

Adriano Duarte da Silva

**Modelagem Preditiva do Comportamento de Operações de
Pista da Aviação Comercial nos Aeroportos Internacionais do
Galeão, Brasília, Guarulhos e Recife.**

Dissertação apresentada à Fundação Getúlio Vargas, para a obtenção de Título de Mestre em Modelagem Matemática da Informação, na Escola de Matemática Aplicada da Fundação Getúlio Vargas.

Orientador: Vincent Gérard Y. Guigues

Co-Orientador: Renato Rocha Souza

Rio de Janeiro

2016

Silva, Adriano Duarte da

Modelagem preditiva do comportamento de operações de pista da aviação comercial nos aeroportos internacionais do Galeão, Brasília, Guarulhos e Recife / Adriano Duarte da Silva. – 2016.

96 f.

Dissertação (mestrado) – Fundação Getulio Vargas, Escola de Matemática Aplicada.

Orientador: Vincent Gérard Y. Guigues.

Coorientador: Renato Rocha Souza.

Inclui bibliografia.

1. Previsão com metodologia de Box-Jenkins. 2. Controle do tráfego aéreo. 3. Análise de séries temporais. 4. Aeronáutica comercial. I. Guigues, Vincent Gérard Yannick. II. Souza, Renato Rocha. III. Fundação Getulio Vargas. Escola de Matemática Aplicada. IV. Título.

CDD – 519.55

ADRIANO DUARTE DA SILVA

**MODELAGEM PREDITIVA DO COMPORTAMENTO DE OPERAÇÕES DE PISTA
DA AVIAÇÃO COMERCIAL NOS AEROPORTOS INTERNACIONAIS DO GALEÃO,
BRASÍLIA, GUARULHOS E RECIFE.**

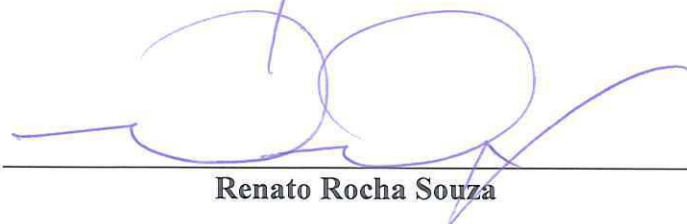
Dissertação apresentada ao Curso de Mestrado em Modelagem Matemática da Informação da Escola de Matemática Aplicada da Fundação Getúlio Vargas para obtenção do grau de Mestre em Modelagem Matemática da Informação.

Data da defesa: 15/07/2016.

ASSINATURA DOS MEMBROS DA BANCA EXAMINADORA



Vincent Gerard Yannick Guiges
Orientador (a)



Renato Rocha Souza



Paulo Cezar Pinto de Carvalho



Ralph S. Silva

AGRADECIMENTO

Agradeço aos meus orientadores que me ajudaram a acreditar e a concretizar este trabalho.

Agradeço aos professores e colegas do mestrado que me ajudaram nessa jornada de acúmulo de conhecimento.

Agradeço aos colegas de trabalho que compreenderam minha necessidade de, muitas vezes, faltar em razão do estudo.

Agradeço aos meus amigos que entenderam minha ausência dos eventos e me cobravam as páginas escritas.

Agradeço aos meus familiares que me incentivaram a todo instante.

Agradeço, em especial, à minha esposa que deu a luz quando iniciei no mestrado, me ajudando a estudar, tolerando minha insegurança e me fortalecendo a encarar os estudos e a paternidade ao mesmo tempo.

RESUMO

A aviação brasileira viveu momentos de crescimentos da oferta e da demanda nos últimos 10 anos, o que gerou a necessidade de planejar e investir no aumento da infraestrutura do Sistema de Controle do Espaço Aéreo Brasileiro (SISCEAB), do parque aeroportuário e dos serviços envolvidos na oferta e na demanda da aviação e das companhias aéreas (exemplo a demanda por Querosene de Aviação – QAV).

O principal objetivo deste trabalho é contribuir para a criação de um modelo preditivo do comportamento das operações de pousos e decolagens dos principais aeroportos brasileiros para o curto e médio prazo. Utilizaremos o modelo Box-Jenkins (ARMA), combinado com covariáveis, para prever a quantidade de operações de pousos e decolagens da aviação comercial em quatro aeroportos brasileiros: Aeroporto do Galeão/RJ, Aeroporto de Brasília/DF, Aeroporto de Guarulhos/SP e Aeroporto de Recife/PE.

Verificamos que os modelos se ajustam bem no curto prazo, mas que no médio prazo poderão necessitar de mais dados que podem interferir na quantidade de pousos e decolagens de forma atípica, como, por exemplo, a redução da malha aérea devido à crise econômica brasileira dos últimos 2 anos.

Os dados utilizados neste trabalho são pertencentes ao Centro de Gerenciamento da Navegação Aérea (CGNA) que é uma unidade militar subordinada ao Departamento de Controle do Espaço Aéreo Brasileiro (DECEA).

Portanto, a previsibilidade da demanda de tráfego aéreo ajudará na alocação de recursos no gerenciamento de tráfego aéreo.

PALAVRAS CHAVES: ARMAX, Aviação brasileira, Modelo Box-Jenkins, Previsão, Tráfego aéreo.

ABSTRACT

The Brazilian aviation went through moments of supply and demand growth in the last decade, which triggered investment and plan decisions aiming at improving the Brazilian airspace control system and airports infrastructure as well as the services offered by airline companies (for instance aviation fuel demand increase).

The main objective of this work is to contribute to the creation of a predictive model of takeoffs and landings for the short and medium run. We use the Box-Jenkins model (ARMA), combined with covariates, to predict the amount of takeoffs and landings of commercial aviation in four Brazilian airports: Galeão Airport (Rio de Janeiro), Brasília Airport (Distrito Federal), Guarulhos Airport (São Paulo) and Recife Airport (Pernambuco).

We find that the models fit well in the short run, but in the medium run in forecast events such as the recent Brazilian economic crises can damage the predictions.

The data analyzed in this work is owned by The Management Center of Air Navigation (CGNA), which is a military unit subordinated to the *Departamento de Controle do Espaço Aéreo* (DECEA).

Therefore, the predictability of the demand for air traffic will help in the allocation of resources for air traffic management.

KEY WORDS: Air Traffic, ARMAX, Brazilian aviation, Box-Jenkins model, Forecasting.

ÍNDICE

AGRADECIMENTO	2
RESUMO	3
ABSTRACT	4
GLOSSÁRIO.....	7
1 INTRODUÇÃO.....	8
1.1 Contexto.....	8
1.2 Objetivo.....	11
1.3 Justificativa.....	12
1.4 Estrutura da Dissertação.....	12
2 HIPÓTESES DO TRABALHO	13
2.1 Modelo Preditivo	13
3 ORIGEM DOS DADOS.....	15
3.1 Fonte de Dados	15
3.2 Descrição dos dados	16
4 ANÁLISE DESCRITIVA DOS DADOS	17
4.1 Variável Dependente	17
4.2 Dia da Semana	19
4.3 Mês do ano	22
4.4 Feriados.....	24
4.5 Grandes Eventos	33
4.6 Ano.....	34
4.7 Meteorologia	34
4.8 Obras na Pista dos Aeroportos	37
5 METODOLOGIA.....	38
5.1 Processo Estocástico Estacionário	38

5.2	Modelos de Autoregressivos e de Médias Móveis (ARMA)	38
5.3	Modelos ARMAX	41
5.4	Função de Autocorrelação e Correlograma.....	41
5.5	Testes de Normalidades e Independência dos Resíduos	42
5.6	Estimação por Máxima Verossimilhança	43
5.7	Critérios para comparar os Modelos (AIC e BIC)	44
5.8	Variável Binária (Dicotômica)	45
5.9	Variáveis Analisadas.....	46
6	RESULTADOS	59
6.1	Variável Dependente	59
6.2	Covariáveis Utilizadas	59
6.3	Índices do ARMA.....	60
6.4	Modelo Aeroporto do Galeão.....	60
6.5	Modelo Aeroporto de Brasília	68
6.6	Modelo Aeroporto de Guarulhos	73
6.7	Modelo Aeroporto de Recife	78
6.8	Programa R (Auxiliar).....	83
7	CONSIDERAÇÕES FINAIS	85
8	FUTUROS TRABALHOS.....	87
9	REFERÊNCIAS BIBLIOGRÁFICAS	88
	ÍNDICE DE FIGURAS	91
	ÍNDICE DE TABELAS	94
	APÊNDICE A	95

GLOSSÁRIO

AIC	Critério de Informação Akaike
AIS	Serviço de Informação Aeronáutica
AISWEB	Fonte de Informações Aeronáuticas em Meio Digital do DECEA
ANAC	Agência Nacional da Aviação Civil
AR(p)	Modelos Autoregressivos de Ordem “p”
ATC	Controle de Tráfego Aéreo
BACEN	Banco Central do Brasil
BIC	Critério de Informação Bayesiana
CGNA	Centro de Gerenciamento da Navegação Aérea
CIRCEA	Circular Normativa de Controle do Espaço Aéreo
COMAER	Comando da Aeronáutica
DECEA	Departamento de Controle do Espaço Aéreo
HOTRAN	Lista da ANAC com os Voos Repetitivos Autorizados
HSTVOOS	Sistema de Banco de Dados da INFRAERO
IEPV	Impresso Especial de Proteção ao Voo
INFRAERO	Empresa Brasileira de Infraestrutura Aeroportuária
IPCA	Índice Nacional de Preços ao Consumidor Amplo
MA(q)	Modelos de Médias Móveis de Ordem “q”
METAR	Informação Meteorológica Horária
NOTAM	Informe para Aero Navegante (Notice to Airmen)
PIB	Produto Interno Bruto
QAV	Querosene de Aviação
R	Programa Estatístico
REDEMET	Rede de Meteorologia do Comando da Aeronáutica
SAS	Programa Estatístico
SGTC	Sistema de Gerenciamento de Torre de Controle de Aeródromo
SISCEAB	Sistema de Controle do Espaço Aéreo Brasileiro
SPECI	Informação Meteorológica Especial
SQR	Soma dos Quadrados dos Resíduos
STDMA	Sistema de Tratamento de Dados de Movimentos de Aeródromos
TATIC	Programa <i>Total Air Traffic Information Control</i> da empresa Saipher ATC
TGL	Toque e Arremetida (do inglês <i>Touch and Go Landing</i>)
TWR	Torre de Controle de um Aeroporto (do inglês <i>Tower</i>)

1 INTRODUÇÃO

1.1 Contexto

Segundo a Empresa Brasileira de Infraestrutura Aeroportuária (INFRAERO), a aviação brasileira cresceu muito nos últimos 10 anos (Figura 1-1¹) e passou por algumas transformações, como a concessão de aeroportos para o setor privado, grandes investimentos em razão de eventos grandiosos que ocorreram e ainda ocorrerão no Brasil (como a Copa do Mundo de Futebol e as Olimpíadas), o surgimento de uma nova companhia aérea (como a Azul Linhas Aéreas) e a absorção de companhias aéreas por outras no mercado (como a Webjet adquirida pela GOL Linhas Aéreas e a TRIP Linhas Aéreas adquirida pela Azul).

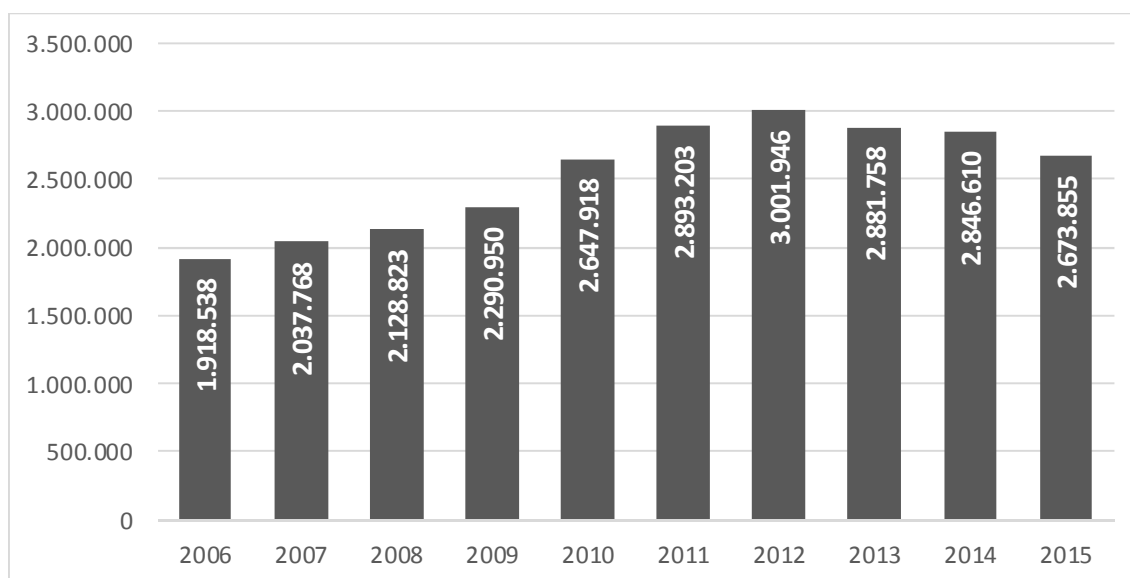


Figura 1-1 - Gráfico do total de pousos e decolagens da aviação comercial e geral de 66 aeroportos do Anuário Estatístico Operacional da INFRAERO de 2010 a 2015.

O Produto Interno Bruto (PIB) do Brasil, divulgado pelo Banco Central do Brasil (BACEN), cresceu até 2014 (com exceção de 2009) indicando aumento da atividade econômica brasileira (Figura 1-2). No ano de 2015, o PIB brasileiro reduziu 3,4%, o que

¹ Os dados para os anos de 2013, 2014 e 2015 dos aeroportos de Guarulhos/SP, Brasília/DF, Galeão/RJ, Campinas/SP (Viracopos), Confins/MG foram estimados pela fonte de Anuário Estatístico de Tráfego Aéreo de 2015 do DECEA.

pode ter impactado a aviação comercial em 2015 e ter afetado o planejamento da malha aérea das companhias aéreas para o ano de 2016 (redução de custos e de demanda por tráfego aéreo). A inflação, medida pelo Índice de Preços ao Consumidor Amplo (IPCA), dos últimos 10 anos, também demonstra uma piora no ano de 2015, o que pode indicar a queda no poder de compra da população brasileira.

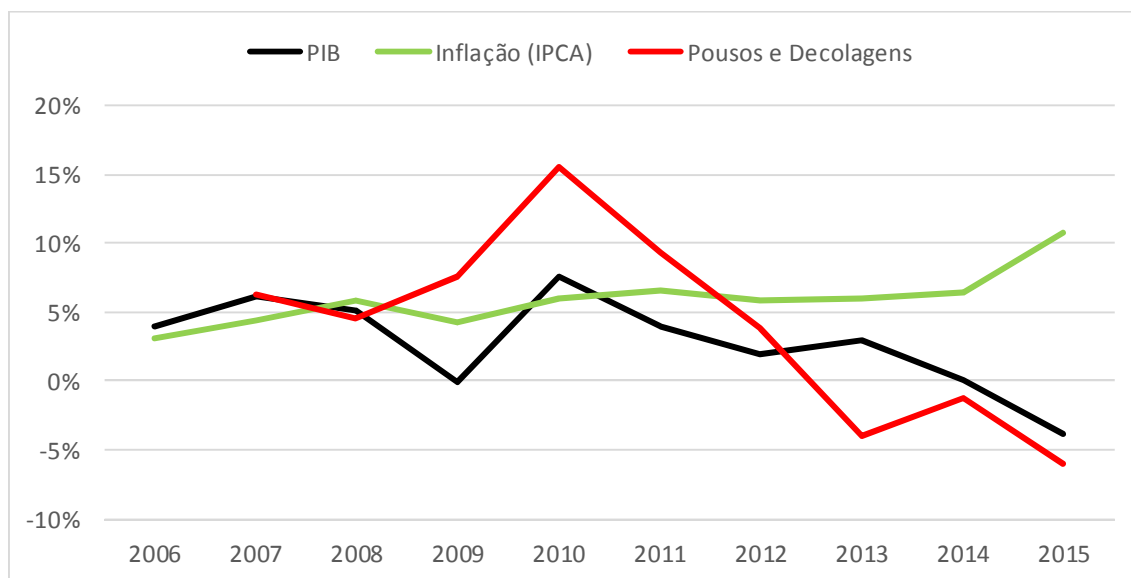


Figura 1-2 - Gráfico com o crescimento anual do PIB brasileiro (BACEN), do índice de inflação IPCA (BACEN) e da quantidade de pousos e decolagens de 66 aeroportos brasileiros (Anuário Estatístico da INFRAERO).

As administradoras dos aeroportos brasileiros, a INFRAERO e as concessionárias planejam constantemente investimento no complexo aeroportuário, seja para melhorar o atendimento aos usuários da aviação, ou para dar condições às companhias aéreas de realizarem o transporte de passageiros de maneira segura, ou ainda para criar oportunidades de aumento da demanda pelo transporte aéreo nacional.

No Brasil, as companhias aéreas comerciais têm que ter autorização da Agência Nacional da Aviação Civil (ANAC) para pousarem e decolarem dos aeroportos brasileiros. Esta autorização é necessária para evitar o congestionamento de aeronaves e de passageiros nos terminais, ou seja, quando uma companhia aérea quer voar do Aeroporto Santos Dumont (no Rio de Janeiro) para o Aeroporto de Brasília (no Distrito Federal), deve enviar uma solicitação à ANAC que faz algumas consultas antes de fornecer a autorização final:

- Consultar a INFRAERO sobre o fluxo de passageiros no terminal de embarque do Aeroporto Santos Dumont;
- Consultar a Concessionária INFRAMERICA S.A. sobre o fluxo de passageiros no terminal de desembarque do aeroporto de Brasília;
- Consultar o Departamento de Controle do Espaço Aéreo Brasileiro (DECEA) sobre o fluxo de aeronaves no tráfego aéreo brasileiro, tanto na decolagem, quanto no voo e no pouso.

Esses pedidos de autorização à ANAC podem ser de Voos Repetitivos (aqueles que se repetem em outras semanas no mesmo dia e horário) ou de Voos Charters (voos pontuais, apenas para uma data específica).

Os maiores (principais) aeroportos brasileiros eram administrados pela INFRAERO até 2012, quando o governo brasileiro deu início a um processo de concessões de grandes aeroportos brasileiros à iniciativa privada. Na primeira rodada de concessões, em julho de 2012, depois do aeroporto de São Gonçalo do Amarante em Natal/RN (que foi concedido enquanto estava sendo construído), os aeroportos de Guarulhos/SP, de Campinas/SP (Viracopos) e de Brasília/DF passaram a ser administrados por concessionárias, possuindo o governo, através da INFRAERO, até 49% de participação. A segunda rodada de concessões contemplou os aeroportos do Galeão/RJ e de Confins/MG que passaram para as concessionárias em janeiro de 2014.

O Anuário Estatístico Operacional da INFRAERO, de 2012, foi o último anuário que ainda contemplava os aeroportos de Guarulhos, Campinas e Brasília (os primeiros a ir para as concessionárias). Neste documento, o aeroporto de Guarulhos é o que apresenta maior quantidade de operações de pousos e decolagens (somando todos os tipos de aviação, menos a militar), o aeroporto de Brasília aparece em terceiro lugar, o aeroporto do Galeão em quarto e o aeroporto de Recife em décimo segundo (dentro de 66 aeroportos administrados pela INFRAERO).

O Aeroporto Internacional Antônio Carlos Jobim do Rio de Janeiro, também conhecido como Galeão, é administrado pela concessionária Rio Galeão S.A.. O fluxo da aviação comercial na cidade do Rio de Janeiro é dividido entre o Galeão e o aeroporto Santos Dumont, sendo este último voltado exclusivamente para o tráfego aéreo doméstico. Outro aeroporto com grande fluxo de aeronaves na mesma cidade é o aeroporto de Jacarepaguá, mas este não opera voos da aviação comercial (de companhias aéreas).

O Aeroporto Internacional André Franco Montoro de São Paulo, também conhecido como Guarulhos, é administrado pela concessionária GRU Airport S.A.. O fluxo da aviação comercial na Grande São Paulo é dividido entre os aeroportos de Guarulhos, Congonhas e Campinas (Viracopos), sendo o aeroporto de Congonhas o único voltado exclusivamente para o tráfego aéreo doméstico. Outro aeroporto com grande fluxo de aeronaves nesta região é o aeroporto de Campo de Marte, mas este não opera voos da aviação comercial.

O Aeroporto Internacional Presidente Juscelino Kubitschek, de Brasília, é administrado pela concessionária INFRAMERICA S.A. (a mesma do aeroporto de São Gonçalo do Amarante em Natal/RN). Este aeroporto é o único dentro do Distrito Federal, e por estar aproximadamente no centro do país serve de *Hub* (centro distribuidor de voos com conexões e escalas) de voos entre o Norte e o Sul do país.

O Aeroporto dos Guararapes em Recife é administrado pela INFRAERO. Este aeroporto foi o segundo mais movimentado (pousos + decolagens + sobrevoos) da região Nordeste, ficando atrás do aeroporto de Salvador e na frente do aeroporto de Fortaleza no ano de 2015 (Anuário Estatístico de Tráfego Aéreo de 2015 do DECEA).

1.2 Objetivo

O principal objetivo deste trabalho é contribuir para a criação de um modelo preditivo do comportamento das operações de pousos e decolagens dos principais aeroportos brasileiros para o curto e médio prazo. Usaremos, como exemplo, os aeroportos internacionais do Galeão/RJ, de Brasília/DF, de Guarulhos/SP e de Recife/PE por serem aeroportos de grande porte, por conterem voos nacionais e internacionais, por funcionarem 24 horas e por terem disponibilidade de dados para serem analisados.

Outro objetivo que pode ser alcançado a partir dos resultados deste trabalho é a previsão de demanda de tráfego aéreo nos órgãos de Controle do Tráfego Aéreo (ATC) para a preparação dos eventos que exigem maior previsibilidade do fluxo de tráfego aéreo, como os Jogos Olímpicos e Paraolímpicos que ocorrerão no Brasil em 2016, e em períodos especiais, como Carnaval, Semana Santa, Natal e Réveillon.

1.3 Justificativa

O Centro de Gerenciamento da Navegação Aérea (CGNA, prevista pelo Decreto nº 6.834, de 30 de abril de 2009) e a Organização do Comando da Aeronáutica (COMAER) têm por finalidade exercer a gestão das ações correntes dos processos de gerenciamento de tráfego aéreo e de infraestrutura relacionada, visando à suficiência e à qualidade dos serviços prestados no âmbito do Sistema de Controle do Espaço Aéreo Brasileiro (SISCEAB) e dos elos afins, em tempo real e a partir das intenções de voo. As organizações dependem de previsões estatísticas de demanda para prover em uma fundação sólida de cenários para muitos processos importantes de tomada de decisão e planejamento.

O transporte de Querosene de Aviação (QAV) para os aeroportos brasileiros, por exemplo, depende de uma melhor previsão da quantidade de operações de pousos e decolagens para a redução de custos e para o aumento da eficácia da distribuição de QAV para os aeroportos, contribuindo assim, para a redução dos custos das empresas aéreas e dos preços das passagens.

Portanto, a previsibilidade da demanda de tráfego aéreo ajudará na alocação de recursos no gerenciamento de tráfego aéreo.

1.4 Estrutura da Dissertação

Este trabalho está dividido em 6 capítulos. No Capítulo 1 está a apresentação do problema, dentro de um contexto econômico e do tráfego aéreo brasileiro. No Capítulo 2 são apresentadas algumas hipóteses que serão usadas e testadas nos modelos preditivos, enquanto que o Capítulo 3 se dedica a apresentar os dados que serão usados nestes modelos. A análise descritiva das principais variáveis que serão utilizadas neste trabalho está no Capítulo 4. No capítulo seguinte, se encontra a metodologia que será utilizada para a estimação dos modelos preditivos propostos neste trabalho. No Capítulo 6 estão os resultados dos modelos para o Aeroporto do Galeão, de Brasília, de Guarulhos e de Recife. Os últimos capítulos são dedicados as considerações finais que foram encontradas com os modelos obtidos e uma discussão de futuros trabalhos.

2 HIPÓTESES DO TRABALHO

A aviação comercial se utiliza, principalmente, dos voos repetitivos, que têm autorização da ANAC para voar no mesmo dia da semana e horário ao longo dos meses. Verificaremos no modelo preditivo se a quantidade de operações de pousos e decolagens varia dependendo do dia da semana.

Os feriados nacionais podem mudar a dinâmica do tráfego aéreo, ou seja, a demanda por motivo de trabalho ou de turismo deve ser diferente devido a um feriado como o Carnaval, por exemplo, que pode aumentar a demanda de tráfego aéreo no início e no final do feriado, mas que durante o período festivo diminui o transporte de pessoas que voam a trabalho. Outros feriados nacionais que precisam ser analisados são: Semana Santa, Tiradentes, Dia do Trabalho, Corpus Christi, Dia da Independência, Dia de Nossa Senhora, Dia de Finados, Dia da Proclamação da República, Natal e Réveillon.

Os feriados estaduais, municipais e os grandes eventos também podem interferir na demanda do tráfego aéreo de determinados aeroportos. A Copa do Mundo de futebol de 2014, realizada no Brasil, pode ter mudado a malha aérea dos principais aeroportos brasileiros. A final realizada no Rio de Janeiro fez com que fossem registradas a maior quantidade de pousos e decolagens do aeroporto do Galeão, como veremos na exploração dos dados de 2014. Outro exemplo é o festival de Parintins, no Pará, que devido à grande festividade pode modificar a malha aérea local durante esse período.

Nos meses de verão a demanda por tráfego aéreo pode aumentar devido às férias escolares, assim como no mês de julho. Da mesma forma, ao longo dos anos a demanda pode ter crescido devido ao aumento do poder de compra nos primeiros anos do período analisado neste estudo e diminuído no último ano.

A meteorologia e obras nas pistas também podem modificar e desviar o tráfego aéreo de um aeroporto para outro. No caso da meteorologia, o aeroporto pode ficar fechado para pousos ou decolagens por um período de tempo e, com isso, desviar o fluxo de aeronaves que pousariam nele para um aeroporto alternativo.

2.1 Modelo Preditivo

O modelo espera prever a quantidade diária de operações de pousos e decolagens da aviação comercial dos aeroportos do Galeão, de Brasília, de Guarulhos e de Recife.

Utilizaremos a mineração de dados para criar modelos preditivos e descritivos precisos com base na análise de grandes quantidades de dados disponíveis, explorando-os e buscando relacionamentos, tendências e anomalias.

Pretendemos adaptar as informações coletadas por meio da criação, seleção e transformação das variáveis, realizando o processo de seleção de modelos com aprendizado estatístico para prever, de forma confiável, um resultado desejado. Serão utilizadas técnicas como regressão linear, séries temporais e testes de hipóteses. Também avaliaremos os dados e a confiabilidade dos resultados do processo de mineração de dados, comparando os modelos encontrados, computando novas estatísticas de ajustes, incluindo e excluindo variáveis na análise.

Além disso, a utilização do programa R disponibiliza ferramentas de visualização avançadas que permitem o exame, de forma rápida e fácil de grandes quantidades de dados em histogramas e gráficos multidimensionais para a comparação dos resultados da modelagem.

3 ORIGEM DOS DADOS

Nesta seção serão apresentadas as fontes de dados e como serão obtidas.

3.1 Fonte de Dados

Os dados utilizados neste trabalho são pertencentes ao CGNA, que é uma unidade militar subordinada ao DECEA.

A torre de controle do aeroporto do Galeão registra todos os pousos e as decolagens, além de toque e arremetida (TGL), e cruzamentos de aeronaves que afetam o funcionamento das pistas. Esses dados ficam armazenados no sistema de informática da torre de controle, que pode ser o SGTC ou o TATIC, eles são extraídos e enviados ao CGNA, conforme especificação dos documentos CIRCEA 100-55 e IEPV 100-34 (documentos do DECEA).

