



Tiago Faria Rocha

**Previsão de demanda de passageiros aéreos
offshore utilizando técnicas de séries
temporais hierárquicas**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para
obtenção do grau de Mestre pelo Programa de Pós-
graduação em Engenharia de Produção, do
Departamento de Engenharia Industrial da PUC-Rio.

Orientador: Prof. Fernando Luiz Cyrino Oliveira

Rio de Janeiro
Junho de 2020



Tiago Faria Rocha

**Previsão de demanda de passageiros aéreos
offshore utilizando técnicas de séries
temporais hierárquicas**

Dissertação apresentada como requisito parcial para
obtenção do grau de Mestre pelo Programa de Pós-
Graduação em Engenharia de Produção da PUC-Rio.
Aprovada pela Comissão Examinadora abaixo:

Prof. Fernando Luiz Cyrino Oliveira

Orientador

Departamento de Engenharia Industrial – PUC Rio

Prof. Paula Medina Maçaíra Louro

Departamento de Engenharia Industrial – PUC Rio

Prof. Erick Meira de Oliveira

Financiadora de Estudos e Projetos

Rio de Janeiro, 08 de junho de 2020

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização do autor, do orientador e da universidade.

Tiago Faria Rocha

Graduou-se em Engenharia Aeronáutica pela Universidade de São Paulo (USP) em 2013. Iniciou suas atividades na Petrobras, no mesmo ano de sua formatura, onde trabalha até os dias de hoje. Atualmente, desempenha suas funções na gerência de Competência e Segurança em Aviação e na gerência executiva de Logística, Manutenção e Serviços.

Ficha Catalográfica

Rocha, Tiago Faria

Previsão de demanda de passageiros aéreos *offshore* utilizando técnicas de séries temporais hierárquicas / Tiago Faria Rocha ; orientador: Fernando Luiz Cyrino Oliveira. – 2020.

57 f. : il. color. ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial, 2020.

Inclui bibliografia

1. Engenharia Industrial - Teses. 2. Previsão de passageiros. 3. Transporte aéreo offshore. 4. Séries temporais hierárquicas. I. Oliveira, Fernando Luiz Cyrino. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Industrial. III. Título.

CDD: 658.5

Agradecimentos

Primeiramente, a Deus, por ter me dado vida, saúde e garra para completar essa importante etapa de minha carreira profissional.

Aos meus familiares, amigos e companheira, pela serenidade de entender minha ausência em certos eventos importantes para me dedicar à elaboração desse trabalho.

À Petrobras, por incentivar que seus funcionários continuem aprimorando seus conhecimentos e por ter patrocinado o presente estudo.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Ao gerente do Transporte Aéreo *Offshore*, pela confiança ao me indicar para o Programa de Mestrado Profissional da PUC.

Ao gerente de Segurança e Competência em Aviação, pela compreensão das limitações que o mestrado me imputava na atividade laboral.

Aos colegas de trabalho, sempre dispostos a colaborar com minhas atividades quando o tempo me era escasso.

A todos aqueles que escreverem os artigos científicos que me suportaram a fundamentar minha tese.

Por fim, ao meu orientador, Prof. Fernando Luiz Cyrino Oliveira, pelas valiosas sugestões e boa vontade de ajudar em todos os aspectos que envolveram esse trabalho.

Resumo

Rocha, Tiago Faria; Oliveira, Fernando Luiz Cyrino (Orientador). **Previsão de demanda de passageiros aéreos *offshore* utilizando técnicas de séries temporais hierárquicas**. Rio de Janeiro, 2020, 57p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

Um bom gerenciamento logístico otimiza as atividades de transporte aéreo *offshore*, tornando-as mais eficientes e diminuindo custos para o contratante. Uma série de decisões estratégicas, por exemplo a contratação de helicópteros e os investimentos em infraestrutura aeroportuária, são dependentes da previsão de demanda de passageiros. O presente trabalho analisou a demanda de transporte aéreo *offshore* da Petrobras para o Estado do Rio de Janeiro, à luz das principais teorias de séries temporais hierárquicas, com o objetivo de identificar qual destas é mais adequada para um horizonte de previsão de doze meses à frente. Foram analisadas as estratégias de *single-level approach* (*bottom-up* e *top-down*), de reconciliação ótima (*ordinary least squares* e *weighted least squares*) e de minimização de traço (covariância da própria amostra e valendo-se do *shrink estimator*), todas utilizando como método de previsão base o amortecimento exponencial. Foram utilizados dados dos anos de 2014 até 2019 de todos os aeródromos usados pela Petrobras no Estado do Rio de Janeiro: Farol de São Tomé, Campos dos Goytacazes, Macaé, Cabo Frio e Jacarepaguá. Os resultados foram avaliados em três métricas distintas de acurácia: RMSE (*Root Mean Square Error*), MAPE (*Mean Absolute Percentage Error*) e MASE (*Mean Absolute Scaled Error*), sendo aplicados para os dois níveis existentes de agregação. Os resultados foram ranqueados para cada técnica, nas três métricas citadas anteriormente, sendo, então, consolidados através de uma média aritmética simples. Ao cabo, concluiu-se que o método de minimização de traço *sample covariance* é o mais preciso em termos globais.

Palavras-chave

Previsão de passageiros; transporte aéreo *offshore*; séries temporais hierárquicas.

Abstract

Rocha, Tiago Faria; Oliveira, Fernando Luiz Cyrino (Advisor). **Forecasting demand for offshore air passengers using hierarchical time series techniques.** Rio de Janeiro, 2020, 57p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

Good logistical management optimizes offshore air transport activities, making them more efficient and reducing costs for the contractor. A series of strategic decisions, such as hiring helicopters and investments in airport infrastructure are dependent on forecasting passenger demand. The present work consisted of analyzing the demand for Petrobras offshore air transport to the State of Rio de Janeiro, based on the main theories of hierarchical time series, with the objective of identifying which of these is more suitable for a twelve-month steps ahead forecast. The strategies of single-level approach (bottom-up and top-down), optimal reconciliation (ordinary least squares and weighted least squares) and trace minimization (sample covariance and shrink estimator) were analyzed, all using exponential smoothing as the basic forecasting method. Data from 2014 to 2019 were gathered for all aerodromes used by Petrobras in the State of Rio de Janeiro: Farol de São Tomé, Campos dos Goytacazes, Macaé, Cabo Frio and Jacarepaguá. The results were evaluated with three different metrics of accuracy: RMSE (Root Mean Square Error), MAPE (Mean Absolute Percentage Error) and MASE (Mean Absolute Scaled Error), applied to the two existing levels of aggregation. The results were ranked for each technique, in the three metrics mentioned above, and then consolidated using a simple arithmetic mean. The overall results indicated that sample covariance trace minimization method provided the most accurate results.

Keywords

Passenger forecast; offshore air transport; hierarchical time series.

