema analisado. A partir do modelo preliminar, é possível traçar as primeiras propostas de aprimoramento e para este apontamento, podem ser utilizadas algumas ferramentas para definir as soluções mais adequadas, que serão posteriormente testadas e, caso sejam comprovadamente eficazes, aplicadas à produção. Assim como mostrado no fluxograma abaixo.

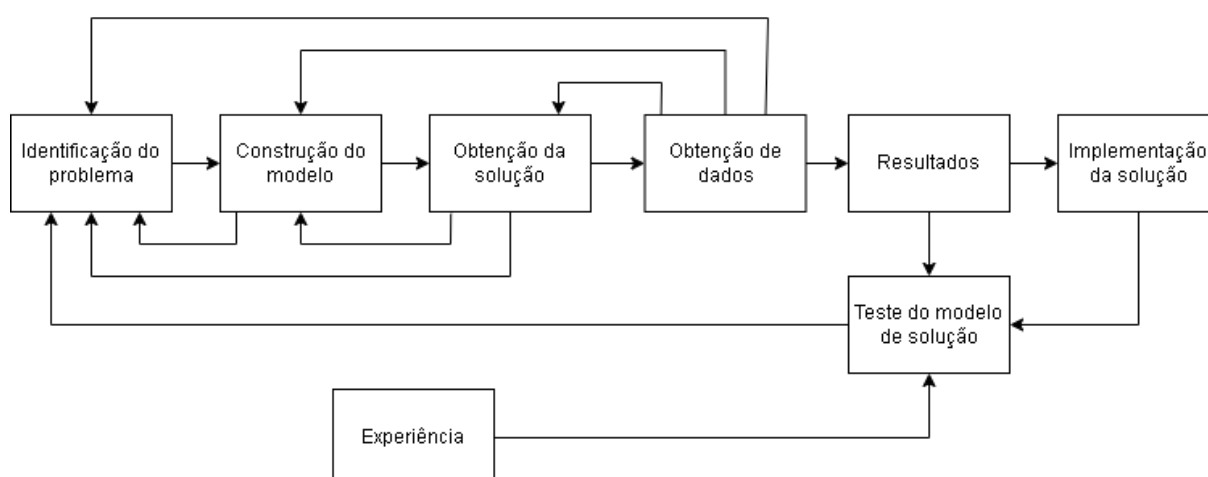


Figura 2 – Análise do processo de Pesquisa Operacional (PO). Fonte: adaptado de Rodrigues (2013)

De acordo com Taha (2008), não existe uma única técnica para solucionar todos os modelos matemáticos existentes e o método escolhido depende da complexidade do sistema. Dentre as principais ferramentas existentes encontram-se: a programação linear, programação inteira, programação dinâmica, otimização de redes, programação não linear, entre outras. Grande parte das técnicas existentes utilizam algoritmos baseados na repetição de regras de cálculos até o alcance da solução ótima. Entretanto, devido a complexidade dos modelos, algumas vezes não é possível resolver a otimização completa, em casos assim, a alternativa seria a busca pela redução da complexidade ou então o uso de heurísticas para o aprimoramento do modelo.

Dentre as principais técnicas de modelagem matemática utilizadas na Pesquisa Operacional, encontra-se a programação linear. Neste caso, o termo “programação” não

está atrelado a questões computacionais, mas sim à ao planejamento sequencial de atividades. Ou seja, pode-se caracterizar a programação linear como um planejamento baseado em funções matemáticas lineares e contínuas com o objetivo de promover uma produção que evite desperdícios e alcance os resultados desejados (LIEBERMAN, 2013).

Um dos modelos matemáticos mais importantes da programação linear é o chamado método simplex, uma técnica intensamente utilizada em problemas lineares de otimização. Desenvolvido por George B. Dantzig no ano de 1947, a ferramenta utiliza-se de uma solução básica inicial para encontrar a solução ótima para o problema, sendo geradas soluções intermediárias (RODRIGUES, 2013).

Além disso, tem-se que a obtenção de uma solução ótima para qualquer sistema está baseada na adequação às restrições impostas pelo próprio processo. Dessa forma, o uso do método simplex passa pela definição de equações de restrição que serão representativas da formação de uma modelagem geométrica dos limites do sistema. A partir dessa modelagem é possível encontrar soluções básicas que serão aprimoradas até o alcance da solução ótima (LIEBERMAN, 2013).

Entretanto, a programação linear não é a única técnica existente para a resolução de problemas utilizando pesquisa operacional. Além do planejamento linear, que baseia-se no uso de funções lineares para a resolução de tarefas, existem também: a programação dinâmica, pautada na divisão de subproblemas para a busca da solução, a otimização de rede e, por fim, a programação inteira.

2.6 Programação inteira

Assim como referenciado no nome da técnica, a programação inteira é um tipo de abordagem que utiliza-se igualmente de funções lineares que, porém, possuem pelo menos uma variável de restrição inteira (LACHTERMACHER, 2016). Na prática, o uso dessa técnica para resolução de problemas de PO é bastante comum, destaca-se a eficiência do algoritmo *branch-and-bound* (BB), que é um ponto de referência no mercado comercial. Entretanto, apesar de sua relevância na aplicação de PO, algoritmos de programação inteira podem apresentar problemas em sua implantação computacional (RODRIGUES, 2013).

Além dos casos em que a programação inteira é útil para a resolução de problemas de otimização de recursos, também existe o uso de modelos de PI para análises de decisão em que as respostas possíveis seriam "sim" ou "não", a chamada programação inteira binária (PIB) (LIEBERMAN, 2013). A partir do entendimento da natureza das variáveis existentes (genérico ou binário), já é possível elaborar uma metodologia de solução para o problema. Nos casos em que a programação inteira é utilizada, diferentemente da programação linear, as variáveis consideradas no problema são mais importantes que o número

de restrições propriamente ditas, para a busca e definição da solução ótima.

Partindo da concepção de que o suporte computacional é praticamente mandatório à medida que o número de variáveis consideradas aumenta, é muito importante saber definir bem o problema a ser resolvido, as variáveis e restrições. Estes são fatores influentes para a determinação da complexidade computacional de um determinado modelo. Sabe-se que parte considerável dos algoritmos de PI utilizam-se de uma correspondência entre o problema de PI e um problema de programação linear correspondente, o chamado relaxamento PL. Neste caso, o mesmo problema é aplicado ao modelo, com exceção da restrição inteira (LIEBERMAN, 2013).

2.7 Complexidade Computacional na Programação Inteira

A elaboração de um modelo que possua um nível de complexidade computacional compatível com as máquinas envolvidas é um fator bastante importante para o projeto e para a construção da metodologia em si, visto que a capacidade de memória dos computadores utilizados para a resolução do algoritmo pode influenciar no resultado final, uma vez que um modelo muito complexo pode refletir na impossibilidade de convergência de um algoritmo.

Pensando na complexidade computacional aplicada ao tipo de problema, de acordo com Lieberman (2013), apesar dos problemas de PI terem um número finito de soluções viáveis, não necessariamente são modelos fáceis de serem resolvidos computacionalmente todas as vezes. Um exemplo citado pelo autor para ilustrar a questão é o caso de problemas de PIB que possuam n variáveis, nesses casos existirão 2^n possíveis soluções. Portanto, mesmo sendo um dos modelos dos mais simples possíveis, cada vez que o número de variáveis for aumentado em uma unidade, o número de cenários a serem considerados dobra. Ou seja, conforme o número de variáveis aumenta, a quantidade de soluções possíveis aumenta de forma exponencial.

Fatores bastante influentes no número de soluções a serem consideradas em problemas de PI são: o número de variáveis, a natureza dessas variáveis (binárias ou genéricas) e também a inserção de estruturas especiais ao problema (LIEBERMAN, 2013). A fim de adequar a exigência computacional de um determinado modelo às máquinas disponíveis existem algumas possibilidades, dentre elas: o aumento das restrições, que pode refletir na redução de variáveis, ou a alteração dos critérios de convergência da solução, que pode refletir na redução no número de iterações no algoritmo.

A ideia de alteração nos critérios de convergência para a solução do problema de PI é bastante importante e pode possibilitar a convergência de diversos algoritmos. Essa tolerância representa uma diferença (absoluta ou relativa) entre o valor da última melhor solução conhecida e a solução atual, sendo a busca iniciada pela diferença entre o valor da

função objetivo de programação linear relaxada e a função do problema de programação inteira (AJAYI, 2021). Ou seja, a implementação de um intervalo de otimalidade representa uma margem de tolerância entre duas soluções consecutivas na busca da solução ótima, quais sejam, a melhor solução conhecida até então e a próxima solução avaliada, partindo da solução “relaxada” para o problema. Tem-se o poder de estimar a distância limite entre as duas propostas.

Ainda de acordo com Ajayi (2021) essa função de intervalo pode ser considerada uma medida da qualidade do modelo. Ou seja, quanto menor for o intervalo escolhido mais próximo da solução ótima o modelo desenvolvido estará. Entretanto, o uso do intervalo de tolerância não deixa de ser um recurso bastante útil em caso de sistemas complexos, uma vez que a distância entre o modelo relaxado e o modelo completo pode ser delimitada pelo próprio pesquisador.

Entende-se que a teoria de complexidade de algoritmos pode ser aplicada em problemas de programação inteira e o intervalo de otimalidade pode ser considerado um parâmetro relevante para a avaliação de complexidade de um determinado modelo. Ainda sobre o tema, de acordo com Ajayi (2021, apud KHACHIYAN, 1980), é possível caracterizar problemas de programação inteira com o uso de intervalos de tolerância como problemas de complexidade NP-difícil.

2.8 Problema da Alocação de Frota - FAP

O tópico em questão tem por objetivo a contextualização do Problema da Alocação de Frota, do inglês *Fleet Assignment Problem (FAP)* e, para entender a alocação, observa-se também o delineamento do planejamento do processo aéreo de uma determinada companhia aérea. Assim como demais projetos, sabe-se que a estruturação do processo aéreo depende de etapas sequenciais, sendo elas: planejamento da frota, planejamento de rotas, alocação da frota, gestão e manutenção da frota e, por fim, alocação da tripulação (MENDES, 2019).

Ou seja, assim como citado no parágrafo anterior, as companhias aéreas precisam apoiar-se em técnicas de análise de decisão para aprimorar o planejamento aéreo. A partir dessa necessidade, vários estudos ou métodos são implementados ao longo do século XX, até que, no ano de 1989, Jeph Abara, desenvolveu um método de integrar a programação linear ao processo de planejamento, que foi o primeiro a ser utilizado e a trazer resultados positivos para uma companhia aérea.

A partir deste primeiro modelo, várias outras propostas foram desenvolvidas e o FAP atualmente ainda é bastante útil para as empresas na etapa de seleção de aeronaves. Basicamente, nesse modelo de pesquisa a alocação de uma aeronave para cumprir determinada rota é tratada como um problema de viabilidade. Ou seja, o objetivo do

FAP é encontrar uma forma de atender todas as rotas abordadas pela companhia, tendo o mínimo de gastos operacionais possível e com manutenções mais simples (MEDAU, 2018).

Ainda segundo Medau (2018), o uso do FAP pode ser utilizado em diferentes etapas do projeto, a primeira delas sendo a alocação de frota a partir de uma malha pré definida, nesse caso o desafio é a seleção de aeronaves para realizar os voos e a segunda forma seria a definição da malha aérea a partir de uma frota já existente. Usualmente, no mercado de transporte aéreo atual, é mais comum que o FAP seja aplicado a partir da frota aérea disponível em cada companhia para que então sejam elaboradas as rotas mais viáveis, o que é logicamente aceitável e economicamente viável para o caso de organizações. Entretanto, na pesquisa atual, o FAP será incorporado a partir de rotas definidas por meio de demandas de usuários/potenciais clientes para então a definição da frota, algo estratégico para indicar novos mercados e consequentemente impactar diferentes regiões brasileiras que estejam distantes dos grandes centros.

Assim, infere-se que a programação linear inteira mista para a resolução da alocação de frota é algo possível, tendo como embasamento e referências principais os trabalhos de Jesus (2019) e Pereira (2021). Assim sendo, a partir dos resultados obtidos pelas pesquisas anteriores, entende-se que as questões a serem resolvidas no modelo pré-existente, são, em sua maioria, de origem computacional.

2.9 Métodos de regressão linear

Uma ferramenta matemática essencial para a análise de dados são os modelos de regressão linear, que podem ser do tipo simples ou múltiplo. Os métodos de regressão podem ser utilizados para a análise de comportamento de um conjunto de variáveis a partir de uma variável de referência, o fenômeno do estudo. Existem várias técnicas de regressão e, para satisfazer os requisitos da programação inteira, optou-se pelo estudo exclusivo de métodos lineares que, segundo Fávero (2015) pode ser definida como: estudo da “relação entre uma ou mais variáveis explicativas, que se apresentam de forma linear, e uma variável dependente quantitativa”.

Considera-se que modelos de regressão linear simples são aqueles em que há apenas uma variável explicativa e para aqueles modelos em que as variáveis consideradas são diversas tem-se o desenvolvimento de um modelo de regressão linear múltiplo (FÁVERO, 2015). Os modelos citados podem ser representados por diferentes equações:

- Regressão linear simples:

$$Y_n = \alpha + \beta.X_i \quad (2.3)$$

- Regressão linear múltipla:

$$Y_i = \alpha + \beta_1 \cdot X_{1_i} + \beta_2 \cdot X_{2_i} + \dots + \beta_k \cdot X_{k_i} + u_i \quad (2.4)$$

Nas equações descritas, tem-se que Y representa a variável dependente quantitativa, α está relacionado ao intercepto (coeficiente linear), β representam os coeficientes angulares, X são as variáveis explicativas e u é a diferença entre valores reais e sugeridos de Y . A consideração dos resíduos de cada modelo é essencial e de acordo com [Fávero \(2015\)](#), a inclusão da diferença entre os valores de Y pode estar atrelada a motivos como falhas no levantamento de dados ou na especificação do modelo e a existência de variáveis atreladas.

A estimação de resíduos pode ser relativamente complexa e, em casos de estimação equivocada, podem representar falhas nos resultados da pesquisa. Portanto, para garantir que o modelo seja capaz de representar um sistema real, segundo [Fávero \(2015\)](#), faz-se necessário que a soma dos resíduos seja igual a zero e que a soma dos resíduos ao quadrado seja a mínima possível. Ou seja, a definição dos coeficientes lineares e angulares podem ser obtidas analiticamente por meio da minimização do quadrado dos resíduos (método de mínimos quadrados).

Após a definição da equação de regressão, é preciso atentar-se para a definição do coeficiente de ajuste R^2 . De acordo com [Fávero \(2015, apud Wooldridge 2012\)](#), o fator do coeficiente de ajuste pode ser definido como “a proporção da variação amostral da variável dependente explicada pelo conjunto de variáveis explicativas, podendo ser utilizado como uma medida do grau de ajuste do modelo proposto”.

3 Metodologia

Este capítulo tem por objetivo a explicação da ferramenta organizacional utilizada no trabalho bem como a explicação dos métodos utilizado para formulação dos artefatos desenvolvidos. São apresentados no capítulo os tópicos: *Design Science Research*, etapas da metodologia proposta, desenvolvimento do modelo, modelo matemático, algoritmo pré existente e alterações no algoritmo pré existente.

3.1 *Design Science Research*

Segundo [Cauchick \(2019\)](#), a utilização de métodos científicos tradicionais quando usados, de forma exclusiva, para resolução de projetos de engenharia podem tornar-se agentes limitadores das pesquisas. Uma vez que essas metodologias foram criadas para serem aplicadas na resolução de questões dos campos das ciências naturais ou sociais que têm, por definição, objetivos centrais bastante diferentes da engenharia. Assim sendo, encontra-se na ferramenta do *Design Science Research* uma perspectiva complementar para a resolução de problemáticas ligadas à engenharia. Visto que, nesse caso, o objetivo do método passa a ser criação de algo e não exclusivamente a definição de um fenômeno ou sintoma.

Idealmente, o uso do *Design Science Research* é recomendado para pesquisas que buscam resolver questões cotidianas a partir da sugestão de melhorias de um modelo existente ou da criação de um novo conceito. Aplicando essa recomendação para a questão central do problema deste trabalho, que é a modelagem de uma malha aérea regional brasileira ótima a partir da pesquisa operacional, infere-se que a finalidade da metodologia é condizente com a problemática abordada no trabalho.

Para além das questões relacionadas à natureza dos problemas, assim como as demais metodologias científicas, o *Design Science Research* também possui normas que regem a conduta de um determinado estudo. De acordo com o método, as principais etapas de um projeto podem ser definidas como: “(i) definição do problema; (ii) proposta de sugestões de solução; (iii) desenvolvimento do artefato; (iv) avaliação do artefato” ([CAUCHICK, 2019](#)). Definindo-se como artefato as criações do projeto.

Mais especificamente, um projeto adepto desta metodologia deve ser desenvolvido sequencialmente por um momento de identificação do problema, seguido pela conscientização e revisão sistemática da literatura existente relacionada à problemática abordada, passando para a determinação dos artefatos a serem desenvolvidos ou trabalhados e as classes do(s) problema(s) a serem resolvidos. Após esse período inicial, o novo ponto a

ser discutido é a determinação do projeto e desenvolvimento de artefatos específicos à temática abordada. Por fim, existe também a etapa de avaliações dos feitos, aprendizados, conclusões gerais e definição de planos de comunicação dos resultados obtidos.

A partir de estudos preliminares, conclui-se que este é o melhor método para a pesquisa, tanto por ser uma abordagem que comporta a natureza de questões da engenharia, quanto pelo foco da metodologia na relevância de um determinado trabalho para o cotidiano das pessoas. Assim sendo, nos tópicos seguintes do capítulo, será desenvolvido um plano de estudo seguindo as principais normas do *Design Science Research*.