No CGNA, os dados são tratados, validados e disponibilizados pelo Sistema de Tratamento de Dados de Movimentos de Aeródromos (STDMA) ou pelo programa estatístico SAS, no formato de planilhas. Foi concedida a permissão de utilização dos dados históricos consolidados por dia da aviação comercial e aviação geral (excluindo a aviação militar), para fins de estudos, como o apresentado por este trabalho.

Outras fontes também foram utilizadas para a validação da quantidade de operações diárias de pousos e decolagens da aviação comercial nos aeroportos que serão analisados neste trabalho: o HSTVOOS (que é um sistema pertencente a INFRAERO), e o HOTRAN (planilha disponibilizada pela ANAC com a lista dos voos repetitivos que tem autorização para decolar e pousar nos aeroportos brasileiros).

As informações de meteorologia foram coletadas na Rede de Meteorologia do Comando da Aeronáutica (REDEMET), que contém as informações de meteorologia dos principais aeroportos brasileiros. As informações meteorológicas são disponibilizadas de hora em hora (METAR), ou em menor tempo (SPECI) quando há mudança significativa das condições meteorológicas. Os dados foram tratados seguindo especificações da Seção de

Meteorologia do CGNA, que foram adicionados ao Relatório de Meteorologia em formato de planilha Excel (programa da Microsoft).

O banco de dados utilizado para encontrar as obras que ocorreram nas pistas dos aeroportos (principalmente em Recife) foi o do Serviço de Informação Aeronáutica (AIS, pelo website AISWEB), e através dos informativos obrigatórios (NOTAM) que são disponibilizados quando há obras nos aeroportos.

3.2 Descrição dos dados

Os dados disponibilizados para o estudo são listas diárias com todas as operações de pousos, decolagens e cruzamentos (que são toque e arremetida e sobrevoos de aeronaves no aeroporto que afetam as operações de pousos ou decolagens).

Cada operação tem o registro de algumas características sobre a aeronave, a rota e o horário do voo:

- Indicativo do Voo - Ex.: TAM3665.
- Tipo de Voo - “S” ou “N” para voos comerciais, “G” para voos da aviação geral (executiva) e “M” para voos da aviação militar.
- Equipamento - são códigos do tipo de equipamento que é a aeronave, exemplo B737 que é um Boing 737.
- Aeroporto de Origem e/ou Destino – se a operação for um pouso aparecerá o aeroporto de origem daquele voo, se for uma decolagem aparecerá o aeroporto de destino.
- Horários - se a operação for um pouso aparecerá o horário do pouso do voo, se for uma decolagem aparecerá o horário da decolagem.
- Pista - o código da cabeceira de pista utilizada na operação de pouso ou decolagem.

Neste trabalho, utilizaremos o total de pousos e decolagens por dia da aviação comercial. Os dados diários são de 01 de janeiro de 2010 à 31 de dezembro de 2015. Mas também serão utilizados os dados de 01 de janeiro de 2016 à 31 de março de 2016, com a finalidade de verificação dos resultados alcançados no modelo final.

4 ANÁLISE DESCRITIVA DOS DADOS

Vamos analisar algumas variáveis que serão testadas no modelo preditivo.

- A variável dependente – nos nossos modelos são as quantidades de operações de pousos e decolagens dos aeroportos analisados;
- Dia da semana – que é a influência do dia da semana na quantidade de operações;
- Mês do ano – que é a influência do mês na quantidade de operações;
- Feriados – que é a influência dos feriados nacionais, estaduais e municipais na quantidade de operações;
- Grandes eventos – que é a influência de grandes eventos como Copa do Mundo de Futebol na quantidade de operações;
- Ano – que é a influência do ano na quantidade de operações;
- Meteorologia – que é a influência das condições meteorológicas na quantidade de operações.
- Obras nas pistas – que é a influência da restrição de voos por motivos de manutenção das pistas de pousos dos aeroportos.

4.1 Variável Dependente

A variável dependente dos quatro modelos preditivos (para cada aeroporto), que esperamos estimar no final do trabalho, são os totais diários de operações de pousos e decolagens da aviação comercial dos aeroportos do Galeão, Brasília, Guarulhos e Recife. Os gráficos da *Figura 4-1* mostram os totais diários de operações, de 1º de janeiro de 2010 até 31 de dezembro de 2015, para cada aeroporto analisado neste trabalho. Note que na escala dos gráficos que indica que o aeroporto com o maior número de operações é o de Guarulhos, seguido por Brasília e Galeão, enquanto que o aeroporto com o menor número é o de Recife.

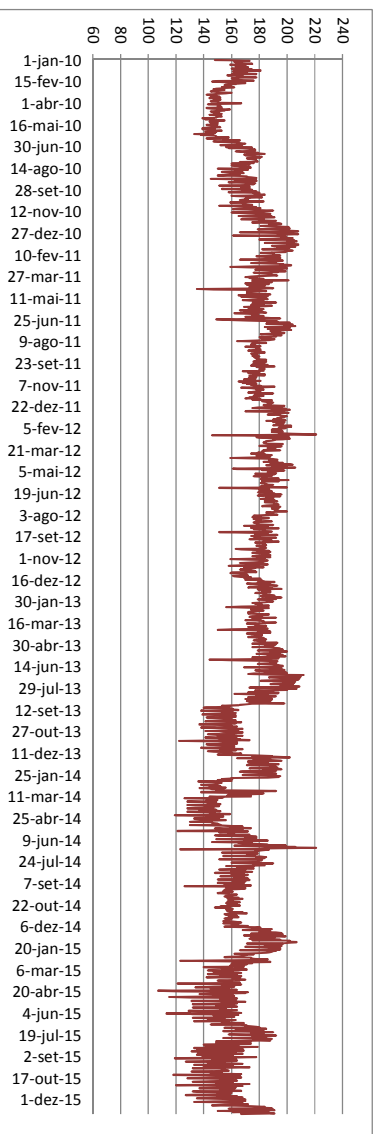
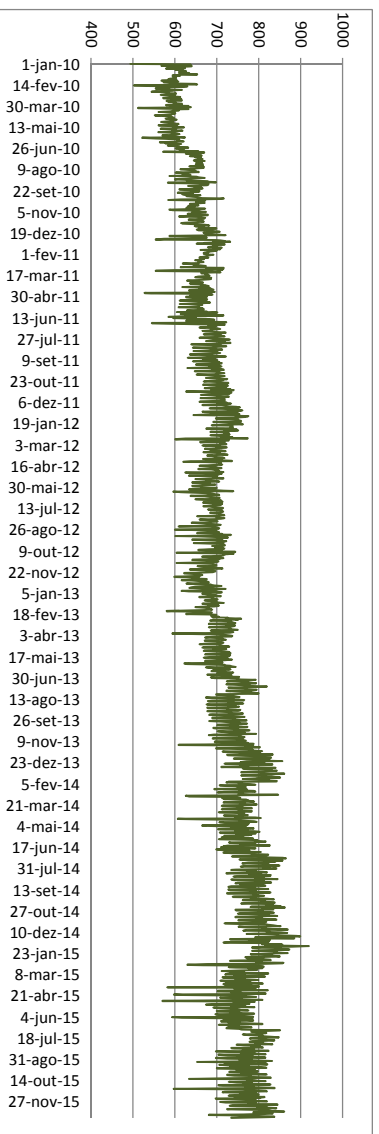
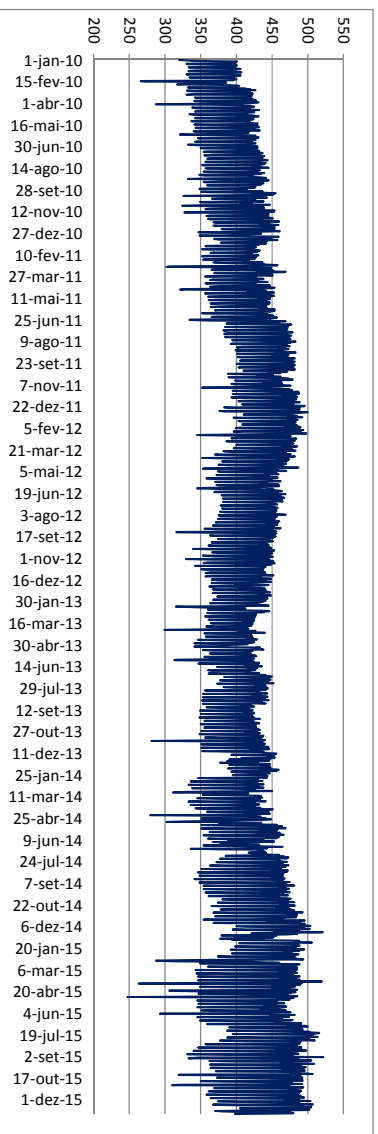
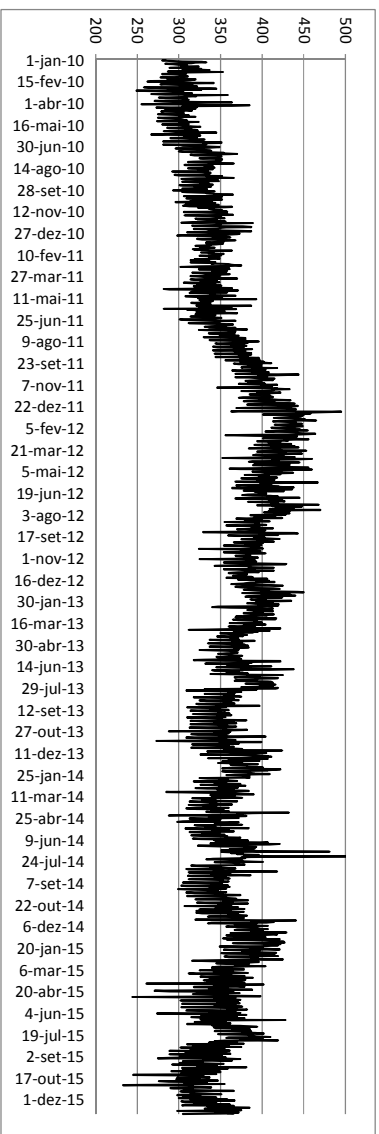


Figura 4-1 - Gráfico do total de pouso e decolagens por dia da aviação comercial de 1º de janeiro de 2010 a 31 de dezembro de 2015 dos aeroportos do Galeão (a), de Brasília (b), de Guarulhos (c) e de Recife (d).

A variável dependente apresentou problemas de consistência nos anos de 2008 e 2009, por essa razão esses anos foram retirados da análise. Algumas datas nos outros anos (uma média de 8 dias por aeroporto, menos de 0,4% dos dados) apresentaram erros de sistema, ocorrendo uma inconsistência na quantidade de operações de pousos e decolagens. Para manter a unidade do dia de semana, substituímos os dados por uma interpolação (média) da quantidade de operações do 7º dia anterior e do 7º dia posterior. Para o caso de uma dessas datas ser dia festivo, próxima ou em feriado, foi usada a quantidade de mais 7 dias de diferença para mais ou para menos da data festiva. Com isso, mantemos a quantidade de voos repetitivos por dia da semana e do período no ano.

4.2 Dia da Semana

A variável de dia da semana pode indicar que há mais operações de pousos e decolagens em alguns dias durante a semana e menos operações nos finais de semana. Também foi observado que existe a possibilidade de uma mudança na distribuição dos voos por dia da semana. Verificou-se que nos primeiros anos da análise deste trabalho determinados dias tinham mais operações, enquanto que nos últimos anos outros dias se tornaram mais importantes.

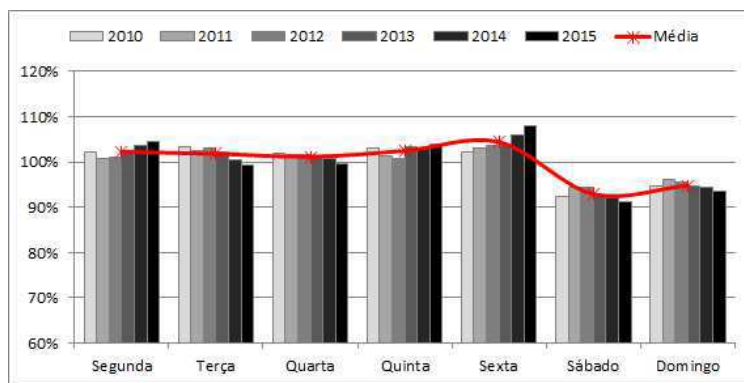
Na Figura 4-2 (a), no gráfico do aeroporto do Galeão, as colunas acima de 100% indicam que naquele dia da semana, em média, houveram mais operações de pousos e decolagens do que a média de todos os dias do ano. A linha vermelha é a média dos seis anos. No ano de 2015, a coluna mais escura do gráfico indica que a segunda, a quinta e a sexta-feira são os dias da semana com mais operações, em média, e também ficam acima da média diária de operações no ano (por estarem acima de 100%). Diferente do ocorrido no ano de 2012, quando a terça-feira teve mais operações do que a segunda e a quinta. Esta observação sugere um cruzamento das variáveis ano e dia da semana para o modelo preditivo.

Na Figura 4-2 (b), no gráfico do aeroporto de Brasília, pode ser observada uma redução nas operações no final de semana e uma distribuição das operações de pousos e

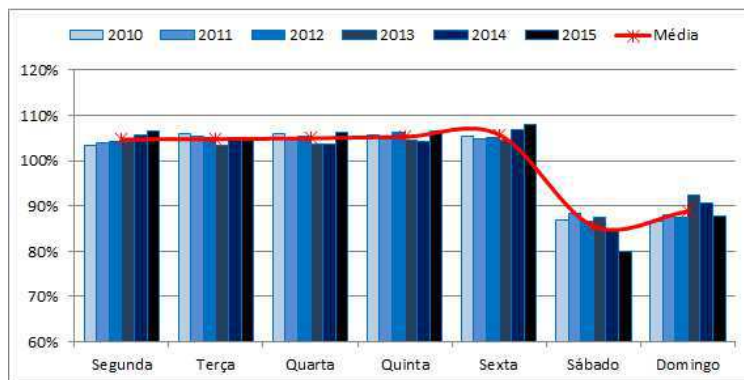
decolagens no ano de 2015, com mais operações na segunda e na sexta e menos operações aos sábados e domingos, semelhante ao ocorrido no aeroporto do Galeão.

Na Figura 4-2 (c), no gráfico do aeroporto de Guarulhos, segue-se análise semelhante aos aeroportos anteriores, onde a maior diferença está na quantidade reduzida no final de semana (que é mais baixa do que nos aeroportos do Galeão e de Brasília), ficando mais próxima dos 100% em todos os anos analisados.

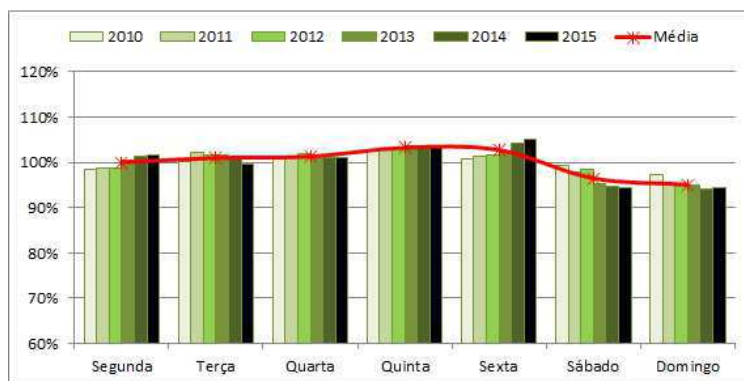
Na Figura 4-2 (d), no gráfico do aeroporto de Recife, temos uma distribuição por dia da semana bem diferente dos demais aeroportos analisados. O sábado aparece com uma quantidade de operações muito próxima aos dias da semana e apenas domingo tem, em média, 13% a menos de operações em relação à média do último ano de análise.



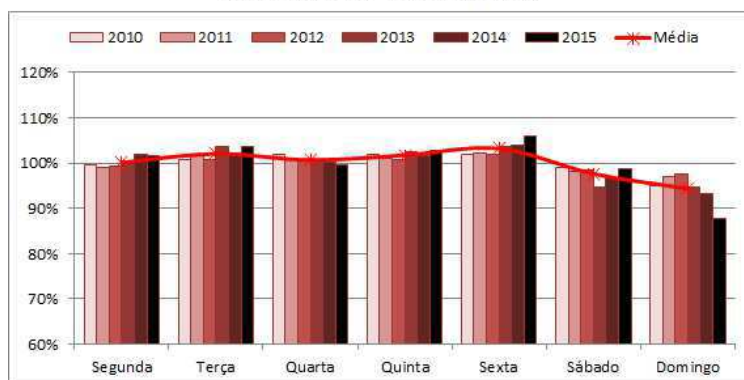
(a) Aeroporto do Galeão



(b) Aeroporto de Brasília



(c) Aeroporto de Guarulhos



(d) Aeroporto de Recife

Figura 4-2 - Gráfico da divisão do total de pousos e decolagens por dia da semana pelo total de pousos e decolagens médio do ano da aviação comercial nos aeroportos do Galeão (a), de Brasília (b), de Guarulhos (c) e de Recife (d).

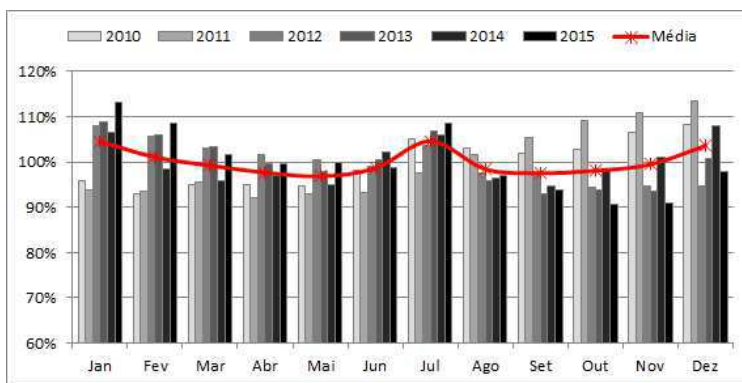
4.3 Mês do ano

A variável de mês pode indicar que há mais operações de pousos e decolagens nos meses de férias escolares como janeiro, julho e dezembro e menos operações nos outros meses do ano. Também foi observado que existe a possibilidade de uma mudança na distribuição dos voos por mês: Nos primeiros anos da análise deste trabalho a aviação cresceu ao longo dos meses, já nos últimos anos analisados a aviação diminuiu no decorrer do ano.

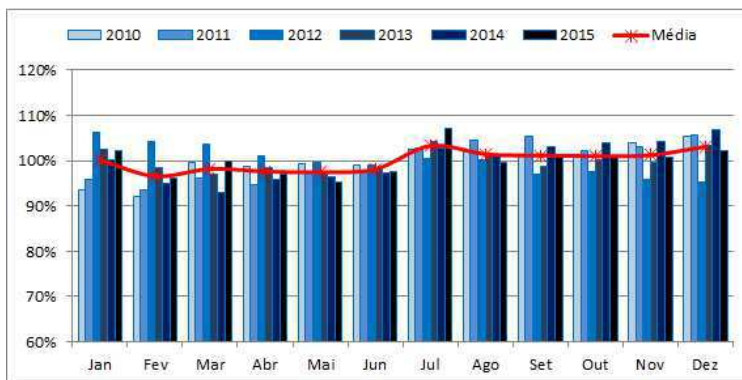
A Figura 4-3 (a), no gráfico do aeroporto do Galeão, as colunas acima de 100% indicam que naquele mês, em média, houveram mais pousos e decolagens do que a média de todos os dias do ano. Isso evidencia que, nos anos de 2010 e 2011, os meses do segundo semestre tiveram mais pousos e decolagens do que do primeiro trimestre, comprovando que esses anos foram de crescimento da aviação comercial neste aeroporto. Já no último ano da análise (2015) a comparação entre os semestres é o contrário, sugerindo um cruzamento das variáveis ano e mês para o modelo preditivo.

A Figura 4-3 (b) e (c), no gráfico do aeroporto de Brasília e Guarulhos, os meses do segundo semestre aparentam ter mais operações de pousos e decolagens do que nos meses do primeiro semestre, com exceção de janeiro dos três últimos anos e do ano de 2012, onde este último mostrou ser o ano mais atípico com relação à distribuição da quantidade de operações nos meses no decorrer dos anos analisados neste trabalho.

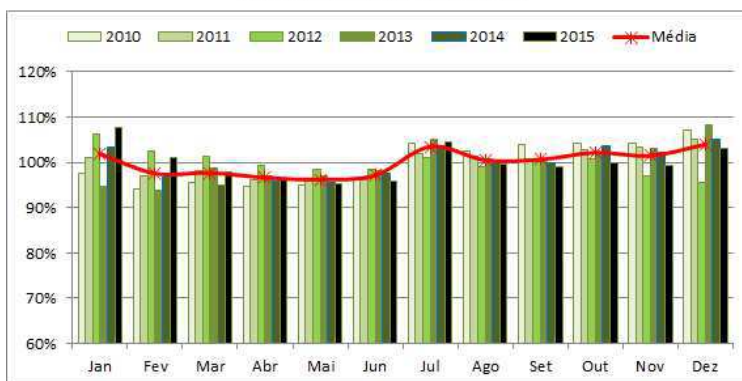
A Figura 4-3 (d), no gráfico do aeroporto de Recife, apresenta com mais evidência a importância das operações nos meses de férias escolares (janeiro, julho e dezembro). Neste aeroporto, o ano que, aparentemente, parece mais atípico é 2010, sendo que 2014 e 2015 mostraram características parecidas e com aparente tendência para o ano seguinte.



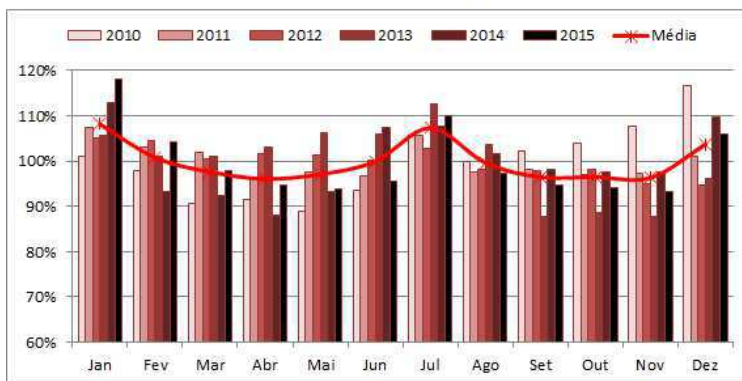
(a) Aeroporto do Galeão



(b) Aeroporto de Brasília



(c) Aeroporto de Guarulhos



(d) Aeroporto de Recife

Figura 4-3 - Gráfico da divisão do total de pousos e decolagens por mês pelo total de pousos e decolagens médio do ano da aviação comercial nos aeroportos do Galeão (a), de Brasília (b), de Guarulhos (c) e de Recife (d).

4.4 Feriados

Os feriados podem ser divididos em diferentes grupos:

- Feriados nacionais fixos no dia da semana: Carnaval (terça), Semana Santa (sexta) e Corpus Christi (quinta);
- Feriados nacionais fixos no dia do mês: Dia de Tiradentes (21/04), Dia do Trabalho (01/05), Dia da Independência (07/09), Dia de Nossa Senhora (12/10), Dia de Finados (02/11) e Dia da Proclamação da República (15/11);
- Feriados de final de ano: Natal (24/12 e 25/12) e Réveillon (31/12 e 01/01);
- Feriados estaduais ou municipais, como o dia de São Sebastião, no Rio de Janeiro, em 20 de janeiro.

A diferença entre os tipos de feriados se deve ao impacto que podem causar no tráfego aéreo brasileiro: ora são feriados fixos no dia da semana, compondo feriados prolongados e se repetindo todos os anos; ora são feriados que variam o dia da semana e que podem ocasionar ou não fins de semanas prolongados, ocorrendo esporadicamente em diferentes meses do ano; e ora são feriados de final de ano, coincidindo com período de férias escolares e recessos de trabalho.

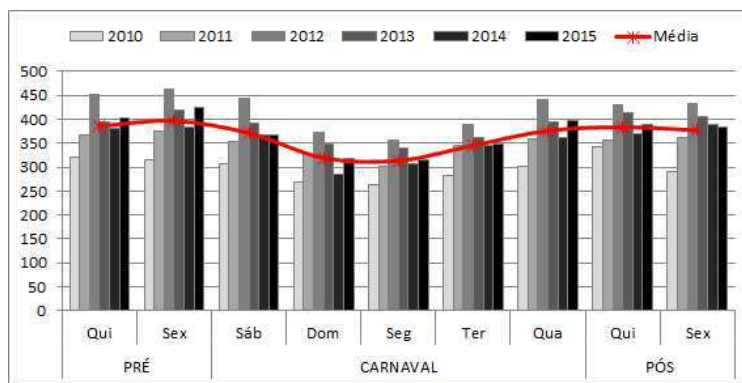
4.4.1 Feriados nacionais fixos no dia da semana

Os feriados nacionais fixos no dia da semana têm, como uma de suas características, o comportamento da demanda por tráfego aéreo mais previsível. Com isso, as companhias aéreas se preparam adicionando voos no início e no final do feriado e cancelando voos no seu interim.

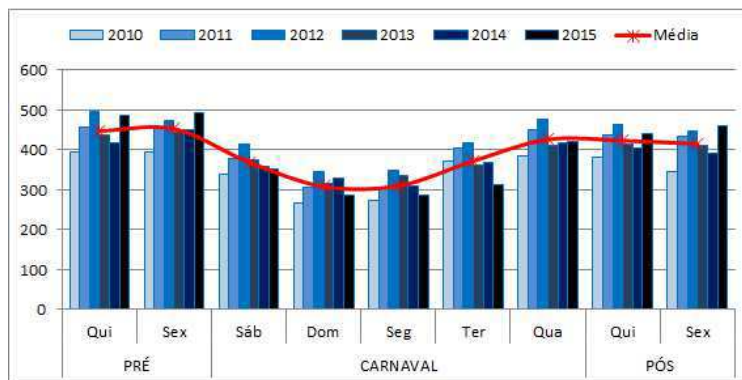
A Figura 4-4, nos gráficos do aeroporto do Galeão, Guarulhos e Recife, há a quantidade de operações diárias durante o carnaval nos anos de 2010 a 2015. Observa-se uma semelhança na quantidade de operações entre o sábado de carnaval e a sexta-feira de pré-carnaval e bem diferente do domingo de carnaval (já vimos na análise por dia da semana que no final de semana essas quantidades, normalmente, são menores e parecidas),

enquanto que na segunda e terça-feira de carnaval o total de pousos e decolagens se aproxima mais do total de domingo do que o total de quarta, quinta ou sexta-feira.

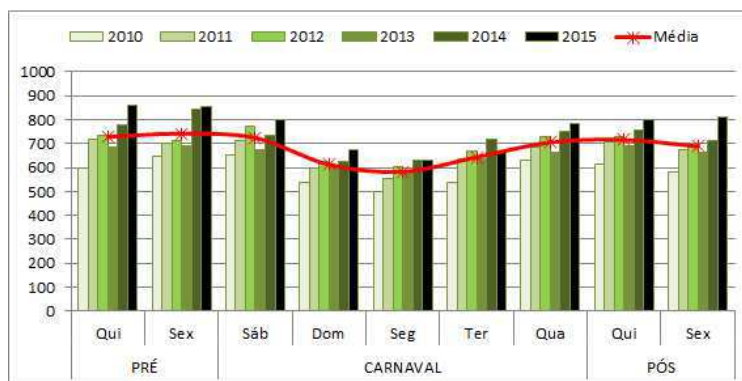
A Figura 4-4 (b), no gráfico do aeroporto de Brasília, tem características parecidas com o do aeroporto do Galeão. A maior diferença está na quantidade de operações que ocorreram no sábado, que registrou uma quantidade intermediária entre a quantidade de domingo (que normalmente seria mais parecida) e a de sexta-feira.