Sumário

1. Introdução	12
2. Revisão da literatura	14
2.1. Previsão de demanda de transporte aéreo <i>offshore</i>	14
2.2. Séries temporais hierárquicas	17
3. Referencial teórico	20
3.1. <i>Single-level approach</i>	21
3.1.1. <i>Bottom-up</i>	22
3.1.2. <i>Top-down</i>	22
3.1.3. <i>Middle-out</i>	23
3.2. Reconciliação	23
3.2.1. <i>Optimal mint reconciliation</i>	24
3.3. Amortecimento exponencial	26
4. Aspectos metodológicos	31
4.1. Quantificação de erro	32
5. Passageiros transportados <i>offshore</i>	34
5.1. Macaé	34
5.2. Farol de São Tomé	37
5.3. Campos do Goytacazes	40
5.4. Cabo Frio	42
5.5. Jacarepaguá	44
6. Resultados e discussões	48
6.1. Análise comparativa	52
7. Conclusão	54

LISTA DE SIGLAS

ANFIS	<i>Adaptative-Network-Based Fuzzy Inference System</i>
ARIMA	<i>Autoregressive Integrated Moving Average</i>
ARMA	<i>Autoregressive Moving Average Model</i>
Aseb	Aeroportos do Sudeste do Brasil
Bagging	<i>Bootstrap Aggregating</i>
Brexit	<i>Britain Exit</i>
EPTA	Estação Prestadora de Telecomunicação Aeronáutica
HTS	<i>Hierarchical Time Series</i> (Séries Temporais Hierárquicas)
HW	<i>Holt-Winters</i>
IFR	<i>Instrumented Flight Rules</i>
Infraero	Empresa Brasileira de Infraestrutura Aeronáutica
KELM	<i>Kernel Extreme Learning Machine</i>
MAPE	<i>Mean Absolute Percentage Error</i>
MASE	<i>Mean Absolute Squared Error</i>
NVARNN	<i>Nonlinear Vector Auto-Regression Neural Network</i>
OLS	<i>Ordinary Least Squares</i>
PAX	Número de passageiros
PIB	Produto Interno Bruto
POB	<i>People on Board</i>
RMSE	<i>Root Mean Squared Error</i>
sMAPE	Erro Percentual Médio Simétrico Absoluto
SSA	<i>Singular Spectrum Analysis</i>
SVR	<i>Support Vector Regression</i>
VMD	<i>Variational Mode Decomposition</i>
WLS	<i>Weighted Least Squares</i>

Lista de figuras

Figura 1 - Série Temporal Hierárquica	20
Figura 2 - Estrutura Hierárquica Aeroportos Rio de Janeiro	34
Figura 3 - PAX Macaé X Brent	36
Figura 4 - PAX Farol de S. Tomé X Brent	39
Figura 5 - PAX Campos dos Goytacazes X Brent	41
Figura 6 - PAX Cabo Frio X Brent	44
Figura 7 - PAX Jacarepaguá X Brench	46

Lista de quadros

Quadro 1 - Classificação Amortecimento Exponencial	27
Quadro 2 - Espaço de Estados para Amortecimento Exponencial	28
Quadro 3 - Resultados Modelo Petrobras	52

Lista de tabelas

Tabela 1 - Número de Passageiros Transportados Macaé	35
Tabela 2 - Número de Passageiros Transportados Farol de São Tomé	38
Tabela 3 - Passageiros Transportados Campos dos Goytacazes	41
Tabela 4 - Passageiros Transportados Cabo Frio	43
Tabela 5 - Passageiros Transportados Jacarepaguá	45
Tabela 6 - Resultados RMSE	48
Tabela 7 - Resultados MAPE	49
Tabela 8 - Resultados MASE	50
Tabela 9 - Resultados Consolidados por Técnica	51

Introdução

Nos últimos anos, o Brasil tornou públicas várias descobertas de petróleo localizadas, principalmente, em águas profundas abaixo da camada de sal. Estudos têm confirmado constantemente o potencial das reservas, e as projeções indicam que o Brasil poderá estar entre os países com as maiores reservas de petróleo do mundo em um futuro próximo (RODRIGUES; SAUER, 2015). Para explorar essas novas descobertas, as demandas de transporte de equipes de perfuração e de plataformas de produção aumentarão. O transporte aéreo também se tornará mais complexo à medida que a distância média entre os campos e a costa do Brasil aumentar (HERMETO; FERREIRA FILHO; BAHIENSE, 2014). Para levar as equipes que trabalham nas instalações *offshore*, as empresas costumam usar helicópteros, por ser uma maneira bastante rápida e eficiente de atender as embarcações, tanto para as pessoas que as operam diariamente, como também para trazer especialistas em manutenção ou tarefas especiais por períodos mais curtos ou mais longos (HALSKAU, 2014).

Pode-se dividir a gestão da frota de helicópteros em duas áreas principais: operacional e estratégica. A área operacional é focada no curto prazo e a estratégica, no médio e no longo prazo. Um dado fundamental que apoia as decisões dos gerentes nesse contexto é a previsão da demanda de passageiros aéreos. As previsões de curto prazo fornecem informações importantes para as decisões diárias de gerenciamento de operações. Isso inclui escala de aeronaves, escala de tripulação, ocorrências de atraso devido a congestionamentos, planejamento de manutenção, gerenciamento de receita, campanhas de publicidade e vendas, entre outros. As previsões de longo e médio prazo fornecem informações importantes para o planejamento da infraestrutura aeroportuária (capacidade do terminal, utilização da pista e decisões de expansão), serviços de navegação aérea, pedidos e planejamento de frota (DANTAS; OLIVEIRA; REPOLHO, 2017).

Previsões inadequadas podem levar ao congestionamento do sistema e, conseqüentemente, a um aumento dos custos para as operadoras, além de acarretar ineficiência e baixos níveis de serviço aos usuários. Já as estimativas de demanda muito altas nas previsões podem produzir excesso de capacidade de infraestrutura e custos desnecessários, drenando fundos de investimento que poderiam ser melhor utilizados em outras finalidades (MARAZZO; SCHERRE; FERNANDES, 2010).

A Petrobras, maior empresa de petróleo do Brasil, opera em vários aeroportos diferentes: Macaé, Cabo Frio, Farol de São Tomé, Campo dos Goytacazes, Jacarepaguá, Vitória, Aracaju, Paracuru, Guamaré, Salvador e Navegantes; abrangendo todo o litoral brasileiro, do Sul ao Norte. Nesses aeroportos, existem 58 helicópteros dispostos em três portes diferentes (pequeno, médio e grande), operados por quatro empresas, suprimindo 122 embarcações. Vale dizer que os aeroportos localizados no estado do Rio de Janeiro concentram aproximadamente 90% dos passageiros, por isso, serão o objeto do presente estudo.