3.1.1 Etapas da metodologia proposta

A primeira etapa da metodologia é dedicada à identificação do problema, em que o pesquisador deve entender completamente a problemática e suas complexidades. No estudo em questão, entende-se que o problema identificado a ser resolvido é a alta complexidade computacional do algoritmo desenvolvido previamente por [Pereira \(2021\)](#). Essa decisão foi tomada partindo da conclusão do autor citado, visto que o modelo pré-existente ainda não é capaz de fazer a alocação de mais de uma aeronave simultaneamente, o que impossibilita a análise do modelo matemático por inteiro.

Após a identificação do problema, parte-se para etapa de contextualização, ou seja, são realizadas investigações buscando o entendimento do contexto geral da temática abordada, este processo deve ser feito de forma paralela à revisão de literatura, espera-se então que seja evidenciada a relevância do estudo pretendido bem como a geração de resultados mais adequados ao problema identificado. Sendo assim, atesta-se que o trabalho atual é uma tentativa de aprimorar os modelos desenvolvidos por [Jesus \(2019\)](#) e [Pereira \(2021\)](#), os pontos de mudança podem ser identificados na Tabela 1.

Tabela 1 – Comparativo entre pesquisas

Ponto Analisado	Jesus (2019)	Pereira (2021)	Trabalho Atual
Objetivo Geral	Modelar o processo de escolha de aeronaves e de uma malha aérea regional, operadas a partir do <i>hub</i> de Goiânia - GO.	Obter uma solução ótima de escolha de alocação de aeronaves para operar potenciais rotas aéreas regionais por meio do Problema da Alocação de Frota (FAP).	Reduzir a complexidade computacional de um modelo já existente de malha aérea regional.
Hubs Abordados	Goiânia - GO	12 <i>hubs</i>	12 <i>hubs</i>
Linguagem de Programação	Lingo	Python	Python

Entende-se também que a contextualização do problema vai além do reconhecimento dos trabalhos anteriores, por isso foi realizada uma revisão narrativa da literatura propriamente dita no Capítulo 2. Além disso, identificou-se a necessidade de estudo do cenário geral objetivando encontrar pontos a serem alterados no modelo matemático ou confirmar cenários estimados anteriormente.

Desta forma, foi realizada uma pesquisa na situação geral da indústria de transporte aéreo brasileiro. O primeiro ponto importante foi compreender se a base de dados de demanda de rotas Horus (2022), utilizada em modelos anteriores, não está defasada, foi realizado um comparativo da demanda e oferta do mercado entre os anos de 2017 até 2021. As datas limite de comparação foram escolhidas de acordo com o período de coleta de dados registrado na matriz de origem e destino e também de acordo com o último registro de oferta e demanda disponibilizada pela ANAC (2022b).

Tabela 2 – Demanda e Oferta do transporte aéreo. Adaptado de ANAC, 2021

Tipo de registro	2017	2018	2019	2020	2021
ASK	112 bi	117 bi	116 bi	61 bi	86 bi
RPK	91bi	95 bi	96 bi	49 bi	69 bi

Os registros da Tabela 2 mostram os índices de oferta de assentos (ASK) e de demanda de passageiros pagos transportados (RPK) anualmente. Percebe-se uma queda considerável tanto na demanda quanto na oferta nos anos 2020 e 2021. Esta variação está atrelada às consequências das medidas de contenção da COVID-19 ANAC (2022a). Em relação ao ano de 2022, acredita-se que, mesmo com a flexibilização das medidas de restrição de voos em um cenário pós-pandêmico, uma recuperação plena do setor de transporte aéreo para patamares pré-COVID ainda pode demorar, devido aos desdobramentos dos conflitos entre Rússia e Ucrânia, o que já levou as companhias aéreas brasileiras a anunciarem uma diminuição de operações de voos pontuais em função do aumento no combustível (AEROIN, 2022). Desta forma, entende-se que a continuação do uso da matriz O&D como fonte de dados é algo aceitável, visto que os valores presentes ainda não são defasados.

A partir da compreensão da identificação do problema e do estudo do contexto geral da temática e dos avanços tecnológicos já desenvolvidos, é possível avançar para a próxima etapa de identificação dos artefatos já existentes e classificação dos problemas propostos. Entende-se que os principais artefatos similares previamente desenvolvidos são os resultados das pesquisas de Pereira (2021), cujo objetivo foi a modelagem de uma malha aérea regional brasileira ótima utilizando ferramentas da Pesquisa Operacional e o trabalho de Jesus (2019), que tem metas semelhantes, porém aplicado apenas na região de Goiânia - Goiás. Quanto à classificação do estudo, trata-se de um Problema de Alocação de Frota (FAP).

Passadas as etapas listadas anteriormente, é possível partir para a proposição de artefatos a serem desenvolvidos na pesquisa, seguido do projeto destes. Assim sendo, buscou-se os dados apresentados na matriz de origem e destino [Horus \(2022\)](#) para a identificação de possíveis rotas aéreas regionais. Em seguida, foi utilizado o Registro de Serviços Aéreos da [ANAC \(2022b\)](#), onde foi possível localizar as aeronaves que são utilizadas atualmente para esse tipo de viagem. Após essas buscas foi possível fazer as adaptações necessárias para o aprimoramento do algoritmo pré-existente. O fluxograma para o projeto dos artefatos está evidenciado na [Figura 3](#).

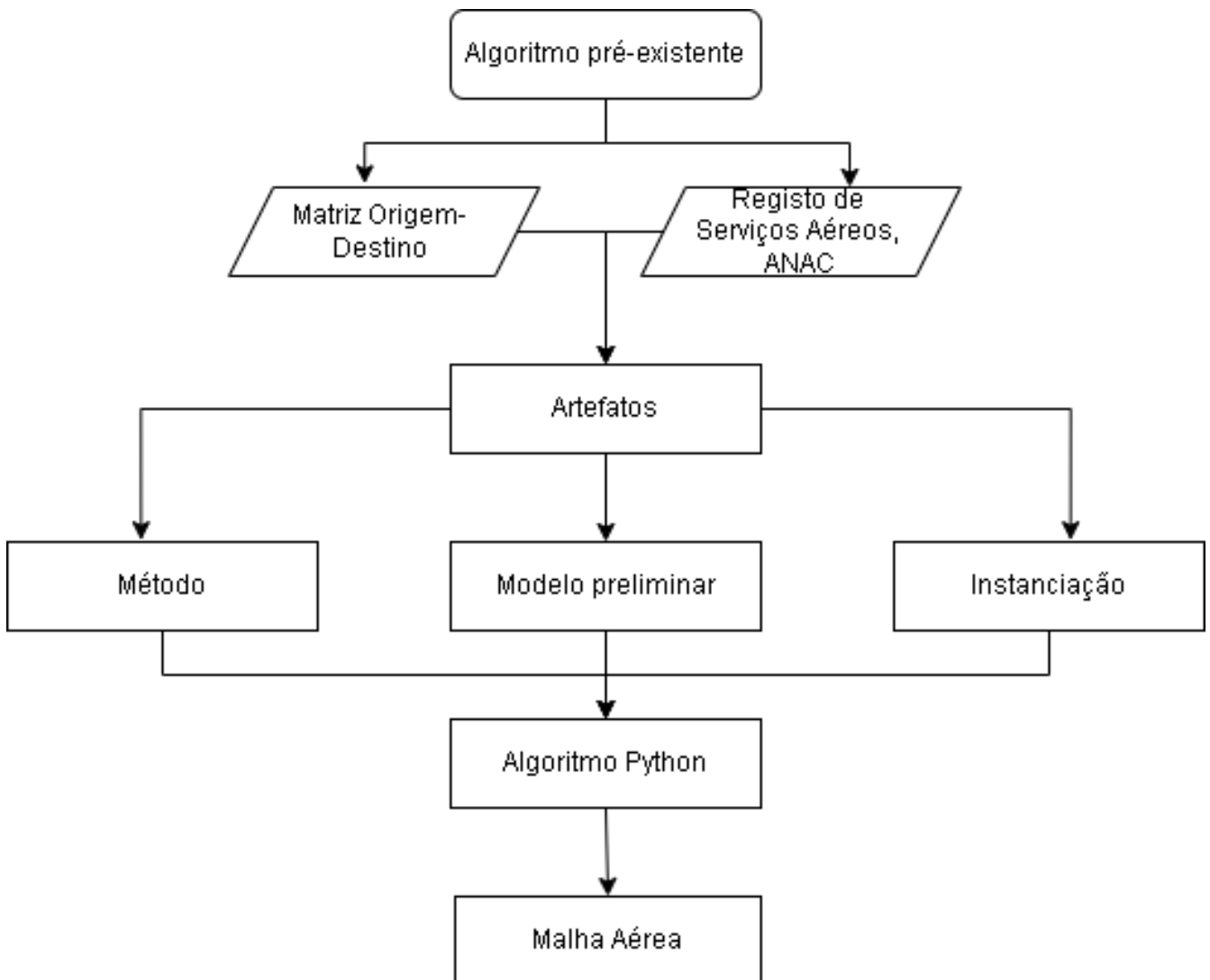


Figura 3 – Projeto de artefatos a serem desenvolvidos.

Após a definição completa do projeto dos artefatos, partiu-se para o desenvolvimento propriamente dito. Para a pesquisa, optou-se pela escolha de três artefatos: um método, um modelo e uma instanciação, a etapa de desenvolvimento dos artefatos está melhor desenvolvida na [Seção 3.2](#).

- Método: o método do trabalho em questão é a própria pesquisa operacional que será aplicada para o FAP;
- Modelo: é a instrumentação matemática e algorítmica utilizada para representar o problema a ser trabalhado. Ou seja, fazem parte do modelo do trabalho toda a etapa de construção ou simulação do cenário estudado.
- Instanciação: a instanciação será a aplicação do modelo propriamente dita.

Sequencialmente ao desenvolvimento, existe a etapa de avaliação dos artefatos dos recursos desenvolvidos. Os critérios para avaliação devem ser estipulados pelos próprios pesquisadores, no caso deste trabalho, os critérios de avaliação tem como base na Seção 1.4. Definidos os critérios, parte-se para a explicitação das aprendizagens e conclusão, este passo está dividido entre os Capítulos 4 e 5, nesse momento cabe ao pesquisador levantar os pontos relevantes aprendidos durante a pesquisa, no que tange aos sucessos e questões a serem melhoradas no conteúdo desenvolvido.

Por fim, a última etapa é chamada de generalização para uma classe de problemas e comunicação dos resultados, é neste momento em que são definidos os canais de divulgação da pesquisa realizada. Para a presente situação, optou-se pela submissão de uma versão resumida da pesquisa no Simpósio de Transporte Aéreo (SITRAER) 2022, sob o título de “Modelagem do Problema da Alocação de Frota Aplicado à Aviação Regional Brasileira a Partir de Ferramentas da Pesquisa Operacional”. Cumpridas todas as etapas citadas de forma eficaz, considera-se que o projeto seguiu a metodologia do *Design Science Research*. A fim de ilustrar melhor como a metodologia foi implementada na pesquisa, foi desenvolvida a figura 4.

3.2 Desenvolvimento do Modelo

A metodologia utilizada para o desenvolvimento dos artefatos pode ser dividida em duas etapas essenciais, representativas das disciplinas de Trabalho de Conclusão de Curso 1 e 2. A primeira etapa foi marcada pela revisão narrativa da literatura e foco na redução da complexidade computacional do algoritmo pré-existente. O segundo momento foi marcado pelas análises dos resultados obtidos com o modelo menos complexos, sendo assim foi possível encontrar pontos de melhoria na formulação do modelo matemático seguido de novas análises dos resultados obtidos após todas as alterações. Um esquemático da metodologia encontra-se registrado na Figura 5.

Assim como descrito na Seção 1.4, o foco principal da presente pesquisa é possibilitar a resolução do modelo matemático pré-existente a partir da redução da complexidade computacional. Ou seja, um dos artefatos a serem desenvolvidos no trabalho é um modelo que no caso é composto por uma modelagem matemática aplicada em algoritmos de

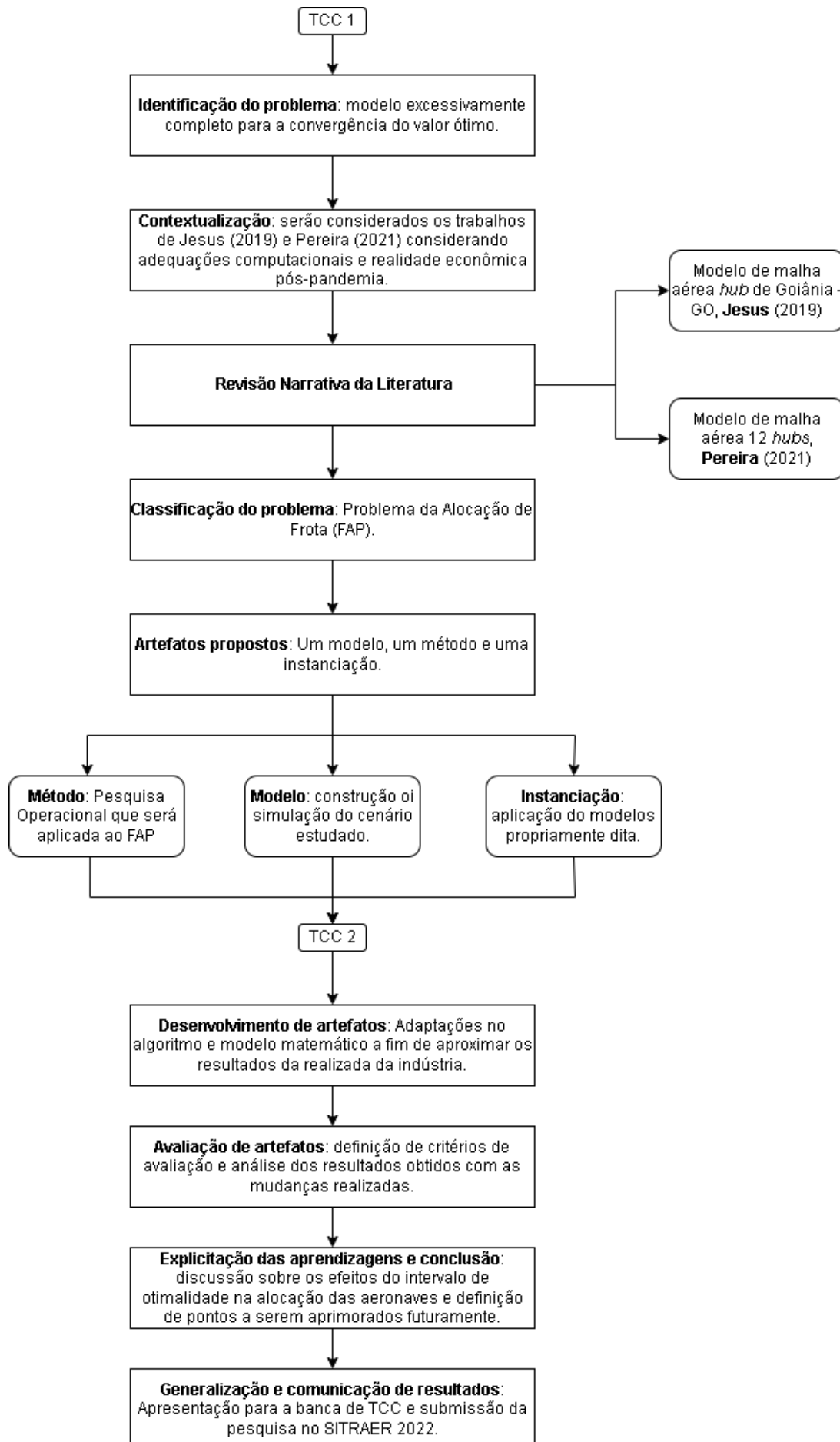


Figura 4 – Metodologia do *Design Science Research* aplicada ao projeto.

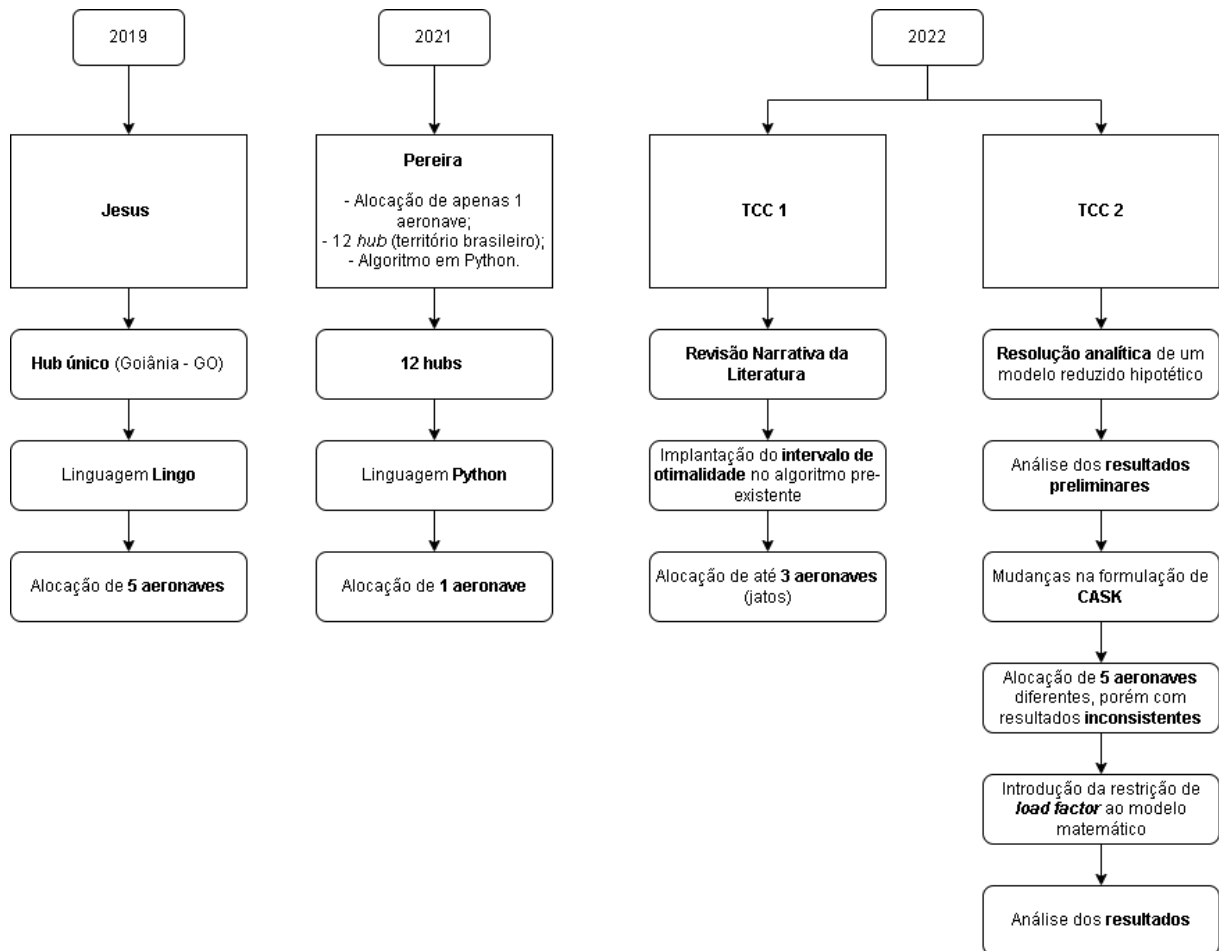


Figura 5 – Etapas do desenvolvimento dos artefatos.