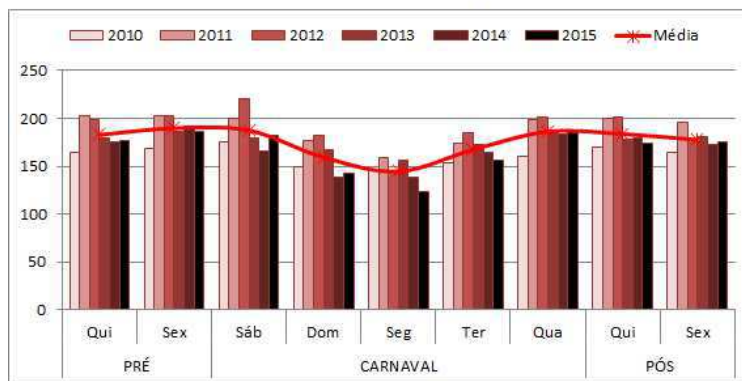
(a) Aeroporto do Galeão



(b) Aeroporto de Brasília



(c) Aeroporto de Guarulhos



(d) Aeroporto de Recife

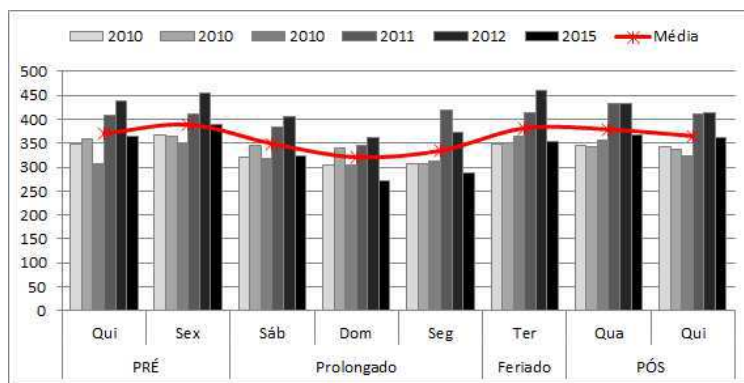
Figura 4-4 - Gráfico do total de pousos e decolagens por dia durante o carnaval e por ano da aviação comercial nos aeroportos do Galeão (a), de Brasília (b), de Guarulhos (c) e de Recife (d).

4.4.2 Feriados nacionais fixos no dia do mês

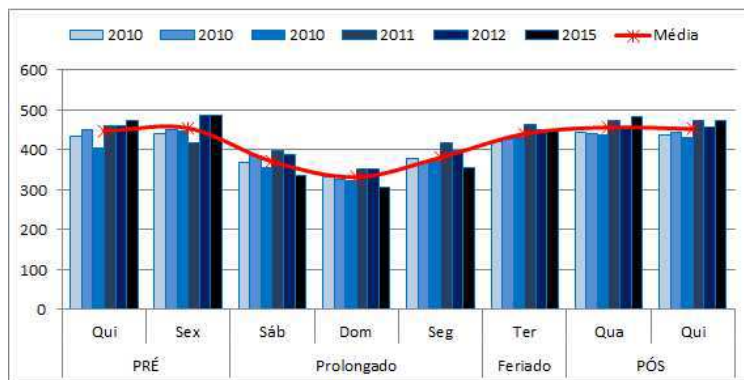
Os feriados nacionais fixos no dia do mês podem impactar o tráfego aéreo de forma diferente, dependendo do dia da semana em que caírem. Sendo numa segunda, terça, quinta ou sexta-feira, podem gerar um feriado prolongado (junto com o final de semana), aumentando a quantidade de passageiros a turismo e diminuindo a quantidade de passageiros a trabalho.

Na Figura 4-5, há o total de operações de pousos e decolagens durante os feriados que caíram numa terça-feira nos anos de 2010, 2011, 2012 e 2015 (não houve feriado deste tipo na terça-feira nos anos de 2013 e 2014) por aeroporto analisado. Na segunda-feira, o total de pousos e decolagens se aproxima mais do total de domingo do que o total de terça ou quarta-feira, diferente do que foi observado no feriado de Carnaval. Os dias de feriados utilizados neste gráfico foram: 07/09/2010 (Dia da Independência), 12/10/2010 (Dia de Nossa Senhora), 02/11/2010 (Dia de Finados), 15/11/2011 (Dia da Proclamação de República), 01/05/2012 (Dia do Trabalho) e 21/04/2015 (Dia de Tiradentes).

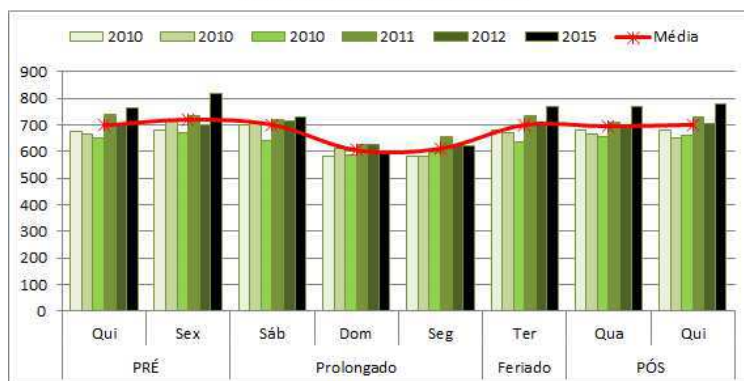
A Figura 4-5, na comparação com o gráfico do Carnaval (Figura 4-4), fica mais evidente que a segunda e a terça-feira de Carnaval indicam uma redução maior de operações de pousos e decolagens.



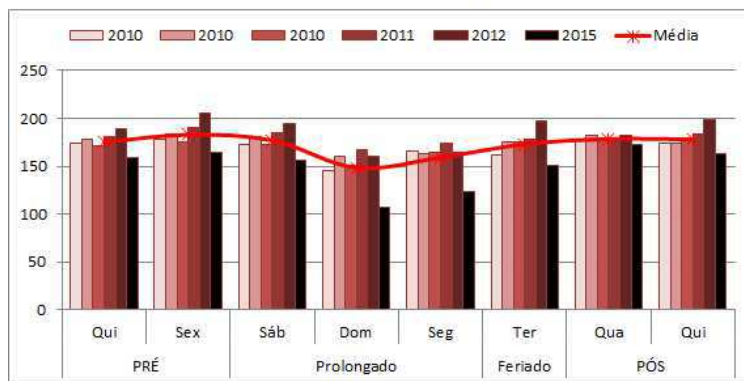
(a) Aeroporto do Galeão



(b) Aeroporto de Brasília



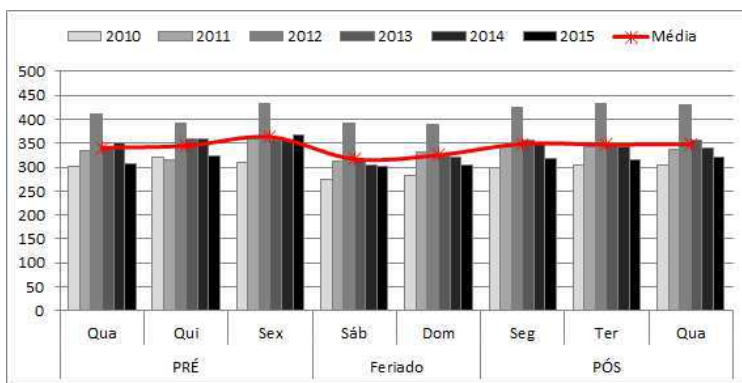
(c) Aeroporto de Guarulhos



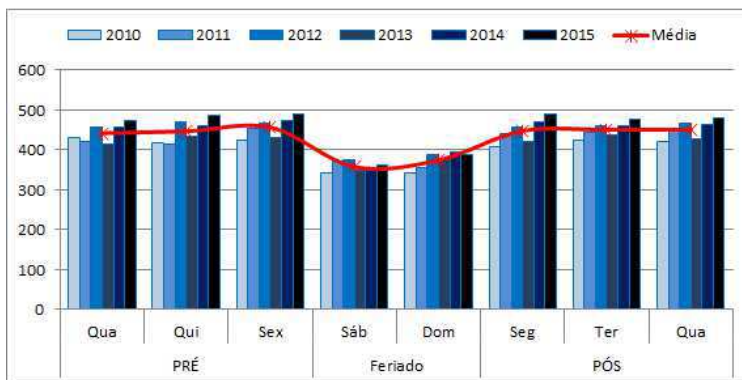
(d) Aeroporto de Recife

Figura 4-5 - Gráfico do total de pousos e decolagens por dia durante feriados que caíram na terça-feira e por ano da aviação comercial nos aeroportos do Galeão (a), de Brasília (b), de Guarulhos (c) e de Recife (d).

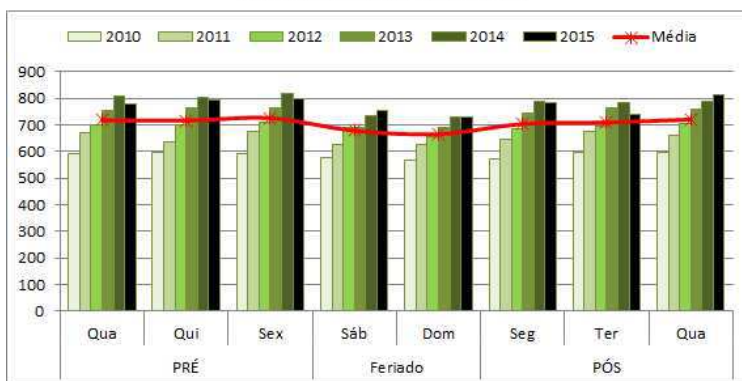
Na Figura 4-6, neste mesmo tipo de feriado nacional fixo no dia do mês, há o total de operações de pousos e decolagens durante feriados que caíram no final de semana nos anos de 2010 a 2015 por aeroporto analisado. O impacto é praticamente imperceptível, sabendo-se que já pode existir uma redução no total de operações no sábado e no domingo. Os dias de feriados utilizados neste gráfico foram: 01/05/2010 (Dia do Trabalho), 01/05/2011 (Dia do Trabalho), 21/04/2012 (Dia de Tiradentes), 02/11/2013 (Dia de Finados), 07/09/2014 (Dia da Independência) e 15/11/2015 (Dia da Proclamação da República).



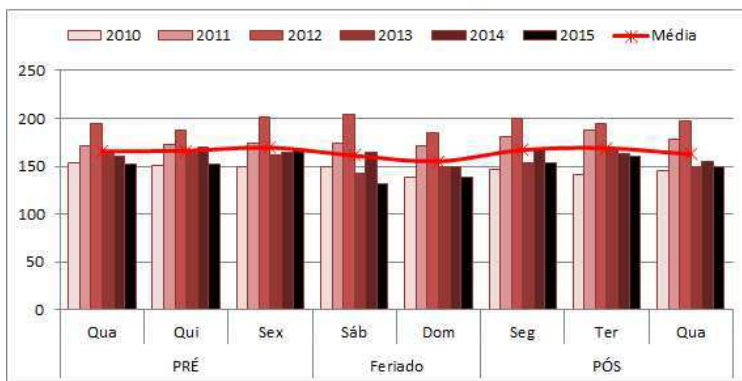
(a) Aeroporto do Galeão



(b) Aeroporto de Brasília



(c) Aeroporto de Guarulhos



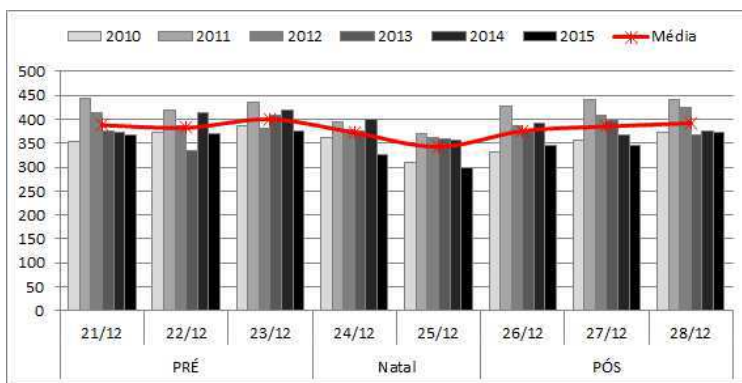
(d) Aeroporto de Recife

Figura 4-6 - Gráfico do total de pousos e decolagens por dia durante feriados que caíram no final de semana e por ano da aviação comercial nos aeroportos do Galeão (a), de Brasília (b), de Guarulhos (c) e de Recife (d).

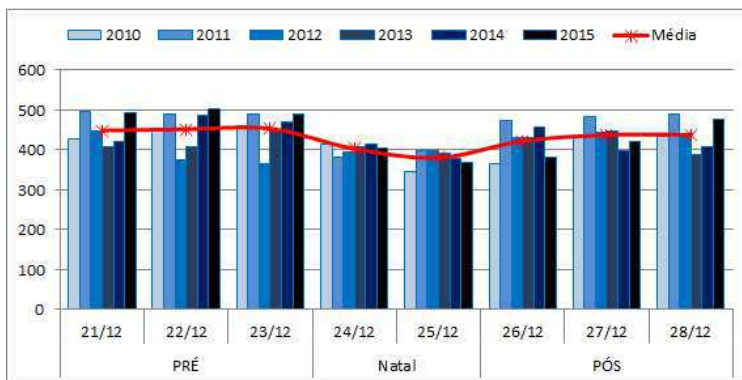
4.4.3 Feriados de final de ano

Os feriados de final de ano têm, como uma de suas características, certa independência do dia da semana em que ocorre. Um dos motivos pode ser o fato de estarem numa época do ano de alta demanda do tráfego aéreo, onde as companhias aéreas se preparam adicionando voos neste período e cancelando voos no meio dos feriados de Natal e Réveillon.

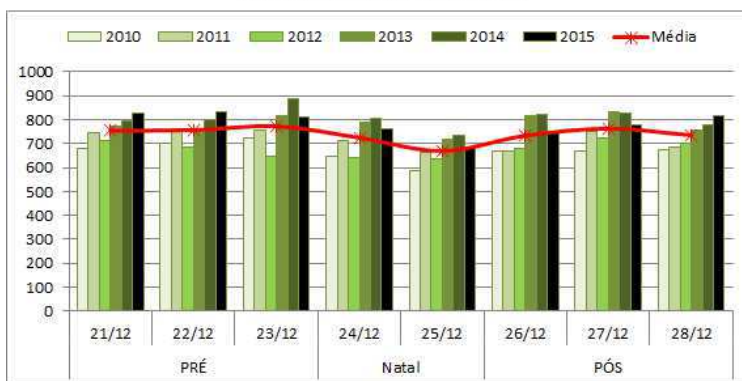
Na Figura 4-7, há a quantidade de operações diárias durante o Natal nos anos de 2010 a 2015 por aeroporto analisado. Observa-se que sempre há uma redução durante o feriado, principalmente no dia 25, enquanto que nos dias “pré” e “pós” (antes e depois do feriado) se mantêm a regularidade, ficando próximo ao que é observado durante os dias da semana. Em 2012, o sábado e domingo ocorreram nos dias 22/12 e 23/12, respectivamente, enquanto que em 2013 ocorreram nos dias 21/12 e 22/12, por essa razão esses dias apresentam reduções na quantidade de operações de pousos e decolagens, conforme já foi observado nos finais de semana.



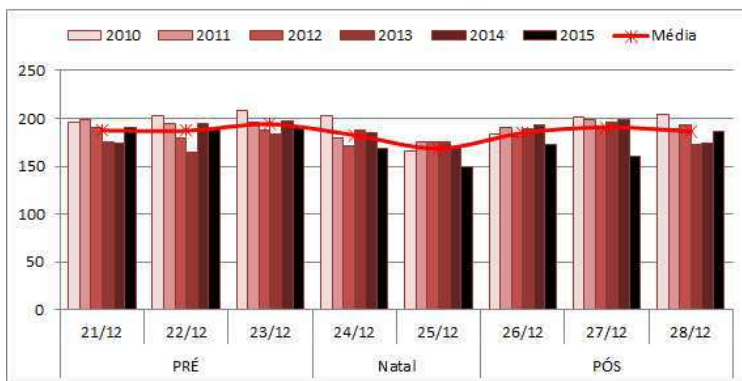
(a) Aeroporto do Galeão



(b) Aeroporto de Brasília



(c) Aeroporto de Guarulhos



(d) Aeroporto de Recife

Figura 4-7 - Gráfico do total de pousos e decolagens por dia durante o Natal por ano da aviação comercial nos aeroportos do Galeão (a), de Brasília (b), de Guarulhos (c) e de Recife (d).

4.4.4 Feriados Municipais e estaduais

Os feriados estaduais e municipais também podem impactar o tráfego aéreo de uma região. Em razão disso, separamos os principais feriados das regiões onde os aeroportos estão localizados e fizemos a análise descritiva igual aos feriados nacionais fixos no dia do mês.

Os feriados analisados para cada aeroporto foram:

- Aeroporto do Galeão: Dia de São Sebastião (20/01), Dia de São Jorge (23/04) e Dia da Consciência Negra (20/11);
- Aeroporto de Brasília: Dia da Fundação de Brasília (21/04) e Dia do Evangélico (30/11);
- Aeroporto de Guarulhos: Dia da Revolução Constitucionalista de 1932 (09/07) e Dia do aniversário de São Paulo (25/01);
- Aeroporto de Recife: Dia de São João (24/06).

Pelos modelos analisados neste trabalho, nenhum desses feriados citados demonstrou impactar na quantidade de pousos e decolagens dos aeroportos em estudo.

4.5 Grandes Eventos

Alguns eventos foram analisados, através do uso de modelos, para ver se impactavam a quantidade de operações de pousos e decolagens da aviação comercial. Eventos como a Copa do Mundo de Futebol (2014), Copa das Confederações (2013) e Jornada Mundial da Juventude (2013). Dentre esses eventos, apenas a Copa do Mundo mostrou impactar a quantidade de operações, principalmente no jogo final no Rio de Janeiro, quando o aeroporto do Galeão registrou o maior número de pousos e decolagens da aviação comercial dos últimos 5 anos (provavelmente da sua história).

Outro evento que mereceu atenção foi o incentivo fiscal dado por alguns estados para diminuir o preço do combustível QAV. A variação, que pode ter ocorrido, no número de operações de pousos e decolagens dos aeroportos, foi diluída ao longo do tempo e tornou

difícil a determinação do período de tempo em que ocorreu, tornando essa análise difícil de ser quantificada nesse trabalho.

4.6 Ano

Os anos serão analisados separadamente, mas a principal contribuição está na análise do ano junto com outras variáveis (interação entre as variáveis) como dias da semana e meses (já analisado). Assim, podemos perceber que há indícios de mudanças do comportamento do número de operações de pousos e decolagens da aviação comercial ao longo dos anos.

4.7 Meteorologia

As condições meteorológicas podem interferir no funcionamento dos aeroportos. São as torres de controle (TWR) que definem quando o aeroporto está impróprio para pousos ou decolagens (em outras palavras, abaixo dos mínimos). As TWR baseiam essa definição usando informações meteorológicas disponíveis (METAR) e adotando alguns critérios como:

- Visibilidade – que é a distância máxima horizontal que a torre consegue enxergar (a olho nu);
- Teto – que é a distância vertical da altura da base das nuvens.

Cada aeroporto segue um patamar de Visibilidade e de Teto, podendo estar funcionando de 3 formas diferentes:

- Visual – quando não há restrições de Visibilidade e Teto;
- Instrumento – quando há restrições de Visibilidade ou Teto, tendo as aeronaves que utilizarem equipamentos especiais para efetuarem o pouso ou a decolagem;
- Abaixo dos mínimos – quando há restrições de Visibilidade ou Teto (maiores do que a instrumento). Nesses casos, os pousos ou as decolagens ficam impróprios (se o piloto quiser tentar pousar ele pode, mas assume o completo risco da manobra).

Os critérios usados nos aeroportos para o funcionamento abaixo dos mínimos podem variar até no mesmo aeroporto, pois dependem do equipamento especial que a aeronave possui, ou se o equipamento especial do aeroporto está em manutenção ou foi substituído. Atualmente, o aeroporto de Guarulhos, por exemplo, pode receber aeronaves com total restrição de Visibilidade e de Teto (0 metros e 0 pés), dependendo dos equipamentos disponíveis na aeronave.

As condições meteorológicas podem impactar tanto o número de operações de pousos e decolagens da aviação comercial no próprio aeroporto, quanto às operações de outros aeroportos que têm algum tipo de relação com o aeroporto em questão. Há dois cenários principais para serem analisado:

- Redução do número de operações no aeroporto “A”: pode ocorrer quando o aeroporto “A” está abaixo dos mínimos; ou quando o aeroporto “B”, que tem grande fluxo de aeronaves com o aeroporto “A”, está abaixo dos mínimos e os voos não decolam para o aeroporto “A”;
- Aumento do número de operações no aeroporto “A”: pode ocorrer quando o aeroporto “B” está abaixo dos mínimos, e os voos que estavam indo para este destino são desviados para o aeroporto “A”; ou quando, em caso parecido com o citado anteriormente, aumenta-se a procura por novos voos do aeroporto “A” por estar próximo do aeroporto “B” (caso do Galeão quando o Santos Dumont está abaixo dos mínimos).

Os critérios adotados para o aeroporto estar funcionando abaixo dos mínimos ou a instrumento variam para cada aeroporto:

- Galeão: Abaixo do mínimo com Visibilidade menor que 400 metros ou Teto menor que 100 pés; Instrumento com Visibilidade entre 400 e 5.000 metros ou Teto entre 100 e 1.500 pés;
- Brasília: Visibilidade menor que 800 metros ou Teto menor que 200 pés; Instrumento com Visibilidade entre 800 e 5.000 metros ou Teto entre 200 e 1.500 pés;
- Guarulhos: Visibilidade menor que 400 metros ou Teto menor que 100 pés; Instrumento com Visibilidade entre 400 e 5.000 metros ou Teto entre 100 e 1.500 pés;

- Recife: Visibilidade menor que 800 metros ou Teto menor que 200 pés; Instrumento com Visibilidade entre 800 e 5.000 metros ou Teto entre 200 e 1.500 pés;
- Santos Dumont/RJ: Visibilidade menor que 1.600 metros ou Teto menor que 300 pés; Instrumento com Visibilidade entre 1.600 e 5.000 metros ou Teto entre 300 e 1.500 pés;
- Congonhas/SP: Visibilidade menor que 1.200 metros ou Teto menor que 200 pés; Instrumento com Visibilidade entre 1.200 e 5.000 metros ou Teto entre 200 e 1.500 pés.

Na Figura 4-8, dentre os aeroportos analisados, o que mais foi impactado pela meteorologia ficando abaixo dos mínimos, foi o aeroporto de Guarulhos. Os aeroportos que também serão analisados, porque podem impactar os que estão sendo analisados neste trabalho, também ficaram alguns dias abaixo dos mínimos (Santos Dumont e Congonhas). O gráfico mostra a quantidade de dias que o aeroporto ficou abaixo dos mínimos por determinado tempo (não necessariamente de forma ininterrupta), onde não se contabilizam os dias em que o aeroporto ficou abaixo dos mínimos em menos de uma hora (por ser considerado um tempo insuficiente para cancelar uma quantidade significativa de voos do total do dia).

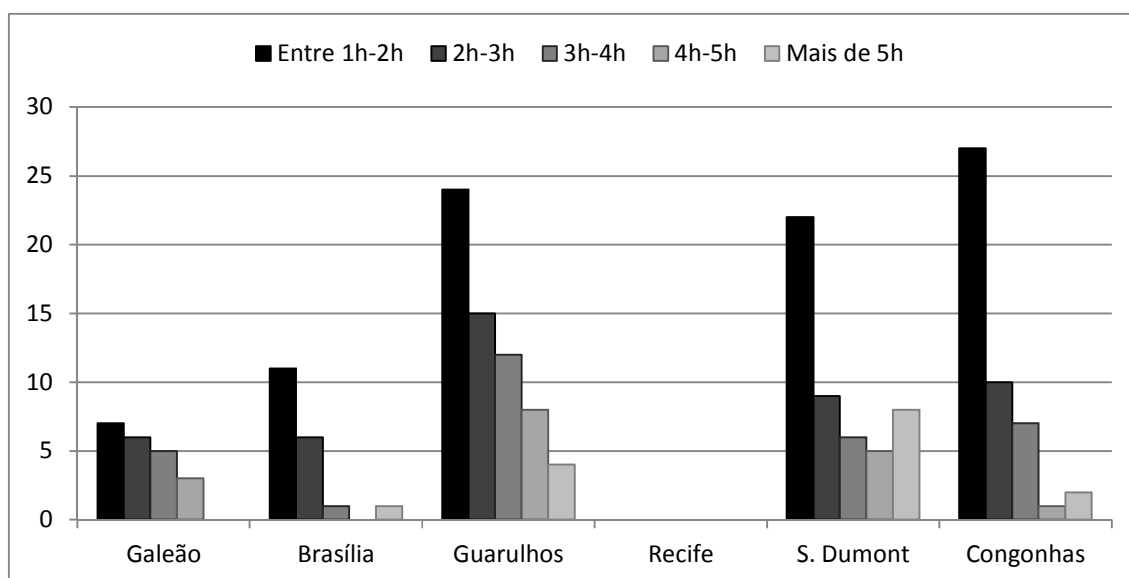


Figura 4-8 - Gráfico com a quantidade de dia, de 1º de janeiro de 2010 a 31 de dezembro de 2015, que os aeroportos ficaram funcionando, com condições meteorológicas, abaixo dos mínimos.

A Figura 4-9, mostra o gráfico da quantidade de dias que os aeroportos analisados mais o Santos Dumont e Congonhas ficaram funcionando a instrumento. Os grupos de tempo que foram considerados para serem analisados foram entre 4 horas e 8 horas, 8 horas e 12 horas e mais de 12 horas.

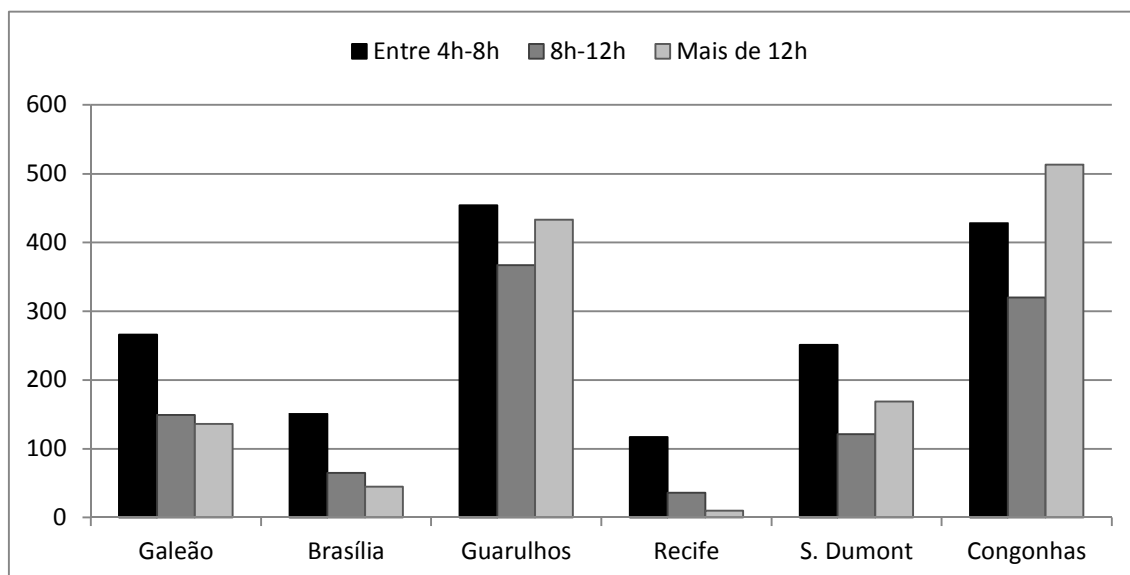


Figura 4-9 - Gráfico com a quantidade de dia, de 1º de janeiro de 2010 a 31 de dezembro de 2015, que os aeroportos ficaram funcionando, com condições meteorológicas, a instrumento.

4.8 Obras na Pista dos Aeroportos

Algumas obras de manutenção nas pistas dos aeroportos podem restringir a quantidade de pousos e decolagens. Esta restrição pode ser parcial (afetando apenas alguns tipos de aeronaves) ou total (fechando a pista para pousos e decolagens).