Esta dissertação tem como objetivo avaliar qual das técnicas de previsão de séries temporais hierárquicas – *bottom-up*, *top-down* ou reconciliação – aplicadas à demanda de transporte aéreo da Petrobras, no período de um ano à frente (curto a médio prazo) nos aeroportos do Rio de Janeiro, possui a melhor acurácia. Também será realizada uma breve comparação com o método corrente utilizado pela companhia. Para calcular o total de passageiros, foi considerada uma aproximação da demanda de transporte aéreo. Os resultados aferidos foram comparados em termos de *Root Mean Squared Error* (RMSE), *Mean Absolute Percentage Error* (MAPE) e *Mean Absolute Squared Error* (MASE). Em todas as técnicas, foram utilizadas, como previsão base, o método de previsão de amortecimento exponencial.

O conteúdo desta dissertação está disposto da seguinte maneira: na seção 2, consta uma breve revisão da literatura sobre previsão de demanda de transporte aéreo *offshore* e séries temporais hierárquicas. A Seção 3 consiste na análise das metodologias a serem utilizadas. Na seção 4, são apresentados dados e métricas de avaliação usados no estudo. A seção 5 é dedicada à aplicação teórica da metodologia. A seção 6 expõe e analisa os resultados obtidos. Na seção 7, há uma breve comparação entre a metodologia aqui utilizada com o método corrente na Petrobras. Por fim, a última seção apresenta as conclusões e sugestões para futuras pesquisas sobre o assunto.

Revisão da literatura

A revisão da literatura, exposta a seguir, foi organizada em duas seções: previsão de demanda de transporte aéreo *offshore* e séries temporais hierárquicas.

2.1

Previsão de demanda de transporte aéreo *offshore*

Existe uma vasta gama de publicações científicas abordando a temática da previsão de demanda de passageiros. Entretanto, nessa revisão de literatura, não foi encontrado nenhum artigo abordando a questão específica de previsão de demandas de passageiros *offshore*. A comunidade acadêmica, no que tange ao assunto de transporte aéreo *offshore*, ocupa-se majoritariamente com problemas de roteirização, segurança e planejamento de frotas.

Foram identificados alguns artigos que utilizam como insumo uma previsão de demanda de passageiros *offshore*, entretanto, essa informação é tratada de maneira superficial, não sendo abordado com detalhes como aqueles números foram aferidos. Nesse caso, podemos citar o artigo de Hermeto, Ferreira Filho e Bahiense (2014), em que é proposto um problema de pesquisa operacional objetivando identificar o número ideal de aeronaves, em seus respectivos portes, e infraestrutura aeronáutica (aeroportos) que minimizassem o custo total por parte da empresa detentora dos recursos. Uma das premissas utilizadas foi a demanda de passageiros *offshore*, obtida através da multiplicação do número esperado de embarcações pelo valor médio do *People on Board* (POB) delas.

Devido à inexistência de artigos específicos contemplando a previsão de demanda de passageiros *offshore*, o melhor referencial teórico disponível é aquele que aborda a aviação geral. É possível resumir os métodos de previsão em três categorias principais: modelos causais, séries temporais e aprendizado de máquina. A primeira consiste em realizar a previsão com base na interação de variáveis que têm relação com o objeto previsto. Já as séries temporais consistem em métodos estatísticos que analisam os dados passados para propor os melhores valores prováveis no futuro. Finalmente, as metodologias de aprendizado de máquina utilizam algoritmos suportados pela alta capacidade computacional para encontrar padrões, às vezes, não compreensíveis para o ser humano.

Como exemplos de aplicação de modelos causais na previsão de demanda de passageiros aéreos, destacam-se as publicações de Marazzo, Scherre e Fernandes (2010), e de Gelhausen, Berster e Wilken (2018), que serão abordadas nos parágrafos a seguir.

Marazzo, Scherre e Fernandes (2010) investigaram a relação existente entre demanda de transporte aéreo e Produto Interno Bruto (PIB), no período de 1996 a 2006, no Brasil. Os resultados sugerem que o PIB e o número de passageiros (PAX) são cointegrados. A análise de resposta ao impulso indica uma forte reação positiva do PAX devido a uma mudança positiva no PIB. No entanto, o PIB reage a uma mudança no PAX de maneira mais lenta e moderada.

Gelhausen, Berster e Wilken (2018) desenvolveram um modelo versátil baseado na teoria da cointegração, que prevê diretamente o volume de passageiros e voos nos aeroportos alemães. O modelo inclui choques de demanda e a elasticidade-PIB estimada é de 1,31. O modelo foi empregado para estimar os efeitos do *Britain Exit* (Brexit) no volume de tráfego nos aeroportos alemães nos anos de 2016 a 2018.

No que diz respeito aos métodos estatísticos de séries temporais, destacam-se os artigos de Andreoni e Postorino (2006), e de Dantas, Oliveira e Repolho (2017), que serão apresentados nos próximos parágrafos.

Andreoni e Postorino (2006) propuseram modelos *Autoregressive Integrated Moving Average* (ARIMA) univariados e multivariados para estimar os níveis de demanda no aeroporto regional de Reggio Calabria (Sul da Itália), com o objetivo de analisar o impacto de modificações recentes onde se espera que o resultado seja o aumento da demanda de transporte aéreo. As calibrações dos modelos foram obtidas usando a metodologia de Box-Jenkins.

Dantas, Oliveira e Repolho (2017) propuseram um método que combina *Bootstrap Aggregating* (Bagging) e Holt-Winters (HW). Os resultados foram comparados com outros cinco métodos de *benchmark* (SARIMA, Holt-Winters, ETS, Bagged.BLD.MBB.ETS e *Seasonal Naive*), usando o erro médio absoluto percentual simétrico (sMAPE) como métrica de comparação. Essa abordagem foi testada usando o conjunto de dados disponível de 14 países e os resultados do método proposto superaram todos os *benchmarks*.

Por fim, usando métodos de aprendizado de máquina, encontramos as publicações de Alekseev e Seixas (2009), de Xiao *et al.* (2014), de Sun *et al.* (2019), e de Jin *et al.* (2020), que serão discutidas a seguir.

Alekseev e Seixas (2009) desenvolveram um modelo de rede neural artificial para realizar previsão de passageiros no transporte aéreo que se vale de um método de pré-processamento que decompõe as informações para revelar recursos relevantes. Verifica-se que o processamento neural supera a abordagem econométrica tradicional e oferece generalização no comportamento das séries temporais, mesmo onde existem apenas amostras pequenas.

Xiao *et al.* (2014) propuseram uma nova metodologia para previsão de demanda de transporte aéreo, combinando *Singular Spectrum Analysis* (SSA), *Adaptative-Network-Based Fuzzy Inference System* (ANFIS) e *Particle Swarm Optimization* aprimorada para previsão a curto prazo. Foram utilizados dados de passageiros aéreos de Hong Kong para validar o modelo, sendo que os resultados empíricos apontaram para a enorme capacidade dessa nova abordagem, já que os principais *benchmarks* foram superados.