Python. Para a confecção do algoritmo foram utilizadas as bibliotecas Pandas e Numpy, além do *framework* PuLP. Partindo do conceito de simplificação do modelo, o algoritmo elaborado por Pereira (2021) foi utilizado como base para o desenvolvimento do projeto.

A fim de entender melhor o contexto e certificar-se da viabilidade dos recursos a serem utilizados, optou-se pela revisão dos dados a serem utilizados. Primeiramente, a matriz OD (Origem e Destino) foi analisada e assim como explicitado na seção 3.1.1, verificou-se que os valores ainda podem ser utilizados. A matriz OD estabelece doze *hubs* principais que serão sempre considerados pontos de origem nas rotas e com destinos em outras UTPs externas às doze principais.

Além do uso dos dados do sistema Hórus, também optou-se por manter, em um primeiro momento o tratamento na matriz OD realizada por Pereira (2021), que fez adaptações voltadas para a adequação da do mercado e também para atender a definição de aviação regional citada na seção 2.4.

3.3 Modelo Matemático

O modelo matemático utilizado na primeira fase da pesquisa foi desenvolvida por Jesus (2019) e Pereira (2021), e tem por objetivo atingir o maior lucro possível para toda a operação da malha aérea. A modelagem matemática foi elaborada seguindo os moldes da Pesquisa Operacional, ou seja, foi elaborada uma função objetivo, Equação 3.1, que nesse caso é responsável por calcular a maximização do lucro, a partir da diferença entre receita obtida com a venda de passagens e os custos operacionais de cada rota. Além disso, também foram desenvolvidas outras equações de restrições que condicionam a resolução da solução final.

$$[MAX] = \sum_{h \in H} \sum_{j \in J} \sum_{a \in A} (T_{h,j} \cdot P_{h,j,a} - C_{h,j,a} \cdot D_{h,j} \cdot S_a \cdot F_{h,j,a}) \quad (3.1)$$

Em que:

- h = origem em qualquer *hub* dentre aqueles considerados;
- j = conjunto de destinos;
- a = aeronave da amostra A.

A equação 3.1 calcula a diferença do somatório, rota a rota, aeronave a aeronave, entre a receita (calculada pela multiplicação da tarifa T pelo número de passageiros transportados P) e o custo operacional da rota (calculado pela multiplicação do CASK C pela distância voada D pelo número de assentos S e pelo fluxo de aeronaves F necessário para atender à demanda da rota).

As funções de restrição estão listadas a seguir:

$$\sum_{a \in A} P_{h,j,a} = Q_{h,j} \quad \forall h \in H, \forall j \in J \quad (3.2)$$

$$P_{h,j,a} \leq S_a \cdot F_{h,j,a} \quad \forall h \in H, \forall j \in J, \forall a \in A; \quad (3.3)$$

$$P_{h,j,a} \geq LF \cdot S_a \cdot F_{h,j,a} \quad \forall h \in H, \forall j \in J, \forall a \in A; \quad (3.4)$$

$$F_{h,j,a} \leq M \cdot BIN1_{h,j,a} \quad \forall h \in H, \forall j \in J, \forall a \in A; \quad (3.5)$$

$$R_a \geq D_j \cdot BIN1_{h,j,a} \quad \forall h \in H, \forall j \in J, \forall a \in A; \quad (3.6)$$

$$F_{h,j,a} \leq M \cdot BIN2_a \quad \forall h \in H, \forall j \in J, \forall a \in A; \quad (3.7)$$

$$\sum_a BIN2_a \geq K \quad \forall h \in H, \forall j \in J; \quad (3.8)$$

Tabela 3 – Correspondências de símbolos utilizados no modelo

Símbolo	Correspondência
$BIN1_{\{h,j,a\}}$	Variável binária que possui valor 0 ou 1;
$BIN2_{\{a\}}$	Variável binária que possui valor 0 ou 1;
$C_{\{h,j,a\}}$	CASK da aeronave $\{a\}$ para o destino $\{j\}$ pelo hub $\{h\}$, $C \in R$
$D_{\{h,j\}}$	Distância total entre o hub $\{h\}$ e o destino $\{j\}$, $D \in R$
$F_{\{h,j,a\}}$	Fluxo total de aeronaves $\{a\}$ para o destino $\{j\}$ pelo hub $\{h\}$, $F \in Z$
K	Constante do número máximo de diferentes tipos de aeronaves
LF	Constante load factor que estabelece a lotação mínima de determinada aeronave
M	Constante do número máximo de fluxos semanais para uma determinada rota
$P_{\{h,j,a\}}$	Número de passageiros transportados pela aeronave $\{a\}$ para o destino $\{j\}$ pelo hub $\{h\}$, $P \in Z$
$Q_{\{h,j\}}$	Demanda de passageiros entre hub $\{h\}$ e o destino $\{j\}$, $Q \in Z$
$R_{\{a\}}$	Alcance máximo de voo da aeronave $\{a\}$, $R \in R$
$S_{\{a\}}$	Constante do número total de assentos da aeronave $\{a\}$
$T_{\{h,j\}}$	Tarifa de passagem entre o hub $\{h\}$ e o destino $\{j\}$, $T \in R$

A equação 3.2 é a restrição que estabelece que toda a demanda existente em cada rota precisa ser atendida em igualdade com o número de passageiros transportados, podendo ser atendida por mais de um tipo de aeronave.

A equação 3.3 é a restrição que estabelece que a oferta de assentos por fluxo tem que ser maior ou igual à demanda expressa em passageiros. A oferta (fluxo * assentos) pode vir de mais de um tipo de aeronave para atender à demanda. Em cada rota, a oferta precisa ser maior que a demanda atendida (passageiros) para cada aeronave.

A equação 3.4 é a restrição que regula a demanda atendida por fluxo de cada aeronave na rota. A oferta (fluxo * assentos) não pode superar demasiadamente a demanda atendida (passageiros). A partir do *load factor* (LF) (fator de ocupação) por rota, estabelece-se um limite máximo que de uma oferta $1/LF$ vezes maior que a demanda na rota.

A equação 3.5 é a restrição que define o fluxo máximo de cada modelo por rota. Um modelo de aeronave só pode atender uma demanda de M vezes a capacidade de assentos. Já a equação 3.6 limita a alocação de um modelo de aeronave às rotas cuja distância é menor que o alcance máximo da aeronave.

As equações 3.7 e 3.8 estabelecem restrições adicionais aos fluxos, sendo que, respectivamente, impede que haja fluxo de modelo de aeronave não incluído e propicia limitar o número máximo de modelos alocados (diferentes cenários estudados).

3.4 Algoritmo pré existente

Primeiramente, faz-se necessário indicar que a modelagem foi dividida em duas partes, sendo elas: a primeira delas para o cálculo das distâncias entre *hubs* (disponível no Anexo A) e outra para a resolução do modelo de PO propriamente dito.

Os resultados do algoritmo do Anexo A bem como os dados de demanda de origem e destino da plataforma Hórus (Horus, 2022) são usados como parte dos dados de entrada para a resolução do problema de PO. Após a leitura dos dados, o próximo passo é o cálculo dos valores de CASK e valor da passagem de cada uma das rotas.

Na segunda parte, foi utilizado o comando “LpProblem” do *framework* Pulp para criar um problema do tipo LP que deve ser acompanhado de dois parâmetros: um para nomear a função e outro para sinalizar o objetivo da modelagem, neste caso, maximizar o lucro, logo foi utilizado o comando “LpMaximize”. Em seguida são registradas ambas função objetivo e restrições definidas para o modelo. Após todos os passos é possível acionar o comando de solução do problema. Posteriormente, são descritos apenas comandos para armazenamento dos resultados. Uma representação gráfica do problema está disponível na Figura 6.

3.5 Alterações no algoritmo pré-existente

Como mencionado anteriormente, a partir dos resultados obtidos por Pereira (2021), foram evidenciados alguns pontos que precisam ser tratados no modelo matemático e no algoritmo. Assim sendo, esta seção é dedicada à explicação de cada um dos pontos alterados no modelo matemático.

3.5.1 O Parâmetro K e a Complexidade Computacional

Um dos parâmetros muito importantes para a resolução do problema é o parâmetro K , que faz referência a quantidade de aeronaves que podem ser testadas e inseridas

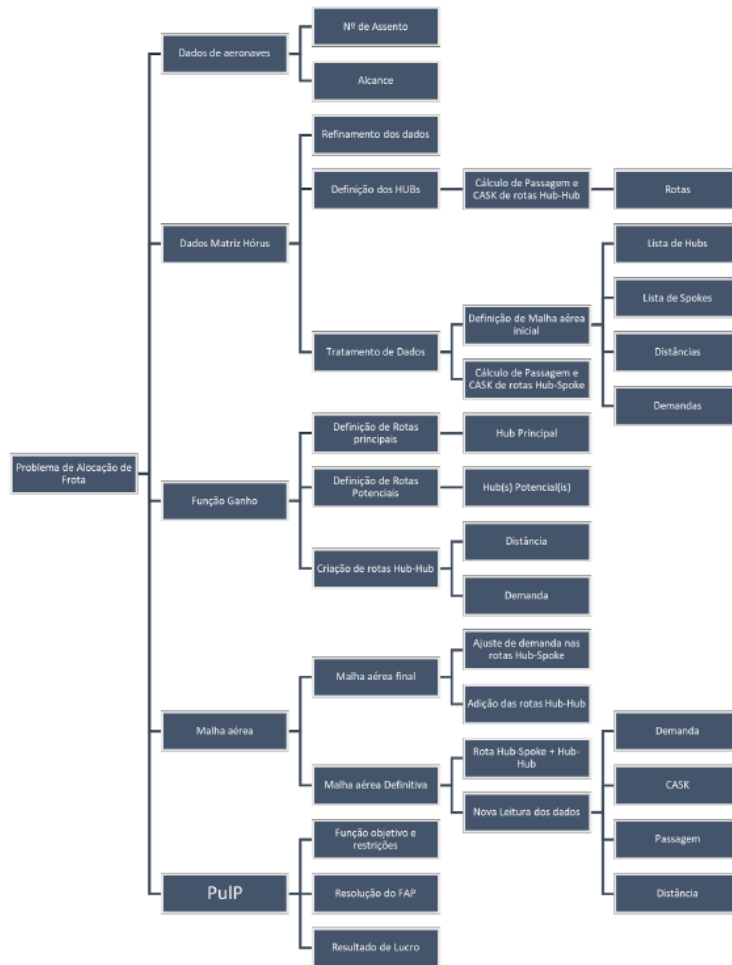


Figura 6 – Mapa Mental usado para a Criação do Código. Fonte: (PEREIRA, 2021)

no modelo (PEREIRA, 2021) de forma simultânea. O valor de K pode variar de acordo com a quantidade de aeronaves testadas. A intenção do modelo pré-existente seria testar o parâmetro K de 1 até 5. No entanto, o principal problema do trabalho anterior foi a impossibilidade de resolução do problema para valores de K acima de 2. Ou seja, considerou-se que, a partir da inserção de mais de duas aeronaves para escolha de alocação, o problema torna-se demasiadamente complexo para a capacidade computacional ao alcance do pesquisador. Para solução desta limitação, foram identificadas duas alternativas: a reformulação completa do modelo ou a redução computacional do modelo existente. Optou-se, assim como mencionado outras vezes neste trabalho, pelo aproveitamento do modelo pré-existente com redução da complexidade computacional.

A alternativa encontrada foi a mudança de parâmetros na etapa de solução do algoritmo, ou seja, ao invés de utilizar o comando do *solver* de forma genérica que condiciona apenas o valor ótimo do problema como solução, foi utilizado o comando *GapRel* (existente no pacote PuLP da linguagem de programação *Python*) que dá ao pesquisador a opção de estipular um intervalo de tolerância para a seleção da solução do problema.

Ou seja, com este comando é possível escolher o quão próxima a solução considerada adequada deve estar do ponto ótimo, seguindo o conteúdo da seção 2.7.

A sintaxe correta para o comando *GapRel* deve indicar qual é a distância entre o ótimo e o viável. Será estipulado um valor em porcentagem que pode ser chamado de intervalo de otimalidade, esse intervalo por ser escolhido de acordo com as necessidades da pesquisa. Foram realizados testes com diferentes valores de *gap* com o objetivo de entender melhor o funcionamento do recurso e quais suas consequências ao modelo.

Traduzindo para a linguagem Python, a estratégia na verdade resume-se a uma mudança no comando de solução do modelo. Ou seja, abandona-se o comando genérico que considera apenas a resposta ótima, *F0.solve()*, para utilizar um novo comando que especifica as margens consideradas adequadas, instituindo o *gapRel* inicialmente como 0.01.

O objetivo a ser alcançado a partir da introdução do intervalo de otimalidade é a redução da complexidade computacional. Isso seria alcançado pois a tolerância estipulada é capaz de reduzir as iterações necessárias e conseqüentemente também reduzir o tempo de simulação. Assim, o modelo torna-se possível para o uso de máquinas convencionais.

É importante pontuar que o tempo de simulação pode variar de máquina para máquina, portanto, todas as simulações realizadas neste trabalho foram feitas utilizando *notebooks* do *Google Colaboratory* (GOOGLE, 2022), um ambiente virtual do *Google*, onde é possível fazer o uso de máquinas virtuais para a programação na linguagem Python. As especificações da máquina disponível gratuitamente ao usuário, que foi escolhida para a realização da simulação do modelo do trabalho, encontram-se na tabela 4.

Processador	Dois núcleos
Memória RAM	12 GBytes
Cache L3	40-50 MBytes

Tabela 4 – Especificações da máquina virtual utilizada para simulações

Com a intenção de comprovar a redução de complexidade computacional a partir do implantação do *gap* de otimalidade, foram realizadas, em uma mesma máquina virtual (tabela 4), simulações considerando apenas a solução ótima e simulações, apenas para testagem do recurso, considerando a tolerância de 1% do cenário ótimo. Com a redução da complexidade, foi possível resolver o problema em questão para cenários mais complexos.

Nos testes realizados foram considerados também diferentes cenários em termos de variação de aeronaves, que no algoritmo pode ser representado pela variação no parâmetro K. No primeiro cenário (K=1), apenas a aeronave mais adequada pode ser alocada ao modelo. Foram realizados testes variando o parâmetro K de 1 até 5, ou seja, foi simulado um modelo com a opção de alocar as duas aeronaves mais adequadas e assim por diante Pereira (2021).

3.5.2 Adaptações para o cálculo de CASK

A partir da redução da complexidade computacional foi possível observar outras inconsistências no modelo, sendo a primeira delas o cálculo de CASK. Os resultados obtidos nos testes iniciais apontaram para a necessidade de mudança no cálculo do indicador, particularmente, a inclusão no algoritmo de uma forma de apontar as diferenças operacionais entre os diferentes tipos de aeronaves, jatos e turboélices.

Assim sendo, foi necessário fazer uma separação na formulação entre os cálculos de CASK para aeronaves do tipo jato e uma diferente para aeronaves do tipo turboélice. Não houve mudanças significativas na forma a qual a curva de CASK foi calculada, apenas uma aproximação que representasse melhor a situação dos custos operacionais de cada um dos modelos. Assim sendo, deixou-se de ter apenas uma equação para o cálculo do indicador, passando-se a ter duas, uma representante para cada tipo de veículo.

$$CASK_{TF} = e^{1.728912 - 0.40453 \cdot \ln(D_{h,j}) - 0.35671 \cdot \ln(S_a)} \quad (3.9)$$

$$CASK_{TP} = e^{-0.481549 - 0.099212 \cdot \ln(D_{h,j}) - 0.356708 \cdot \ln(S_a)} \quad (3.10)$$

O caminho até a formulação listada acima foi realizado por meio de aproximações entre as bases de dados disponíveis e utilizadas para a elaboração do modelo pré-existente elaborado por [Pereira \(2021\)](#) e os resultados mais intuitivos possíveis. Como os dados de CASK no estudo original foram obtidos por consulta a especialistas da indústria aeronáutica, voltou-se a consultar especialistas, junto dos quais constatou-se ser possível revisar as estimativas de CASK para as aeronaves turboélices. Desta forma, foi possível aproximar a representação matemática do indicador para representar as características operacionais de cada tipo de aeronave.