A principal obra observada nos aeroportos analisados neste trabalho ocorreu em Recife, onde a pista passou por manutenção de setembro de 2013 à dezembro de 2014, com alguns intervalos na obra, principalmente no verão e durante a Copa do Mundo de futebol em 2014. Durante a obra, a pista do aeroporto ficou fechada na madrugada, afetando a quantidade de pousos e decolagens do aeroporto, como pode ser observado na Figura 4-1 (d).

5 METODOLOGIA

Nesta seção, passaremos pelos modelos que serão utilizados para fazer a previsão proposta neste trabalho, assim como, veremos os testes estatísticos empregados e as variáveis binárias que terão a sua significância analisada pelo ajuste final do modelo.

5.1 Processo Estocástico Estacionário

O modelo de série temporal que queremos criar é um processo estocástico, pois são variáveis aleatórias ordenadas no tempo (Gujarati 2006). Para podermos usar os modelos Box-Jenkins (que serão explicados em seguida), a série temporal deve ser estacionária (de segunda ordem), ou seja, com média e variância constantes ao longo do tempo e covariância dependente apenas da distância de tempo entre as variáveis da série temporal (defasagem).

Para os modelos preditivos que estamos buscando, a estacionariedade será alcançada quando utilizarmos covariáveis junto com o modelo de Box-Jenkins (ARMAX).

5.2 Modelos de Autoregressivos e de Médias Móveis (ARMA)

Os modelos que usaremos neste trabalho são conhecidos como *metodologia de Box-Jenkins* (Box e Jenkins 1976), que permite usar o passado da série de dados para prever o futuro da mesma.

Podemos separar em dois tipos de modelos:

- Autoregressivos (AR): que usa valores defasados da própria série de dados; e
- Médias Móveis (MA): que usa os erros estimados da própria série de dados.

Analisaremos cada um dos modelos separadamente e depois mostraremos que eles também podem ser utilizados juntos e com covariáveis (variáveis explicativas independentes), que colocaremos no modelo para explicar a tendência, as sazonalidades e os eventos especiais e ainda, para satisfazer as restrições que estes modelos sejam

estacionários, inversíveis e que, portanto, seja possível a inferência estatística com estas ferramentas.

5.2.1 Modelos Autoregressivos (AR)

Os modelos Autoregressivos de primeira ordem, AR(1), podem ser escrito conforme a equação (1).

$$Y_t = \mu + \alpha_1 * Y_{t-1} + u_t \quad (1)$$

Onde Y_t é o valor da série temporal no tempo t , Y_{t-1} é o valor da série temporal a um período atrás, no tempo $t-1$ e μ é o valor da constante do modelo. Assim, no nosso caso, Y_t é o total de operações de pousos e decolagens da aviação comercial em um aeroporto analisado no dia t e Y_{t-1} é a quantidade do dia anterior.

O erro associado a este modelo, representado por u_t , deve ter distribuição normal, média zero, variância constante e não ter autocorrelação (ruído branco). O coeficiente α_1 é a proporção da série no tempo $t-1$, logo, o valor da série no tempo t é uma proporção no tempo $t-1$ mais uma variável aleatória do tipo ruído branco.

Os modelos Autoregressivos podem ser expandidos para usarem mais tempo para trás, ou seja, podem ser de ordem p , AR(p), e escritos da seguinte forma na equação (2).

$$Y_t = \mu + \alpha_1 * Y_{t-1} + \alpha_2 * Y_{t-2} + \alpha_3 * Y_{t-3} + \dots + \alpha_p * Y_{t-p} + u_t \quad (2)$$

A interpretação é parecida com a dada em AR(1), só que agora o Y_t , valor da série temporal no tempo t , é o resultado da soma de proporções α_i dos valores da série, onde i varia de 1 até p , mais um erro u_t (ruído branco). No nosso caso, seria o total de operações de pousos e decolagens no dia t , sendo a soma de proporções da quantidade de operações de dias anteriores até o dia p mais uma perturbação aleatória.

5.2.2 Modelos Médias Móveis (MA)

Os modelos Médias Móveis de primeira ordem, MA(1), podem ser escrito conforme a equação (3).

$$Y_t = \theta + u_t + \beta_1 * u_{t-1} \quad (3)$$

Onde Y_t é o valor da série temporal no tempo t , θ é um valor constante, e os erros continuam sendo do tipo ruído branco, u no tempo t e $t-1$. É usado uma proporção do erro do passado, essa proporção é β_1 .

No nosso trabalho seria como estimar a quantidade de operações de pousos e decolagens através de uma constante mais uma proporção dos erros aleatórios do mesmo dia e do dia anterior.

Os modelos de Médias Móveis podem ser expandidos para usarem mais tempo para trás, ou seja, podem ser de ordem q , MA(q), e escritos da seguinte forma na equação (4).

$$Y_t = \theta + u_t + \beta_1 * u_{t-1} + \beta_2 * u_{t-2} + \beta_3 * u_{t-3} + \dots + \beta_q * u_{t-q} \quad (4)$$

A interpretação é parecida com a dada em MA(1), onde é utilizada a soma de mais proporções de erros aleatórios do passado, até o dia q do passado.

5.2.3 Modelos Autoregressivos e de Médias Móveis (ARMA)

O modelo pode ser uma combinação dos dois modelos explicados anteriormente, onde um ARMA(1, 1) pode ser escrito conforme a equação (5).

$$Y_t = \theta + \alpha_1 * Y_{t-1} + u_t + \beta_1 * u_{t-1} \quad (5)$$

A explicação dos termos é igual à dada nas equações anteriores. Y_t , que é o valor da série temporal no tempo t , é a combinação de uma constante θ com a proporção de Y_{t-1} , valor da série no tempo $t-1$, e dos erros aleatórios u , nos tempos t e $t-1$.

Este modelo também pode ser expandido para mais tempo para trás, ou seja, pode ser de ordem ARMA(p , q), onde p é da parte Auto-regressiva e q da parte de Médias Móveis, e escrito da seguinte forma na equação (6).

$$Y_t = \theta + \alpha_1 * Y_{t-1} + \dots + \alpha_p * Y_{t-p} + u_t + \beta_1 * u_{t-1} + \dots + \beta_q * u_{t-q} \quad (6)$$

O passado pode ser usado com tempos diferentes, onde o Y_t é o valor da série temporal no tempo t e seu valor é estimado através da soma de:

- Constante θ ;

- Soma das proporções da própria série temporal (Y_{t-i}) em tempos passados ($i = 1$ até p dias do passado, “dias” é o intervalo de tempo usado neste trabalho); e
- Soma das proporções dos erros (u_{t-j}) do presente e do passado ($j = 0$ até q dias do passado).

5.3 Modelos ARMAX

A principal diferença entre o modelo ARMAX e o modelo ARMA é a presença de covariáveis junto ao modelo ARMA.

As covariáveis (variáveis binárias e uma variável ordenada de 1 até 2.191) entram no modelo para tentar explicar qualquer tendência ou sazonalidade na série de dados. Removidas estas tendências, a série de dados passa a ser estacionária e, portanto, pode ser analisada pelos modelos ARMA (estacionários).

Este tipo de estimação pode ser feito em dois passos, mas a abordagem ARMAX combina os dois procedimentos em um único e isto melhora as propriedades estatísticas dos estimadores destes parâmetros do modelo.

O modelo final ARMAX(p,q), complementando a equação (6), pode ser visto na equação (7).

$$Y_t - \sum_{j=1}^p \alpha_j Y_{t-j} = \sum_{j=1}^q \beta_j u_{t-j} + \sum_{j=1}^M \gamma_j d_{t,j} + u_t \quad (7)$$

Onde o primeiro componente (Y_t) é o valor da série temporal no tempo t , o segundo componente é a parte do AR(p), o terceiro é a parte MA(q) e o quarto e última parte são as covariáveis (total de “M”), sendo $d_{t,j}$ o valor da j -ésima covariável no tempo no tempo t (a constante do modelo está aparecendo como uma covariável).

Os critérios adotados para a análise dos coeficientes e de seus desvios ainda serão explicadas ao longo deste trabalho.

5.4 Função de Autocorrelação e Correlograma

Para verificarmos que a série temporal que vamos analisar é estacionária, podemos usar a função de autocorrelação, onde a função é definida como (Gujarati 2006) na equação (8).

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (8)$$

Onde γ_k é a covariância com defasagem k e γ_0 é a variância. Vamos definir a seguir como calculamos cada uma delas.

Como utilizamos uma amostra da série temporal para estimá-la, calcularemos a função de autocorrelação amostral ($\hat{\rho}_k$) como sendo na equação (9).

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \quad (9)$$

Onde $\hat{\gamma}_k$ e $\hat{\gamma}_0$ podem ser calculados como nas equações (10) e (11), onde \bar{Y} é a média aritmética.

$$\hat{\gamma}_k = \frac{\sum(Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{n} \quad (10)$$

$$\hat{\gamma}_0 = \frac{\sum(Y_t - \bar{Y})^2}{n} \quad (11)$$

O correlograma é uma tabela com os resultados da equação ($\hat{\rho}_k$), onde podemos verificar que se a função de autocorrelação amostral diminuir depois de uma defasagem indica que a série é estacionária (nos casos de Médias Móveis).

O conceito que também aparece no correlograma é a função de autocorrelação parcial, que é a correlação entre Y_t e Y_{t-k} , depois de removida a influência dos Y_{t-i} intermediários, onde $i = 1, \dots, (k-1)$. Onde podemos verificar que se a função de autocorrelação parcial amostral diminuir depois de uma defasagem indica que a série é estacionária (nos casos de Autoregressivos).

5.5 Testes de Normalidades e Independência dos Resíduos

Vamos descrever alguns testes que utilizaremos para verificar os ajustes dos modelos e para atender as restrições de uso dos modelos preditivos.

5.5.1 Teste de Jarque-Bera

Utilizaremos este teste para verificar se a série analisada segue uma distribuição normal. A estatística de teste utilizada está na equação (12).

$$JB = n \left[\frac{S^2}{6} + \frac{(K - 3)^2}{24} \right] \quad (12)$$

Onde n é o tamanho da amostra, S é o coeficiente de assimetria e K é o coeficiente de curtose. Sob a hipótese nula de que a série é normalmente distribuída, a estatística JB segue, assintoticamente, a distribuição de qui-quadrado com 2 graus de liberdade (Gujarati 2006).

5.5.2 Teste de Ljung e Box

O teste de Ljung e Box verifica a presença de autocorrelação na série, onde a hipótese nula é de que as autocorrelações são zero (Hill 2010), até um número grande de defasagens. Onde a estatística de teste é Q' como está na equação (13).

$$Q' = T(T + 2) \sum_{s=1}^{\infty} \frac{\hat{\rho}_s^2}{T - s} \quad (13)$$

A distribuição dessa estatística de teste segue distribuição qui-quadrado com “m” graus de liberdade. Se rejeitarmos a hipótese nula de que todas as correlações sejam zero, e aceitarmos a alternativa de que uma ou mais delas são diferentes de zero, logo, não são independentes.

Esta estatística de teste (Q') da equação (13) aparece no correlograma citado neste trabalho.

5.6 Estimação por Máxima Verossimilhança

O método de estimação por máxima verossimilhança trata o problema de estimação da seguinte forma: baseado nos resultados obtidos pela amostra, devemos determinar qual a

distribuição, dentre todas aquelas definidas pelos possíveis valores de seus parâmetros, com maior probabilidade de ter gerado tal amostra.

Seja θ o vetor de parâmetros dos modelos a serem estimados e (Y_1, Y_2, \dots, Y_n) a amostra das observações temporais. Então, a função de máxima verossimilhança é dada pela equação (14), como a função de densidade conjunta de Y_1, Y_2, \dots, Y_n dependendo de θ .

$$L(\theta) = f(Y_1, Y_2, \dots, Y_n | \theta) \quad (14)$$

O estimador de máxima verossimilhança é dado pelo valor $\hat{\theta}$ que maximiza a função $L(\theta)$, como na equação (15).

$$\hat{\theta}_{EMV} = \max_{\theta} L(\theta) \quad (15)$$

Logo, o estimador de máxima verossimilhança escolhe o parâmetro que melhor explica a amostra observada.

Para a estimação do modelo ARMA por máxima verossimilhança utilizaremos os métodos numéricos presentes no programa R (R Core Team 2014) e seus pacotes.

Os erros padrões das estimativas são obtidas pelos resultados clássicos da estimação por máxima verossimilhança através da matriz hessiana. Isto já é fornecido pelo programa R.

5.7 Critérios para comparar os Modelos (AIC e BIC)

Utilizaremos 2 critérios para comparar os diferentes modelos ajustados, o principal será o Critério de Informação Akaike (AIC, abreviação do inglês) e o outro será o Critério de Informação Bayesiana ou de Schwarz (BIC, abreviação do inglês). O segundo critério ajudará a escolha em caso de empate (ou proximidade de valores) do primeiro critério. O critério AIC usa a informação da soma dos quadrados dos resíduos (SQR) e impõe uma penalidade ao acrescentar mais regressores (variáveis usadas para fazer a estimação da variável dependente) ao modelo (Akaike 1974), que é definido na equação (16).

$$AIC = -2\ln(L(\hat{\theta})) + 2p \quad (16)$$

Onde p é o número de regressores do modelo, $L(\hat{\theta})$ é a função de máximo verossimilhança do modelo no ponto $\hat{\theta}$.

Quanto menor o valor do AIC, melhor será o modelo, pois o critério tem uma parte que gera qualidade ao critério e outra que gera uma penalidade pela quantidade de regressores utilizados ($2p$).

O outro critério para a escolha é o BIC, que é definido na equação (17).

$$BIC = -2\ln(L(\hat{\theta})) + p \ln(p) \quad (17)$$

A diferença deste critério para o primeiro é no fator de punição ao modelo que tem mais regressores (Gujarati 2006), onde mudam de $2p$ na equação (16), para $p \ln(p)$ na equação (17).

5.8 Variável Binária (Dicotômica)

As variáveis binárias podem ser usadas dentro de modelos para indicar a presença de uma característica com referência à ausência dela. Onde uma variável binária (D) pode ser escrita como:

- $D = 1$, se a característica estiver presente
- $D = 0$, se a característica estiver ausente

Um exemplo dentro do nosso trabalho pode ser a presença de um feriado ou não para estimar o número de operações de pousos e decolagens para a data em questão.

As variáveis binárias usadas dentro de modelos de regressão ou de previsão também são conhecidas como variáveis *dummys*, que pode usar uma variável binária ou um conjunto de variáveis binárias para a presença ou ausência de características relacionadas. Podemos usar, como exemplo, o uso de 6 binárias para caracterizar os dias da semana no modelo, podendo escrever como:

- $D1 = 1$, se o dia for terça-feira, e $D1 = 0$, se o dia não for terça-feira;
- $D2 = 1$, se o dia for quarta-feira, e $D2 = 0$, se o dia não for quarta-feira;
- $D3 = 1$, se o dia for quinta-feira, e $D3 = 0$, se o dia não for quinta-feira;
- $D4 = 1$, se o dia for sexta-feira, e $D4 = 0$, se o dia não for sexta-feira;
- $D5 = 1$, se o dia for sábado, e $D5 = 0$, se o dia não for sábado;
- $D6 = 1$, se o dia for domingo, e $D6 = 0$, se o dia não for domingo.

A segunda-feira não tem uma variável específica porque ela é a referência dessas outras, ou seja, quando as 6 binárias citadas são iguais a zero é o dia segunda-feira no modelo, logo o coeficiente ligado a uma dessas 6 binárias é em relação à segunda-feira. Se no modelo o coeficiente da variável binária D1 for 20, significa que a quantidade de operações de pousos e decolagens da aviação comercial na terça-feira tem 20 operações a mais do que na segunda-feira. Do ponto de vista de modelagem é essencial retirar uma delas na presença de uma constante no modelo porque o mesmo apresenta colinearidade perfeita e, portanto, o modelo ficaria não estimável.

5.9 Variáveis Analisadas

Testamos a importância de determinadas variáveis no modelo, onde algumas se mostraram importantes e outras não.

Segue a lista de variáveis e uma breve explicação sobre cada uma delas com o grau de importância para o modelo.

5.9.1 Total de operações de pousos e decolagens da aviação comercial do aeroporto do Galeão

Uma das variáveis dependente a ser modelada é o total de operações de pousos e decolagens da aviação comercial do aeroporto do Galeão, que é uma variável de contagem, como mostra a Figura 5-1. Entretanto, por se ter contagens altas, uma aproximação para um modelo com erros normais está sendo sugerida para todos os aeroportos.

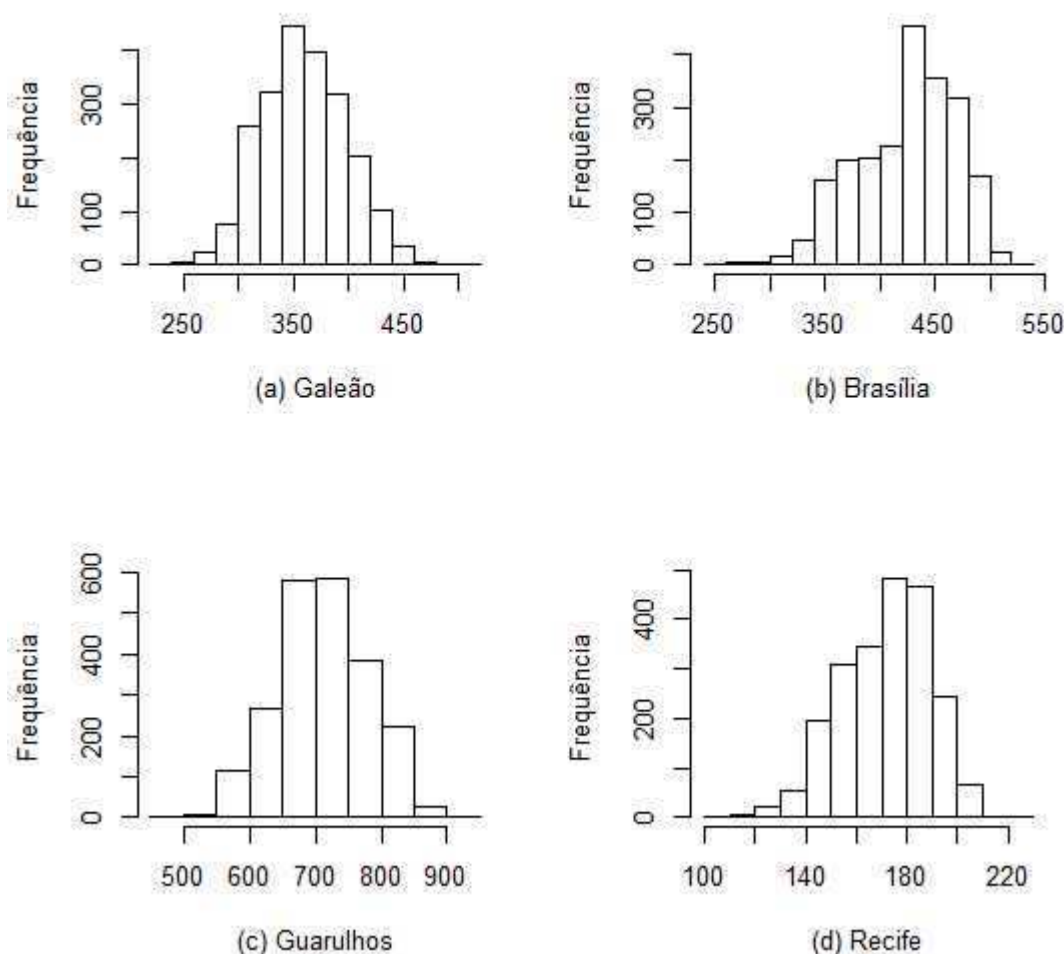


Figura 5-1 - Histograma da quantidade de operações de pousos e decolagens da aviação comercial nos aeroportos do Galeão (a), de Brasília (b), de Guarulhos (c) e de Recife (d).

O total de operações no aeroporto do Galeão apresentou inconsistência nos dados para alguns dias no período analisado neste trabalho, entre 2010 e 2015. Essas datas foram substituídas seguindo o critério descrito na subseção 4.1 (Variáveis Dependentes) deste mesmo trabalho.

As 15 datas que foram substituídas foram: 14/03/2010, 29/03/2010, 05/05/2010, 19/06/2010, 12/08/2010, 19/08/2010, 27/09/2010, 10/11/2010, 05/04/2011, 08/08/2011, 14/09/2011, 20/10/2011, 07/11/2011, 17/05/2013 e 25/07/2013.

5.9.2 Total de operações de pousos e decolagens da aviação comercial do aeroporto de Brasília

O total de operações no aeroporto de Brasília apresentou inconsistência nos dados para alguns dias no período analisado neste trabalho, entre 2010 e 2015. Essas datas foram substituídas seguindo o critério descrito na subseção 4.1 (Variáveis Dependentes) deste mesmo trabalho.

As 6 datas que foram substituídas foram: 18/04/2010, 19/04/2010, 28/05/2011, 15/12/2011, 16/12/2011 e 06/05/2014.

5.9.3 Total de operações de pousos e decolagens da aviação comercial do aeroporto de Guarulhos

O total de operações no aeroporto de Guarulhos apresentou inconsistência nos dados para alguns dias no período analisado neste trabalho, entre 2010 e 2015. Essas datas foram substituídas seguindo o critério descrito na subseção 4.1 (Variáveis Dependentes) deste mesmo trabalho.

As 5 datas que foram substituídas foram: 08/05/2011, 07/06/2011, 01/02/2012, 05/02/2012 e 27/04/2012.

5.9.4 Total de operações de pousos e decolagens da aviação comercial do aeroporto de Recife

O total de operações no aeroporto de Recife apresentou inconsistência nos dados para alguns dias no período analisado neste trabalho, entre 2010 e 2015. Essas datas foram substituídas seguindo o critério descrito na subseção 4.1 (Variáveis Dependentes) deste mesmo trabalho.

As 6 datas que foram substituídas foram: 22/02/2011, 06/05/2011, 25/07/2011, 09/08/2011, 10/08/2011 e 09/04/2012.

5.9.5 Número Ordenado

A verificação se há tendência linear na quantidade de operações de pousos e decolagens da aviação comercial para cada aeroporto analisado será através de uma variável chamada “Tende”, que começará com 1 para o primeiro dia da série (1º de janeiro de 2010), irá aumentando em uma unidade para cada data da série, até chegar em 2.191 no último dia da série (31 de dezembro de 2015).

5.9.6 Dia da semana

Os dias da semana serão testados de diferentes formas, sendo a principal delas através das variáveis binárias para cada dia da semana, usando como referência a segunda-feira. Para detalhar melhor: serão 6 variáveis binárias, uma para cada dia da semana (de terça-feira a domingo), onde, por exemplo, na variável de terça todos os dias que caírem na terça terão valor igual a 1 e nos outros dias terão valor igual a 0. A segunda-feira será a referência dessas binárias, com isso, a interpretação dos coeficientes dessas variáveis será em comparação com a segunda-feira.

As 6 variáveis binárias serão chamadas de:

- “Ter” - para as terças-feiras;
- “Qua” - para as quartas-feiras;
- “Qui” - para as quintas-feiras;
- “Sex” - para as sextas-feiras;
- “Sab” - para os sábados;
- “Dom” - para os domingos.

As binárias por dia da semana serão combinadas com outras variáveis (interação entre variáveis), com isso, verificaremos se há diferença na quantidade de operações de pousos e decolagens por dia da semana e em anos diferentes.

A outra combinação que faremos com o dia da semana será com os feriados. Essa análise não será por interações entre as variáveis (como com os anos), a explicação para isso está na análise das variáveis dos feriados nessa seção.

5.9.7 Meses do ano

Serão analisadas variáveis por meses do ano, sendo 11 variáveis do tipo binária para cada mês do ano usando como referência o mês de janeiro.

Um agrupamento de binárias por meses será testada, uma variável binária para os meses de férias escolares janeiro, julho e dezembro, usando como referência os outros meses. Esta variável será analisada separada das binárias por meses do ano para não dar problema de colinearidade entre as variáveis.

As 11 variáveis binárias serão chamadas de:

- “Fev” - para o mês de fevereiro;
- “Mar” - para o mês de março;
- “Abr” - para o mês de abril;
- “Mai” - para o mês de maio;
- “Jun” - para o mês de junho;
- “Jul” - para o mês de julho;
- “Ago” - para o mês de agosto;
- “Set” - para o mês de setembro;
- “Out” - para o mês de outubro;
- “Nov” - para o mês de novembro;
- “Dez” - para o mês de dezembro.

As binárias por meses serão combinadas com outras variáveis (interação). Com isso verificaremos se há diferença na quantidade de operações de pousos e decolagens por mês e em anos diferentes.

5.9.8 Anos

Serão analisadas variáveis por ano, sendo 5 variáveis do tipo binária para cada ano, usando como referência o ano de 2010.

As 11 variáveis binárias serão chamadas de:

- “Ano_11” - para o ano de 2011;
- “Ano_12” - para o ano de 2012;
- “Ano_13” - para o ano de 2013;
- “Ano_14” - para o ano de 2014;
- “Ano_15” - para o ano de 2015.

Agrupamentos de binárias por ano serão testadas para fazer combinações (interações) com outras variáveis. Por exemplo, vamos agrupar dois anos numa variável binária e combinar com as binárias por mês. Os anos agrupados que se mostraram interessantes para cada aeroporto foram:

- “AnoA” - no Galeão os anos de 2010 e 2011;
- “AnoA” - em Brasília os anos de 2010 e 2011;
- “AnoA” - em Guarulhos os anos de 2010, 2011 e 2013; e
- “AnoA” - em Recife os anos de 2011, 2012 e 2013.

5.9.9 Feriados

Os feriados possuem características diferentes das outras variáveis explicadas até agora. Um feriado pode acontecer num domingo em um ano e numa terça em outro ano, outro feriado pode cair sempre na terça-feira, mas em meses diferentes dependendo do ano.

Com isso, separamos os feriados em grupos diferentes para podermos verificar a influência mais detalhada de cada um no tráfego aéreo de cada aeroporto.

5.9.10 Feriados fixos por dia da semana

Foram testadas variáveis binárias para diferentes dias dos feriados e para cada feriado separadamente. Essas binárias foram baseadas na análise descritiva dos dados. São três feriados nesse grupo: Carnaval, Semana Santa e Corpus Christi.

Para o Carnaval serão analisadas 4 binárias diferentes (ver Tabela 5-1):

- “Carna4” - para a quinta-feira e a sexta-feira anteriores à data;
- “Carna3” - para o sábado de carnaval;
- “Carna2” - para o domingo e a segunda-feira;
- “Carna1” - para a terça-feira de carnaval.

Para a Semana Santa serão analisadas 3 binárias diferentes (ver Tabela 5-1):

- “Santa1” - para a sexta-feira e o sábado no feriado;
- “Santa2” - para o domingo no feriado; e
- “Santa3” - para a segunda-feira depois do feriado.

Para o Corpus Christi serão analisadas 3 binárias diferentes (ver Tabela 5-1):

- “Christi1” - para a sexta-feira no feriado prolongado;
- “Christi2” - para o sábado no feriado prolongado; e
- “Christi3” - para a segunda-feira após o feriado.

5.9.11 Feriados fixos no dia do mês

Todos os feriados foram analisados juntos e verificamos que independente de qual é o feriado, é o dia da semana que influencia mais a variação das operações de pousos e decolagens da aviação comercial nos aeroportos analisados. Dessa forma, foram testadas binárias diferentes dependendo de qual dia da semana ocorrem os feriados. São seis feriados nesse grupo: o Dia de Tiradentes (21/04), o Dia do Trabalho (01/05), o Dia da Independência (07/09), o Dia de Nossa Senhora (12/10), o Dia de Finados (02/11) e o Dia da Proclamação da República (15/11).

Para os feriados que ocorreram às quartas, aos sábados e aos domingos não foram observadas mudanças no comportamento das operações de pousos e decolagens na

semana, então não serão analisadas binárias para os feriados que aconteceram nesses dias da semana.

Para os feriados que ocorreram às segundas, às terças, às quintas e às sextas foram observadas mudanças na semana devido a possibilidade de emendar o feriado ao final de semana (feriado prolongado). Os dias serão separados em 6 binárias (ver Tabela 5-1).