Sun *et al.* (2019) apresentaram uma *MIV-based nonlinear vector autoregression neural network* (NVARNN) para prever fluxo de passageiros aéreos. Para ilustrar e verificar a eficácia da abordagem proposta, foram utilizados dados de fluxos de passageiros do Aeroporto Internacional de Pequim. Os resultados de desempenho da previsão mostram que a abordagem proposta MIV-NVARNN supera consistentemente modelos univariados e outras abordagens híbridas.

E Jin *et al.* (2020) propuseram uma abordagem híbrida VMD-ARMA / KELM-KELM para previsão de curto prazo, que consiste na combinação de *variational mode decomposition* (VMD), *autoregressive moving average model* (ARMA) e *kernel extreme learning machine* (KELM). Para provar e verificar a viabilidade e robustez da abordagem proposta, as demandas de passageiros dos aeroportos de Pequim, Guangzhou e Pudong foram utilizadas para testar o desempenho do modelo. Os resultados experimentais mostram que a nova abordagem tem uma vantagem, em termos de precisão e robustez, em relação aos *benchmarks*.

Os resultados da revisão de literatura, tanto quanto sabe-se, indicam que não existe publicação que analise a questão específica da previsão da demanda de transporte aéreo *offshore*.

2.2

Séries Temporais hierárquicas

Algumas séries temporais podem oferecer uma desagregação natural em uma estrutura hierárquica. Essa estrutura hierárquica pode ser útil para melhorar a qualidade da previsão. A estratégia com estruturas hierárquicas é conhecida como análise de séries temporais hierárquicas (HTS) (KARMY; MALDONADO, 2019). As principais publicações acadêmicas sobre esse assunto serão apresentadas nos parágrafos a seguir.

No artigo Athanasopoulos, Hyndman e Ahmed (2009) exploraram a natureza hierárquica das séries temporais de demanda turística para produzirem previsões de curto prazo para a demanda de turismo doméstico australiano. Os dados e previsões foram organizados em uma hierarquia baseada nas regiões geográficas e nos objetivos da viagem. Foram consideradas cinco estratégias para a previsão hierárquica: duas variações da abordagem *top-down*, o método *bottom-up*, uma versão recentemente proposta de análise *top-down* em que as previsões de nível superior são desagregadas de acordo com as proporções previstas nas séries de nível inferior e, por fim, a *optimal combination approach*. A avaliação de desempenho da previsão mostra que a abordagem de *top-down* com base nas proporções entre as previsões e o *optimal combination approach* apresenta melhor desempenho entre as estratégias analisadas.

Athanasopoulos e Hyndman (2014) aplicaram um método de reconciliação para fazer com que a soma das previsões de nível inferior e superior seja coerente. Os autores argumentaram que os procedimentos tradicionais de *bottom-up* e *top-down* falham ao utilizar os dados disponíveis. Foi sugerida, então, uma nova abordagem, na qual são tomadas médias ponderadas das previsões de diferentes níveis, em um método denominado *optimally reconciliation forecast*.

Liu *et al.* (2016), estudaram o problema de estimativa de valores ausentes de tráfego *web*, onde o mesmo é organizado em várias estruturas hierárquicas, como geográficas e de sites. Foi desenvolvido um algoritmo eficiente, o HTSImpute, para estimar com precisão o valor ausente em séries temporais de tráfego *web*. O HTSImpute é capaz de utilizar as informações de dependência temporal em cada série individualmente, explorar as inter-relações entre elas, bem como garantir a satisfação de restrições de consistência hierárquica. Os resultados obtidos,

utilizando como base de dados três conjuntos de dados sintéticos e três conjuntos de dados hierárquicos de tráfego *web* reais, demonstraram que sua o HTSInput é capaz de fornecer estimativas mais precisas e hierarquicamente consistentes do que os *benchmarks*.

Huber, Gossman e Stuckenschmidt (2017) propuseram um sistema de suporte à decisão, em relação a bens perecíveis, fornecendo previsões hierárquicas em diferentes níveis organizacionais, com base nos dados obtidos dos pontos de venda. Foram utilizados modelos ARIMA multivariados para prever a demanda diária para apoiar as decisões operacionais. A análise de cluster aplicada revelou que itens substituíveis têm um padrão de vendas intra-dia semelhante, o que torna razoável prever a demanda em um nível agregado.

Athanasopoulos *et al.* (2017), introduziram o conceito de hierarquias temporais para previsão de séries temporais. No artigo, estabelecem que uma hierarquia temporal pode ser construída para qualquer série temporal por meio de agregação temporal não sobreposta. As previsões construídas em todos os níveis de agregação são combinadas com a estrutura proposta para resultar em previsões temporalmente reconciliadas, precisas e robustas. A metodologia proposta é independente dos modelos de previsão, podendo incorporar previsões gerenciais de alto nível, que incorporam informações complexas e não estruturadas com previsões estatísticas de nível inferior. Os resultados mostram que a previsão com hierarquias temporais aumenta a precisão em relação à previsão convencional.

Karmy e Maldonado (2019) propuseram uma nova estratégia para previsão de vendas usando *Support Vector Regression* (SVR) e séries temporais hierárquicas. Foram utilizadas três abordagens hierárquicas diferentes de séries temporais: *SVR bottom-up*, *SVR top-down* e *SVR middle-out*, analisando a previsão de vendas para o setor de varejo de viagens. Várias estruturas hierárquicas foram propostas a fim de obter previsões precisas. Os resultados demonstraram as virtudes das séries temporais hierárquicas baseadas em SVR, em termos de desempenho preditivo, quando comparadas às abordagens tradicionais ARIMA e Holt-Winters.

Athanasopoulos *et al.* (2019), exploraram as estruturas de agregação inerentes a variáveis macroeconômicas. Foram utilizados métodos de reconciliação para gerar previsões coerentes com as restrições hierárquicas. Também foram geradas previsões pontuais e probabilísticas, pela primeira vez, no cenário

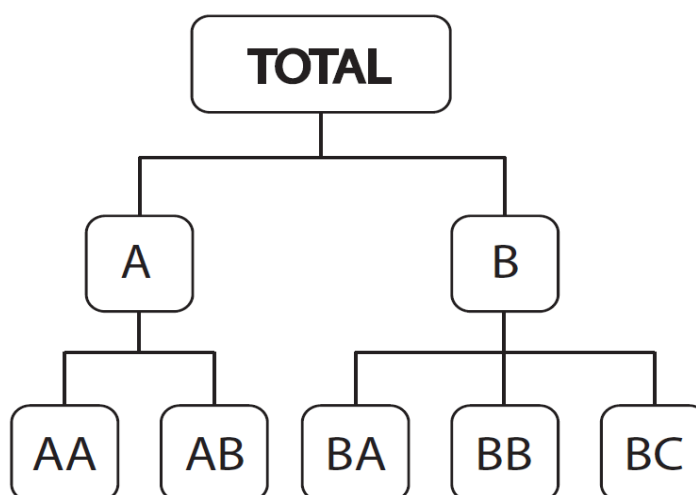
macroeconômico. Usando o PIB australiano, foi demonstrado que a reconciliação, além de resultar em previsões coerentes, também melhora a precisão geral das previsões, tanto para estruturas pontuais quanto para probabilísticas.