3.5.3 Inserção do fator de ocupação (*load factor*)

Além das adaptações citadas anteriormente, também houve a inserção de um fator mínimo de ocupação, a fim de garantir que nenhuma aeronave seja alocada para uma demanda muito menor do que sua capacidade total de assentos. Foram instituídos um fator de ocupação de 50% do total de assentos em casos em que a distância entre origem e destino for menor que 1300 km e de 25% do total nos casos em que esta distância é maior que isso. Esta restrição não foi utilizada para o cenário em que havia apenas uma aeronave disponível para alocação ($K = 1$), visto que nesse caso não foi possível encontrar um único veículo, dentre os disponíveis, que fosse capaz de atender simultaneamente a condição de ocupação e toda a variedade de distâncias existentes na base de dados utilizadas. Portanto, para o caso em questão, foi necessário abrir mão do critério de ocupação.

3.5.4 Algoritmo atualizado

Após as mudanças citadas nas subseções anteriores, obteve-se um novo código em Python capaz de resolver o problema de Pesquisa Operacional estudado, o algoritmo é, portanto, uma adaptação do código desenvolvido por [Pereira \(2021\)](#). Esta subseção tem por objetivo a explicação da nova versão do algoritmo que pode ser encontrado em sua plenitude no Apêndice [A](#). Assim como todas as outras rotinas de programação utilizadas no trabalho, o código foi desenvolvido em um *notebook* na ferramenta do *Google Colaboratory*.

Inicialmente, são feitas as preparações necessárias para a leitura dos dados. Portanto, primeiro é utilizado o comando *reset* para limpar o banco de resultados, em seguida, são realizadas as importações de todas as bibliotecas e *frameworks* necessários para a resolução do problema. Em seguida, é feita a leitura do conjunto de dados previamente montado com as informações sobre rotas e aeronaves que serão utilizadas posteriormente.

Após a preparação do *notebook*, parte-se para o cálculo da tarifa para cada uma das 710 rotas existentes no banco de dados do Projeto Hórus que satisfazem os critérios de aviação regional explicitados na Seção [2.4](#). O cálculo do preço das passagens varia apenas de acordo com a distância entre origem e destino de cada rota, a Equação [3.11](#) registra a função matemática utilizada para obter o valor desejado.

$$T_{h,j} = e^{-0.73953 \cdot \ln(D_{h,j}) + 4.44446} D_{h,j} \quad (3.11)$$

Em seguida, parte-se para o cálculo do CASK que, diferente da tarifa, depende da distância voada por rota e também da quantidade de assentos disponíveis. O detalhamento e as funções utilizadas podem ser encontradas na Subseção [3.5.2](#). Nota-se também, a existência de um laço de repetição que condiciona que em casos em que a distância voada seja maior que o alcance da aeronave, o indicador recebe o valor automática de 10, o que condicionaria um aumento elevado nos custos operacionais para tal rota, evitando com que algum avião seja alocado para uma viagem cuja distância entre *hub* e *spoke* seja maior que as possibilidades do veículo. O passo seguinte foi dedicado a instituição do fator de ocupação *load factor* para cada uma das rotas, assim como a instituição anterior, a explicação concreta do ponto encontra-se na Subseção [3.5.3](#).

Em seguida, são definidos os parâmetros M e K, o primeiro deles representa a quantidade máxima de viagens que uma determinada aeronaves pode fazer para atender uma única rota e o segundo é referente a quantidade de aeronaves que podem ser alocadas de forma simultânea. Desta forma, todos os parâmetros necessários para a resolução do problema de PO estão definidos.

As próximas fases são dedicadas a pesquisa operacional propriamente dita, o primeiro ponto é a criação do problema propriamente dito. Isso é realizado por meio do

comando *LpProblem* que realiza a definição de um problema do tipo LP. Devem ser estipulados no comando, o nome dados ao problema e o objetivo do modelo, neste caso, maximizar o lucro. Devem também ser definidas as variáveis por meio do comando *LpVariable*. No modelo estudado, são consideradas variáveis o número de passageiros alocados em cada aeronaves para cada rota e também o fluxo para cada aeronave em cada rota, bem como as variáveis binárias usadas para o auxílio nas restrições.

A etapa seguinte é destinada para a definição da função objetivo do modelo, que encontra-se registrada na Equação 3.1, e posteriormente todas as restrições do problema. Por fim, encontra-se na etapa de resolução do modelo propriamente dito, para os casos em que o problema é muito complexo do ponto de vista computacional, tem-se a opção de inserir o intervalo de otimalidade por meio do comando *GapRel*, assim como explicitado na 3.5.1. Caso o cenário analisado seja mais simples é possível utilizar o comando de resolução sem nenhum intervalo, *solver()*. Por fim, a parte final do algoritmo tem objetivo de recolher os dados do resultado em uma planilha.

4 Resultados e Discussão

O capítulo a seguir apresenta os resultados obtidos durante a pesquisa, bem como as análises dos resultados e sugestões de pontos a serem melhorados. O capítulo é dividido entre os seguintes tópicos: seleção de aeronaves, resultados preliminares e verificação do modelo, impactos do intervalo de otimalidade, alterações no modelo matemático e resultados definitivos.

4.1 Resultados preliminares e verificação do modelo

A fim de entender melhor como o modelo matemático funciona na prática, optou-se pelo desenvolvimento analítico do problema de PO para um cenário hipotético reduzido. Assim, foram escolhidas 3 rotas dentre as 710 disponíveis na base de Dados Hórus (Horus, 2022) e 2 aeronaves dentre as consideradas. Uma melhor descrição do cenário pode ser encontrada na Tabela 5.

Tabela 5 – Detalhamento do cenário reduzido.

Rotas			
Hub	Spoke	Demanda	Distância
284	49	700	250
278	311	45	2100
53	34	370	926
Aeroplanos			
Modelo	Assentos	Alcance [km]	Tipo
Turboélice 3	75	1600	Turboprop
Jato 3	145	6800	Turbofan

Para que a verificação seja realizada de forma correta, optou-se por seguir os passos do algoritmo utilizado para a resolução da malha aérea completa, o código encontra-se no Apêndice A. Assim sendo, o próximo passo foi o cálculo da tarifa por rota, seguindo a Equação 3.11 e os resultados podem ser encontrados na Tabela 6.

Tabela 6 – Variação da tarifa de acordo com a distância entre um *hub* h e um *spoke* j .

Hub	Spoke	Distância	Tarifa [Dólar Americano]
284	49	250	358,75
278	311	2100	624,52
53	34	926	504,57

Em seguida, faz-se necessário calcular o indicador CASK que, assim como explicitado nas Equações 3.9 e 3.10, é variável de acordo com a distância e o número de assentos

disponível, como considerado neste modelo matemático. Assim sendo, foi calculado um valor de CASK para cada rota a ser voada para as duas aeronaves: turboélice 3 e jato 3. Os resultados encontram-se na Tabela 7.

Tabela 7 – Variação de CASK de acordo com a rota e a aeronave alocada.

Hub	Spoke	CASK (Turboélice 3)	CASK (Jato 3)
284	49	0,08	0,10
278	311	10,00	0,04
53	34	0,07	0,06

É interessante observar que para a rota com *hub* 278 e *spoke* 311 o valor de CASK é indicado como 10,00. Isso ocorre pois a aeronave Turboélice 3 possui alcance máximo de 1600 km enquanto a distância a ser voada seria de 2100 km, o que torna o avião inviável para a operação. O modelo força o valor suficientemente alto de CASK para induzir o modelo a não alocação, assim como explicitado na Subseção 3.5.4. O próximo passo foi a atribuição do fator de ocupação (*load factor*), seguindo os critérios estabelecidos na Subseção 3.5.3 e explicitados na Tabela 8.

Tabela 8 – Atribuição de *Load factor* de acordo com cada rota.

Hub	Spoke	Distância [km]	Load Factor
284	49	250	0,5
278	311	2100	0,25
53	34	370	0,5

Para este cenário reduzido, optou-se para que o valor máximo de viagens (parâmetro M) fosse de 10 e que o número de aeronaves que podem ser alocados de forma simultânea é de 2 (parâmetro K). Desta forma, todos os parâmetros prévios necessários estão definidos e, assim, é possível partir para a solução do Problema de Alocação de Frota propriamente dito. Aplicando a Equação 3.1 ao cenário em questão passa-se a ter como função objetivo a Equação 4.1.

$$\begin{aligned}
 L_{Max} = & 358,75(P_{284,49,T} + P_{284,49,J}) + 624,52(P_{278,311,T} + P_{278,311,J}) + \\
 & 504,57(P_{53,34,T} + P_{53,34,J}) - \\
 & 75(20F_{284,49,T} + 21000F_{278,311,T} + 64,82F_{53,34,T}) - \\
 & 145(20F_{284,49,J} + 21000F_{278,311,J} + 64,82F_{53,34,J}) \quad (4.1)
 \end{aligned}$$

Após a definição da função objetivo, tem-se o apontamento das restrições. A primeira delas, Equação 3.2, aplicada ao cenário analisado para verificação, geraria as seguintes equações.

$$\begin{aligned}
P_{284,49,T} + P_{284,49,J} &= Q_{284,49} \\
P_{284,49,T} + P_{284,49,J} &= 700
\end{aligned} \tag{4.2}$$

$$\begin{aligned}
P_{278,311,T} + P_{278,311,J} &= Q_{278,311} \\
P_{278,311,T} + P_{278,311,J} &= 45
\end{aligned} \tag{4.3}$$

$$\begin{aligned}
P_{53,34,T} + P_{53,34,J} &= Q_{53,34} \\
P_{53,34,T} + P_{53,34,J} &= 370
\end{aligned} \tag{4.4}$$

$$\begin{aligned}
Q_{284,49} + Q_{278,311} + Q_{53,34} &= Q_{total} \\
700 + 45 + 370 &= 1115
\end{aligned} \tag{4.5}$$

Em seguida, é possível desenvolver as equações específicas para o cenário trabalhado para a segunda restrição, mostrada na Equação 3.3.

$$P_{284,49,T} \leq 75F_{284,49,T} \tag{4.6}$$

$$P_{278,311,T} \leq 75F_{278,311,T} \tag{4.7}$$

$$P_{53,34,T} \leq 75F_{53,34,T} \tag{4.8}$$

$$P_{284,49,J} \leq 145F_{284,49,J} \tag{4.9}$$

$$P_{278,311,J} \leq 145F_{278,311,J} \tag{4.10}$$

$$P_{53,34,J} \leq 145F_{53,34,J} \quad (4.11)$$

Quanto a próxima restrição, referente a Equação 3.4, tem-se:

$$P_{284,49,T} \geq 37,5F_{284,49,T} \quad (4.12)$$

$$P_{278,311,T} \geq 18,75F_{278,311,T} \quad (4.13)$$

$$P_{53,34,T} \geq 37,5F_{53,34,T} \quad (4.14)$$

$$P_{284,49,J} \geq 72,5F_{284,49,J} \quad (4.15)$$

$$P_{278,311,J} \geq 36,25F_{278,311,J} \quad (4.16)$$

$$P_{53,34,J} \geq 72,5F_{53,34,J} \quad (4.17)$$

A restrição seguinte, evidenciada pela Equação 3.5, tem-se:

$$F_{284,49,T} \leq 10 \cdot BIN1_{284,49,T}; \quad (4.18)$$

$$F_{278,311,T} \leq 10 \cdot BIN1_{278,311,T}; \quad (4.19)$$

$$F_{53,34,T} \leq 10 \cdot BIN1_{53,34,T}; \quad (4.20)$$

$$F_{284,49,J} \leq 10 \cdot BIN1_{284,49,J}; \quad (4.21)$$

$$F_{278,311,J} \leq 10 \cdot BIN1_{278,311,J}; \quad (4.22)$$

$$F_{53,34,J} \leq 10 \cdot BIN1_{53,34,J}; \quad (4.23)$$

Partindo para a restrição seguinte, Equação 3.6, tem-se:

$$\begin{aligned} R_T &\geq D_{284,49} \cdot BIN1_{284,49,T} \\ 1600 &\geq D_{250} \cdot BIN1_{284,49,T} \end{aligned} \quad (4.24)$$

$$\begin{aligned} R_T &\geq D_{278,311} \cdot BIN1_{278,311,T} \\ 1600 &\geq D_{2100} \cdot BIN1_{278,311,T} \end{aligned} \quad (4.25)$$

$$\begin{aligned} R_T &\geq D_{53,34} \cdot BIN1_{53,34,T} \\ 1600 &\geq D_{926} \cdot BIN1_{53,34,T} \end{aligned} \quad (4.26)$$

$$\begin{aligned} R_J &\geq D_{284,49} \cdot BIN1_{284,49,J} \\ 6800 &\geq D_{250} \cdot BIN1_{284,49,J} \end{aligned} \quad (4.27)$$

$$\begin{aligned} R_J &\geq D_{278,311} \cdot BIN1_{278,311,J} \\ 6800 &\geq D_{2100} \cdot BIN1_{278,311,J} \end{aligned} \quad (4.28)$$

$$\begin{aligned} R_J &\geq D_{53,34} \cdot BIN1_{53,34,J} \\ 6800 &\geq D_{926} \cdot BIN1_{53,34,J} \end{aligned} \quad (4.29)$$

A penúltima restrição, Equação 3.7, tem-se para o problema reduzido:

$$F_{284,49,T} \leq 10 \cdot BIN2_T \quad (4.30)$$

$$F_{278,311,T} \leq 10 \cdot BIN2_T \quad (4.31)$$

$$F_{53,34,T} \leq 10 \cdot BIN2_T \quad (4.32)$$

$$F_{284,49,J} \leq 10 \cdot BIN2_J \quad (4.33)$$

$$F_{278,49,J} \leq 10 \cdot BIN2_J \quad (4.34)$$

$$F_{53,34,J} \leq 10 \cdot BIN2_J \quad (4.35)$$

Por fim, a partir da Equação 3.8, tem-se:

$$BIN2_T + BIN2_J \geq 2; \quad (4.36)$$

Percebe-se então que ao escolher uma base de dados com 2 aeronaves passíveis de alocação e e 3 rotas para serem voadas gerou um número de 20 variáveis, sendo 6 delas referentes ao número de passageiros alocados em uma determinada aeronave para voar determinada rota, 6 referentes ao quantidade de vezes que cada aeronave deve ser alocada para uma mesma rota a fim de transportar o número total de passageiros destinados e 8 variáveis binárias. Fazendo uma analogia ao cenário completo gerado a partir da base de dados do Projeto Hórus, em que existem 710 rotas e 6 aeronaves passivas de alocação, quando utilizamos o parâmetro K com o valor de 5, seriam geradas mais de 8.500 variáveis, além da quantidade ainda maior de equações de restrições, o que explica o motivo da dificuldade em encontrar uma solução ótima sem o auxílio de ferramentas de redução computacional.

Já para a resolução do problema, optou-se por, inicialmente, investigar a situação referente a rota 2, que liga o *hub* 278 ao *spoke* 311. Para isso, o primeiro passo foi analisar a Equação 4.29, percebe-se que para que a inequação seja satisfeita é necessário que o valor do binário presente seja 0. A partir disso, parte-se para analisar a Equação 4.31 e infere-se também que o fluxo de viagens da aeronave turboélice 3 na rota analisada também deve ser zero. Logo, é possível analisar a restrição número 1, Equação 4.3, infere-se que o número de passageiros transportado pela aeronave jato 3 para esta rota deve ser, obrigatoriamente, igual a demanda total, ou seja, 45 passageiros, que podem ser transportados em uma única viagem, logo o fluxo desta aeronave nesta rota será equivalente a 1.

Após a resolução do problema para a primeira rota analisada, parte-se para a resolução da primeira rota, responsável por ligar o *hub* 284 ao *spoke* 49. Para esta rota, parte-se da análise da função que calcula os lucros por cada rota. Entende-se que, desconsiderando as duas variáveis (passageiros transportados e fluxo de viagens), tem-se, dentre

os valores constantes previamente determinados, que a tarifa contribui para a receita, enquanto CASK, distância e quantidade de assentos disponíveis fazem parte das despesas. Nesta rota, a aeronave jato 3 gera mais despesas que a aeronave turboélice 3, devido ao maior CASK e maior número de assentos disponíveis (Tabelas 7 e 5). Além disso, levando em consideração a curva de CASK de cada aeronave, evidenciada na Figura 8, infere-se que o avião do tipo *turboprop* é a escolha mais vantajosa para alocação devido a distância de apenas 250 km entre *hub* e *spoke*. Logo, infere-se que o número de passageiros transportados e o fluxo de viagens realizadas pelo jato 3 deve ser 0.

Assim, entende-se que o número de passageiros transportados pelo turboélice 3 para a rota analisada deve ser igual a demanda total do trajeto de 700 pessoas. Partindo da determinação de que a quantidade de assentos disponíveis na aeronave é de 145 pessoas, infere-se que o fluxo de viagens a ser realizado deve ser de 10, já que as variáveis analisadas devem ser números inteiros. Assim, é resolvido o cenário para a rota de *hub* 284 ao *spoke* 49.

Por fim, é necessário analisar o caso da terceira rota, do *hub* 53 ao *spoke* 34. Neste caso, observa-se que a diferença existente entre os valores de CASK para a mesma rota considerando diferentes aeronaves é mínima (Tabela 7). Assim sendo, a opção de alocação não é tão intuitiva. Observa-se que, devido à semelhança no indicador CASK, a quantidade de assentos disponíveis passa a ser um maior contribuinte para a escolha de alocação. A questão da proximidade do valor do CASK fica evidenciado na Figura 8, onde se percebe que a distância a ser voada encontra-se praticamente no limite entre qual das duas aeronaves seria mais vantajoso alocar. Assim sendo, passa-se a analisar as demais restrições disponíveis.