- “Feriado1” - para o dia após os feriados prolongados quando estes terminaram na segunda ou terça-feira, e para os dias anteriores ao feriado prolongado quando estes começaram na quinta ou sexta-feira. Por exemplo, se o feriado é na terça-feira a binária terá valor 1 na quarta desta semana. Outro exemplo é quando o feriado ocorre na sexta-feira, nesse caso essa binária terá valor 1 na quinta-feira desta semana.
- “Feriado2” - para o dia anterior ao feriado, dentro do feriado prolongado, quando estes terminaram na segunda ou terça-feira, e para os dias posteriores ao feriado, ainda dentro do feriado prolongado, quando estes começaram na quinta ou sexta-feira. Por exemplo, se o feriado é na segunda-feira então a binária terá valor 1 no domingo deste feriado prolongado. Outro exemplo é se o feriado ocorre na quinta-feira, nesse caso essa binária terá valor 1 na sexta-feira deste feriado prolongado.
- “Feriado3” - para o dia dos feriados que ocorreram na segunda ou terça-feira.
- “Feriado4” - para o dia dos feriados que ocorreram na quinta ou sexta-feira.
- “Feriado5” - para o domingo dos feriados prolongados, quando estes terminaram na terça-feira, e para o sábado dos feriados prolongados, quando estes começaram na quinta-feira.
- “Feriado6” - para a sexta-feira (antes do feriado) e o sábado dos feriados prolongados quando estes terminaram na terça, e para o domingo e segunda-feira (depois do feriado) dos feriados prolongados quando estes começam na quinta.

	Feriados Fixos no Dia da Semana			Feriados Fixos no Dia do Mês			
	Carnaval	Semana Santa	Corpus Christi	Feriados na Segunda	Feriados na Terça	Feriados na Quinta	Feriados na Sexta
- 5 dias	Carna4						
- 4 dias	Carna4				Feriado6		
- 3 dias	Carna3				Feriado6		
- 2 dias	Carna2				Feriado5		
- 1 dia	Carna2			Feriado2	Feriado2	Feriado1	Feriado1
Feriado	Carna1	Santa1		Feriado3	Feriado3	Feriado4	Feriado4
+ 1 dia		Santa1	Christi1	Feriado1	Feriado1	Feriado2	Feriado2
+ 2 dias		Santa2	Christi2			Feriado5	
+ 3 dias		Santa3				Feriado6	
+ 4 dias			Christi3			Feriado6	
+ 5 dias							

Tabela 5-1 - Quadro de binárias para os feriados que serão testadas para todos os aeroportos. Com exceção dos feriados de Natal e Réveillon.

5.9.12 Feriados de Final de Ano

Foi identificada, pela análise descritiva, que os feriados de Natal e Réveillon possuem características que independem do dia da semana que ocorrem. A época do ano influencia para que dificulte a identificação de um padrão de variabilidade no número de operações de pousos e decolagens da aviação comercial nos aeroportos analisados durante estes feriados. Essas binárias foram baseadas na análise descritiva dos dados.

Para o Natal e Réveillon serão analisadas 4 binárias diferentes:

- “FimAno1” - para o dia do feriado (25/12 e 1/1);
- “FimAno2” - para a véspera do feriado (24/12 e 31/12);
- “FimAno3” - para o dia antes da véspera do feriado e o dia depois do feriado;
- “FimAno4” - para o dia que fica dois dias antes da véspera do feriado e dois dias depois do feriado.

5.9.13 Eventos

Alguns eventos foram analisados de forma descritiva para verificarmos impactos nas operações de pousos e decolagens da aviação comercial nos aeroportos analisados. Estes

impactos podem ser nas datas dos eventos ou próximo a essas datas, por isso se torna necessário a análise descritiva para tentarmos identificar os possíveis dias impactados e aprofundarmos as análises.

A Copa do Mundo de Futebol, realizada no Brasil em 2014, teve impactos diferentes nos aeroportos, dependendo de ter havido ou não jogo na cidade onde está localizado o aeroporto (ou cidades vizinhas), a importância do jogo (do Brasil ou, até, da Argentina) e o impacto no tráfego aéreo no dia anterior e posterior ao jogo. Foram criadas 4 binárias:

- “CopaM1” - para os dias especiais na Copa do Mundo (jogo de abertura, final da Copa e jogos da seleção brasileira).
- “CopaM2” - para os dias dos jogos (no aeroporto de Recife com exceção do dia 29 de junho que houve mais restrição com relação aos voos na região);
- “CopaM3” - para os dias anteriores aos jogos;
- “CopaM4” - para os dias posteriores aos jogos

Essas binárias são diferentes das outras até agora citadas, porque são em dias diferentes para os aeroportos analisados.

Em Recife dois eventos importantes podem ter impactado na quantidade de operações de pousos e decolagens do aeroporto, uma foi a queda do avião da companhia aérea NOAR (Nordeste Aviação Regional Linhas Aéreas) em 13 de julho de 2011 que encerrou as operações de pousos e decolagens dessa companhia aérea; a outra foi a obra na pista de pousos do aeroporto que ocorreu em 5 etapas, a primeira começou em 2 de setembro de 2013 e se encerrou no dia 6 de dezembro de 2014, as obras interromperam a utilização da pista durante a madrugada e só foram interrompidas durante as altas demandas do verão (de 16 de dezembro de 2013 a 2 de fevereiro de 2014), do carnaval (27 de fevereiro a 11 de março de 2014) e da Copa do Mundo de Futebol (10 de maio a 22 de setembro de 2014). Foram criadas 6 binárias:

- NoarRF1 – para a interação dos dias com mais voos da NOAR (de 9 de setembro de 2010 a 13 de julho de 2011) e os dias da semana (segunda a sexta-feira);
- NoarRF2 – para a interação dos dias com voos da NOAR (de 9 de setembro de 2010 a 13 de julho de 2011) e os dias do final de semana (sábado e domingo);
- ObraRF1 – para os dias em que o aeroporto ficou com obras na pista durante a madrugada durante os dias de 2 de setembro a 15 de dezembro de 2013;

- ObraRF2 – para os dias em que o aeroporto ficou com obras na pista durante a madrugada durante os dias de 3 a 26 de fevereiro de 2014;
- ObraRF3 – para os dias em que o aeroporto ficou com obras na pista durante a madrugada durante os dias de 12 de março a 9 de maio de 2014;
- ObraRF4 – para os dias em que o aeroporto ficou com obras na pista durante a madrugada durante os dias de 23 de setembro a 2 de dezembro de 2014.

Outros eventos foram analisados, mas não demonstraram nenhum impacto no número de operações da aviação comercial nos aeroportos. Os eventos analisados foram: Copa das Confederações (2013), Jornada Mundial da Juventude (Rio de Janeiro, 2013) e Incentivo fiscal ao consumo de QAV (a partir de 2013).

5.9.14 Meteorologia

Serão analisadas variáveis binárias indicando se houve uma quantidade de tempo do dia que o aeroporto teve restrições nos pousos ou nas decolagens por motivos meteorológicos. Também serão testadas variáveis binárias desse mesmo tipo, só que de outros aeroportos, que podem ter relação com o aeroporto analisado.

Segue a lista de binárias que serão analisadas e em quais modelos de aeroportos serão testadas:

- GLa1h, GLa2h, GLa3h, GLa4h e GLa5h são as binárias para os dias em que o aeroporto do Galeão funcionou abaixo dos mínimos por mais de 1 hora, 2 horas, 3 horas, 4 horas e 5 horas, respectivamente. Serão testadas nos modelos preditivos do Galeão e de Guarulhos (grande fluxo);
- GLi4h, GLi8h e GLi12h são as binárias para os dias em que o aeroporto do Galeão funcionou a instrumento por mais de 4 horas, 8 horas e 12 horas, respectivamente. Serão testadas nos modelos preditivos do Galeão e de Guarulhos (grande fluxo);
- BRa1h, BRa2h, BRa3h, BRa4h e BRa5h são as binárias para os dias em que o aeroporto de Brasília funcionou abaixo dos mínimos por mais de 1 hora, 2 horas, 3 horas, 4 horas e 5 horas, respectivamente. Serão testadas nos modelos preditivos do Galeão (grande fluxo), de Brasília e de Guarulhos (grande fluxo);

- BRi4h, BRi8h e BRi12h são as binárias para os dias em que o aeroporto de Brasília funcionou a instrumento por mais de 4 horas, 8 horas e 12 horas, respectivamente. Serão testadas nos modelos preditivos do Galeão (grande fluxo), de Brasília e de Guarulhos (grande fluxo);
- GRa1h, GRa2h, GRa3h, GRa4h e GRa5h são as binárias para os dias em que o aeroporto de Guarulhos funcionou abaixo dos mínimos por mais de 1 hora, 2 horas, 3 horas, 4 horas e 5 horas, respectivamente. Serão testadas nos modelos preditivos do Galeão (grande fluxo), de Brasília (grande fluxo), de Guarulhos e de Recife (grande fluxo);
- GRi4h, GRi8h e GRi12h são as binárias para os dias em que o aeroporto de Guarulhos funcionou a instrumento por mais de 4 horas, 8 horas e 12 horas, respectivamente. Serão testadas nos modelos preditivos do Galeão (grande fluxo), de Brasília (grande fluxo), de Guarulhos e de Recife (grande fluxo);
- RFa1h, RFa2h, RFa3h, RFa4h e RFa5h são as binárias para os dias em que o aeroporto de Recife funcionou abaixo dos mínimos por mais de 1 hora, 2 horas, 3 horas, 4 horas e 5 horas, respectivamente. Serão testadas nos modelos preditivos de Recife;
- RFi4h, RFi8h e RFi12h são as binárias para os dias em que o aeroporto de Recife funcionou a instrumento por mais de 4 horas, 8 horas e 12 horas, respectivamente. Serão testadas nos modelos preditivos de Recife;
- RJa1h, RJa2h, RJa3h, RJa4h e RJa5h são as binárias para os dias em que o aeroporto de Santos Dumont/RJ funcionou abaixo dos mínimos por mais de 1 hora, 2 horas, 3 horas, 4 horas e 5 horas, respectivamente. Serão testadas nos modelos preditivos do Galeão (aeroporto alternativo/grande proximidade) e de Brasília (grande fluxo);
- R Ji4h, R Ji8h e R Ji12h são as binárias para os dias em que o aeroporto de Santos Dumont/RJ funcionou a instrumento por mais de 4 horas, 8 horas e 12 horas, respectivamente. Serão testadas nos modelos preditivos do Galeão (aeroporto alternativo/grande proximidade) e de Brasília (grande fluxo);
- SPa1h, SPa2h, SPa3h, SPa4h e SPa5h são as binárias para os dias em que o aeroporto de Congonhas/SP funcionou abaixo dos mínimos por mais de 1 hora, 2 horas, 3 horas, 4 horas e 5 horas, respectivamente. Serão testadas nos modelos

preditivos de Brasília (grande fluxo) e de Guarulhos (aeroporto alternativo/grande proximidade);

- SPi4h, SPi8h e SPi12h são as binárias para os dias em que o aeroporto de Congonhas/SP funcionou a instrumento por mais de 4 horas, 8 horas e 12 horas, respectivamente. Serão testadas nos modelos preditivos de Brasília (grande fluxo) e de Guarulhos (aeroporto alternativo/grande proximidade).

6 RESULTADOS

Foram testados alguns modelos de regressão e ARMAX, dentre os quais os modelos ARMAX geraram melhores resultados.

A análise será separada em 4 partes para chegarmos ao modelo final proposto: Variável Dependente, Covariáveis Utilizadas, índices do ARMA (parte autoregressiva e de médias móveis) e Modelos Selecionados.

6.1 Variável Dependente

As variáveis dependentes dos quatro modelos são as quantidades de operações de pousos e decolagens da aviação comercial nos aeroportos do Galeão, de Brasília, de Guarulhos e de Recife.

Algumas transformações foram utilizadas para melhorar o modelo, mas não se mostraram necessárias até a última análise. Transformações testadas: Logaritmo Neperiano, Raiz Quadrada, Inversa e Boxcox (λ).

6.2 Covariáveis Utilizadas

As covariáveis do modelo preditivo ajudaram a explicar as variações sazonais e pontuais das operações de pousos e decolagens da aviação comercial, e foram tornando o modelo preditivo mais preciso e explicativo.

O critério adotado para indicar se uma covariável é importante para o modelo ou não, será a questão do módulo do valor do coeficiente ser maior do que duas vezes o valor do seu desvio padrão. Este critério é a construção do intervalo de confiança, verificando se a respectiva variável é estatisticamente significativa.

Existe o caso em que covariável deveria ficar fora do modelo (segundo o critério adotado), mas foi mantida porque deixa o valor do AIC menor do que sem o seu uso, além dessas covariáveis estarem muito próximas do critério para mantê-las no modelo.

6.3 Índices do ARMA

Foram testados 50 modelos ARMAXs para cada conjunto covariáveis que tiveram suas importâncias testadas no modelo preditivo de cada aeroporto deste trabalho. Os modelos foram: ARMAX(i , 0), onde $i = 1, \dots, 15$; ARMAX(0, j), onde $j = 1, \dots, 10$; e ARMAX(a , b), onde $a = 1, \dots, 5$ e $b = 1, \dots, 5$.

Para escolhermos os melhores modelos para cada aeroporto foram utilizados os critérios AIC (Critério de Akaike) e BIC (Critério de Schwarz), sendo considerado melhor o modelo que apresentou os menores critérios de AIC e BIC. Nos casos de um modelo ter o AIC menor e outro modelo ter o BIC menor, adotaremos o modelo com o menor AIC.

Para os modelos dos quatro aeroportos analisados, os índices selecionados foram iguais, ARMAX(14, 0). Esse resultado reforça a hipótese levantada por este trabalho sobre a importância do dia da semana, tendo os maiores pesos nas componentes do modelo AR(14) nos dias: 1º dia, 7º dia e 14º dia do passado, sendo os dois últimos exatamente os mesmos dias da semana que o dia que está sendo estimado.

A seleção do modelo final para cada aeroporto passou pelos seguintes passos:

- Foi utilizado todas as covariáveis para calcular os 50 AICs e BICs dos modelos ARMAXs citados. Destes modelos, os que tinham os 5 menores valores foram selecionados.
- Dos 5 modelos, foram sendo retiradas as covariáveis menos significativas enquanto o valor do AIC diminuísse ou se mantivesse estável.
- O modelo com menor AIC, e em seguida com menos covariáveis foi selecionado.

6.4 Modelo Aeroporto do Galeão

O modelo preditivo encontrado para a quantidade de operações de pousos e decolagens no aeroporto do Galeão é o ARMAX(14, 0). Os critérios foram: AIC = 17.510,26 e BIC = 17.994,09.

6.4.1 Coeficientes do Modelo do Galeão

Os coeficientes que se mostraram importantes:

Variável	Coef	DP	Sig	Variável	Coef	DP	Sig
ar1	0,251	0,0218	***	Ago	-20,079	3,4883	***
ar2	0,063	0,0229	**	Set	-20,228	4,3880	***
ar3	0,118	0,0229	***	Out	-17,221	4,7256	***
ar4	0,105	0,0226	***	Nov	-17,741	4,5824	***
ar5	0,047	0,0227	**	Dez	-9,460	3,8893	**
ar6	0,066	0,0230	**	CopaM1	106,350	9,2061	***
ar7	0,197	0,0228	***	CopaM2	30,833	6,1857	***
ar8	0,005	0,0228		CopaM3	29,988	5,4964	***
ar9	-0,030	0,0228	*	CopaM4	13,713	6,0418	**
ar10	0,020	0,0230		Anos10	-76,180	16,8764	***
ar11	0,021	0,0226		Anos11	-46,324	9,8385	***
ar12	-0,025	0,0227	*	MarA	18,658	5,8337	***
ar13	0,004	0,0224		AbrA	18,213	7,1317	**
ar14	0,127	0,0221	***	MaiA	19,163	7,1385	**
intercept	393,636	9,5317	***	JunA	15,947	5,9320	**
Qua	-3,030	1,1384	**	AgoA	19,051	6,1688	***
Sex	7,369	1,1196	***	SetA	27,138	8,1320	***
Sab	-31,083	1,2532	***	OutA	37,003	9,1932	***
Dom	-24,377	1,1826	***	NovA	46,585	9,6277	***
Feriado1	12,621	2,7884	***	DezA	36,770	9,4974	***
Feriado2	-43,133	2,8436	***	Ter15	-17,994	2,8341	***
Feriado3	5,734	3,6955	*	Qua15	-12,587	3,0500	***
Feriado4	-11,395	4,2808	**	Sex15	7,201	2,8599	**
Feriado5	-19,586	4,0961	***	Sab15	-13,862	3,3108	***
Feriado6	14,853	3,1664	***	Dom15	-9,709	2,8906	***
Carna1	-28,211	4,9654	***	GLa2h	-12,032	5,3899	**
Carna2	-45,807	3,8645	***	GLa3h	-32,463	5,8048	***
Carna3	23,764	5,0051	***	GLa4h	-35,182	7,3141	***
Carna4	9,900	3,8701	**	GLi8h	3,799	1,1537	***
Santa1	-27,194	3,8480	***	GLi12h	6,030	1,4096	***
Santa2	9,703	5,0676	*	GRa1h	5,553	2,4765	**
Santa3	34,105	5,0270	***	GRa3h	6,607	3,4887	*
Christi1	-56,987	5,0746	***	GRa4h	12,480	4,3079	**
Christi2	-19,358	5,1778	***	GRa5h	23,377	5,9758	***
Christi3	15,188	4,9159	***	RJa1h	8,988	2,6250	***
FimAno1	-41,469	3,7761	***	RJa2h	6,943	4,0969	*
FimAno2	-26,713	3,9000	***	RJa3h	26,255	5,2496	***
Fev	-4,255	2,7980	*	RJa4h	30,699	5,7111	***
Mar	-14,968	4,0202	***	RJa5h	22,811	4,6444	***
Abr	-18,814	4,4266	***	RJi4h	2,833	0,8592	***
Mai	-21,211	4,2589	***	RJi8h	2,489	1,2004	**
Jun	-16,820	3,4657	***	RJi12h	7,585	1,3028	***

Tabela 6-1 – Coeficientes das variáveis do modelo (Coef), desvio padrão (DP) e significância (Sig, onde “***” = $Coef > 3xDP$; “**” = $Coef > 2xDP$; e “*” $1xDP$) para o aeroporto do Galeão.

Para entender o modelo vamos interpretar os coeficientes estimados:

- AR(14): os coeficientes da parte Auto-regressiva do modelo indicam as proporções dadas ao passado da série que foi ajustada (quantidade de operações no Galeão). Então os maiores coeficientes são: $0,251*Y_{t-1}$; $0,197*Y_{t-7}$; $0,127*Y_{t-14}$; $0,118*Y_{t-3}$; e $0,105*Y_{t-4}$ (os outros somam 0,171, com isso todos somam 0,970);
- Intercept: o valor do intercepto do modelo é 393,636, que é importante por ser o coeficiente do valor inicial de referência de todas as dummies (segunda-feira, janeiro, dia sem feriado ou eventos, sem tendência, do ano de 2015);
- Dias da semana: Os coeficientes dos dias da semana indicam que comparado à segunda-feira, terça e quinta não se diferenciaram. Enquanto que quarta-feira tem menos 3,030 operações, sexta-feira tem mais 7,369, sábado tem menos 31,083 e domingo tem menos 24,377 operações;
- Dias da semana de 2015: Os coeficientes dos dias da semana para o ano de 2015 tem de ser adicionado aos dias da semana normal (todos os anos). Logo, a interpretação desses coeficientes servirá para os dias da semana de 2015 (e para 2016 dentro da previsão) e indica que comparado à segunda-feira, quinta-feira não apresentou diferença. Enquanto que terça-feira tem menos 17,994 operações, quarta-feira tem menos 15,617 (soma dos coeficientes de “Qua” e “Qua15”), sexta-feira tem mais 14,570, sábado tem menos 44,945 e domingo tem menos 34,086 operações;
- Meses do ano: Os coeficientes dos meses indicam que comparado à janeiro, julho não apresentou diferença. Enquanto que fevereiro tem menos 4,255 operações, março tem menos 14,968, abril tem menos 18,814, maio tem menos 21,211, junho tem menos 16,820, agosto tem menos 20,079, setembro tem menos 20,228, outubro tem menos 17,221, novembro tem menos 17,741 e dezembro tem menos 9,460 operações;
- Meses do ano de 2010 e 2011: Os coeficientes dos meses para os anos de 2010 e 2011 (AnoA) indicam que comparado à janeiro normal (de 2010 a 2015), março tem mais 3,690 (soma dos coeficientes de “Mar” e “MarA”), abril tem menos 0,600, maio tem menos 2,048, junho tem menos 0,873, agosto tem menos 1,028,

setembro tem mais 6,910, outubro tem mais 19,782, novembro tem mais 28,844 e dezembro tem menos 27,310 operações;

- Copa do Mundo de 2014: Os coeficientes para o período da Copa do Mundo indicam que durante os jogos, no Rio de Janeiro, as operações aumentaram 30,833, os dias anteriores aos jogos aumentaram 29,988, os dias posteriores aos jogos aumentaram 13,712 e os dias da final da Copa do Mundo e o dia seguinte aumentaram 106,350 operações;
- Feriados de Natal e Réveillon: Os coeficientes para os dias do Natal (25/12) e Réveillon (1/1) indicam que as operações diminuem 41,469 e as vésperas desses dias (24/12 e 31/12) diminuem 26,713 operações;
- Feriado de Carnaval: Os coeficientes para as terças-feiras de Carnaval indicam que as operações diminuem 28,211, para as segundas-feiras e os domingos diminuem 45,807, para os sábados aumentam 23,764 e para as sextas-feiras e as quintas-feiras (Pré-Carnaval) aumentam 9,900 operações;
- Feriado de Semana Santa: Os coeficientes para as sextas-feiras e os sábados da Semana Santa indicam que as operações diminuem 27,194, para os domingos aumentam 9,703 e para as segundas-feiras (Pós-Semana Santa) aumentam 34,105 operações;
- Feriado de Corpus Christi: Os coeficientes para as sextas-feiras do feriado prolongado indicam que as operações diminuem 56,987, para os sábados diminuem 19,358 e para as segundas-feiras aumentam 15,188 operações;
- Feriados fixos no dia do mês: Os coeficientes para os dias desses feriados (que estão detalhados na subseção 5.9.11 deste trabalho) “Feriado1” indicam que as operações aumentam 12,621, para os “Feriado2” diminuem 43,133, para os “Feriado3” aumentam 5,734, para os “Feriado4” diminuem 11,395, para os “Feriado5” diminuem 19,586 e para “Feriado6” aumentam 14,853 operações;
- Meteorologia do Galeão: Os coeficientes para os dias em que o Galeão funciona abaixo dos mínimos entre 2 e 3 horas indica que as operações diminuem 12,032, entre 3 e 4 horas diminuem 32,463 e 4 e 5 horas diminuem 12,480 operações. Enquanto que os dias em que funciona a instrumento entre 8 e 12 horas aumentam 3,799 e mais de 12 horas aumentam 6,030 operações;
- Meteorologia de Guarulhos: Os coeficientes para os dias em que Guarulhos funciona abaixo dos mínimos entre 1 e 2 horas indica que as operações aumentam

5,553, entre 3 e 4 horas diminuem 6,607, entre 4 e 5 horas aumentam 12,480 e mais de 5 horas aumentam 23,377 operações;

- Meteorologia de Santos Dumont: Os coeficientes para os dias em que o Santos Dumont funciona abaixo dos mínimos entre 1 e 2 horas indica que as operações aumentam 8,988, entre 2 e 3 horas aumentam 6,943, entre 3 e 4 horas diminuem 26,255, entre 4 e 5 horas aumentam 30,699 e mais de 5 horas aumentam 22,811 operações. Enquanto que os dias em que funciona a instrumento entre 4 e 8 horas aumentam 2,833, entre 8 e 12 horas aumentam 2,489 e mais de 12 horas aumentam 7,585 operações.

6.4.2 Análise dos Resíduos do Modelo do Galeão

Para entendermos se o modelo está bem ajustado temos que verificar os resíduos. Estes deveriam ser “ruídos brancos gaussianos”, ou seja, terem distribuição normal com média zero e desvio padrão constante ao longo do tempo, enquanto que no nosso caso a distribuição não é normal, está mais parecida com a distribuição t-Student, o que precisa ser analisado.

Essa análise vai se repetir para os outros 3 aeroportos: o teste usado para checar a normalidade foi o Jarque-Bera, que rejeitou a hipótese nula de normalidade dos resíduos.

Na Figura 6-1 estão os gráficos dos resíduos do modelo do aeroporto do Galeão, com os resíduos num histograma, num box-plot, num QQ-plot e numa série em linha.

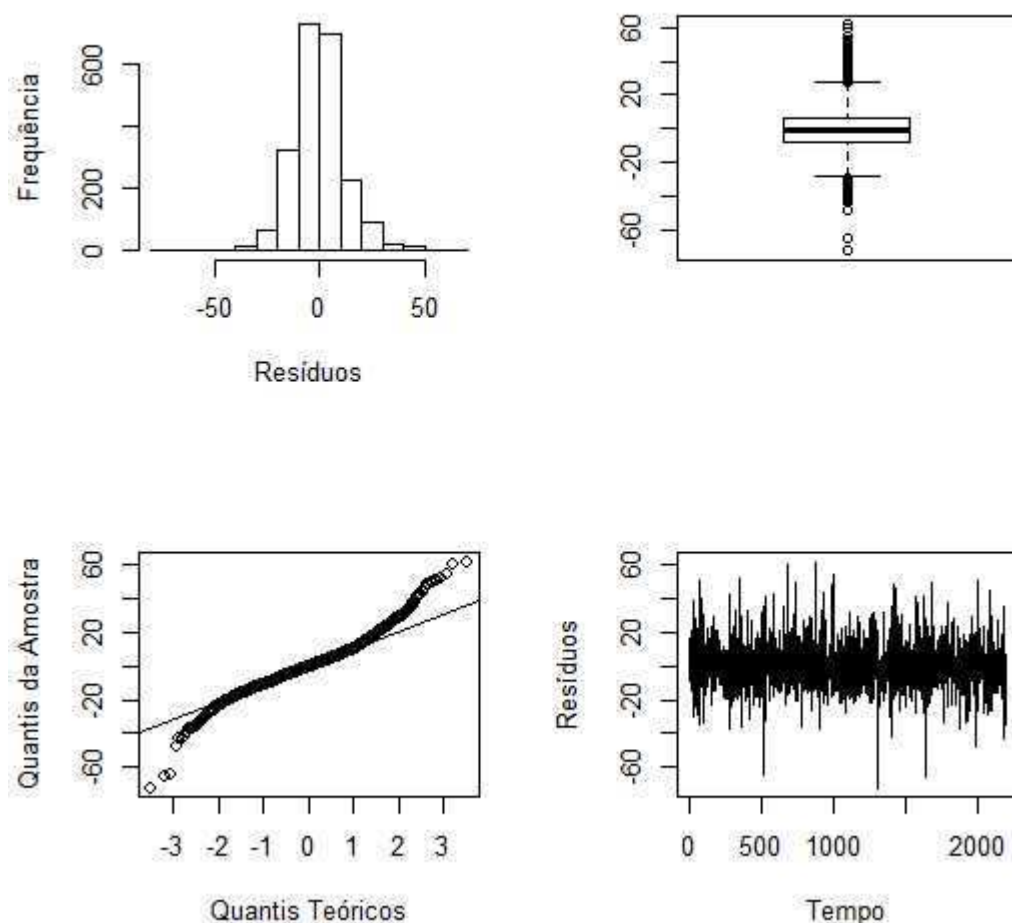


Figura 6-1 - Em cima vemos o histograma e o boxplot, embaixo o Q-Q plot e a série dos resíduos do modelo do aeroporto do Galeão.

6.4.3 Análise da Previsão do Modelo do Galeão

Ajustado o modelo, o próximo passo será analisar o resultado da previsão. Usaremos as covariáveis padronizadas como no modelo preditivo para os três primeiros meses de 2016.