Os resultados de nossa revisão de literatura indicam que não existe publicação sobre aplicações de previsão hierárquica na demanda de transporte aéreo *offshore*.

Referencial teórico

Para melhor ilustração da metodologia sobre séries temporais, imaginemos a seguinte estrutura hierárquica:

Figura 1 - Série Temporal Hierárquica



Fonte: Baseado em Athanasopoulos *et al.*, 2019.

Cada quadrado do esquema acima representa uma série temporal, e cada um que se encontra acima é a soma direta dos que estão abaixo. Denotamos, portanto, que o valor total se encontra no maior nível de agregação, assim como a linha mais abaixo, no menor nível de agregação.

Cada uma das séries temporais a seguir será denotada por $y_{i,t}$. Consideremos, então, o vetor $y_{i,t} = (y_{Total,t}, y_{A,t}, y_{B,t}, y_{AA,t}, y_{AB,t}, y_{BA,t}, y_{BB,t}, y_{BC,t})'$ como aquele que abrange todas as séries em todos os níveis hierárquicos. Da mesma forma, tomemos $b_t = (y_{AA,t}, y_{AB,t}, y_{BA,t}, y_{BB,t}, y_{BC,t})'$ como sendo o vetor que contém todas as séries temporais do menor nível de agregação.

Em uma série temporal hierárquica, o modelo de previsão deverá seguir a seguinte premissa de agregação: $y_{Total,t} = y_{A,t} + y_{B,t}$. A equação descrita anteriormente também pode ser reescrita na forma matricial conforme equação 3.1:

$$y_{i,t} = S \cdot b_t \quad (3.1)$$

Em que:

$$\mathbf{S} = \begin{bmatrix} \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ & & \mathbf{I}_5 & & \end{bmatrix}$$

A matriz \mathbf{S} é conhecida como *summing matrix*, sendo \mathbf{I}_5 uma matriz identidade cinco por cinco. Pode-se generalizar a equação (1), que foi deduzida a partir do caso específico da figura 1, para o caso geral onde há n elementos na hierarquia com m elementos no nível máximo de desagregação. Nesse contexto, teríamos $\mathbf{y}_{Total,t} \in \mathbb{R}^n$ e $\mathbf{b}_t \in \mathbb{R}^m$, sendo $m > n$. \mathbf{S} , então, passaria a ser uma matriz $n \times m$.

Uma característica notável da matriz \mathbf{S} é que suas colunas formam um *subspace* vetorial de \mathbb{R}^n , onde residem os resultados que respeitam as premissas de agregação. Tal conjunto de vetores recebe a nomenclatura de *coherent subspace* (ATHANASOPOULOS *et al.*, 2019) e é denominado pela sigla \mathbf{s} .

A seguir, necessitamos definir um conceito importante para a explanação das futuras estratégias: um conjunto de previsões h -passos à frente $\tilde{\mathbf{y}}_{T+h|T}$, ordenadas na mesma sequência de \mathbf{y}_t e geradas utilizando as informações disponíveis até o tempo T são denominadas como sendo coerentes se $\tilde{\mathbf{y}}_{T+h|T} \in \mathbf{s}$. Ou seja, previsões coerentes são aquelas que respeitarão as condições de agregação.

Podemos dividir as estratégias de modelagem de séries temporais hierárquicas em duas correntes principais: *single-level approach* e reconciliação, que serão abordadas nas seções a seguir.

3.1

Single-level approach

Nesse grupo de estratégias, as previsões temporais são realizadas para um nível e depois extrapoladas através de combinações lineares para os demais. As abordagens existentes são classificadas de acordo com a escolha do nível onde as previsões são elaboradas inicialmente: *bottom-up*, começando pelo nível mais desagregado, *top-down*, pelo mais agregado e, por fim, *middle-out*, por um nível intermediário.

3.1.1

Bottom-up

Na estratégia *bottom-up*, inicialmente, são elaboradas as previsões para o nível mais desagregado a partir dos dados históricos disponíveis para o mesmo, ou seja, \mathbf{b}_t . Portanto, podemos afirmar que é necessário, primeiramente, criar um conjunto de previsões h-passos à frente $\hat{\mathbf{b}}_{T+h|T} \in \mathbb{R}^m$.

Para o caso específico da figura 1, temos que $\hat{\mathbf{b}}_{T+h|T} = (\hat{\mathbf{y}}_{AA,T+h|T}, \hat{\mathbf{y}}_{AB,T+h|T}, \hat{\mathbf{y}}_{BA,T+h|T}, \hat{\mathbf{y}}_{BB,T+h|T}, \hat{\mathbf{y}}_{BC,T+h|T})$, onde $\hat{\mathbf{y}}_{i,T+h|T}$ é a previsão h-passos a frente para o nó i . O conjunto de previsões coerentes, para toda a estrutura hierárquica, é gerado através da multiplicação da *summing matrix* por $\hat{\mathbf{b}}_{T+h|T}$.

$$\tilde{\mathbf{y}}_{T+h|T}^{BU} = \mathbf{S} \cdot \hat{\mathbf{b}}_{T+h|T} \quad (3.2)$$

A principal vantagem da estratégia *bottom-up* é que não ocorre perda de informação no processo de agregação. Entretanto, os dados disponíveis no nível mais desagregado podem ser altamente voláteis ou muito ruidosos, tornando o processo de previsão desafiador *subspace* (ATHANASOPOULOS *et al.*, 2019).

3.1.2

Top-down

Diferentemente da estratégia *bottom-up*, a análise *top-down* se inicia construindo as previsões para o nível mais agregado para, então, realizar os desdobramentos para os níveis mais desagregados. Genericamente, podemos afirmar que previsões coerentes, para todo nível hierárquico, são obtidas através da seguinte equação:

$$\tilde{\mathbf{y}}_{T+h|T}^{TD} = \mathbf{S} \cdot \mathbf{p} \cdot \hat{\mathbf{y}}_{Total,T+h|T} \quad (3.3)$$

onde \mathbf{S} é a *summing matrix*, $\hat{\mathbf{y}}_{Total,T+h|T}$ é a previsão h-passos à frente para o nível mais agregado da hierarquia e \mathbf{p} é um vetor m-dimensional contendo as proporções de desdobramento entre os níveis mais agregado e mais desagregado da estrutura, ou seja:

$$\mathbf{p} \cdot \hat{\mathbf{y}}_{Total,T+h|T} = \hat{\mathbf{b}}_{T+h|T} \quad (3.4)$$

Existem várias maneiras de se determinar o vetor \mathbf{p} , a maioria deles se baseando nas proporções dos valores históricos disponíveis. Nesse sentido, Gross e Sohl (1990) apresentaram e propuseram vinte e uma diferentes abordagens possíveis para o cálculo de \mathbf{p} .