Pela primeira restrição sabe-se que o total de passageiros a serem alocados, contando as duas aeronaves deve ser de no máximo 370 pessoas. Para a continuação da análise, é feita uma distinção entre os cálculos para cada uma das aeronaves. Iniciando pelo turboélice 3, tem-se, a partir das restrições citadas previamente, a Equação 4.37:

$$37, 5F_{53,34,T} \leq P_{53,34,T} \leq 75F_{53,34,T} \quad (4.37)$$

Quanto ao fluxo de viagens, também é sabido que o número mínimo de viagens que pode ser realizadas pela aeronave em questão é 0 (caso todos os passageiros sejam alocados para o jato 3) e o valor máximo seria 5 (caso todos os passageiros da rota sejam alocados para o turboélice 3). Portanto tem-se a Equação 4.38:

$$0 \leq F_{53,34,T} \leq 5 \quad (4.38)$$

Já quanto a análise do jato 3, a seguinte situação é imposta (Equação 4.39):

$$72, 5F_{53,34,J} \leq P_{53,34,J} \leq 145F_{53,34,T} \quad (4.39)$$

Enquanto as restrições do fluxo deve estar entre zero (para o caso em que nenhum jato 3 seja alocado para o transporte da rota) e 3 (caso todos os passageiros da rota sejam transportados pelo jato 3), como mostrado na Equação 4.40.

$$0 \leq F_{53,34,J} \leq 3 \quad (4.40)$$

A partir dessa situação e, considerando que quanto mais passageiros forem alocados para um mesmo voo maior será a receita, sem haver nenhuma alteração nas despesas, tem-se 3 possíveis cenários para a escolher. O primeiro seria alocar todos os passageiros em 5 voos de turboélice 3 (cenário 1), a segunda opção seria alocar parte dos passageiros em 3 voos de turboélice 3 e mais 1 voo de jato 3 (cenário 2) e, por fim, alocar todos os passageiros em três viagens de jato 3. O lucro resultante de cada um dos cenários pode ser encontrado na Tabela 9.

Tabela 9 – Variação de lucro de acordo com a mudança no número de vezes que uma determinada aeronave foi alocada.

Fluxo Turboélice 3	Fluxo Jato 3	Lucro [Dólar Americano]
5	0	162.383,4
3	1	164.050,2
0	3	162.522,3

Após a análise da Tabela 9, entende-se que a melhor alocação possível é alocar 3 viagens para o turboélice 3 e 1 para o jato 3. Sendo assim, tem-se que devem ser transportados 225 passageiros com o turboélice 3 e 145 com o jato 3. Considerando os valores das alocações atribuídas às três rotas analisadas, o lucro a ser obtido neste cenário teste, deve ser de U\$ 416.098,67.

Com a intenção de verificação de resultados do algoritmo e também de checagem do código desenvolvido em Python, para ser utilizado para resolução do cenário completo com as 710 rotas disponibilizadas pela matriz Hórus de origem e destino, o mesmo cenário teste reduzido foi utilizado como dado de entrada para o algoritmo e os resultados podem ser encontrados na Tabela 10.

Tabela 10 – Resultado do cenário teste utilizado o código em Python.

Origem	Destino	Passageiros	Aeronaves	Fluxo	Lucro por Rota
284	49	700	Turboelice_3	10	236.768,66
278	311	45	Jato_3	1	14.935,01
53	34	225	Turboelice_3	3	99.515,69
53	34	145	Jato_3	1	65.075,31
Total de Passageiros Alocados		1115	Lucro Total Gerado		416.294,67

Por fim, verifica-se que os resultados obtidos tanto por meio da resolução analítica quanto com auxílio do algoritmo foram iguais, o que permite concluir que o algoritmo

está verificado e pode ser utilizado para cenários maiores com auxílio de ferramentas computacionais, já que nestes casos seria inviável resolver de forma analítica. Desta forma, a mesma rotina de programação será utilizada para as próximas seções do trabalho.

4.2 Impactos do intervalo de otimalidade

Assim como demonstrado na seção 3.4, a estratégia utilizada para o alcance da resolução completa do modelo foi a exploração do intervalo de otimalidade na solução do problema. Neste caso, o uso do comando *GapRel* foi usado para encontrar uma solução adequada que fosse o mais próxima possível da solução ótima. A forma encontrada para a comprovação da redução da complexidade computacional foi a medição do tempo de simulação necessário para a obtenção de resultados. Portanto, a próxima seção será dedicada à comparação entre o tempo que cada modelo computacional levou para ser solucionado, tendo como parâmetro principal a exploração do *gap* de otimalidade.

Inicialmente, optou-se por utilizar um intervalo de otimalidade de 0.01, ou seja, os resultados obtidos estão registrados na Tabela 11 e são 99% próximos da solução ótima. Nos registros constam a quantidade de aeronaves disponíveis para alocação, por meio da mudança do parâmetro *K*, o lucro total da malha aérea em dólar americano e por fim, o tempo de simulação gasto para a obtenção da solução.

Tabela 11 – Resultados do cenário teste, *gap* = 0.01

Quantidade de Aeronaves	Lucro (Dólar Americano)	Tempo de simulação [s]
K=1	554.216.365.47	1,75
K=2	554.253.825.99	0,97
K=3	554.224.161.83	0,99
K=4	554.224.161,83	1,66
K=5	554.224.161,83	0,94

Para a simulação com *K*=1, a aeronave considerada foi o "jato 3", ou seja, um modelo análogo às aeronaves *Boeing* 737-700 ou *Airbus* A319. Para a simulação com *K* = 2, foram alocadas as aeronaves "jato 3" e "jato 1". Já com a mudança para os parâmetros *K* = 3, os três modelos de aeronaves do tipo jato foram utilizadas. Nas simulações com *K* = 4 e *K* = 5 as mesmas 3 aeronaves foram alocadas.

Por razões comparativas, o modelo também foi simulado sem utilizar o *gap* de otimalidade. Com as configurações computacionais descritas na tabela 4, foi possível encontrar resultados apenas para o modelo com parâmetro *K* = 1, cujo resultado de pesquisa foi gerado em aproximadamente 24 segundos. Entretanto, a partir da inserção de mais de uma aeronave para a alocação no modelo, o tempo de simulação ultrapassou as 6 horas, sem encontrar uma solução ótima. Assim, considerou-se que a convergência de solução não foi atingida.

Ou seja, comparando os resultados obtidos pelas diferenças entre os tempos de simulação para os cenários com e sem o *gap* de otimalidade deste trabalhos e os resultados apontados por [Pereira \(2021\)](#), é possível confirmar que a complexidade computacional do algoritmo é o principal problema do modelo, mas pode ser superada a partir do momento em que o intervalo de otimalidade é inserido.

Verifica-se que a primeira parte do objetivo principal da pesquisa foi cumprido, porém a partir da análise do modelo com a variação do parâmetro K , foi possível identificar que os resultados obtidos ainda não refletem algo condizente com a realidade. Existem alguns pontos a serem considerados, sendo eles: a não alocação de nenhuma aeronave do tipo turboélice e a repetição da solução para os cenários de de $K = 3, 4$ e 5 .

A partir dos resultados da Tabela 11, optou-se por testar os limites do algoritmo quanto ao intervalo de otimalidade e como este pode modificar os cenários finais obtidos. Desta forma, foram realizados testes alterando o valor do intervalo de otimalidade para 0.001 e posteriormente para 0.0005, os resultados estão registrados nas Tabelas 12 e 13.

Tabela 12 – Resultados da variação do valor de intervalo de otimalidade.

Intervalo de otimalidade de 0.001		
Quantidade de aeronaves	Lucro [Dólar Americano]	Tempo de simulação [s]
$K = 1$	554.003.892,77	12,0
$K = 2$	554.541.473,43	27,0
$K = 3$	554.258.946,27	18,3
$K = 4$	554.258.946,27	19,0
$K = 5$	554.258.946,27	17,9
Intervalo de otimalidade de 0.0005		
Quantidade de aeronaves	Lucro [Dólar Americano]	Tempo de simulação [s]
$K = 1$	554.307.858,54	15,9
$K = 2$	554.541.473,43	27,0
$K = 3$	554.529.634,26	26,5
$K = 4$	554.529.634,26	26,2
$K = 5$	554.529.634,26	23,5

Tabela 13 – Variação de aeronaves alocadas.

Intervalo de otimalidade = 0.001				
$K = 1$	$K = 2$	$K = 3$	$K = 4$	$K = 5$
		Jato 1	Jato 1	Jato 1
Jato 3	Jato 1	Jato 2	Jato 2	Jato 2
	Jato 3	Jato 3	Jato 3	Jato 3
Intervalo de otimalidade = 0.0005				
$K = 1$	$K = 2$	$K = 3$	$K = 4$	$K = 5$
		Jato 1	Jato 1	Jato 1
Jato 3	Jato 3	Jato 2	Jato 2	Jato 2
	Jato 2	Jato 3	Jato 3	Jato 3

De acordo com os resultados mostrados nas Tabelas 12 e 13, verifica-se que, apesar das mudanças no lucro e também nas aeronaves alocadas, como consequência da variação do intervalo de otimalidade, os dois principais problemas observados anteriormente ainda permanecem. Ainda há uma repetição da solução a partir da alocação de 3 aeronaves e também nenhum turboélice ainda foi alocado. Ou seja, apesar do problema de PO não ser mais demasiadamente complexo, o modelo matemático ainda não representa o suficiente da realidade da indústria de transporte aéreo. Conclui-se que então que o *gap* de otimalidade é uma solução viável para a redução computacional, porém a representação matemática ainda precisa ser melhor desenvolvida.

4.3 Alterações no modelo matemático

A primeira alteração matemática realizada está mais relacionada com as possibilidades que surgem a partir da redução da complexidade computacional. A partir do uso do intervalo de otimalidade passa-se a ser possível utilizar a malha aérea completa disponibilizada pela base de dados, não havendo mais necessidade de utilizar a chamada função ganho criada por Pereira (2021), o que deixa o modelo matemático mais próximo das demandas identificadas pelo próprio projeto Hórus (Horus, 2022).

Assim como citado na Seção 3.5.2, a não inclusão de aeronaves do tipo turboélice não é algo a se estranhar após a análise da modelagem pré-existente. Isso acontece pois a formulação existente não foi capaz de contemplar as características operacionais de todas as seis aeronaves. Acredita-se que ao utilizar a mesma formulação para todos os tipos de veículos aéreos, os jatos acabam ficando em posição de vantagem para a alocação devido a serem aeronaves com um número maior de assentos e com maior alcance consequentemente passível de maiores receitas, sendo alocadas em todas as rotas.

Entretanto, esse resultado é pouco intuitivo e não contempla a realidade. Isto se justifica porque os assentos disponíveis e a distância voada não são os únicos fatores que afetam o CASK em casos reais. Vários outros recursos podem afetar o custo operacional de uma aeronave, como a quantidade de pessoas alocadas e consequentemente o peso de determinada aeronave, a quantidade de combustível necessária, o preço de manutenção e do combustível e alguns outros fatores. Sendo assim, entende-se que os resultados obtidos nas primeiras simulações são um reflexo da modelagem do indicador CASK.

Entende-se também que o método de obtenção do indicador CASK é bastante complexo, podendo até mesmo ser considerado um ponto de vantagem estratégica pelas companhias aéreas, logo, tratado por muitas delas como segredo industrial. Sendo assim, considera-se que a estimação do cálculo de CASK considerando todos os fatores que o afetam seria bastante trabalhoso, senão inviável.

Sendo assim, optou-se por continuar utilizando apenas as variáveis de assentos

disponíveis e distância voada como variáveis para compor a função de cálculo do indicador, porém, desta vez fazendo a separação entre os diferentes tipos de modelo. Ou seja, foram estipuladas diferentes curvas de CASK por tipo de veículo. Logo, foram inseridas em uma nova versão do algoritmo duas equações, representadas na Seção 3.5.2, uma que calcula o indicador para aeronaves turboélice e outra para aeronaves do tipo *turbofan*. As Figuras 7 e 8 apresentam representações gráficas entre as curvas de CASK para cada aeronave antes e após a mudança na formulação. Nessa representação em particular, não foram levados em consideração os alcances de cada uma das aeronaves, foram considerados apenas o número total de assentos e a distância a ser voada. As demais restrições são aplicadas ao modelo posteriormente nas restrições do modelo de PO.

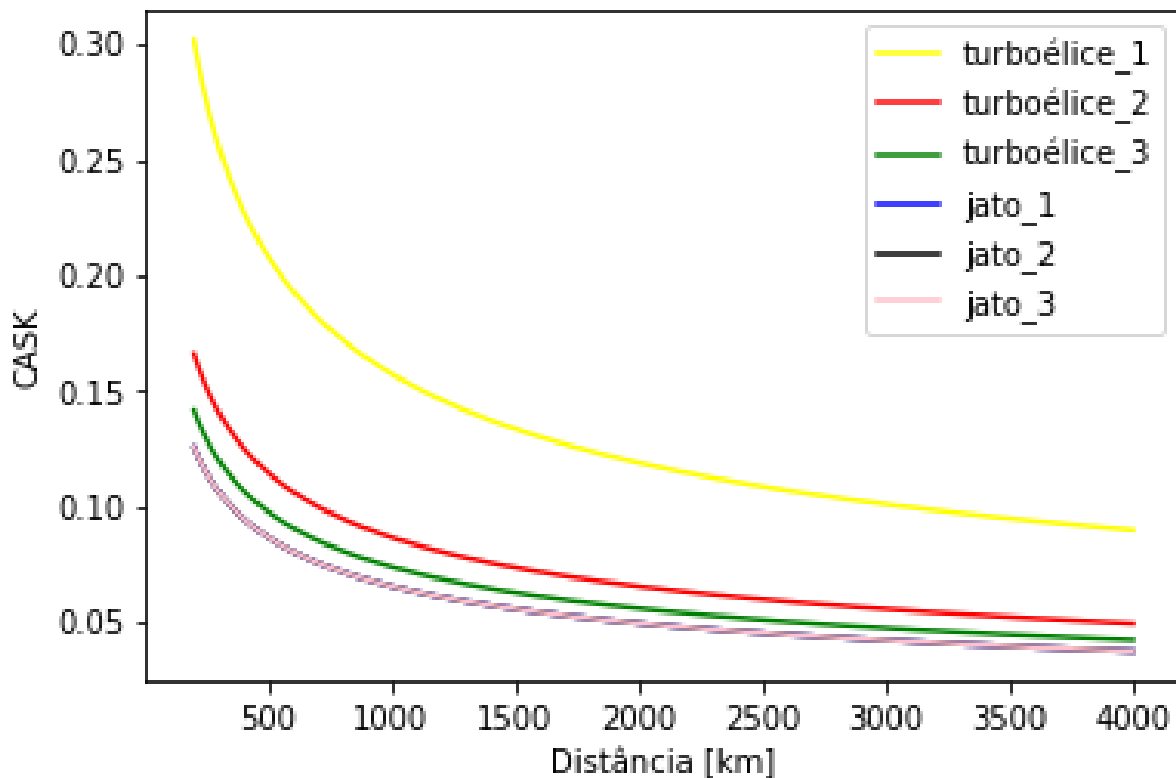


Figura 7 – Variação da curva de CASK de acordo com as diferentes aeronaves antes da alteração nos valores de CASK.

Por meio das Figuras 7 e 8 é encontrada uma verificação da razão que faz com que as aeronaves do tipo turboélices não tenham sido escolhidas inicialmente. Assim como pode ser identificado nas curvas mostradas na Figura 7. Observa-se que independente da distância voada, neste caso, os aviões a jato sempre têm um custo operacional menor do que os demais modelos. Isto vai em desencontro do que acontece em casos reais, visto que, para distâncias menores, o mais intuitivo é que aviões turboélice sejam mais vantajosas do ponto de vista operacional.

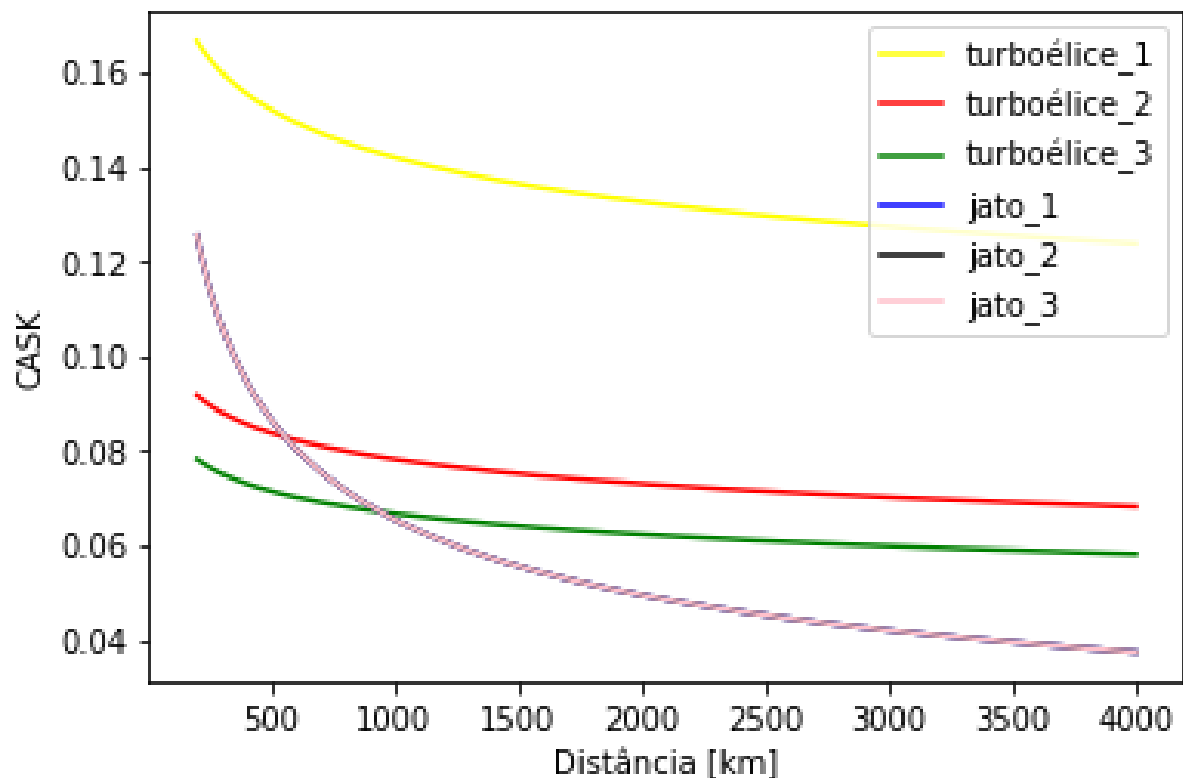


Figura 8 – Variação da curva de CASK de acordo com as diferentes aeronaves após a alteração nos valores de CASK.