Na Figura 6-2, do aeroporto do Galeão, podemos analisar a série (linha preta) com a quantidade de operações de pousos e decolagens da aviação comercial de 1º de outubro de 2015 até 31 de março de 2016. Os 3 últimos meses de 2015 foram usados no modelo e os 3 primeiros meses de 2016 não foram. Nesta mesma figura podemos ver a série

projetada pelo modelo (linha vermelha) para o primeiro trimestre de 2016, dentro dos limites de 2 vezes os desvios padrões (área pontilhada). Podemos ver que no primeiro mês ficaram parecidas, mas que a partir de fevereiro a série projetada pelo modelo está acima e em março a série real chega a estar abaixo do intervalo de variação da série projetada pelo modelo. Uma explicação para o que foi analisado é a crise econômica de 2015/2016 que pode estar diminuindo o tráfego aéreo do Galeão de forma atípica, o que não conseguimos colocar dentro do modelo estimado.

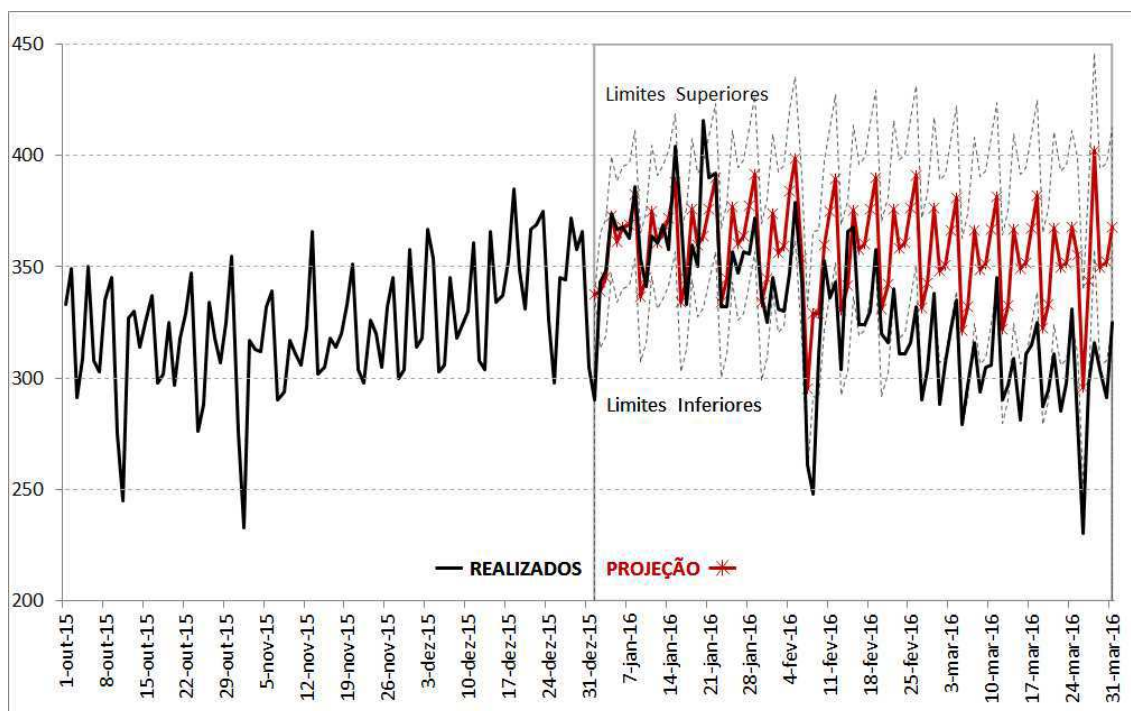


Figura 6-2 - Gráfico do total de pousos e decolagens por dia com a projeção de 3 meses a partir de 1º de janeiro de 2016 da aviação comercial no aeroporto do Galeão.

Na Figura 6-3, tem a mesma informação da figura anterior só que com o foco em um mês, antes da previsão e o primeiro mês de 2016 com a série real e a previsão segundo o modelo proposto. Repare que no curto prazo (1 mês), o modelo proposto se adequa a quantidade de pousos e decolagens da aviação comercial no aeroporto do Galeão.

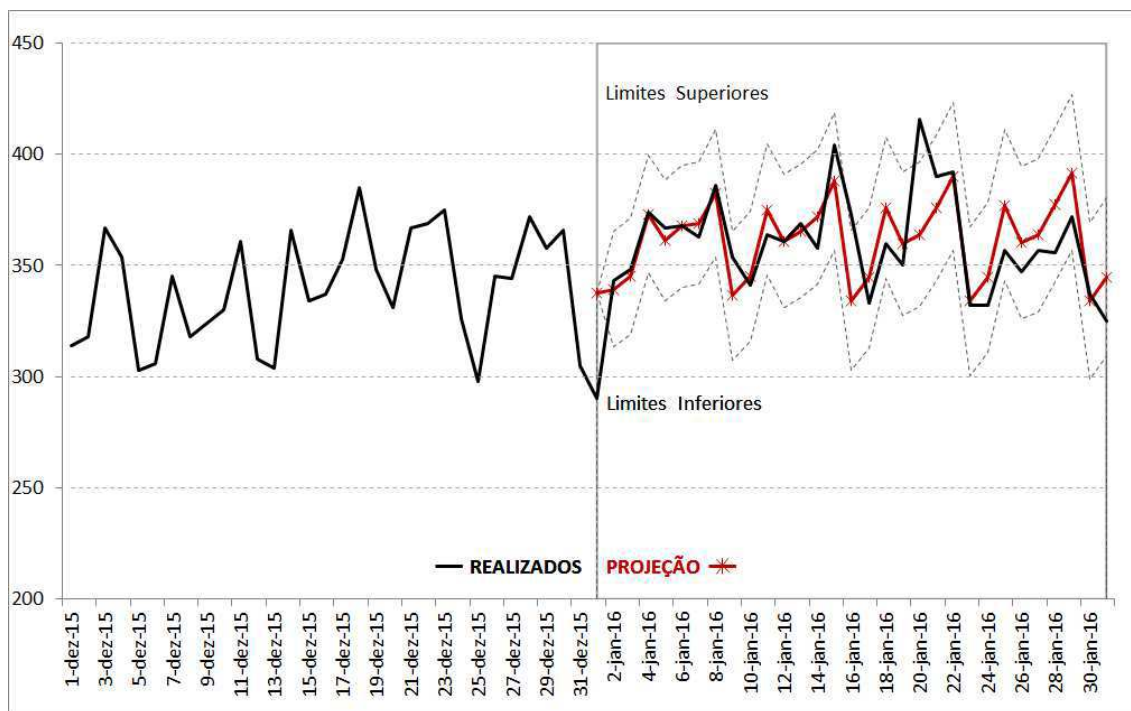


Figura 6-3 - Gráfico do total de pousos e decolagens por dia com a projeção de 1 mês a partir de 1º de janeiro de 2016 da aviação comercial no aeroporto do Galeão.

Um segundo cenário pode ser criado para verificarmos, graficamente, o poder preditivo do modelo. Na Figura 6-4, o modelo foi ajustado a partir dos dados de 1º de janeiro de 2010 até 31 de julho de 2014 (seria até 30 de junho, mas devido a Copa do Mundo de futebol nós tivemos que mudar), e fizemos a previsão do que restou do ano de 2014. Nesta figura podemos ver o que realmente aconteceu e a previsão do modelo com o intervalo de segurança de 2 desvios padrões, e com isso, confirmar a qualidade do modelo ajustado neste trabalho. Com relação as covariáveis do modelo, tivemos que retirar as que tinham relação direta com o ano de 2015 e com a binária AnoA.

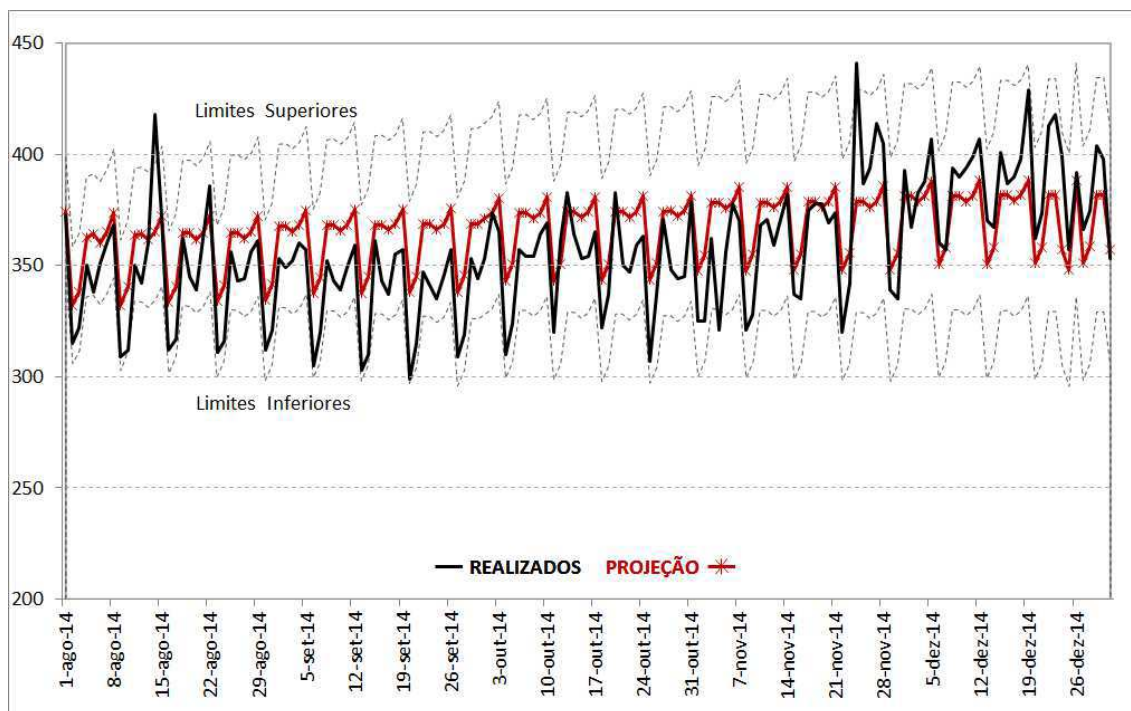


Figura 6-4 - Gráfico do total de pousos e decolagens por dia com a projeção de 5 meses a partir de 1º de agosto de 2014 da aviação comercial no aeroporto do Galeão.

6.5 Modelo Aeroporto de Brasília

O modelo preditivo encontrado para a quantidade de operações de pousos e decolagens da aviação comercial do aeroporto de Brasília é o ARMAX(14, 0). Os critérios foram: AIC = 17.094,99 e BIC = 17.527,59.

6.5.1 Coeficientes do Modelo de Brasília

Diferente do modelo do Galeão, a covariável “Tende” se mostrou significativa e com tendência negativa (-0,061), enquanto que mais binárias dos anos se mostraram importantes como Ano10, Ano11, Ano12, Ano13 e Ano14. A meteorologia mostrou impactar menos neste aeroporto, sendo o principal aeroporto que mostrou importância foi a condição climática do Santos Dumont.

Os coeficientes que se mostraram importantes:

Variável	Coef	DP	Sig	Variável	Coef	DP	Sig
ar1	0,289	0,0219	***	Set	17,428	5,4907	***
ar2	0,115	0,0234	***	Out	20,505	6,0566	***
ar3	0,064	0,0229	**	Nov	26,768	6,6060	***
ar4	0,033	0,0230	*	Dez	37,390	7,1157	***
ar5	-0,039	0,0227	*	CopaM1	36,704	7,2556	***
ar6	0,031	0,0226	*	CopaM2	18,835	5,6084	***
ar7	0,287	0,0229	***	CopaM4	17,023	4,9237	***
ar8	-0,074	0,0227	***	Anos10	-177,419	38,9335	***
ar9	-0,055	0,0227	**	Anos11	-119,429	31,3297	***
ar10	-0,014	0,0231		Anos12	-73,592	23,3429	***
ar11	0,043	0,0230	*	Anos13	-89,421	15,7496	***
ar12	-0,017	0,0229		Anos14	-49,777	8,4735	***
ar13	-0,025	0,0225	*	MarA	7,639	4,4527	*
ar14	0,181	0,0220	***	AbrA	7,330	4,9923	*
intercept	587,620	40,2030	***	MaiA	16,845	5,6141	***
Sex	5,972	1,2052	***	JunA	11,778	4,9371	**
Sab	-73,315	1,3596	***	AgoA	12,707	5,2457	**
Dom	-63,915	1,2984	***	SetA	15,417	6,1885	**
Tende	-0,061	0,0211	**	OutA	18,574	6,4368	**
Feriado1	5,558	2,5108	**	NovA	16,543	6,4288	**
Feriado2	-65,393	2,5559	***	DezA	10,402	5,9276	*
Feriado3	-9,853	3,2854	**	Fev12	9,479	5,3998	*
Feriado4	-45,400	3,8315	***	Jul12	-25,072	6,5714	***
Feriado5	-39,364	3,5905	***	Ago12	-18,701	7,8400	**
Carna1	-58,564	4,6540	***	Set12	-26,575	8,3087	***
Carna2	-96,714	3,6455	***	Out12	-31,688	8,4459	***
Carna4	8,406	3,5974	**	Nov12	-39,499	8,4867	***
Santa1	-56,795	3,3218	***	Dez12	-50,542	8,1629	***
Christi1	-83,935	4,4936	***	Ter15	-8,857	2,7773	***
Christi2	-27,859	4,5004	***	Sab15	-41,478	3,1748	***
FimAno1	-56,065	3,7307	***	Dom15	-14,094	3,1040	***
FimAno2	-53,466	3,9291	***	RJa1h	-4,208	2,2564	*
FimAno3	-7,805	2,8720	**	RJa2h	-7,748	3,5534	**
FimAno4	3,816	2,6191	*	RJa3h	-8,762	4,2956	**
Fev	-7,102	2,3645	***	RJa5h	-8,461	3,7543	**
Mai	-7,480	2,4528	***	BRa2h	-10,666	4,2994	**
Jul	22,649	3,6792	***	SPa2h	-6,484	3,3580	*
Ago	14,259	4,8244	**				

Tabela 6-2 – Coeficientes das variáveis do modelo (Coef), desvio padrão (DP) e significância (Sig, onde “***” = $Coef > 3xDP$; “**” = $Coef > 2xDP$; e “*” $1xDP$) para o aeroporto de Brasília.

6.5.2 Análise dos Resíduos de Brasília

Na Figura 6-5 estão os gráficos dos resíduos do modelo do aeroporto de Brasília, com os resíduos num histograma, num box-plot, num QQ-plot e numa série em linha.

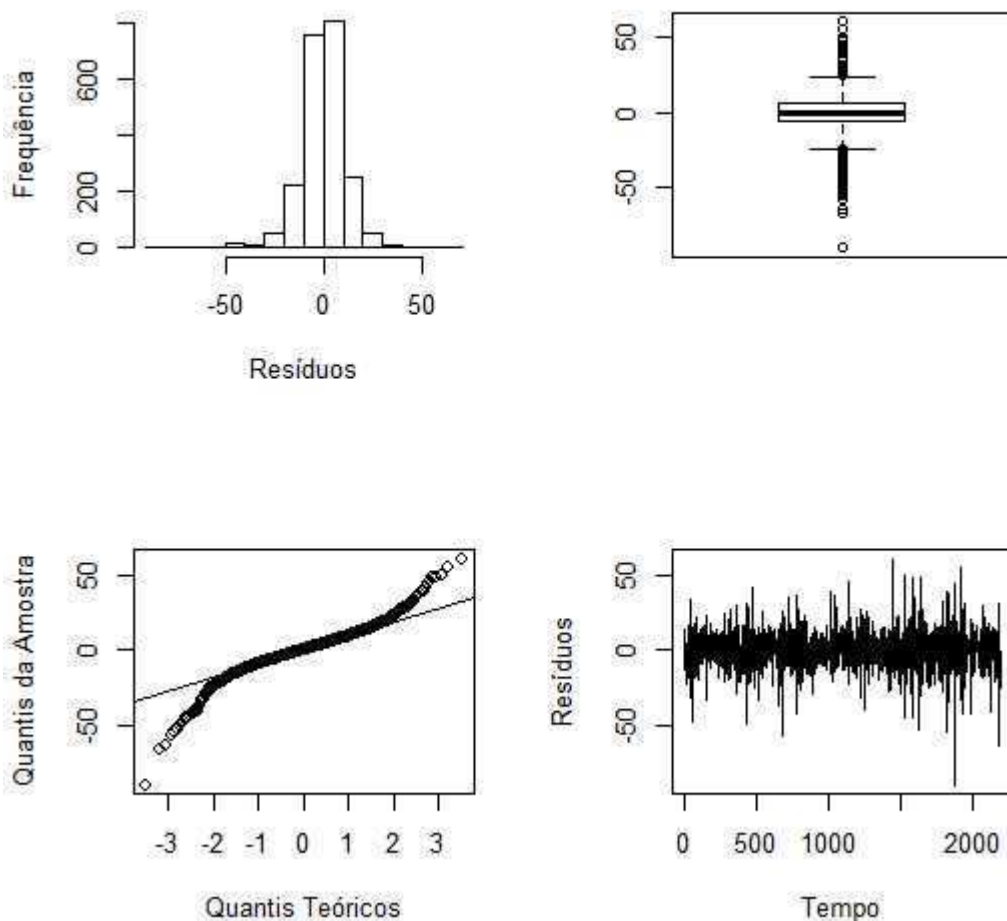


Figura 6-5 - Em cima vemos o histograma e o boxplot, embaixo o Q-Q plot e a série dos resíduos do modelo do aeroporto de Brasília.

6.5.3 Análise da Previsão do Modelo de Brasília

Ajustado o modelo, o próximo passo será analisar o resultado da previsão. Usaremos as covariáveis padronizadas como no modelo preditivo para os três primeiros meses de 2016.

Na Figura 6-6, do aeroporto de Brasília, podemos analisar a série (linha azul) com a quantidade de operações de pousos e decolagens da aviação comercial de 1º de outubro

de 2015 até 31 de março de 2016. Os 3 últimos meses de 2015 foram usados no modelo e os 3 primeiros meses de 2016 não foram. Nesta mesma figura podemos ver a série projetada pelo modelo (linha vermelha) para o primeiro trimestre de 2016, dentro dos limites de 2 vezes os desvios padrões (área pontilhada). Podemos ver que no primeiro mês não ficaram parecidas, pois a série real ficou bem acima da projetada. A partir de fevereiro a série projetada pelo modelo fica próxima da série real, principalmente durante o carnaval e no mês de março.

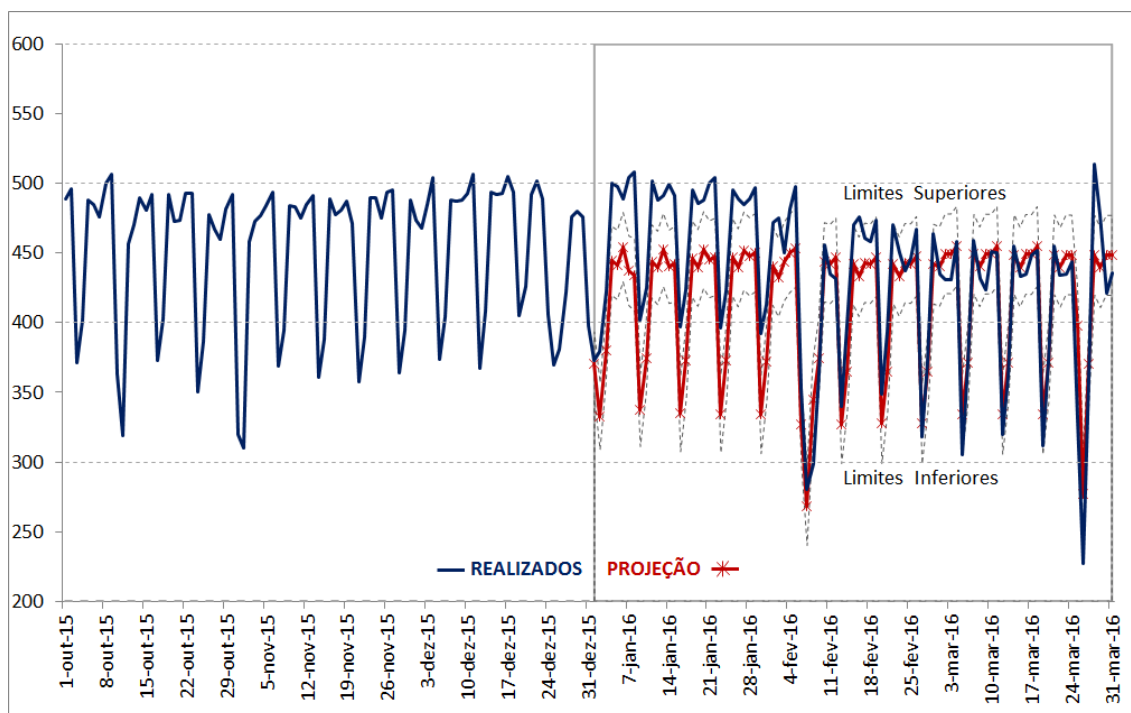


Figura 6-6 - Gráfico do total de pousos e decolagens por dia com a projeção de 3 meses a partir de 1º de janeiro de 2016 da aviação comercial no aeroporto de Brasília.

Na Figura 6-7, tem a mesma informação da figura anterior só que com o foco em um mês, antes da previsão e o primeiro mês de 2016 com a série real e a previsão segundo o modelo proposto. Repare que no curto prazo (1 mês), o modelo proposto se adequa a quantidade de pousos e decolagens da aviação comercial no aeroporto de Brasília.

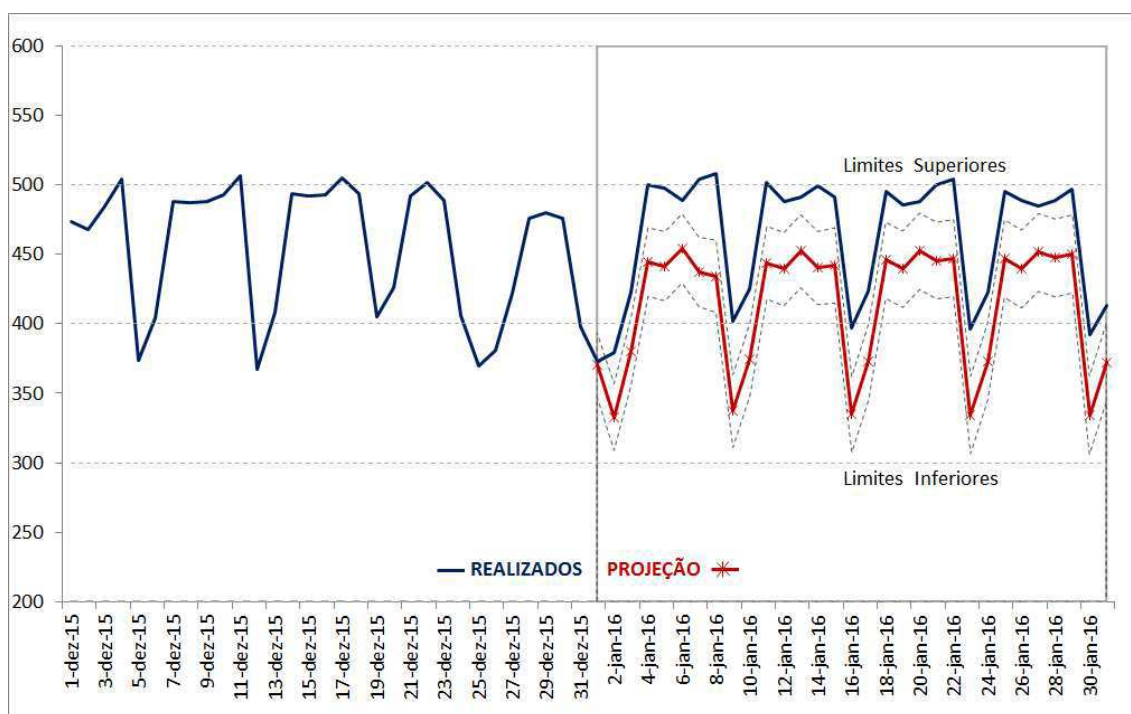


Figura 6-7 - Gráfico do total de pousos e decolagens por dia com a projeção de 1 mês a partir de 1º de janeiro de 2016 da aviação comercial no aeroporto de Brasília.

Na Figura 6-8, o modelo foi ajustado a partir dos dados de 1º de janeiro de 2010 até 31 de julho de 2014 para o aeroporto de Brasília, e fizemos a previsão do que restou do ano de 2014. Nesta figura podemos ver o que realmente aconteceu e a previsão do modelo com o intervalo de segurança de 2 desvios padrões, e com isso, confirmar a qualidade do modelo ajustado neste trabalho. Com relação as covariáveis do modelo, tivemos que retirar as que tinham relação direta com o ano de 2015 e com a binária AnoA.

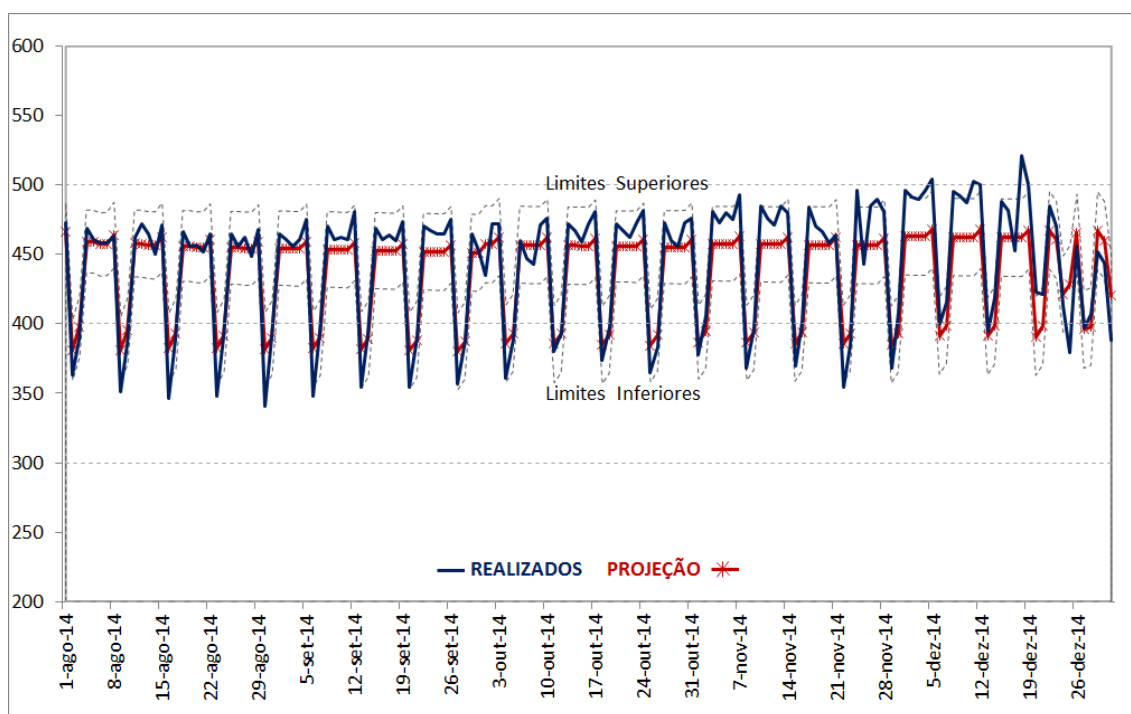


Figura 6-8 - Gráfico do total de pousos e decolagens por dia com a projeção de 5 meses a partir de 1º de agosto de 2014 da aviação comercial no aeroporto de Brasília.

6.6 Modelo Aeroporto de Guarulhos

O modelo preditivo encontrado para a quantidade de operações de pousos e decolagens da aviação comercial do aeroporto de Guarulhos é o ARMAX(14, 0). Os critérios foram: AIC = 18.808,54 e BIC = 19.246,83.

6.6.1 Coeficientes do Modelo de Guarulhos

Assim como no modelo de Brasília, existe duas interações com as binárias dos meses, uma com a covariável “AnoA” e outra com a covariável (Ano12). A meteorologia de Guarulhos e de Congonhas se mostraram importantes para o modelo deste aeroporto.