Athanasopoulos, Hyndman e Ahmed (2009), almejando superar a perda de informação decorrente do processo de agregação da estratégia *top-down*, introduziram uma nova abordagem para o cálculo do vetor de ponderações \mathbf{p} , onde são utilizados dados das previsões ao invés das informações históricas. Além disso, comprovaram a eficácia do método quando comparado aos principais *benchmarks*.

Por fim, é importante ressaltar que todos os métodos *top-down* têm a desvantagem de introduzirem um viés na previsão, mesmo quando a série histórica em si é não enviesada (ATHANASOPOULOS *et al.*, 2019).

3.1.3

Middle-out

A estratégia *middle-out* pode ser considerada como uma combinação da metodologia *top-down* e *bottom-up*. Primeiramente, escolhe-se um nível intermediário para realizar as previsões. Após essa etapa, são realizadas as previsões dos níveis hierárquicos acima do escolhido pela abordagem *bottom-up* e abaixo do escolhido pela *top-down*.

A desvantagem desse método é a mesma da estratégia *top-down*, uma vez que a contempla em uma de suas etapas, ou seja, a introdução de viés na previsão mesmo quando os dados em si são não enviesados.

3.2

Reconciliação

Diferentemente da abordagem *single-level*, em que são realizadas as previsões em um único nível hierárquico para, posteriormente, através de combinações lineares, estender para os demais níveis; as técnicas de reconciliação baseiam-se em elaborar as previsões em todos os níveis hierárquicos (base) para, então, fazer com que elas se tornem coerentes.

As técnicas de reconciliação podem ser entendidas simplesmente como a projeção das previsões base no *subspace* coerente \mathbf{s} . Tal projeção será equivalente

à multiplicação das soluções base pela *summing matrix* (condição de agregamento) e por uma matriz G .

$$\hat{\mathbf{y}}_{T+h|T} = \mathbf{S} \cdot \mathbf{G} \cdot \hat{\mathbf{y}}_{T+h|T} \quad (3.5)$$

As estratégias de reconciliação utilizadas nesse estudo são restritas àquelas que garantem soluções não enviesadas quando $\hat{\mathbf{y}}_{T+h|T}$ também é não enviesado, ou seja, $(\mathbf{S} \cdot \mathbf{G} \cdot \mathbf{S}) = \mathbf{S}$, e se baseiam na minimização de traço (MinT).

É importante ressaltar que os *single-level approaches* podem ser vistos como casos particulares de reconciliação, sendo que para o caso específico da estratégia *top-down* não é possível manter a condição de ausência de viés.

3.2.1

Optimal minT reconciliation

As estratégias de reconciliação por minimização de traço se baseiam na premissa de que as séries temporais base, ou seja, aferidas sem as condições de agregação e coerência, são não enviesadas. Portanto,

$$\mathbf{E}_{1:T}(\hat{\mathbf{y}}_{T+h|T}) = \mathbf{E}_{1:T}(\mathbf{y}_{T+h|y_1}, \dots, \mathbf{y}_{T+h|y_T}) \quad (3.6)$$

Logo, o valor esperado das previsões é igual ao valor esperado do conjunto de dados.

O erro do conjunto de previsão base pode ser denotado por

$$\hat{\mathbf{e}}_{T+h|T} = \mathbf{y}_{T+h|T} - \hat{\mathbf{y}}_{T+h|T} \quad (3.7)$$

A variância do conjunto de erros de previsão é denotado por \mathbf{W}_h :

$$\mathbf{W}_h = \mathbf{Var}(\hat{\mathbf{e}}_{T+h|T}) \quad (3.8)$$

O erro do conjunto de previsão coerente pode ser denotado por

$$\tilde{\mathbf{e}}_{T+h|T} = \mathbf{y}_{T+h|T} - \tilde{\mathbf{y}}_{T+h|T} \quad (3.9)$$

Wickramasuriya, Athanasopoulos e Hyndman (2019) mostraram que existe uma solução única para G , satisfazendo a condição de não enviesamento $(\mathbf{S} \cdot \mathbf{G} \cdot \mathbf{S}) = \mathbf{S}$, que minimiza o traço da matriz de erros das previsões coerentes:

$$\mathbf{G} = (\mathbf{S}' \cdot \mathbf{W}_h^{-1} \cdot \mathbf{S})^{-1} \cdot \mathbf{S}' \cdot \mathbf{W}_h^{-1} \quad (3.10)$$

O grande desafio da reconciliação é a estimativa da variação do conjunto de erros de previsão \mathbf{W}_h , principalmente para os casos onde $h > 1$. Nesse sentido, Wickramasuriya, Athanasopoulos e Hyndman (2018) demonstraram as principais estratégias para seu cálculo. A primeira delas consiste, basicamente, em estimar que

a variância do conjunto de erros de previsão é diretamente proporcional à matriz identidade, ou seja, $\mathbf{W}_h = \mathbf{K}_h \cdot \mathbf{I}_n$. Essa consideração resulta em $(\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'$. Isso pode ser compreendido como a projeção ortogonal das previsões base no *subspace* coerente \mathbf{s} , que minimiza a distância euclidiana entre os conjuntos de previsões $\hat{\mathbf{y}}_{T+h|T}$ e o conjunto de previsões coerente $\tilde{\mathbf{y}}_{T+h|T}$.

Vale ressaltar que Hyndman *et al.* (2011) obtiveram o mesmo resultado através de uma interpolação linear para o conjunto de previsões, em que ela seria equivalente à condição de agregação multiplicada pelo valor esperado dos dados de maior nível de desagregação em adição a um erro de previsões coerentes com média zero e variância V .

$$\hat{\mathbf{y}}_{T+h|T} = \mathbf{S} \cdot \boldsymbol{\beta}_{T+h|T} + \boldsymbol{\varepsilon}_{T+h|T} \quad (3.11)$$

A minimização quadrática do erro, *Ordinary Least Squares* (OLS), acarreta na mesma matriz de projeção $\mathbf{G} = (\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'$. A principal desvantagem desse método é não levar em conta as correlações existentes entre os níveis hierárquicos.