Já quando a Figura 8 passa a ser o objeto de análise, percebe-se que com as mudanças na formulação do indicador observado (Equações 3.9 e 3.10) já é possível considerar que as aeronaves do tipo turboélice passam a serem competitivas até certas distâncias. Como demonstrado na figura, até a distância de 500 km as aeronaves Turboélice 2 e 3 são mais vantajosas para alocação que os jatos, por exemplo, assim como costuma acontecer na alocação de recursos na indústria de transporte aéreo.

Observa-se ainda que o único modelo que continua não sendo atrativo para alocação com a nova formulação é o Turboélice 1. Entretanto, este cenário não representa grande discrepância com a realidade, visto que este modelo, representativo de um Cessna Caravan 208, realmente possui custos operacionais altos. No entanto, é uma aeronave capaz de voar um número menor de pessoas (nove), além de conseguir realizar pousos e decolagens em aeroportos de menor porte. Ou seja, é uma aeronave que é pouco competitiva quando pensada para uma variedade de rotas. Apesar do limite da infraestrutura aeroportuária não ter sido considerado no modelo deste trabalho, no mundo real em localidades de difícil acesso, esta aeronave Turboélice 1 passa a ser, por vezes, até mesmo a única opção.

Outro ponto observado nos resultados dos primeiros testes é a alocação de ae-

ronaves com uma grande disponibilidade de assentos para rotas com baixa demanda, o que reflete em rotas com baixíssima ocupação, algo que não é necessariamente vantajoso do ponto de vista comercial. Sendo assim, optou-se pela inclusão de um fator mínimo de ocupação para que uma aeronave seja escolhida para cada rota. Ficou estipulado que para que uma aeronave fosse alocada em uma rota de com distância de até 1300 km, pelo menos 50% de sua capacidade total de assento seja ocupada, enquanto para distâncias maiores, pelo menos 30% dos lugares na aeronave sejam ocupados. Essa restrição serve para impedir que existam rotas com baixa demanda alocadas às maiores aeronaves incluídas no modelo.

4.4 Resultados Definitivos

Após essas mudanças partiu-se para as simulações em definitivo do novo modelo matemático, explicado na Seção 3.3 e cujo código em Python encontra-se no Apêndice A. Foram realizadas simulações com a quantidade de aeronaves disponíveis para alocação variando de 1 até 5, com o *gap* de otimalidade variando de forma dependente da complexidade computacional exigida por cada modelo.

Para a simulação com a alocação de apenas uma aeronave disponível para alocação, tem-se um modelo pouco complexo, logo, optou-se por não utilizar o intervalo de otimalidade, já que existe a possibilidade de obtenção da solução ótima. Entretanto, para que fosse possível gerar um resultado adequado, foi necessário desconsiderar o fator de ocupação *loadfactor*. A malha aérea utilizada como base de dados para a resolução do problema de PO possui uma variação de distâncias e demandas muito grande, sendo 711 rotas com distâncias variando entre 250 km e 3811 km de distância entre origem e destino, com demandas variantes entre 27 e 32.844 passageiros semanais.

Para os cenários correspondentes ao número de aeronaves disponíveis para alocação variando entre duas, quatro e cinco aeronaves ($K = 2, 4$ e 5) foi possível resolver o problema de PO utilizando um intervalo de otimalidade de 0.0005, logo o resultado obtido é uma solução 0,05% distante da solução ótima. Já para o caso do cenário com três aeronaves passíveis de alocação ($K = 3$), foi necessário que o *gap* fosse aumentado para 0.005. Isso porque o grau de complexidade computacional não depende exclusivamente do número de aeronaves passíveis de alocação, mas também da estrutura do problema. Isso acontece pois com um número menor de variedades disponíveis, passa a haver um limite menor entre as rotas que são consideradas mais adequadas para cada uma das aeronaves. Para todos os cinco modelos utilizados foram alocados um total de 890.646 passageiros. Na tabela 14 é possível verificar um resumo dos resultados obtidos de acordo com os cenários analisados.

Além dos valores exibidos ainda é importante entender a participação de cada aero-

Tabela 14 – Resumo dos resultados obtidos a partir das simulações definitivas.

K	Intervalo de otimalidade	Aeronaves Alocadas	Lucro [Dólar Americano]
1	0	Jato 3	371.212.297,8
2	0,0005	Jato 1 Turboélice 2	368.643.211,1
3	0,005	Jato 1 Turboélice 2 Turboélice 3	371.769.572,6
4	0,0005	Jato 1 Jato 3 Turboélice 2 Turboélice 3	374.900.673,0
5	0,0005	Jato 1 Jato 2 Jato 3 Turboélice 1 Turboélice 3	374.961.012,1

naves nos cenários obtidos, assim sendo, a Tabela 15 reflete a porcentagem de participação de cada aeronave nos lucros obtidos bem como na alocação de passageiros.

Tabela 15 – Participação de cada aeronave na alocação de passageiros e nos lucros.

K	Aeronave Alocada	Participação na Demanda	Participação no Lucro
1	Jato 3	100%	100%
2	Jato 1	53,06%	57,94%
2	Turboélice 2	46,94%	42,05%
3	Jato 1	25,12%	29,32%
3	Turboélice 2	1,94%	1,85%
3	Turboélice 3	72,94%	68,83%
4	Jato 1	3,53%	4,00%
4	Jato 3	44,76%	49,54%
4	Turboélice 1	0,27%	0,22%
4	Turboélice 3	51,43%	46,23%
5	Jato 1	3,00%	3,39%
5	Jato 2	3,98%	4,44%
5	Jato 3	41,96%	46,25%
5	Turboélice 1	0,26%	0,21%
5	Turboélice 3	50,89%	45,70%

A partir da análise completa do modelo é possível perceber que em momentos distintos as seis aeronaves foram alocadas, o que demonstra a importância da variedade de frota em uma empresa aérea. Além disso, observa-se que a solução em que o parâmetro K é igual a 1, ou seja, há apenas uma única aeronave disponível para atender as 711 linhas aéreas diferentes, houve a alocação do jato 3, avião com o maior alcance e maior número

de assentos disponíveis. Logo, a partir da necessidade de suprir uma grande variedade de viagens, a escolha do algoritmo foi pelo veículo com o menor CASK, justamente por possuir um alcance maior e mais assentos disponíveis para a ocupação.

Já quanto à análise dos demais cenários é possível verificar que o aumento do número de aeronaves disponíveis também aumenta a especificidade com a qual o algoritmo faz a alocação. Isso é demonstrado por meio da escolha da aeronave Turboélice 1 (aeronave análoga ao Cessna 208 Caravan) que, apesar de possuir um custo operacional mais elevado, atende perfeitamente a demanda de uma pequena parcela de rotas (cerca de 0,2% dos cenários com $K = 4$ e $K = 5$), passando a ser o modelo mais indicado para alocação.

A escolha da aeronave Turboélice 1 não é imediatamente intuitiva, considerando a curva de CASK da aeronave. No entanto, observa-se que esta alocação tem uma baixa participação na alocação de passageiros e do lucro, tendo ocorrido possivelmente pelo fato da solução não ser ótima e ter ficado presa a um máximo local satisfatório (atendendo ao gap de otimalidade selecionado).

Partindo para a observação do cenário mais completo, com cinco aeronaves disponíveis para alocação ($K = 5$), obteve-se uma alocação de frota bastante semelhante às frotas de companhias brasileiras, como a Azul Linhas Aéreas que utiliza aeronaves análogas a quase todos os modelos genéricos retratados na simulação, apesar do veículo Embraer E-190 já não ser mais operado pela companhia. Ou seja, os resultados obtidos, a partir das mudanças realizadas ao modelo matemático e ao algoritmo, passam a ser satisfatórios, retornando um cenário minimamente realista.

Entretanto, outro ponto que deve ser analisado é a variação do lucro de acordo com a mudança na disponibilidade de aeronaves para alocação. Intuitivamente, espera-se que a medida com a qual o número de aeronaves passíveis de alocação aumenta, o lucro obtido na totalidade da operação deve aumentar, uma vez que as rotas passam a serem alocadas por aeronaves mais específicas para cada tipo de viagem, com uma melhor alocação de recurso e a maximização dos lucros obtidos.

No entanto, a partir da análise da Figura 9, verifica-se que há uma discrepância entre os resultados obtidos pela simulação e o resultado esperado na comparação do lucro entre os cenários em que $K = 1$ e $K = 2$, pois o segundo momento registrou um lucro menor que o resultado para apenas uma aeronave alocada. Do ponto de vista conceitual, esse resultado pode ser considerado insatisfatório ou irrealista, uma vez que, com o aumento da quantidade de veículos disponíveis, esperava-se que a seleção de aeronaves ocorresse de forma a utilizar os recursos de forma mais otimizada possível, gerando um maior lucro.

No entanto, assim como registrado no início desta seção, os cenários comparados foram simulados em condições diferentes, devido à implantação do intervalo de otimalidade apenas a partir da mudança do parâmetro K para 2 e também devido a ausência da

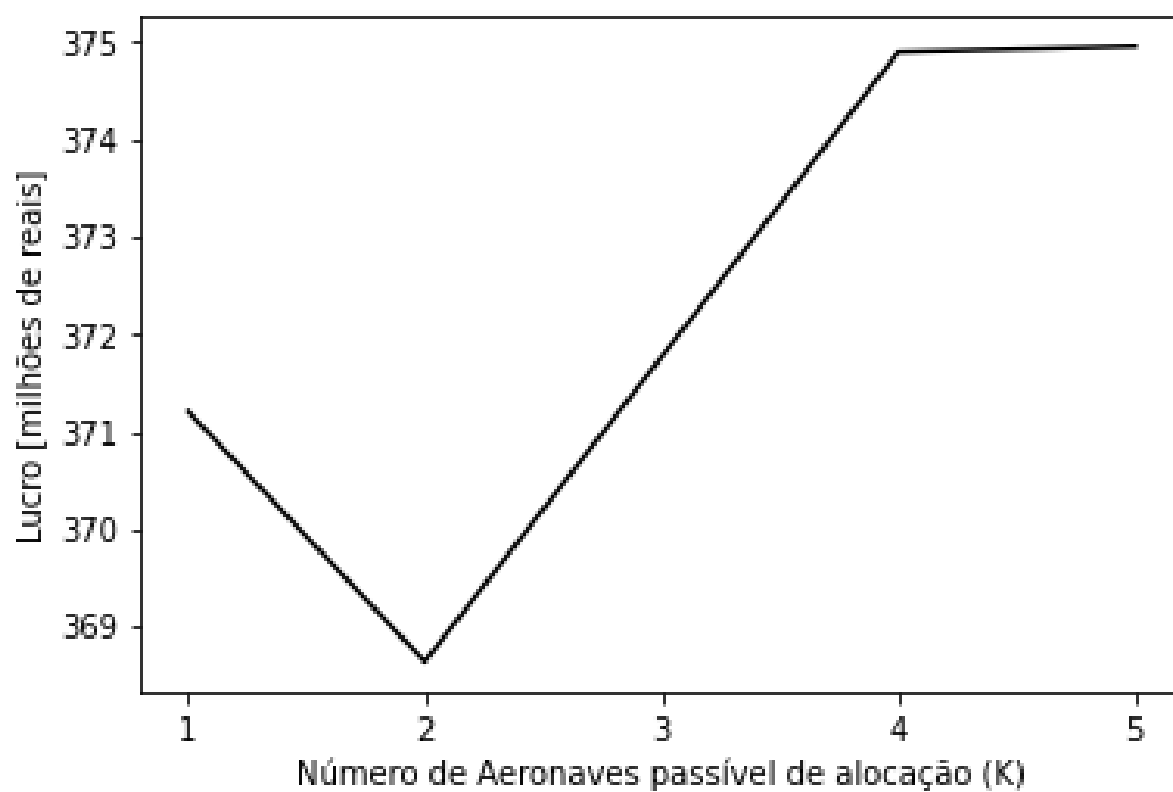


Figura 9 – Variação de lucro de acordo com as diferentes aeronaves

restrição de ocupação (*loadfactor*). Ou seja, enquanto o primeiro cenário apresenta uma solução ótima sem a restrição de *loadfactor*, o segundo apresenta um resultado para o problema apenas satisfatório (com gap de otimalidade) e com restrição do *loadfactor*, o que explica a inadequação conceitual entre os resultados simulados.

5 Conclusão

Ao final do trabalho verifica-se que os objetivos do trabalho foram alcançados. Com relação ao objetivo geral do projeto, este foi cumprido uma vez que foi possível gerar uma malha aérea que pode ser considerada ótima a partir de um modelo matemático capaz de alocar até cinco aeronaves diferentes de forma simultânea. Além disso, os cenários gerados por meio da simulação do algoritmo não trazem grandes discrepâncias com a realidade da indústria do transporte aéreo brasileiro, apesar de ainda haver pontos a serem aprimorados e melhor explicados, como a diminuição no lucro quando o número de aeronaves disponíveis para a alocação passa de um para dois.

Quanto aos dois primeiros objetivos específicos, foi possível, por meio da análise dos modelos anteriores e a partir da redução da complexidade computacional, apontar alterações capazes de fornecer resultados mais realistas ao modelo matemático. Sendo essas alterações: a eliminação da função ganho, permitindo a utilização da malha aérea original disponibilizada pelo projeto Hórus e a alocação de mais passageiros em diferentes rotas; a alteração na formulação do indicador CASK que refletiu em resultados mais adequados a indústria estudada; e também a inserção do fator de ocupação (*load factor*) ao modelo matemático, o que eliminou casos em que aeronaves muito grandes realizem voos com um número muito baixo de passageiros alocados.

O terceiro objetivo específico está relacionado às ferramentas que deveriam ser utilizadas na pesquisa, a Pesquisa Operacional, utilizando o método de Programação Inteira Linear Mista (MILP - *Mixed Integer Linear Programming*). Assim como explicitado nas seções anteriores, verifica-se que é possível elaborar um modelo matemático capaz de gerar uma malha aérea considerada ótima com os recursos de PO previamente determinados. Entretanto, faz-se necessário pontuar que o uso de ferramentas mais poderosas do ponto de vista computacionais (p.ex. metaheurísticas) poderiam retornar resultados mais adequados à realidade utilizando menos recursos, ou pelo menos de maneira mais precisa. Ou seja, apesar da comprovação de possibilidade, não é possível colocar os recursos utilizados no trabalho como os mais adequados para realizar a tarefa.

Desta forma, os resultados finais obtidos são suficientemente satisfatórios. Mesmo que ainda tenham sido apontados pontos de melhorias para o modelo, foi possível atingir os objetivos iniciais propostos. No entanto, entende-se que o modelo matemático e a pesquisa como um todo ainda podem ser aprimorados, com uma nova modelagem de aeronaves utilizadas pela indústria do transporte aéreo brasileiro atualmente, incluindo a diversidade de oferta de serviços e tarifas, com a inclusão de aeronaves do tipo eVTOL ao problema, bem como também com a verificação do algoritmo e a aplicação de novas

ferramentas, investigação de pontos a serem explorados no quesito de custos operacionais e cálculo de tarifas de modo a solucionar as divergências entre os resultados obtidos nas simulações e os esperados, aproximando ainda mais a representação do problema da realidade.

Referências

- AEROIN. *Alta no querosene leva Azul a reduzir voos; rota para Fernando de Noronha foi afetada*. 2022. Disponível em: <<https://aeroin.net/alta-no-querosene-leva-azul-a-reduzir-voos-rota-para-fernando-de-noronha-foi-afetada/?amp>>. Acesso em: 28 de abril de 2022. Citado 2 vezes nas páginas 25 e 41.
- AJAYI, T. The gap function: Evaluating integer programming models over multiple right-hand sides. Houston, USA, 2021. Citado na página 36.
- ANAC. *Demanda e oferta no mercado doméstico em 2021 tiveram recuo ante período pré-pandemia*. 2022. Disponível em: <<https://www.gov.br/anac/pt-br/noticias/2022/demanda-e-oferta-no-mercado-domestico-em-2021-tiveram-recuo-ante-periodo-pre-pandemia>>. Acesso em: 29 de Janeiro 2022. Citado 3 vezes nas páginas 25, 30 e 41.
- ANAC. *Registro de Serviços Aéreos*. 2022. Disponível em: <<https://sistemas.anac.gov.br/sas/downloads/view/frmDownload.aspx>>. Acesso em: 29 de Janeiro 2022. Citado 2 vezes nas páginas 41 e 42.
- BAZARGAN, M. *Airline Operations and Scheduling*. [S.l.]: Ashgate Publishing Limited, 2010. ISBN 978-0-7546-7900-4. Citado na página 30.
- BETTINI, H. F. de A. J. Um retrato da aviação regional no Brasil. *Acervo Científico do Núcleo de Estudos em Competição e Regulação do Transporte Aéreo (NECTAR)*., São José dos Campos, Brasil, n. 18, p. 18, 2007. Citado 4 vezes nas páginas 25, 27, 29 e 30.
- CAUCHICK, P. *Metodologia Científica para Engenharia*. [S.l.]: Grupo Editorial Nacional, 2019. ISBN 978-85-951-508-05. Citado na página 39.
- FREGNANI, J. A. T. G.; FERREIRA, N. dos S.; GRIEBELER, M. de C. Análise de custos logísticos do transporte aéreo regional. *Journal of Transport Literature*, v. 3, n. 2, p. 7–24, 2009. Citado na página 30.
- FREGNANI, J. A. T. G. et al. The impact on production and costs aer the acquisition of an airline by a competitor in Brazil: a Cobb-Douglas function application. *Revista Transporte*, v. 27, n. 1, p. 172–183, 2019. Citado na página 30.
- FÁVERO, L. P. *Análise de Dados: Modelos de Regressão com Excel®, Stata® e SPSS*. [S.l.]: Elsevier, 2015. Citado 2 vezes nas páginas 37 e 38.
- GOOGLE. *Google Colaboratory*. 2022. Disponível em: <https://colab.research.google.com/?utm_source=scs-index>. Citado na página 50.
- Horus. *Matriz Origem Destino*. 2022. Disponível em: <<https://horus.labtrans.ufsc.br/gerencial/?auth=s#MatrizOd>>. Acesso em: 24 de fevereiro 2022. Citado 7 vezes nas páginas 30, 31, 41, 42, 48, 55 e 65.
- INFRAESTRUTURA, M. da. *Governo Federal investe na expansão da aviação regional no Brasil*. 2022. Disponível em: <<https://www.gov.br/infraestrutura/pt-br/assuntos/noticias/2021/3/governo-federal-investe-na-expansao-da-aviacao-regional-no-brasil>>. Acesso em: 28 de abril de 2022. Citado na página 29.