Os coeficientes que se mostraram importantes foram:

Variável	Coef	DP	Sig	Variável	Coef	DP	Sig
ar1	0,265	0,0220	***	Mar	-44,719	6,1994	***
ar2	0,053	0,0235	**	Abr	-43,183	6,5661	***
ar3	0,075	0,0228	***	Mai	-56,746	6,3620	***
ar4	0,131	0,0229	***	Jun	-39,514	5,5609	***
ar5	-0,012	0,0229		Ago	-22,046	5,6397	***
ar6	0,048	0,0229	**	Set	-20,316	5,3444	***
ar7	0,317	0,0227	***	CopaM1	27,427	11,4312	**
ar8	-0,055	0,0229	**	CopaM2	-15,951	7,4374	**
ar9	-0,052	0,0231	**	CopaM4	32,380	7,6248	***
ar10	0,012	0,0231		Anos10	-195,251	10,7185	***
ar11	-0,012	0,0229		Anos11	-127,215	9,5120	***
ar12	-0,084	0,0227	***	Anos12	-58,405	9,7326	***
ar13	0,028	0,0227	*	Anos13	-81,003	10,3416	***
ar14	0,164	0,0225	***	FevA	14,407	7,4508	*
intercept	797,059	5,9024	***	MarA	42,128	9,3958	***
Ter	10,071	2,6377	***	AbrA	41,881	10,3309	***
Qua	7,480	3,0423	**	MaiA	57,196	10,5398	***
Qui	20,818	2,3869	***	JunA	41,389	10,2486	***
Sex	20,929	2,3916	***	JulA	23,218	8,8707	**
Sab	-17,607	3,1547	***	AgoA	38,546	10,6593	***
Dom	-31,637	2,6494	***	SetA	45,823	10,5666	***
Feriado2	-82,406	3,6427	***	OutA	36,596	8,6735	***
Feriado4	-19,441	5,5614	***	NovA	42,585	8,3307	***
Feriado5	-45,774	5,3543	***	DezA	51,066	7,6589	***
Feriado6	20,701	4,1705	***	Jul12	-36,140	9,1020	***
Carna1	-54,824	6,6028	***	Ago12	-20,758	9,7960	**
Carna2	-81,530	5,2057	***	Out12	-34,039	10,1673	***
Carna3	38,824	6,7410	***	Nov12	-35,380	12,1026	**
Carna4	22,255	5,2250	***	Dez12	-41,165	12,2101	***
Santa1	-71,745	4,8377	***	Ter15	-24,695	4,5487	***
Santa3	12,691	6,4037	*	Sab15	-33,563	4,9141	***
Christi1	-110,940	6,6076	***	Dom15	-9,238	4,9859	*
Christi2	-53,929	6,5219	***	GRa2h	-15,266	4,0368	***
Christi3	11,134	6,4728	*	GRa3h	-13,808	4,5171	***
FimAno1	-87,411	5,1997	***	GRa4h	-11,462	6,0281	*
FimAno2	-47,958	5,3702	***	GRa5h	-26,145	7,7186	***
FimAno4	16,628	3,5995	***	SPa1h	-6,763	3,0364	**
Fev	-29,076	5,1230	***	SPa4h	-29,491	16,7863	*

Tabela 6-3 – Coeficientes das variáveis do modelo (Coef), desvio padrão (DP) e significância (Sig, onde “***” = $Coef > 3xDP$; “**” = $Coef > 2xDP$; e “*” $1xDP$) para o aeroporto de guarulhos.

6.6.2 Análise dos Resíduos do Modelo de Guarulhos

Na Figura 6-9 estão os gráficos dos resíduos do modelo do aeroporto de Guarulhos, com os resíduos num histograma, num box-plot, num QQ-plot e numa série em linha.

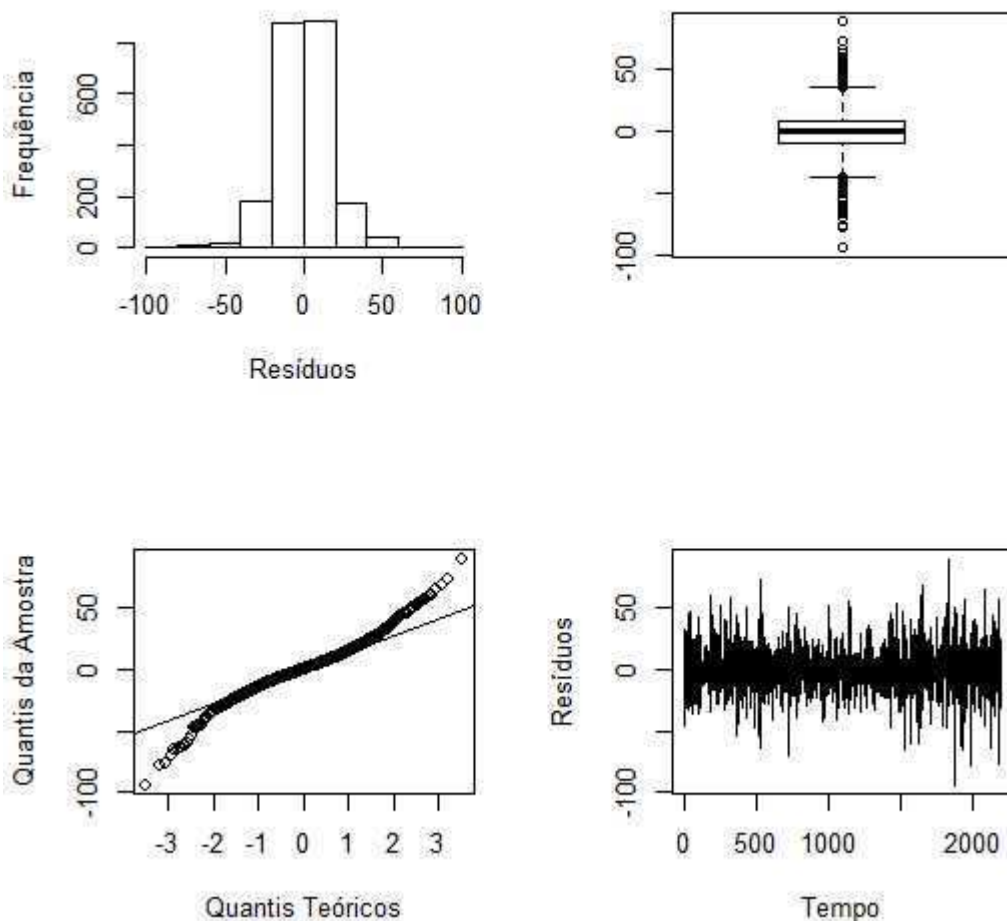


Figura 6-9 - Em cima vemos o histograma e o boxplot, embaixo o Q-Q plot e a série dos resíduos do modelo do aeroporto de Guarulhos.

6.6.3 Análise da Previsão do Modelo de Guarulhos

Ajustado o modelo, o próximo passo será analisar o resultado da previsão. Usaremos as covariáveis padronizadas como no modelo preditivo para os três primeiros meses de 2016.

Na Figura 6-10, do aeroporto de Guarulhos, podemos analisar a série (linha verde) com a quantidade de operações de pousos e decolagens da aviação comercial de 1º de outubro

de 2015 até 31 de março de 2016. Os 3 últimos meses de 2015 foram usados no modelo e os 3 primeiros meses de 2016 não foram. Nesta mesma figura podemos ver a série projetada pelo modelo (linha vermelha) para o primeiro trimestre de 2016, dentro dos limites de 2 vezes os desvios padrões (área pontilhada). Podemos ver que no primeiro mês ficaram parecidas, mas que a partir de fevereiro a série projetada pelo modelo está acima e em março a série real chega a estar abaixo do intervalo de variação da série projetada pelo modelo. Uma explicação para o que foi analisado é a crise econômica de 2015/2016 que pode estar diminuindo o tráfego aéreo de Guarulhos de forma atípica, o que não conseguimos colocar dentro do modelo estimado.

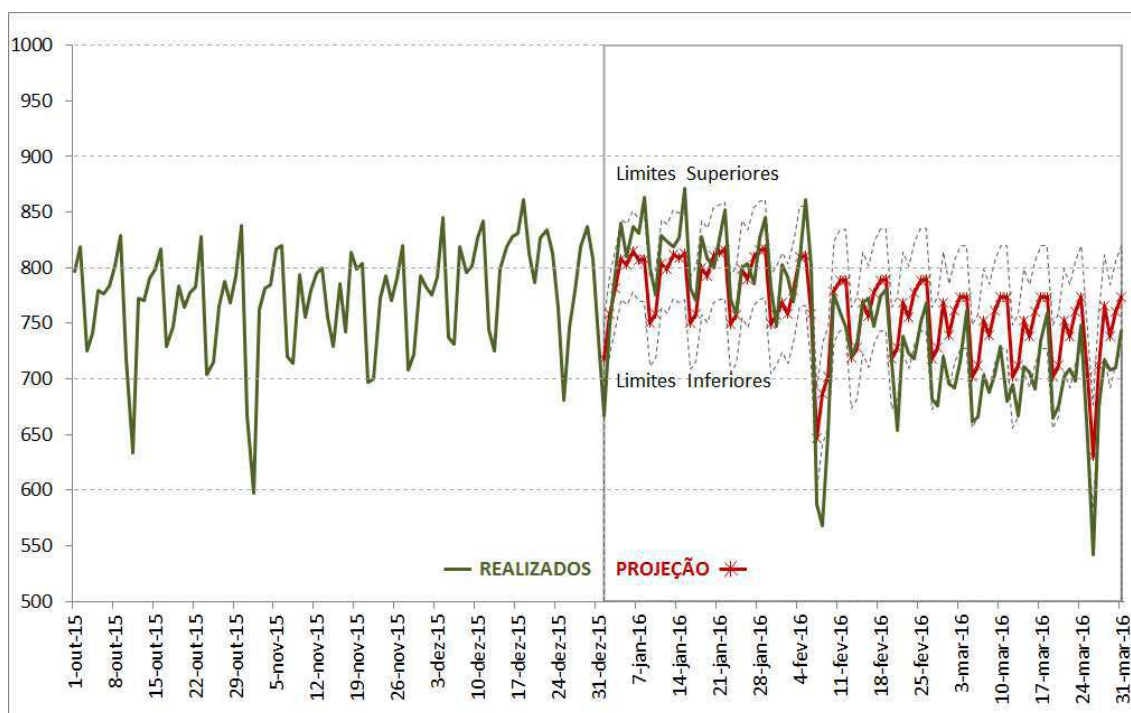


Figura 6-10 - Gráfico do total de pousos e decolagens por dia com a projeção de 3 meses a partir de 1º de janeiro de 2016 da aviação comercial no aeroporto de Guarulhos.

Na Figura 6-11, tem a mesma informação da figura anterior só que com o foco em um mês, antes da previsão e o primeiro mês de 2016 com a série real e a previsão segundo o modelo proposto. Repare que no curto prazo (1 mês), o modelo proposto se adequa a quantidade de pousos e decolagens da aviação comercial no aeroporto de Guarulhos.

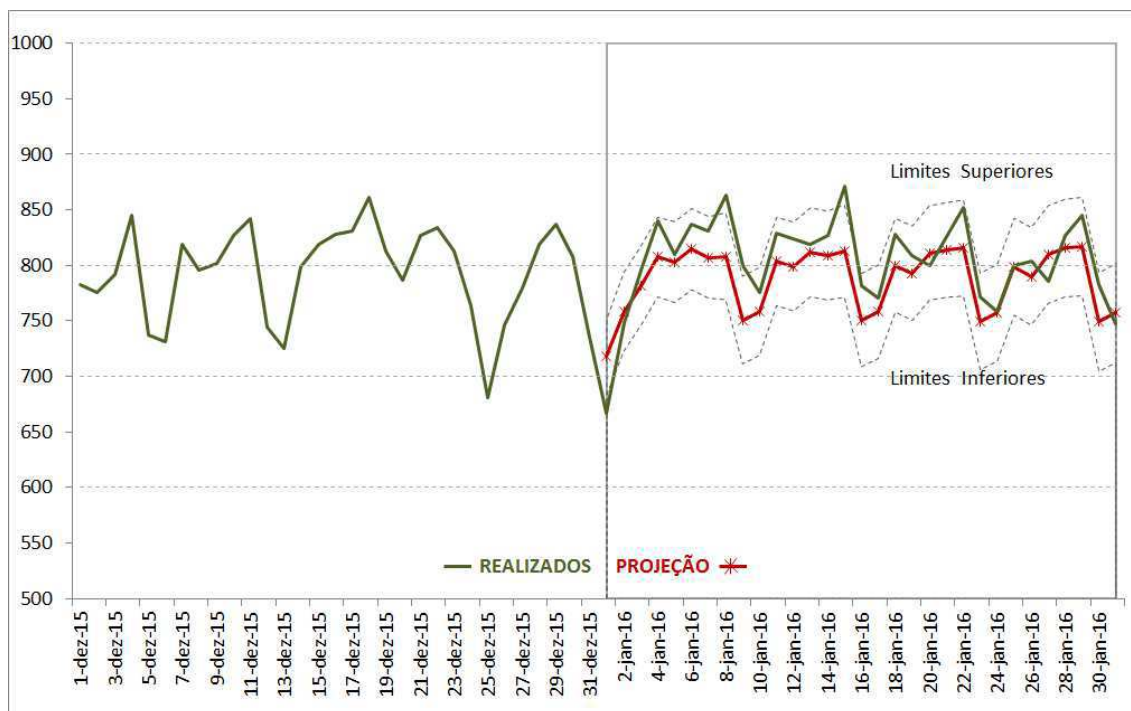


Figura 6-11 - Gráfico do total de pousos e decolagens por dia com a projeção de 1 mês a partir de 1º de janeiro de 2016 da aviação comercial no aeroporto de Guarulhos.

Na Figura 6-12, o modelo foi ajustado a partir dos dados de 1º de janeiro de 2010 até 31 de julho de 2014 para o aeroporto de Guarulhos, e fizemos a previsão do que restou do ano de 2014. Nesta figura podemos ver o que realmente aconteceu e a previsão do modelo com o intervalo de segurança de 2 desvios padrões, e com isso, confirmar a qualidade do modelo ajustado neste trabalho. Com relação as covariáveis do modelo, tivemos que retirar as que tinham relação direta com o ano de 2015 e com a binária AnoA.

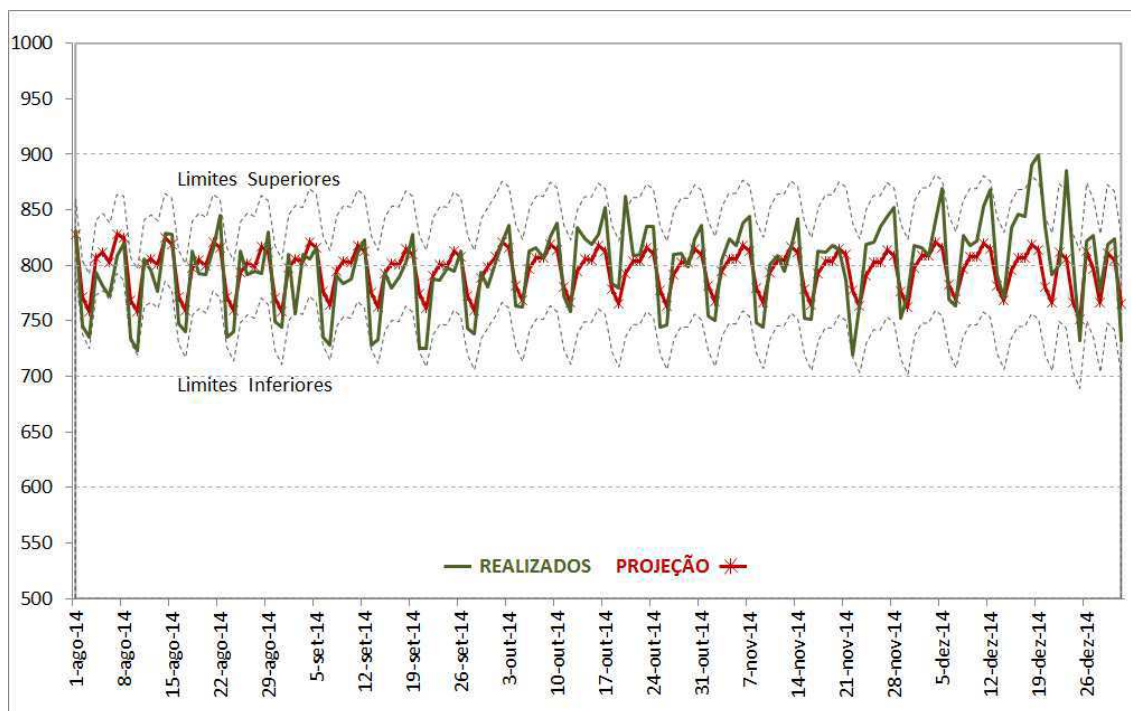


Figura 6-12 - Gráfico do total de pousos e decolagens por dia com a projeção de 5 meses a partir de 1º de agosto de 2014 da aviação comercial no aeroporto de Guarulhos.

6.7 Modelo Aeroporto de Recife

O modelo preditivo encontrado para a quantidade de operações de pousos e decolagens da aviação comercial do aeroporto de Recife é o ARMAX(14, 0). Os critérios foram: AIC = 14.211,30 e BIC = 14.643,90.

6.7.1 Coeficientes do Modelo de Recife

Assim como no modelo de Brasília e de Guarulhos, existe duas interações com as binárias dos meses, uma com a covariável “AnoA” e outra, diferentemente dos outros modelos, com a covariável (Ano10). As binárias das obras na pista do aeroporto de Recife se mostraram significativas.

Os coeficientes que se mostraram importantes foram:

Variável	Coef	DP	Sig	Variável	Coef	DP	Sig
ar1	0,295	0,0216	***	Fev	-10,150	2,3664	***
ar2	0,067	0,0235	**	Mar	-16,425	2,3658	***
ar3	0,079	0,0225	***	Abr	-17,568	2,8263	***
ar4	0,119	0,0228	***	Mai	-19,416	2,8153	***
ar5	0,043	0,0227	*	Jun	-13,199	2,4741	***
ar6	0,017	0,0226		Ago	-7,454	1,8618	***
ar7	0,330	0,0223	***	Set	-8,145	2,4250	***
ar8	-0,110	0,0225	***	Out	-4,532	2,1833	**
ar9	-0,055	0,0232	**	CopaM1	41,852	6,1808	***
ar10	0,008	0,0227		CopaM2	9,543	3,7220	**
ar11	-0,043	0,0227	*	CopaM3	10,612	3,0099	***
ar12	-0,030	0,0228	*	CopaM4	24,061	3,1204	***
ar13	0,017	0,0224		Anos10	-23,475	5,6835	***
ar14	0,195	0,0215	***	FevA	7,461	3,0480	**
intercept	190,927	4,9894	***	MarA	12,585	3,4255	***
Ter	2,840	0,6966	***	AbrA	16,390	3,9469	***
Qui	1,595	0,7113	**	MaiA	20,572	4,0743	***
Sex	5,376	0,7804	***	JunA	16,012	3,9150	***
Sab	-4,261	0,7801	***	JulA	10,875	3,0136	***
Dom	-7,727	0,7667	***	AgoA	9,996	3,4340	**
Tende	-0,008	0,0035	**	SetA	10,933	3,4814	***
Feriado1	5,389	1,2551	***	OutA	5,549	2,9081	*
Feriado2	-18,463	1,2858	***	Fev10	6,250	3,2970	*
Feriado3	-3,343	1,6776	*	Abr10	9,753	3,8287	**
Feriado4	-4,043	1,9451	**	Mai10	9,565	4,4374	**
Feriado5	-14,382	1,8681	***	Jun10	7,121	3,9527	*
Feriado6	4,459	1,4842	***	Set10	8,424	4,0418	**
Carna1	-12,444	2,2971	***	Out10	12,181	5,0275	**
Carna2	-22,746	1,8661	***	Nov10	15,305	5,2742	**
Carna3	10,598	2,3531	***	Dez10	24,365	5,5860	***
Carna4	5,945	1,8903	***	Dom15	-12,051	1,7014	***
Santa1	-18,601	1,7621	***	GRa4h	-3,922	1,9225	**
Santa3	3,926	2,2149	*	RFi12h	6,778	1,8701	***
Christi1	-33,923	2,2983	***	ObraRF1	-24,224	2,6199	***
Christi2	-12,725	2,2768	***	ObraRF2	-23,736	2,9602	***
FimAno1	-17,245	1,7936	***	ObraRF3	-10,326	2,4920	***
FimAno2	-9,268	1,8431	***	ObraRF4	-7,636	2,5592	**
FimAno4	3,358	1,2945	**				

Tabela 6-4 – Coeficientes das variáveis do modelo (Coef), desvio padrão (DP) e significância (Sig, onde “***” = $Coef > 3xDP$; “**” = $Coef > 2xDP$; e “*” $1xDP$) para o aeroporto de Recife.

6.7.2 Análise dos Resíduos do Modelo de Recife

Na Figura 6-13 estão os gráficos dos resíduos do modelo do aeroporto de Recife, com os resíduos num histograma, num box-plot, num QQ-plot e numa série em linha.

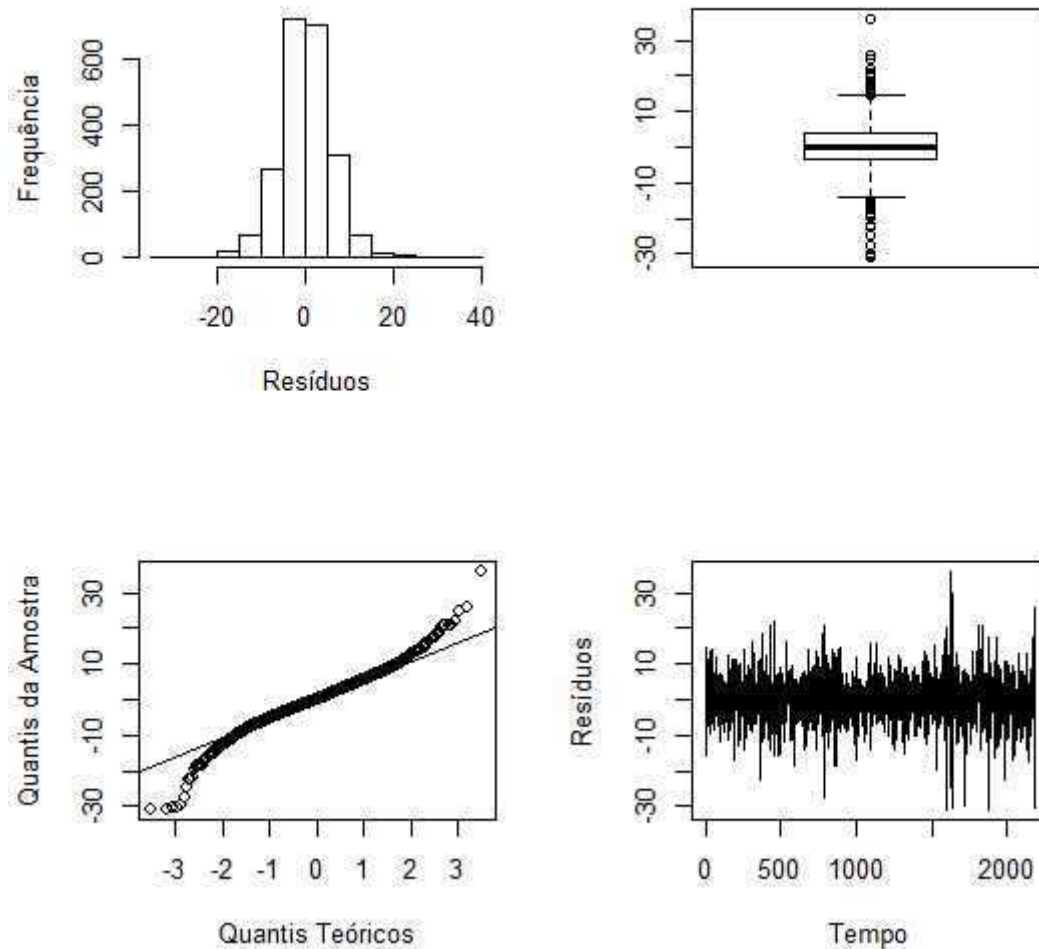


Figura 6-13 - Em cima vemos o histograma e o boxplot, embaixo o Q-Q plot e a série dos resíduos do modelo do aeroporto de Recife

6.7.3 Análise da Previsão do Modelo de Recife

Ajustado o modelo, o próximo passo será analisar o resultado da previsão. Usaremos as covariáveis padronizadas como no modelo preditivo para os três primeiros meses de 2016.

Na Figura 6-14, do aeroporto de Recife, podemos analisar a série (linha vermelha) com a quantidade de operações de pousos e decolagens da aviação comercial de 1º de outubro de 2015 até 31 de março de 2016. Os 3 últimos meses de 2015 foram usados no modelo e os 3 primeiros meses de 2016 não foram. Nesta mesma figura podemos ver a série projetada pelo modelo (linha preta) para o primeiro trimestre de 2016, dentro dos limites de 2 vezes os desvios padrões (área pontilhada). Podemos ver que a série projetada ficou parecida com a série real, mas que durante o carnaval a série real chegou a estar abaixo do intervalo de variação da série projetada pelo modelo.

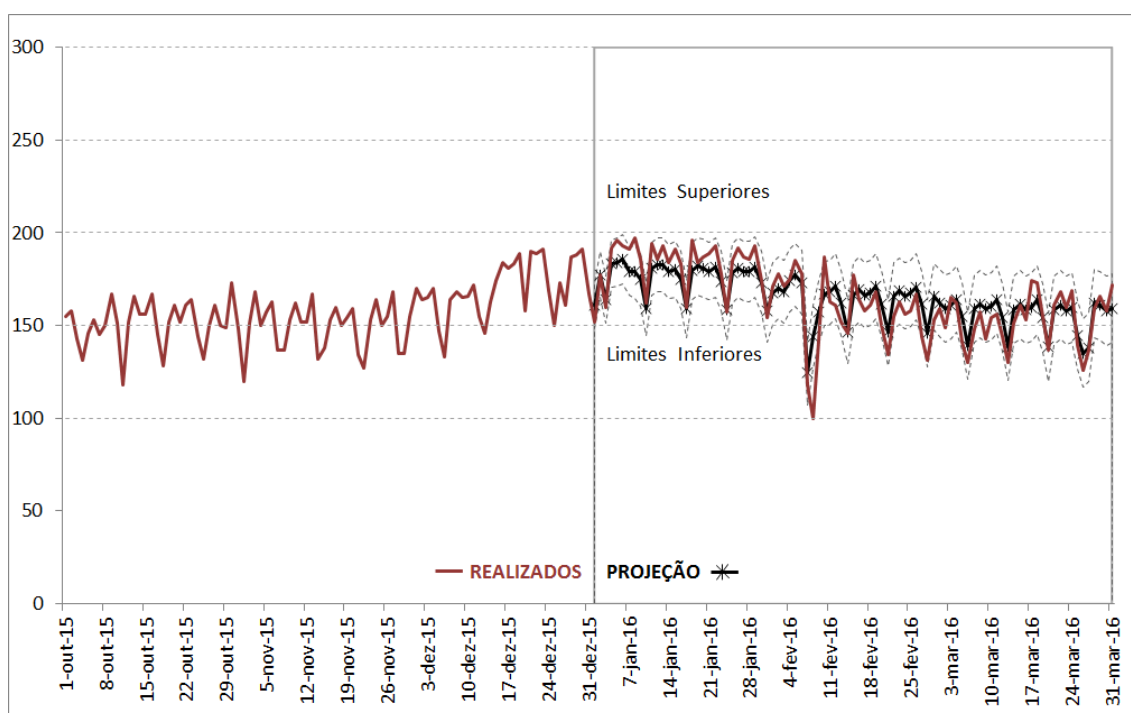


Figura 6-14 - Gráfico do total de pousos e decolagens por dia com a projeção de 3 meses a partir de 1º de janeiro de 2016 da aviação comercial no aeroporto de Recife.

Na Figura 6-15, tem a mesma informação da figura anterior só que com o foco em um mês, antes da previsão e o primeiro mês de 2016 com a série real e a previsão segundo o modelo proposto. Repare que no curto prazo (1 mês), o modelo proposto fica um pouco abaixo da quantidade de pousos e decolagens da aviação comercial no aeroporto de Recife.