Outra maneira de se estimar a variância do conjunto de erros de previsão é supondo que este é proporcional à diagonal da ponderação dos erros quadráticos, ou seja, $\mathbf{W}_h = \mathbf{K}_h \cdot \text{diag}(\widehat{\mathbf{W}}_1)$, onde

$$\widehat{\mathbf{W}}_1 = \frac{1}{T} \sum_{t=1}^T \hat{\mathbf{e}}_t \hat{\mathbf{e}}_t' \quad (3.12)$$

é o estimador da matriz de covariância dos erros de previsão in-sample um passo à frente. Isso, por sua vez, faz com que o estimador MinT, com as simplificações adotadas, colapse para um estimador de mínimos quadrados ponderados. Athanasopoulos *et al.* (2017) propõem uma outra abordagem de mínimos quadrados ponderados, onde $\mathbf{W}_h = \mathbf{K}_h \cdot \text{diag}(\mathbf{S}_1)$, sendo $\mathbf{1}$ um vetor coluna n -dimensional. Portanto, os pesos são proporcionais ao número de variáveis necessárias para formar um nível de agregação.

Para ilustrar, tomemos o caso da figura 1, na qual os números de elementos para formar um nível de agregação são, respectivamente, para total A e B, de 5, 2 e 3. Logo, as ponderações utilizadas dependem exclusivamente da estrutura agregativa das séries hierárquicas.

A principal vantagem desse método em relação ao OLS é que a premissa de equivariância dos erros se restringe ao nível mais desagregado (*bottom level*) e não a toda estrutura hierárquica.

Além desses, devemos citar o método no qual a variância do conjunto de erros de previsão é diretamente proporcional à variância dos erros de previsão in-sample um passo a frente, ou seja, $\mathbf{W}_h = \mathbf{K}_h \cdot \widehat{\mathbf{W}}_1$. Apesar da simplicidade da sua suposição, tal estratégia acarreta ótimos resultados quando a estrutura hierárquica é pequena. Entretanto, sua eficácia diminui gradativamente quando o tamanho da estrutura aumenta em relação ao número de dados disponíveis. Esse método é conhecido como *MinT (Sample) estimator*.

Por fim, é necessário citar a metodologia em que consideramos $\mathbf{W}_h = \mathbf{K}_h \cdot \widehat{\mathbf{W}}_1^D$, onde $\widehat{\mathbf{W}}_1^D = \lambda_D \cdot \text{diag}(\widehat{\mathbf{W}}_1) + (1 - \lambda_D) \widehat{\mathbf{W}}_1$

$$\hat{\lambda}_D = \frac{\sum_{i \neq j} \widehat{VAR}(\hat{r}_{ij})}{\sum_{i \neq j} \hat{r}_{ij}^2} \quad (3.13)$$

$\widehat{\mathbf{W}}_1^D$ é o estimador de redução (*shrinkage estimator*) e $\hat{\lambda}_D$ é o parâmetro de intensidade de redução (*shrinkage intensity parameter*) e \hat{r}_{ij} é o elemento (i,j) de $\widehat{\mathbf{R}}_1$, sendo esta a matriz de correlação para um passo à frente, conforme explicitada por Schäfer e Strimmer (2005).

3.3 Amortecimento exponencial

Como fica explicitado nas metodologias anteriores, a reconciliação, em abordagens hierárquicas, é conduzida tendo como ponto de partida previsões-base, isto é, previsões de cada uma das séries pertencentes à hierarquia feitas de forma independente.

Nesse trabalho, vamos nos limitar a utilizar apenas o método de amortecimento exponencial, que será explicado brevemente a seguir.

O fundamento da técnica em questão consiste, basicamente, em explorar as componentes de tendência e sazonalidade em uma série temporal, na qual as previsões são geradas a partir de uma média ponderada das observações passadas, em que o peso de tais observações decai exponencialmente ao longo do tempo, ou seja, a previsão é mais influenciada pelos resultados mais recentes.

Uma representação bastante utilizada para se classificar formulações de amortecimento exponencial é a abordagens por formas de espaço de estado proposta

por Hyndman et al. (2002). Nesta, as formulações de amortecimento exponencial são classificadas de acordo com a natureza das componentes de erro, tendência e sazonalidade. O resumo de tais classificações encontra-se no quadro 1.

Quadro 1 - Classificação Amortecimento Exponencial

Tendência	Sazonalidade		
	N (Nenhum)	A (Aditivo)	M (Multiplicativo)
N (Nenhum)	N, N	N, A	N, M
A (Aditivo)	A, N	A, A	A, M
A_{am} (Aditivo amortecido)	A_{am} , N	A_{am} , A	A_{am} , M
M (Multiplicativo)	M, N	M, A	M, M
M_{am} (Multiplicativo Amortecido)	M_{am} , N	M_{am} , A	M_{am} , M

Fonte: Adaptado de Athanasopoulos e Hyndman, 2014.

As equações utilizadas para realizar previsões para cada uma das classificações acima estão resumidas no quadro 2, denominado de espaço de estados.

Quadro 2 - Espaço de Estados para Amortecimento Exponencial

(continua)

Tendência	Sazonalidade		
	N(Nenhum)	A (Aditivo)	M (Multiplicativo)
N	$\hat{y}_{t+h t} = l_t$	$\hat{y}_{t+h t} = l_t + s_{t-m+h_m^+}$	$\hat{y}_{t+h t} = l_t s_{t-m+h_m^+}$
	$l_t = \alpha y_t + (1 - \alpha)l_{t-1}$	$l_t = \alpha(y_t - I_t) + (1 - \alpha)l_{t-1}$	$l_t = \alpha \left(\frac{y_t}{s_{t-m}} \right) + (1 - \alpha)l_{t-1}$
		$s_t = \gamma(y_t - I_t) + (1 - \gamma)s_{t-m}$	$s_t = \gamma \left(\frac{y_t}{l_{t-1}} \right) + (1 - \gamma)s_{t-m}$
A	$\hat{y}_{t+h t} = l_t + hb_t$	$\hat{y}_{t+h t} = l_t + mb_t + s_{t-m+h_m^+}$	$\hat{y}_{t+h t} = (l_t + hb_t)s_{t-m+h_m^+}$
	$l_t = \alpha y_t + (1 - \alpha)(l_{t-1} - b_{t-1})$	$l_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + b_{t-1})$	$l_t = \alpha \left(\frac{y_t}{s_{t-m}} \right) + (1 - \alpha)(l_{t-1} - b_{t-1})$
	$b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}$	$b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}$	$b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}$
A _{am}	$\hat{y}_{t+h t} = l_t + \varphi_h b_t$	$\hat{y}_{t+h t} = l_t + \varphi_h b_t + s_{t-m+h_m^+}$	$\hat{y}_{t+h t} = (l_t + \varphi_h b_t)s_{t-m+h_m^+}$
	$l_t = \alpha y_t + (1 - \alpha)(l_{t-1} - \varphi b_{t-1})$	$l_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + \varphi b_{t-1})$	$l_t = \alpha \left(\frac{y_t}{s_{t-m}} \right) + (1 - \alpha)(l_{t-1} - \varphi b_{t-1})$
	$b_t = \beta(l_t - l_{t-1}) + (1 - \beta)\varphi b_{t-1}$	$b_t = \beta(l_t - l_{t-1}) + (1 - \beta)\varphi b_{t-1}$	$b_t = \beta(l_t - l_{t-1}) + (1 - \beta)\varphi b_{t-1}$
		$s_t = \gamma(y_t - l_{t-1} - \varphi b_{t-1}) + (1 - \gamma)s_{t-m}$	$s_t = \gamma \left(\frac{y_t}{l_{t-1} + \varphi b_{t-1}} \right) + (1 - \gamma)s_{t-m}$