JESUS, T. D. de. Definição da malha aérea e aeronaves para uma nova companhia aérea regional com hub em goiânia. Brasília, Brasil, p. 78, 2019. Citado 5 vezes nas páginas [27](#), [37](#), [40](#), [41](#) e [46](#).

LACHTERMACHER, G. *Pesquisa Operacional na Tomada de Decisões*. [S.l.]: Grupo Editorial Nacional, 2016. ISBN 978-85-216-304-94. Citado na página [34](#).

LEDERER, R. S. N. P. J. Airline network design. *Operations Research*, 1998. Citado 3 vezes nas páginas [15](#), [31](#) e [32](#).

LIEBERMAN, F. S. H. G. J. *Introdução à Pesquisa Operacional*. [S.l.]: AMGH Editora Ltda, 2013. ISBN 978-85-805-511-98. Citado 2 vezes nas páginas [34](#) e [35](#).

MEDAU, N. D. F. G. J. C. Alocação de aeronaves a voos considerando restrições operacionais, de manutenção e de desempenho das aeronaves. *Revista Transportes*, 2018. Citado na página [37](#).

MENDES, J. M. B. Comparação de modelos para o fleet assignment problem (problema de afetação de frotas). Lisboa, Portugal, 2019. Citado na página [36](#).

PEREIRA, G. H. de S. Estudo de metodologia de otimização para o problema de alocação de frota na aviação regional. Brasília, Brasil, p. 119, 2021. Citado 20 vezes nas páginas [15](#), [26](#), [27](#), [30](#), [31](#), [32](#), [37](#), [40](#), [41](#), [45](#), [46](#), [48](#), [49](#), [50](#), [51](#), [52](#), [64](#), [65](#), [87](#) e [89](#).

RODRIGUES, R. *Pesquisa Operacional*. [S.l.]: Soluções Educacionais Integradas, 2013. ISBN 9788595020054. Citado 3 vezes nas páginas [15](#), [33](#) e [34](#).

SILVA, L. R. et al. Análise da mobilidade interurbana no brasil por meio de uma matriz de origem destino de dados da telefonia móvel. *Geografia e Ordenamento do Território*, n. 22, p. 52–82, 2021. Citado na página [31](#).

TAHA, H. A. *Pesquisa Operacional*. [S.l.]: Pearson Education Inc, 2008. ISBN 978-85-7605-150-3. Citado na página [33](#).

APÊNDICE A – Código do Modelo Matemático Atualizado

```

%reset
!pip install —upgrade pulp
from pulp import *
import pandas as pd
import numpy as np
from pandas import ExcelWriter
#import time

#inicio = time.time()
#Function para calcular o pre o da passagem
def Tarifa(dist):
    return np.exp(-0.73953 * np.log(dist) + 4.44446)* dist

#Function para calcular o CASK TF e CASK TP
def Cask(dist, seats, tipo):
    if tipo == 'TF':
        return np.exp(1.728912 - 0.40453 * np.log(dist) -
            0.35671 * np.log(seats))
    elif tipo == 'TP':
        return np.exp(-0.481549 - 0.099212 * np.log(dist) -
            0.356708 * np.log(seats))
from google.colab import drive
drive.mount('/content/drive')
%cd /content/drive/My Drive/Colab Notebooks/
organizacao_resultados_sitraer/sitraer
# Leitura dos dados sobre aeronaves

DADOS3 = pd.read_excel('Dados_TCC_2022.xlsx', sheet_name = 'Aeronaves')
DADOS3.dropna(inplace=True)
DADOS3.reset_index(drop=True, inplace=True)
DADOS3.drop_duplicates(inplace=True)

```

```

AERONAVE = DADOS3[ 'AERONAVE' ]. astype( str )
N_ANV = len(AERONAVE)

ASSENTOS = DADOS3[ 'ASSENTOS' ]. astype( int )

ALCANCE = DADOS3[ 'ALCANCE' ]. astype( int )

TIPO = DADOS3[ 'TIPO' ]. astype( str )
#Leitura dos dados para constru o da malha

DADOS1 = pd.read_excel( 'Dados_TCC_2022.xlsx', sheet_name = 'Dados' )
DADOS1.dropna( inplace=True )
DADOS1.reset_index( drop=True, inplace=True )

HUB = DADOS1[ 'Hub' ]. astype( int )
HUBN = HUB.drop_duplicates()
numero_HUB = len(HUBN)
#HUB = HUB.values.tolist()

SPOKE = DADOS1[ 'Spoke' ]. astype( int )
SPOKEN = SPOKE.drop_duplicates()
numero_SPOKE = len(SPOKEN)
#SPOKE = SPOKE.values.tolist()

DEMANDA = round(DADOS1[ 'Demanda' ]). astype( int )
#DEMANDA = DEMANDA.values.tolist()

DISTANCIA = round(DADOS1[ 'Distancia' ]). astype( int )
# #DISTANCIA = DISTANCIA.values.tolist()

# # Defini o do n mero de vertices (edges)
N = len(DISTANCIA)
# C lculo de TARIFA por rota

DADOS2 = pd.DataFrame( columns=[ 'Passagem' ] )
for i in range(N):
    x = Tarifa( DISTANCIA[ i ] )
    DADOS2 = DADOS2.append( { 'Passagem': x }, ignore_index=True )

```



```

print((
    f "Hub({HUB[ i ]}) – Spoke({SPOKE[ i ]}):_Distancia=_="
    f "{DISTANCIA[ i ]} ,_Passagem=_={DADOS2.Passagem[ i ]: ,.2 f}"
))

DADOS2.reset_index(drop=True, inplace=True)
PASSAGEM = DADOS2[ 'Passagem' ]
#PASSAGEM = PASSAGEM.values.tolist()
# C lculo do CASK por aeronave e rota

DADOS4 = pd.DataFrame(columns=[ 'Cask' ])
for i in range(N_ANV):
    print( f 'CASK_da_Aeronave_{AERONAVE[ i ]}' )
    for j in range(N):
        if DISTANCIA[ j ] > ALCANCE[ i ]:
            x = 10
        else :
            x = Cask(DISTANCIA[ j ], ASSENTOS[ i ], TIPO[ i ])
        DADOS4 = DADOS4.append( { 'Cask': x }, ignore_index=True)
        print( f 'Hub({HUB[ j ]}) – Spoke({SPOKE[ j ]}):_Distancia=_='
            f '{DISTANCIA[ j ]} ,_CASK=_={x: ,.6 f}' )

DADOS4.reset_index(drop=True, inplace=True)
CASK = DADOS4[ 'Cask' ]
#CASK = CASK.values.tolist()
# C lculo do Load Factor por dist ncia
# Estabelecendo o load factor por aeronave na rota
# A oferta (Fluxo * ASSENTOS) n o pode superar demasiadamente
a demanda atendida (PASSAGEIROS)
# Para rotas longas (DISTANCIA > 1300km), aceita-se oferta 4 vezes
maior que demanda (LF = 0.25)
# Para rotas curtas (DISTANCIA <= 1300), aceita-se oferta 2 vezes
maior que demanda (LF = 0.50)

LF = []
for j in range(N):
    if DISTANCIA[ j ] <= 1300:
        LF.append(0.50)
    else :

```

```

        LF.append(0.25)
M = 1000
K = 3

FO = LpProblem('Malha_PAN', LpMaximize)

FLUXO = LpVariable.dicts('Fluxos', [(a, j, HUB[j], SPOKE[j])]
for j in range(N) for a in range(N_ANV)], 0, None, cat='Integer')
PASSAGEIROS = LpVariable.dicts('PASSAGEIROS', [(a, j, HUB[j], SPOKE[j])]
for j in range(N) for a in range(N_ANV)], 0, None, cat='Integer')

BINARIO1 = LpVariable.dicts('Bin1', [(a, j)
for j in range(N) for a in range(N_ANV)], 0, None, cat='Binary')

BINARIO2 = LpVariable.dicts('Bin2', [(a)
for a in range(N_ANV)], 0, None, cat='Binary')

#Função Objetivo
FO += lpSum(PASSAGEM[j]*
PASSAGEIROS[(a, j, HUB[j], SPOKE[j])] - CASK[((N*a)+j)]*ASSENTOS[a]*
DISTANCIA[j]*FLUXO[(a, j, HUB[j], SPOKE[j])])
for a in range(N_ANV) for j in range(N)

# R.1 - Num de pax transportado igual a demanda
# A demanda de cada rota pode ser atendida
por mais de um tipo de aeronave

for j in range(N):
    FO += lpSum(PASSAGEIROS[(a, j, HUB[j], SPOKE[j])])
    for a in range(N_ANV) - DEMANDA[j] == 0

# R.2 - Oferta de assentos por Fluxo maior ou igual
# demanda expressa em PASSAGEIROS (R.1)
# A oferta (Fluxo * ASSENTOS) pode vir de mais de
#um tipo de aeronave para atender demanda
# Em cada rota, a oferta precisa ser maior que a
#demanda atendida (PASSAGEIROS) para cada aeronave

for a in range(N_ANV):

```

```

for j in range(N):
    FO += PASSAGEIROS[(a, j, HUB[j], SPOKE[j])] -
    ASSENTOS[a]*FLUXO[(a, j, HUB[j], SPOKE[j])] <= 0

# R.2.1 - Demanda atendida por Fluxo de cada
# aeronave na rota (NOVA)
# A oferta (Fluxo * ASSENTOS) n o pode superar demasiadamente
# a demanda atendida (PASSAGEIROS)
# A partir do load factor (LF) por rota
# Estabelece um limite m ximo que de uma oferta
# 1/LF vezes maior que a demanda

for a in range(N_ANV):
    for j in range(N):
        FO += PASSAGEIROS[(a, j, HUB[j], SPOKE[j])] -
        LF[j]*ASSENTOS[a]*FLUXO[(a, j, HUB[j], SPOKE[j])] >= 0

# R.3 - Definindo o Fluxo m ximo de cada
# modelo por rota
# Um modelo de aeronave s pode atender uma demanda
# de M vezes a capacidade de ASSENTOS

for a in range(N_ANV):
    for j in range(N):
        FO += FLUXO[(a, j, HUB[j], SPOKE[j])] - M*BINARIO1[(a, j)] <= 0

# R.5 - Alcance de aeronave (MODIFICADA)
# Anteriormente, estava excluindo TP
# quando existia DISTANCIA > ALCANCE

for a in range(N_ANV):
    for j in range(N):
        FO += ALCANCE[a] -
        DISTANCIA[j]*BINARIO1[(a, j)] >= 0

# R.6 - Ajustando o Fluxo somente para
# os modelos inclu dos (MODIFICADA)
# Impede que haja fluxo de modelo de
# aeronave n o inclu do

```

```

for a in range(N_ANV):
    for j in range(N):
        FO += FLUXO[(a, j, HUB[j], SPOKE[j])] - M*BINARIO2[(a)] <= 0
# R.7 - Num Max de modelos de aeronave que podem ser inclu dos

for j in range(N):
    FO += lpSum(BINARIO2[(a)] for a in range(N_ANV)) - K <= 0
#inicio = time.time()
FO.solve(PULP_CBC_CMD(gapRel = 0.001))
print('Status:', LpStatus[FO.status])
#fim = time.time()
print('\n')

for v in FO.variables():
    if v.varValue != 0 and v.varValue != M:
        if 'Bin2' in v.name:
            print(f'{v.name}: {v.varValue}')

lucro = value(FO.objective)
print('\n')
print('O lucro total de US$%.2f' % lucro)
#Alocando Dados finais em uma planilha Excel,
#para melhor visualiza o.
lista_index = []
lista_profit = []
lista_origem = []
lista_destino = []
lista_passageiros = []
lista_aero = []
lista_fluxo = []

for j in range(N):
    for a in range(N_ANV):
        profit = PASSAGEM[j]*
        PASSAGEIROS[(a, j, HUB[j], SPOKE[j]).varValue - CASK[((N*a)+j)]*
        ASSENTOS[a]*DISTANCIA[j]*
        FLUXO[(a, j, HUB[j], SPOKE[j]).varValue
        if profit != 0:

```

```

        lista_profit.append(profit)
        lista_index.append(j)
        lista_origem.append(HUB[j])
        lista_destino.append(SPOKE[j])
        lista_passageiros
        .append(PASSAGEIROS[(a, j, HUB[j], SPOKE[j])].varValue)
        lista_aero.append(AERONAVE[a])
        lista_fluxo
        .append(FLUXO[(a, j, HUB[j], SPOKE[j])].varValue)

Excel_1 = { 'INDEX': lista_index,
            'ORIGEM': lista_origem,
            'DESTINO': lista_destino,
            'PASSAGEIROS': lista_passageiros,
            'AERONAVE': lista_aero,
            'FLUXO': lista_fluxo,
            'LUCRO_ROTA': lista_profit }
DF_EXCEL = pd.DataFrame(Excel_1, columns=['ORIGEM',
                                         'DESTINO',
                                         'PASSAGEIROS',
                                         'AERONAVE',
                                         'FLUXO',
                                         'LUCRO_ROTA'])

Excel_2 = pd.DataFrame({'Caracteristica': ['Lucro (US$)', 'Status'],
                        'Valor': ['%.2f' % lucro, LpStatus[FO.status]]})

with ExcelWriter("Resultados_modelo_k3_gap0.005_22072022.xlsx")
as writer:
    #DATA_HUB.to_excel(writer,
                        sheet_name = 'Rota_Hub-Hub',
                        index=False)
    #DADOS1.to_excel(writer,
                     sheet_name = 'Rotas',
                     index=False)
    #MALHA_DEM_AJ.to_excel(writer,
                            sheet_name = 'Malha_aju_com_zeros',
                            index=False)

```

```
#MALHA_DEM_AJ_NOZEROS.to_excel(writer ,  
                                sheet_name = 'Malha_aju_sem_zeros' ,  
                                index=False)  
  
DADOS1.to_excel(writer ,  
                 sheet_name = 'MALHA_DEFINITIVA' ,  
                 index=False)  
  
DF_EXCEL.to_excel(writer ,  
                  sheet_name = 'RESULTADOS' ,  
                  index=False)  
  
Excel_2.to_excel(writer ,  
                 sheet_name='Config_Result_Simulacao' ,  
                 index=False)  
  
#writer.save()
```

Anexos

ANEXO A – Cálculo de Distâncias *Hub - Hub*

Código Elaborado por [Pereira \(2021\)](#)

```

from geopy.distance import geodesic
import pandas as pd

df = pd.read_excel('Dados_TCC_2021.xlsx', sheet_name = 'Hubs')
lat={}
long={}
coord_i = {}
Hubs = df['Hub']
for a in range(len(df)):
    lat[df['Hub_Dist'][a]] = df['Lat_O'][a]
    long[df['Hub_Dist'][a]] = df['Long_O'][a]
    coord_i[df['Hub_Dist'][a]] =
        [long[df['Hub_Dist'][a]], lat[df['Hub_Dist'][a]]]
for i in df['Hub_Dist']:
    for a in range(len(df)):
        df2 = df
        df2.at[a, i] = round(geodesic(coord_i[df['Hub_Dist'][a]],
            coord_i[i]).km)

### CHECK DE VALORES ###
coordenada_belem = [-48.4589, -1.415]
coordenada_BH = [-43.9892, -19.9092]
voo = geodesic(coordenada_belem, coordenada_BH)