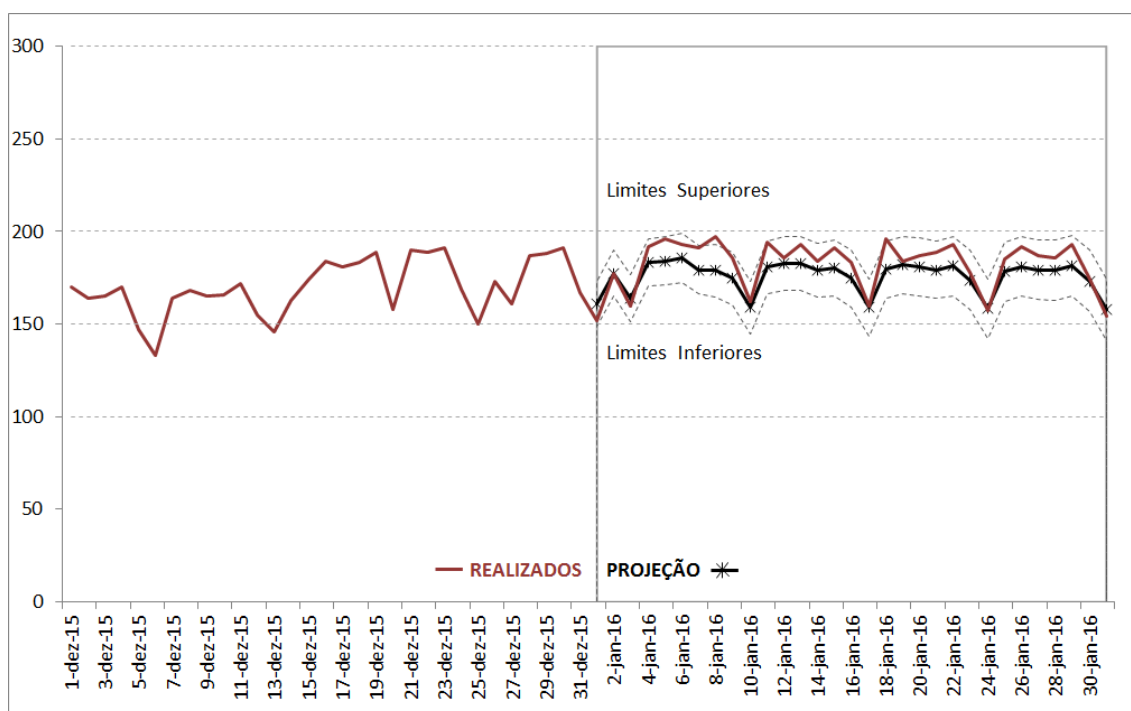


Figura 6-15 - Gráfico do total de pousos e decolagens por dia com a projeção de 1 mês a partir de 1º de janeiro de 2016 da aviação comercial no aeroporto de Recife.

Na Figura 6-16, o modelo foi ajustado a partir dos dados de 1º de janeiro de 2010 até 31 de julho de 2014 para o aeroporto de Recife, e fizemos a previsão do que restou do ano de 2014. Nesta figura podemos ver o que realmente aconteceu e a previsão do modelo com o intervalo de segurança de 2 desvios padrões, e com isso, confirmar a qualidade do modelo ajustado neste trabalho. A previsão no início de dezembro de 2014 mostra uma significativa mudança de comportamento da série, isso se deve a obras na pista do aeroporto, que o modelo já diagnosticou devido a outras obras já quantificadas no modelo de Recife. Com relação às covariáveis do modelo, tivemos que retirar as que tinham relação direta com o ano de 2015 e com a binária AnoA.

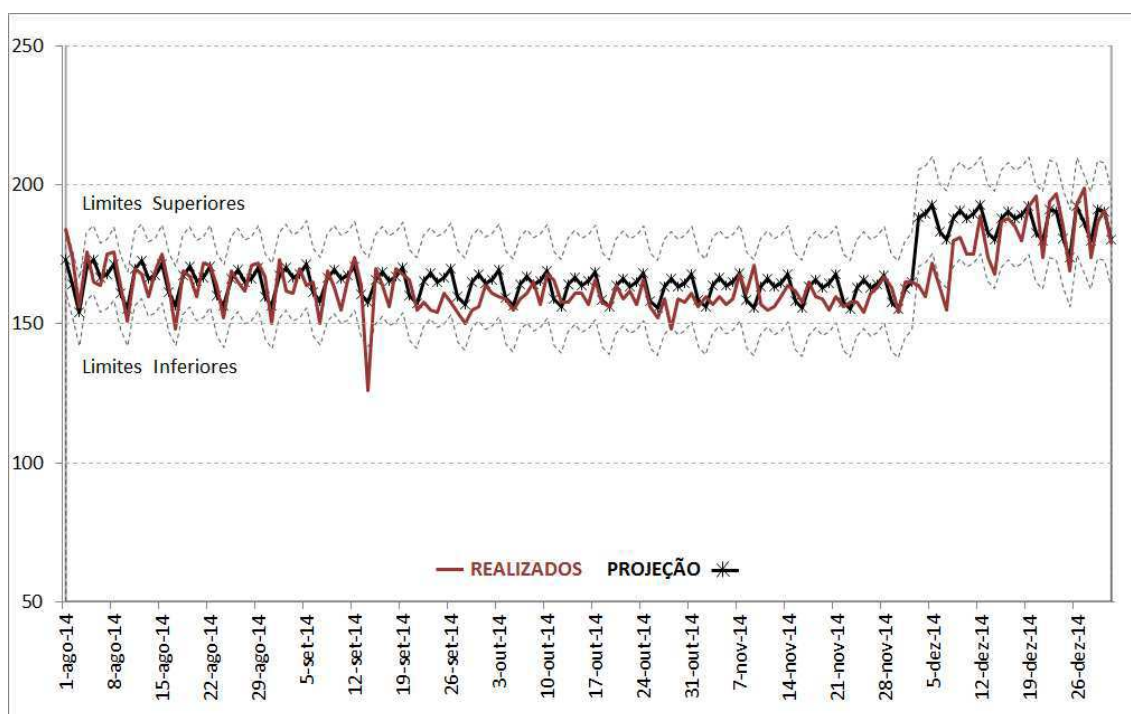


Figura 6-16 - Gráfico do total de pousos e decolagens por dia com a projeção de 5 meses a partir de 1º de agosto de 2014 da aviação comercial no aeroporto de Recife.

6.8 Programa R (Auxiliar)

O programa utilizado para ajudar a implementar o modelo, testar os coeficientes, analisar os resíduos e fazer a previsão de 3 meses é o programa R que possui pacotes já implementados com estatísticas de séries temporais, gráficos e testes estatísticos.

A programação utilizada no programa R neste trabalho está disponível no Apêndice A. As principais funcionalidades utilizadas foram do pacote *Stats Package*, segue alguns exemplos da programação utilizada:

- acf – Gráfico da função de Autocorrelação;
- AIC – Critério de Informação Akaike;
- arima – Modelo de séries temporais;
- BIC – Critério de Informação Bayesiana;
- boxplot – Gráfico de Boxplot;
- coef – Extrai os coeficientes dos modelos;
- hist – Gráfico de Histograma;

- `jarque.bera.test` – Teste de Jarque Bera (pacote *tseries*);
- `predict` – Faz a previsão do modelo;
- `qqnorm` – Gráfico Qqnormal;
- `read.csv` – Faz a importação de dados de um arquivo `.csv`;
- `res` – Extrai os resíduos do modelo;
- `se` – Extrai os desvios padrões dos modelos;
- `ts.plot` – Gráfico de uma série temporal;

7 CONSIDERAÇÕES FINAIS

Este trabalho alcançou seu objetivo de realizar a previsão de operações de pousos e decolagens da aviação comercial para quatro aeroportos brasileiros no curto prazo.

As diversas variáveis binárias e a tendência temporal se mostraram estatisticamente significativas para explicar o número de pousos e decolagens da aviação comercial, assim como a presença de uma dependência temporal explicada por fatores do modelo ARMA. Entretanto, trabalhos com o modelo ARMAX sobre a hipótese de normalidade dos erros idiossincráticos e esta hipótese, infelizmente não pôde ser totalmente corroborada pela análise dos resíduos que demonstrou que as caudas da distribuição geradora dos resíduos é diferente da normal. Uma possibilidade é a modelagem de modelos ARMAX com erros t-Student, mas isso está na agenda dos trabalhos futuros. Mesmo com esta violação, os ajustes mostraram bom poder de previsão, o que nos anima a seguir as análises nesta linha de pensamento.

Para o médio prazo, alguns fatores que podem ser desenvolvidos é a chegada de encomendas de aeronaves (que aumentam a quantidade de oferta de voo de forma atípica) e o ajuste de malha aérea por causa da crise (que também diminui a quantidade de oferta de voo de forma atípica). Esses fatores podem ser previsíveis devido ao tempo necessário, para receber uma aeronave e para ajustar um conjunto de rotas que ficam disponíveis nas autorizações dadas pela ANAC.

Aeroportos com grande quantidade de operações como Guarulhos, Galeão e Brasília podem ser mais afetados pela crise econômica que o Brasil está passando desde 2015. Com isso, a previsão a partir do segundo mês pode ter sido afetada por fatores atípicos que não estavam no modelo (como reestruturação da malha aérea das companhias aéreas). Enquanto que o aeroporto de Recife, por possuir menos operações (mesmo sendo segundo maior aeroporto da Região Nordeste), conseguiu ter sua quantidade de operações previstas muito próximo das operações reais (conforme Figura 6-14) para os 3 primeiros meses do ano de 2016.

Os principais fatores que impactam a quantidade diária de operações é o dia da semana, o mês, a presença de feriado (no dia ou em dia próximo), a presença de um grande evento

nacional e a quantidade de operações de 14 dias anteriores (Série Temporal Autoregressiva de ordem 14, AR(14)).

8 FUTUROS TRABALHOS

A aplicação da metodologia, utilizada neste trabalho, para prever a quantidade de operações de pousos e decolagens de outros aeroportos, pode ajudar a aperfeiçoar a técnica e trazer resultados importantes para um conjunto maior de aeroportos e o tráfego aéreo brasileiro.

O primeiro passo é tentar analisar estes dados com modelos ARMAX que tenham a cauda mais pesada que a distribuição normal, como por exemplo a t-Student. Além disso, seria interessante incluir distribuições assimétricas para os erros (e.g. t-Student assimétrica).

Entretanto, ainda assim estaríamos tratando os dados de forma aproximada por se tratar de dados de contagem. Podemos, então, utilizar modelos mais apropriados para dados de contagem com os modelos que seguem distribuição Poisson e Binomial Negativa. Estes modelos são amplamente utilizados em regressão para dados de contagem. Porém, temos claramente dependência temporal nos dados que também devem ser tratados de forma adequada. Por isso, estamos interessados em utilizar outras classes de modelos de séries temporais – os modelos em espaço de estados – combinados com Poisson. Também incluiremos variações deste modelo para tratar sobredispersão dos dados. Assim como também poderemos incorporar o estudo estatístico em modelos de otimização sob incerteza para o planejamento da rede de produção e distribuição de QAV.

9 REFERÊNCIAS BIBLIOGRÁFICAS

1. AISWEB. **NOTAM: SBRF B1434/2013**. < <http://www.aisweb.aer.mil.br/>>. Acesso em: 07 jun. 2016.
2. AISWEB. **NOTAM: SBRF B1579/2013**. < <http://www.aisweb.aer.mil.br/>>. Acesso em: 07 jun. 2016.
3. AISWEB. **NOTAM: SBRF B0177/2014**. < <http://www.aisweb.aer.mil.br/>>. Acesso em: 07 jun. 2016.
4. AISWEB. **NOTAM: SBRF B0346/2014**. < <http://www.aisweb.aer.mil.br/>>. Acesso em: 07 jun. 2016.
5. AISWEB. **NOTAM: SBRF B1415/2014**. < <http://www.aisweb.aer.mil.br/>>. Acesso em: 07 jun. 2016.
6. BACEN, Banco Central do Brasil. **Indicadores Econômicos Consolidados, I.4 - Índices de preços ao consumidor**. Disponível em: <<http://www.bcb.gov.br/pec/Indeco/Port/indeco.asp>>. Acesso em: 30 mai. 2016.
7. BACEN, Banco Central do Brasil. **Indicadores Econômicos Consolidados, I.51 - Produto Interno Bruto e taxas médias de crescimento**. Disponível em: <<http://www.bcb.gov.br/pec/Indeco/Port/indeco.asp>>. Acesso em: 30 mai. 2016.
8. BARROS, J. S. **Análise e Previsão de Taxas de Retornos Mensais de Fundos de Investimento**. Rio de Janeiro: UFRJ/IM, 2014.
9. BOX, G. E. P.; JENKINS, G. M. **Time Series Analysis Forecast and Control**. Edição revisada. San Francisco: Holden Day, 1976.
10. BRASIL. **Decreto n. 6.834 de 30 de abril de 2009**. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2009/Decreto/D6834.htm>. Acesso em: 31 mai. 2016.
11. BRASIL. **Decreto n. 7.069 de 20 de janeiro de 2010**. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2010/Decreto/D7069.htm#art1>. Acesso em: 31 mai. 2016.
12. CNT. **Transporte e economia – transporte aéreo de passageiros**. Brasília: Confederação Nacional do Transporte, 2015.
13. DECEA, CGNA. **Anuário Estatístico de Tráfego Aéreo 2013**. Disponível em: <<http://portal.cgna.gov.br/>>. Acesso em: 24 mai. 2016.

14. DECEA, CGNA. **Anuário Estatístico de Tráfego Aéreo 2014**. Disponível em: <<http://portal.cgna.gov.br/>>. Acesso em: 24 mai. 2016.
15. DECEA, CGNA. **Anuário Estatístico de Tráfego Aéreo 2015**. Disponível em: <<http://portal.cgna.gov.br/>>. Acesso em: 24 mai. 2016.
16. DECEA. **CIRCEA 100-55 - Instruções para Processamento de Dados Estatísticos de Tráfego Aéreo e Preenchimento dos IEPV 100-34, 100-35, 100-36, 100-39 e 100-40**. Disponível em: <<http://publicacoes.decea.gov.br/?i=publicacao&id=4073>>. Acesso em: 24 mai. 2016.
17. GUJARATI, D. N. **Econometria Básica**. Rio de Janeiro: Elsevier, 2006.
18. HILL, R. C.; JUDGE, G. G.; GRIFFITHS, W. E. **Econometria**. 3 ed. São Paulo: Saraiva, 2010.
19. INFRAERO. **Anuário Estatístico Operacional 2010**. Disponível em: <<http://www.infraero.gov.br/index.php/br/estatistica-dos-aeroportos.html>>. Acesso em: 24 mai. 2016.
20. INFRAERO. **Anuário Estatístico Operacional 2011**. Disponível em: <<http://www.infraero.gov.br/index.php/br/estatistica-dos-aeroportos.html>>. Acesso em: 24 mai. 2016.
21. INFRAERO. **Anuário Estatístico Operacional 2012**. Disponível em: <<http://www.infraero.gov.br/index.php/br/estatistica-dos-aeroportos.html>>. Acesso em: 24 mai. 2016.
22. INFRAERO. **Anuário Estatístico Operacional 2013**. Disponível em: <<http://www.infraero.gov.br/index.php/br/estatistica-dos-aeroportos.html>>. Acesso em: 24 mai. 2016.
23. INFRAERO. **Anuário Estatístico Operacional 2014**. Disponível em: <<http://www.infraero.gov.br/index.php/br/estatistica-dos-aeroportos.html>>. Acesso em: 24 mai. 2016.
24. INFRAERO. **Anuário Estatístico Operacional 2015**. Disponível em: <<http://www.infraero.gov.br/index.php/br/estatistica-dos-aeroportos.html>>. Acesso em: 24 mai. 2016.
25. KICINGER, R.; CHEN, J. T. **Airport Capacity Prediction with Explicit Consideration of Weather Forecast Uncertainty**. Journal of Air Transportation, vol. 24, n. 1. Jan. 2016.

26. LINHARES, G. C. **Planejamento aeroportuário no Brasil: discussão sobre contribuições da prospectiva estratégica.** Journal of Transport Literature, vol. 6, n. 4, pp. 147-163. Out. 2012.
27. R Core Team. **R: A language and environment for statistical computing** [Computer software manual]. Vienna, Austria. Acesso em: 10 fev. 2016.
28. REDEMET. **Rede de Meteorologia do Comando da Aeronáutica** <www.redemet.aer.mil.br>. Acesso em: 06 jun. 2016.

ÍNDICE DE FIGURAS

Figura 1-1 - Gráfico do total de pousos e decolagens da aviação comercial e geral de 66 aeroportos do Anuário Estatístico Operacional da INFRAERO de 2010 a 2015.....	8
Figura 1-2 - Gráfico com o crescimento anual do PIB brasileiro (BACEN), do índice de inflação IPCA (BACEN) e da quantidade de pousos e decolagens de 66 aeroportos brasileiros (Anuário Estatístico da INFRAERO).	9
Figura 4-1 - Gráfico do total de pousos e decolagens por dia da aviação comercial de 1º de janeiro de 2010 a 31 de dezembro de 2015 dos aeroportos do Galeão (a), de Brasília (b), de Guarulhos (c) e de Recife (d).	18
Figura 4-2 - Gráfico da divisão do total de pousos e decolagens por dia da semana pelo total de pousos e decolagens médio do ano da aviação comercial nos aeroportos do Galeão (a), de Brasília (b), de Guarulhos (c) e de Recife (d).	21
Figura 4-3 - Gráfico da divisão do total de pousos e decolagens por mês pelo total de pousos e decolagens médio do ano da aviação comercial nos aeroportos do Galeão (a), de Brasília (b), de Guarulhos (c) e de Recife (d).	23
Figura 4-4 - Gráfico do total de pousos e decolagens por dia durante o carnaval e por ano da aviação comercial nos aeroportos do Galeão (a), de Brasília (b), de Guarulhos (c) e de Recife (d).	26
Figura 4-5 - Gráfico do total de pousos e decolagens por dia durante feriados que caíram na terça-feira e por ano da aviação comercial nos aeroportos do Galeão (a), de Brasília (b), de Guarulhos (c) e de Recife (d).	28
Figura 4-6 - Gráfico do total de pousos e decolagens por dia durante feriados que caíram no final de semana e por ano da aviação comercial nos aeroportos do Galeão (a), de Brasília (b), de Guarulhos (c) e de Recife (d).	30
Figura 4-7 - Gráfico do total de pousos e decolagens por dia durante o Natal por ano da aviação comercial nos aeroportos do Galeão (a), de Brasília (b), de Guarulhos (c) e de Recife (d).	32
Figura 4-8 - Gráfico com a quantidade de dia, de 1º de janeiro de 2010 a 31 de dezembro de 2015, que os aeroportos ficaram funcionando, com condições meteorológicas, abaixo dos mínimos.....	36
Figura 4-9 - Gráfico com a quantidade de dia, de 1º de janeiro de 2010 a 31 de dezembro de 2015, que os aeroportos ficaram funcionando, com condições meteorológicas, a instrumento.	37

Figura 5-1 - Histograma da quantidade de operações de pousos e decolagens da aviação comercial nos aeroportos do Galeão (a), de Brasília (b), de Guarulhos (c) e de Recife (d).	47
Figura 6-1 - Em cima vemos o histograma e o boxplot, embaixo o Q-Q plot e a série dos resíduos do modelo do aeroporto do Galeão.	65
Figura 6-2 - Gráfico do total de pousos e decolagens por dia com a projeção de 3 meses a partir de 1º de janeiro de 2016 da aviação comercial no aeroporto do Galeão.....	66
Figura 6-3 - Gráfico do total de pousos e decolagens por dia com a projeção de 1 mês a partir de 1º de janeiro de 2016 da aviação comercial no aeroporto do Galeão.	67
Figura 6-4 - Gráfico do total de pousos e decolagens por dia com a projeção de 5 meses a partir de 1º de agosto de 2014 da aviação comercial no aeroporto do Galeão.	68
Figura 6-5 - Em cima vemos o histograma e o boxplot, embaixo o Q-Q plot e a série dos resíduos do modelo do aeroporto de Brasília.	70
Figura 6-6 - Gráfico do total de pousos e decolagens por dia com a projeção de 3 meses a partir de 1º de janeiro de 2016 da aviação comercial no aeroporto de Brasília.....	71
Figura 6-7 - Gráfico do total de pousos e decolagens por dia com a projeção de 1 mês a partir de 1º de janeiro de 2016 da aviação comercial no aeroporto de Brasília.....	72
Figura 6-8 - Gráfico do total de pousos e decolagens por dia com a projeção de 5 meses a partir de 1º de agosto de 2014 da aviação comercial no aeroporto de Brasília.	73
Figura 6-9 - Em cima vemos o histograma e o boxplot, embaixo o Q-Q plot e a série dos resíduos do modelo do aeroporto de Guarulhos.	75
Figura 6-10 - Gráfico do total de pousos e decolagens por dia com a projeção de 3 meses a partir de 1º de janeiro de 2016 da aviação comercial no aeroporto de Guarulhos.....	76
Figura 6-11 - Gráfico do total de pousos e decolagens por dia com a projeção de 1 mês a partir de 1º de janeiro de 2016 da aviação comercial no aeroporto de Guarulhos.	77
Figura 6-12 - Gráfico do total de pousos e decolagens por dia com a projeção de 5 meses a partir de 1º de agosto de 2014 da aviação comercial no aeroporto de Guarulhos.	78
Figura 6-13 - Em cima vemos o histograma e o boxplot, embaixo o Q-Q plot e a série dos resíduos do modelo do aeroporto de Recife	80
Figura 6-14 - Gráfico do total de pousos e decolagens por dia com a projeção de 3 meses a partir de 1º de janeiro de 2016 da aviação comercial no aeroporto de Recife.....	81
Figura 6-15 - Gráfico do total de pousos e decolagens por dia com a projeção de 1 mês a partir de 1º de janeiro de 2016 da aviação comercial no aeroporto de Recife.....	82

Figura 6-16 - Gráfico do total de pousos e decolagens por dia com a projeção de 5 meses a partir de 1º de agosto de 2014 da aviação comercial no aeroporto de Recife. 83

ÍNDICE DE TABELAS

Tabela 5-1 - Quadro de binárias para os feriados que serão testadas para todos os aeroportos. Com exceção dos feriados de Natal e Réveillon.	54
Tabela 6-1 – Coeficientes das variáveis do modelo (Coef), desvio padrão (DP) e significância (Sig, onde “***” = Coef > 3xDP; “**” = Coef > 2xDP ;e “*” 1xDP) para o aeroporto do Galeão.....	61
Tabela 6-2 – Coeficientes das variáveis do modelo (Coef), desvio padrão (DP) e significância (Sig, onde “***” = Coef > 3xDP; “**” = Coef > 2xDP ;e “*” 1xDP) para o aeroporto de Brasília.....	69
Tabela 6-3 – Coeficientes das variáveis do modelo (Coef), desvio padrão (DP) e significância (Sig, onde “***” = Coef > 3xDP; “**” = Coef > 2xDP ;e “*” 1xDP) para o aeroporto de guarulhos.	74
Tabela 6-4 – Coeficientes das variáveis do modelo (Coef), desvio padrão (DP) e significância (Sig, onde “***” = Coef > 3xDP; “**” = Coef > 2xDP ;e “*” 1xDP) para o aeroporto de Recife.....	79

APÊNDICE A

Programação utilizada no Programa R para rodar os modelos utilizados neste trabalho, usando como exemplo o aeroporto do Galeão.

```
# IMPORTAR DADOS PARA O PRGRAMA R.
```

```
s2 = read.csv("../DadosGL_2010.csv",header=TRUE,sep=";")
```

```
xreg2 <- cbind( Qua=s2$QuaA, Sex=s2$SexA, Sab=s2$SabA, Dom=s2$DomA,  
              Feriado1=s2$Feriado1A, Feriado2=s2$Feriado2A, Feriado3=s2$Feriado3A,  
              Feriado4=s2$Feriado4A, Feriado5=s2$Feriado5A, Feriado6=s2$Feriado6A,  
              Carna1=s2$Carna1A, Carna2=s2$Carna2A, Carna3=s2$Carna3A, Carna4=s2$Carna4A,  
              Santa1=s2$Santa1A, Santa2=s2$Santa2A, Santa3=s2$Santa3A,  
              Christi1=s2$Christi1A, Christi2=s2$Christi2A, Christi3=s2$Christi3A,  
              FimAno1=s2$FimAno1A, FimAno2=s2$FimAno2A,  
              Fev=s2$FevA, Mar=s2$MarA, Abr=s2$AbrA, Mai=s2$MaiA,  
              Jun=s2$JunA, Ago=s2$AgoA, Set=s2$SetA, Out=s2$OutA,  
              Nov=s2$NovA, Dez=s2$DezA,  
              CopaM1=s2$CopaM1A, CopaM2=s2$CopaM2A, CopaM3=s2$CopaM3A,  
              CopaM4=s2$CopaM4A,  
              Anos10=s2$Ano_10, Anos11=s2$Ano_11,  
              MarA=s2$MarA*s2$AnoA, AbrA=s2$AbrA*s2$AnoA,  
              MaiA=s2$MaiA*s2$AnoA, JunA=s2$JunA*s2$AnoA,  
              AgoA=s2$AgoA*s2$AnoA, SetA=s2$SetA*s2$AnoA, OutA=s2$OutA*s2$AnoA,  
              NovA=s2$NovA*s2$AnoA, DezA=s2$DezA*s2$AnoA,  
              Ter15=s2$TerA*s2$Ano_15, Qua15=s2$QuaA*s2$Ano_15,  
              Sex15=s2$SexA*s2$Ano_15,  
              Sab15=s2$SabA*s2$Ano_15, Dom15=s2$DomA*s2$Ano_15,  
              GLa2h=s2$GLa2h, GLa3h=s2$GLa3h, GLa4h=s2$GLa4h,  
              GLi8h=s2$GLi8h, GLi12h=s2$GLi12h,  
              GRa1h=s2$GRa1h, GRa3h=s2$GRa3h, GRa4h=s2$GRa4h, GRa5h=s2$GRa5h,  
              RJa1h=s2$RJa1h, RJa2h=s2$RJa2h, RJa3h=s2$RJa3h, RJa4h=s2$RJa4h,  
              RJa5h=s2$RJa5h, R Ji4h=s2$R Ji4h, R Ji8h=s2$R Ji8h, R Ji12h=s2$R Ji12h)
```

```
# Modelo ARIMAX do aeroporto do Galeão.
```

```
J10D2 = arima(s2$Voo_ComAjust, order=c(14,0,0), xreg=xreg2)
```

```
AIC(J10D2)
```

```
BIC(J10D2)
# GRÁFICOS DOS RESÍDUOS.
res <- J10D2$res
par(mfrow=c(2,2))
hist(res, main = "", xlab = "Resíduos", ylab = "Frequência")
boxplot(res)
qqnorm(res, main = "", xlab = "Quantis Teóricos", ylab = "Quantis da Amostra")
qqline(res)
ts.plot(res, xlab = "Tempo", ylab = "Resíduos")
# EXPORTAR OS RESÍDUOS.
tabela <- data.frame(residuo = res, voos = s2$Voo_ComAjust)
write.csv(tabela, "ResiduosGL.csv", row.names = TRUE)
# EXPORTAR COEFICIENTES E DESVIOS EM FORMA DE TABELA.
tabela2 <- data.frame(coef = J10D2$coef, desv = sqrt(diag(J10D2$var.coef)))
write.csv(tabela2, "CoefsGL.csv", row.names = TRUE)
# PREVISÃO DO MODELO ENCONTRADO.
s2 = read.csv("../PrevGL_2016.csv",header=TRUE,sep=";")
# REPETE O COMANDO DA TABULAÇÃO USADA NO MODELO COM XREG2 MUDANDO PARA
XREG22.
PredJ10D2 = predict(J10D2, n.ahead=91, newxreg=xreg22)
# EXPORTAR PREVISÃO.
tabela3 <- data.frame(voos = PredJ10D2$pred, desv = PredJ10D2$se)
write.csv(tabela3, "PredictGL.csv", row.names = TRUE)
```