```


ANEXO B – Código do Modelo Matemático Pre-existente

Adaptado de (PEREIRA, 2021) e utilizado como comparativo para os resultados finais.

```
%reset

!pip install pulp

from pulp import *
import pandas as pd
import numpy as np
from pandas import ExcelWriter
import time

#inicio = time.time()

#Function para calcular o preco da passagem
def Tarifa(dist):
    return np.exp(-0.73953 * np.log(dist) + 4.44446)* dist

#Function para calcular o CASK
def Cask(dist, seats):
    return np.exp(1.728912 - 0.40453 * np.log(dist)
                  - 0.35671 * np.log(seats))

from google.colab import drive
drive.mount('/content/drive')

%cd /content/drive/My Drive/Colab Notebooks/TCC

DADOS3 = pd.read_excel('Dados_TCC_2021.xlsx',
                      sheet_name = 'Aeronaves')
DADOS3.dropna(inplace=True)
DADOS3.reset_index(drop=True, inplace=True)
DADOS3.drop_duplicates(inplace=True)
```

```

AERONAVE = DADOS3[ 'AERONAVE' ]
N_ANV = len(AERONAVE)

ASSENTOS = DADOS3[ 'ASSENTOS' ]

ALCANCE = DADOS3[ 'ALCANCE' ]

DADOS5 = pd.read_excel( 'Dados_TCC_2021.xlsx' ,
                        sheet_name='Hubs' )

Dist_hub = {}
tarifa_hub = {}
cask_hub = {}

for i in DADOS5[ 'Hub_Dist' ]:
#for i in HUBN:
    dic1 = {}
    dic2={}
    dic3 = {}
    #for j in range(len(DADOS5[ 'Hub Dist' ])):

    for j in range(len(DADOS5[ 'Hub_Dist' ])):
        dic1 [DADOS5[ 'Hub_Dist' ][j]] = DADOS5[i][j]

        if i!= DADOS5[ 'Hub_Dist' ][j]:
            dic2 [DADOS5[ 'Hub_Dist' ][j]] = Tarifa(DADOS5[i][j])

            dic4={}
            for anv in range(N_ANV):
                dic4 [DADOS3[ 'AERONAVE' ][anv]] =
                    Cask(DADOS5[i][j] , ASSENTOS[anv])
            dic3 [DADOS5[ 'Hub_Dist' ][j]] = dic4

    Dist_hub[i] = dic1
    tarifa_hub[i] = dic2
    cask_hub[i] = dic3

#Leitura dos dados para construcao da malha

```

```

DADOS1 = pd.read_excel('Dados_TCC_2021.xlsx',
                       sheet_name = 'Dados')
DADOS1.dropna(inplace=True)
DADOS1.reset_index(drop=True, inplace=True)

HUB = DADOS1['Hub'].astype(int)
HUBN = HUB.drop_duplicates()
numero_HUB = len(HUBN)
#HUB = HUB.values.tolist()

SPOKE = DADOS1['Spoke'].astype(int)
SPOKEN = SPOKE.drop_duplicates()
numero_SPOKE = len(SPOKEN)
#SPOKE = SPOKE.values.tolist()

DEMANDA = round(DADOS1['Demanda']).astype(int)
#DEMANDA = DEMANDA.values.tolist()

DISTANCIA = round(DADOS1['Distancia']).astype(int)
# #DISTANCIA = DISTANCIA.values.tolist()

## Definicao do numero de vertices (edges)
N = len(DISTANCIA)

# Calculo de TARIFA por rota:
DADOS2 = pd.DataFrame(columns=['Passagem'])
for i in range(N):
    x = Tarifa(DISTANCIA[i])
    DADOS2 = DADOS2.append({'Passagem': x}, ignore_index=True)
DADOS2.reset_index(drop=True, inplace=True)
PASSAGEM = DADOS2['Passagem']
#PASSAGEM = PASSAGEM.values.tolist()

# Calculo do CASK por aeronave e rota:
DADOS4 = pd.DataFrame(columns=['Cask'])
for i in range(N_ANV):
    for j in range(N):
        x = Cask(DISTANCIA[j], ASSENTOS[i])
        DADOS4 = DADOS4.append({'Cask': x}, ignore_index=True)

```

```

DADOS4.reset_index(drop=True, inplace=True)
CASK = DADOS4[ 'Cask' ]
#CASK = CASK.values.tolist()

Ganho = PASSAGEM - DISTANCIA
print( 'Ganho: ' )
print( Ganho )

INDEX = {}

for spoke in SPOKEN:
    print( '\n└─Spoke=', spoke )
    ref= -100000 #Valor de referencia o menor possivel
    c = DADOS1[ 'Hub' ][ DADOS1[ 'Spoke' ]==spoke ].index
    for index in c:
        #print( ' ──index=', index )
        if Ganho[index] >= ref:
            old = ref
            ref = Ganho[index] #novo valor de referencia
            #print( '   ref=', old, ' // nova ref=', ref )
            INDEX[spoke] = index

        #else:
            #print( '   Nao muda index ' )
    #print( 'INDEX final do spoke', spoke, ' - ', INDEX[spoke] )

# Criando ligacao HUB->HUB com soma de demanda

DATA = pd.DataFrame( columns=[ 'Hub', 'Spoke',
                               'Demanda', 'Distancia' ] )
#DF para armazenar a lista de TODAS AS ROTAS
HUB_POTENCIAL para cada SPOKE

for spoke in SPOKEN:
    #print( '\n─SPOKE=', spoke )
    c = DADOS1[ 'Hub' ][ DADOS1[ 'Spoke' ]==spoke ].index
    for index in c:
        #print( ' ──INDEX=', index )
        #print( ' ──INDEX[ ', spoke, ' ]= ',

```

```

INDEX[spoke], '&_HUB[ ', index, ' ]_=' ,
                HUB[index])

if HUB[INDEX[spoke]] != HUB[index]:
    DATA2 = pd.DataFrame({"Hub": [HUB[index]],
                          "Spoke": [HUB[INDEX[spoke]]],
                          "Demanda": [DEMANDA[index]],
                          "Distancia": [Dist_hub[HUB[index]]
                                         [HUB[INDEX[spoke]]]]})
    DATA = DATA.append(DATA2, ignore_index=True)
    #print('Adicionando a DATA...')
    #print(DATA)
#else:
    #print('HUB[INDEX[spoke]] = HUB[index]---',
          HUB[INDEX[spoke]], '=', HUB[index])

#print('DATA:')
#print(DATA)
#print('-----')

DATA_HUB = pd.DataFrame(columns=['Hub',
                                'Spoke', 'Demanda', 'Distancia'])

lista_HH = DATA['Hub'].drop_duplicates()
print('lista_HH\n', lista_HH)

for hub in lista_HH:
    #print('HUB=', hub)
    df1 = DATA[DATA['Hub']==hub]
#DF para todos os dados que tem como origem o HUB=hub.
    lista_spoke =
    df1['Spoke'].drop_duplicates().reset_index(drop=True)
    #print('df1', df1)
    #print('lista_spoke', lista_spoke)

    for j in lista_spoke:
        df2 = df1[df1['Spoke']==j]
#DF para selecionar os dados que tem o mesmo Destino
(HUB_PRINCIPAL).

```

```

    #print('j=', j)
    #print(df2)

    soma = 0
    for i in df2['Demanda']:
        soma += i
    #print('soma', soma)

#Criação de um novo DF para armazenar os valores de
#demanda somada em uma unica rota
    df3 = pd.DataFrame({'Hub': [hub],
                        'Spoke': [j],
                        'Demanda': [soma],
                        'Distancia': Dist_hub[hub][j]})

    #print('df3=')
    #print(df3)
    #print('\n Adicionando ao df3 ao DATA_HUB:')
    DATA_HUB = DATA_HUB.append(df3, ignore_index=True)

    #print('DATA_HUB:')
    #print(DATA_HUB)

# Atribuindo novos valores de DEMANDA para a malha
# SPOKE-HUB_PRINCIPAL: novo valor sera a soma de
#todas as outras demandas para aquele SPOKE
# SPOKE-HUB_POTENCIAL: novo valor sera zerado —> Para nao
ser contabilizado no PULP

NEW_DEMANDA = DEMANDA.copy()
print('DEMANDAS_INICIAIS:')
print(DEMANDA)
print(NEW_DEMANDA)
print('—————')
#print(NEW_DEMANDA, type(NEW_DEMANDA))

for spoke in SPOKEN:
    Inicial_hub = INDEX[spoke]
    c = DADOS1['Hub'][DADOS1['Spoke']==spoke].index
    demanda = 0

```

```

    for index in c:
        demanda += DEMANDA[index]
        if index != Inicial_hub:
            #DEMANDA[index] = 0
            NEW_DEMANDA[index] = 0
    NEW_DEMANDA.iat[Inicial_hub] = demanda
    #print('type:', type(NEW_DEMANDA))

print('POS_TRANSFORMACAO:')
print(DEMANDA)
print(NEW_DEMANDA)
print('DEMANDA=NEW_DEMANDA')
DEMANDA = NEW_DEMANDA
print(DEMANDA)

# Criar uma versao de DADOS1 com o DF do DATA_HUB
#anexada ao final
print('DADOS1_ANTIGO:')
print(DADOS1)
# print('type', type(DADOS1['Demanda'][0]))
# print('type', type(DADOS1['Distancia'][2]))

# Ajuste do Type dos dados de DATA_HUB
DATA_HUB['Demanda'] =
    DATA_HUB['Demanda'].astype(np.float64)
DATA_HUB['Distancia'] =
    DATA_HUB['Distancia'].astype(np.float64)

for i in range(len(DADOS1['Demanda'])):
    a = DEMANDA[i]
    DADOS1['Demanda'] =
        DADOS1['Demanda'].replace([ DADOS1['Demanda'][i] ], a)

MALHA_DEM_AJ = DADOS1 #Armazenamento da info sobre as
#Rotas da malha, antes de serem excluidas as rotas potenciais.
MALHA_DEM_AJ_NOZEROS =
    MALHA_DEM_AJ[MALHA_DEM_AJ['Demanda']!=
    0].reset_index(drop=True)
#print('type', type(DATA_HUB['Distancia'][0]))

```

```

#print('type', type(DATA_HUB['Demanda'][0]))

#Juncao de DADOS1 com DATA_HUB
DADOS1 = DADOS1.append(DATA_HUB, ignore_index=True)
print('DADOS1_□_ATUALIZADO_□(com_□demanda_□ZERO):')
print(DADOS1)

#Removendo Linhas com Valor de Demanda ZERO(0)
#-> Diminui o gasto computacional do calculo
DADOS1 = DADOS1[DADOS1['Demanda']!= 0].reset__index(drop=True)
print('NOVO_□DADOS1', DADOS1)

# print('type', type(DADOS1['Demanda'][0]))
# print('type', type(DADOS1['Distancia'][2]))

# Atualizacao dos Parametros usando novo DADOS1

HUB = DADOS1['Hub']
SPOKE = DADOS1['Spoke']

DEMANDA = DADOS1['Demanda'].astype(int)
#DEMANDA = DEMANDA.values.tolist()

DISTANCIA = DADOS1['Distancia'].astype(int)
#DISTANCIA = DISTANCIA.values.tolist()

# Definicao do numero de vertices (edges)
N = len(DISTANCIA)

# Atualizando a PASSAGEM:
DADOS2 = pd.DataFrame(columns=['Passagem'])
for i in range(N):
    x = Tarifa(DISTANCIA[i])
    DADOS2 = DADOS2.append({'Passagem': x}, ignore_index=True)
DADOS2.reset_index(drop=True, inplace=True)
PASSAGEM = DADOS2['Passagem']
#PASSAGEM = PASSAGEM.values.tolist()

```

```

# Atualizando CASK por aeronave e rota
DADOS4 = pd.DataFrame(columns=['Cask'])
for i in range(N_ANV):
    for j in range(N):
        x = Cask(DISTANCIA[j], ASSENTOS[i])
        DADOS4 = DADOS4.append({'Cask': x}, ignore_index=True)
DADOS4.reset_index(drop=True, inplace=True)
CASK = DADOS4['Cask']
#CASK = CASK.values.tolist()

#Programacao Linear - PULP
M = 787
K = 1

FO = LpProblem('Malha_PAN', LpMaximize)

#Declaracao das Variaveis do Problema:

FLUXO =
    LpVariable.dicts('Fluxos', [(a, j, HUB[j], SPOKE[j])
        for j in range(N) for a in range(N_ANV)],
        0,
        None,
        cat='Integer')
PASSAGEIROS =
    LpVariable.dicts('PASSAGEIROS', [(a, j, HUB[j], SPOKE[j])
        for j in range(N) for a in range(N_ANV)],
        0,
        None,
        cat='Integer')
#PASSAGEIROS =
    LpVariable.dicts('PASSAGEIROS', [(a, j)
        for j in range(N) for a in range(N_ANV)],
        0,
        None,
        cat='Integer')
BINARIO1 =
    LpVariable.dicts('Bin1', [(a, j)

```

```

    for j in range(N) for a in range(N_ANV)] ,
        0,
        None,
        cat=' Binary ')
BINARIO2 =
    LpVariable.dicts( ' Bin2 ',
    [(a) for a in range(N_ANV)] ,
        0,
        None,
        cat=' Binary ')

#Funcao Objetivo
FO += lpSum(PASSAGEM[j]*
            PASSAGEIROS[(a,j,HUB[j],SPOKE[j])] -
            CASK[((N*a)+j)]*ASSENTOS[a]
            *DISTANCIA[j]*
            FLUXO[(a,j,HUB[j],SPOKE[j])]
            for a in range(N_ANV) for j in range(N))

# R.1 - Num de pax transportado igual a demanda
for j in range(N):
    FO += lpSum(PASSAGEIROS[(a,j,HUB[j],SPOKE[j])]
                for a in range(N_ANV)) - DEMANDA[j] == 0
    #Se deixar ==, obriga a ter rotas com prejuizo
    #Se deixa <=, rotas podem ser retiradas com prejuizo.

# R.2 - Demanda atendida pelo Fluxo
# Mesmo codigo do R.2.2 - Deixar apenas uma
for a in range(N_ANV):
    for j in range(N):
        FO +=
            PASSAGEIROS[(a,j,HUB[j],SPOKE[j])] -
            ASSENTOS[a]*
            FLUXO[(a,j,HUB[j],SPOKE[j])] <= 0

# R.2.2 - Demanda atendida pelo Fluxo
#COD_ESTUDO_PAN: PAX - ASSENTO*FLUXO.
# Mesmo codigo do R.2 - Deixar apenas uma
# for j in range(N):

```

```

#      FO += lpSum(ASSENTOS[a]*FLUXO[(a,j)]
           for a in range(N_ANV)) - DEMANDA[j] >= 0

# R.3 - Definindo o Fluxo maximo
for a in range(N_ANV):
    for j in range(N):
        FO += FLUXO[(a,j,HUB[j],SPOKE[j])] -
              M*BINARIO1[(a,j)] <= 0

# R.4 - Numero maximo de aeronave no modelo
for j in range(N):
    FO += lpSum(BINARIO1[(a,j)]
                for a in range(N_ANV)) - K <= 0

# R.5 - Alcance de aeronave
for a in range(N_ANV):
    for j in range(N):
        FO += ALCANCE[a] -
              DISTANCIA[j]*BINARIO2[(a)] >= 0

# R.6 - Ajustando o Fluxo ao alcance
for a in range(N_ANV):
    for j in range(N):
        FO += FLUXO[(a,j,HUB[j],SPOKE[j])] -
              M*BINARIO2[(a)] <= 0

# R.7 - Num Max de anv por rota
for j in range(N):
    FO += lpSum(BINARIO2[(a)]
                for a in range(N_ANV)) - K <= 0

# Solucao
inicio = time.time()

FO.solve(PULP_CBC_CMD(gapRel = 0.001))

print('Status:', LpStatus[FO.status])
print('\n')

```

```

for v in FO.variables():
    if v.varValue != 0 and v.varValue != M:
        print(v.name, '-', v.varValue)
        #print(v.name, '-', v.varValue)
lucro = value(FO.objective)

print('\\n')
print('O lucro total é de US$%.2f' % lucro)
print('Status:', LpStatus[FO.status])

fim = time.time()
print('Tempo de simulacao: %.2fs' % (fim - inicio))

#Alocando Dados finais em uma planilha Excel
lista_index = []
lista_profit = []
lista_origem = []
lista_destino = []
lista_passageiros = []
lista_aero = []
lista_fluxo = []

for j in range(N):
    for a in range(N_ANV):
        profit =
            PASSAGEM[j]*
            PASSAGEIROS[(a, j, HUB[j], SPOKE[j])].varValue -
            CASK[((N*a)+j)]*
            ASSENTOS[a]*
            DISTANCIA[j]*
            FLUXO[(a, j, HUB[j], SPOKE[j])].varValue
        if profit != 0:
            lista_profit.append(profit)
            lista_index.append(j)
            lista_origem.append(HUB[j])
            lista_destino.append(SPOKE[j])
            lista_passageiros.
            append(PASSAGEIROS[(a, j, HUB[j], SPOKE[j])].varValue)

```

```

        lista_aero.append(AERONAVE[a])
        lista_fluxo.append(FLUXO[(a,j,HUB[j],SPOKE[j])].varValue)

Excel_1 = {'INDEX': lista_index,
          'ORIGEM': lista_origem,
          'DESTINO': lista_destino,
          'PASSAGEIROS': lista_passageiros,
          'AERONAVE': lista_aero,
          'FLUXO': lista_fluxo,
          'LUCRO_ROTA': lista_profit}
DF_EXCEL = pd.DataFrame(Excel_1, columns=['ORIGEM',
                                       'DESTINO',
                                       'PASSAGEIROS',
                                       'AERONAVE',
                                       'FLUXO',
                                       'LUCRO_ROTA'])

Excel_2 = pd.DataFrame({'Caracteristica': ['Lucro_(US$)',
      'Tempo_de_Simulcao_(s)', 'Status'],
      'Valor': ['%.2f' % lucro, (fim - inicio),
      LpStatus[FO.status]]})

with ExcelWriter("Resultados_Finais_TCC_Gust_k4_gap_versao2.xlsx")
                as writer:

    DATA_HUB.to_excel(writer,
                       sheet_name = 'Rota_Hub-Hub',
                       index=False)
    MALHA_DEM_AJ.to_excel(writer,
                          sheet_name = 'Malha_aju_com_zeros',
                          index=False)
    MALHA_DEM_AJ_NOZEROS.to_excel(writer,
                                   sheet_name = 'Malha_aju_sem_zeros',
                                   index=False)
    DADOS1.to_excel(writer,
                    sheet_name = 'MALHA_DEFINITIVA',
                    index=False)
    DF_EXCEL.to_excel(writer,

```

```
        sheet_name = 'RESULTADOS',
        index=False)
Excel_2.to_excel(writer,
                 sheet_name='Config□Result□Simulacao',
                 index=False)

#writer.save()
```