Quadro 2 - Espaço de Estados para Amortecimento Exponencial

(conclusão)

Tendência	Sazonalidade		
	N(Nenhum)	A (Aditivo)	M (Multiplicativo)
M	$\hat{y}_{t+h t} = l_t + b_t^h$	$\hat{y}_{t+h t} = l_t b_t^h + s_{t-m+h_m^+}$	$\hat{y}_{t+h t} = (l_t b_t^h) s_{t-m+h_m^+}$
	$l_t = \alpha y_t + (1 - \alpha)(l_{t-1} b_{t-1})$	$l_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} b_{t-1})$	$l_t = \alpha \left(\frac{y_t}{s_{t-m}} \right) + (1 - \alpha)(l_{t-1} b_{t-1})$
	$b_t = \beta \left(\frac{l_t}{l_{t-1}} \right) + (1 - \beta)b_{t-1}$	$b_t = \beta \left(\frac{l_t}{l_{t-1}} \right) + (1 - \beta)b_{t-1}$	$b_t = \beta \left(\frac{l_t}{l_{t-1}} \right) + (1 - \beta)b_{t-1}$
		$s_t = \gamma(y_t - l_{t-1} b_{t-1}) + (1 - \gamma)s_{t-m}$	$s_t = \gamma \left(\frac{y_t}{l_{t-1} b_{t-1}} \right) + (1 - \gamma)s_{t-m}$
M _{am}	$\hat{y}_{t+h t} = l_t + b_t^{\varphi h}$	$\hat{y}_{t+h t} = l_t b_t^{\varphi h} + s_{t-m+h_m^+}$	$\hat{y}_{t+h t} = (l_t b_t^{\varphi h}) s_{t-m+h_m^+}$
	$l_t = \alpha y_t + (1 - \alpha)(l_{t-1} b_{t-1}^{\varphi})$	$l_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} b_{t-1}^{\varphi})$	$l_t = \alpha \left(\frac{y_t}{s_{t-m}} \right) + (1 - \alpha)(l_{t-1} b_{t-1}^{\varphi})$
	$b_t = \beta \left(\frac{l_t}{l_{t-1}} \right) + (1 - \beta)b_{t-1}^{\varphi}$	$b_t = \beta \left(\frac{l_t}{l_{t-1}} \right) + (1 - \beta)b_{t-1}^{\varphi}$	$b_t = \beta \left(\frac{l_t}{l_{t-1}} \right) + (1 - \beta)b_{t-1}^{\varphi}$
		$s_t = \gamma(y_t - l_{t-1} b_{t-1}^{\varphi}) + (1 - \gamma)s_{t-m}$	$s_t = \gamma \left(\frac{y_t}{l_{t-1} b_{t-1}^{\varphi}} \right) + (1 - \gamma)s_{t-m}$

Fonte: Elaborado pelo autor, 2020.

No quadro 3, l_t denota o nível; b_t , a tendência; s_t , a componente sazonal no período t ; m compreende a sazonalidade anual; α, β, γ e φ são constantes, onde $\alpha > 0$, γ e $\varphi < 1$ e $0 < \beta < \alpha$. Por fim, $\hat{y}_{t+h|t}$ denota a previsão h passos à frente, $\varphi_h = \varphi + \varphi^2 + \dots + \varphi^h$ e $h_m^+ = [(h - 1) \bmod m] + 1$.

Por fim, além das variações apresentadas para as componentes de tendência e sazonalidade, é importante frisar que a componente de erro pode ter natureza aditiva ou multiplicativa. Com isso, um total de 30 formulações de amortecimento exponencial podem ser obtidas no contexto das formas de espaço de estado proposta por Hyndman *et al.* (2002).

Para mais detalhes sobre o método, consultar *Forecasting: principles and practice*, de Hyndman e Athanasopoulos (2018).

Aspectos metodológicos

O primeiro passo para aplicar qualquer metodologia de obtenção de previsão de séries temporais hierárquicas é realizar cada uma delas de maneira individual. Nesse estudo em particular, como frisado anteriormente, optou-se por utilizar apenas a rotina automática de seleção de formulações de amortecimento exponencial por formas de espaço de estado (ETS - um acrônimo que significa ExponenTial Smoothing, ou ainda Error, Trend and Seasonality, as três componentes que podem variar de natureza nas formulações). Tal escolha se justifica, principalmente, pelo objetivo da presente dissertação em se concentrar em avaliar qual das estratégias de séries temporais hierárquicas melhor se adequa, em termos de métricas de erros, ao transporte de passageiros offshore, de tal sorte que a metodologia da previsão base pouco influencia no resultado final, conforme explicitado por Athanasopoulos *et al.* (2019).

As séries temporais hierárquicas foram estruturadas, bem como as previsões realizadas, utilizando o pacote *hts* implementado em R por Hyndman *et al.* (2018). Foi utilizado como *set* de treinamento os meses disponíveis entre os anos de 2014 e 2018. Já os dados de 2019 foram utilizados como *set* de teste.

Todos as previsões foram realizadas utilizando a função *forecast()* do pacote HTS, para 12 passos à frente ($h=12$), onde a entrada '*method*' foi utilizada nas seguintes siglas '*bu*', '*tdfp*', '*tdgsf*' ou '*comb*', representando respectivamente as estratégias *bottom-up*, *top-down* baseada em proporções históricas, *top-down* baseada em proporções obtidas através das previsões e, por fim, *comb* denota os métodos de reconciliação ótima.

A entrada '*fmethod*', para todos os casos, foi utilizada como sendo '*ets*', que significa *exponential time smoothing*, ou seja, para todos as estratégias de séries temporais os casos base foram obtidos através de amortecimento exponencial. Já a entrada '*weights*' é utilizada apenas para os casos em que a reconciliação ótima foi escolhida, tendo sido utilizado respectivamente nas seguintes categorias: OLS, WLS, *nseries* para *structural scalling* e, por último, *mint* para minimização de traço. Por fim, a entrada '*covariance*' é utilizada apenas nos casos em que for escolhida a estratégia de minimização de traço, tendo sua covariância calculada via dois métodos distintos: '*sam*', através da amostra, e '*shr*', utilizando o *shrinkage estimator*.

A seguir, serão detalhados os métodos utilizados para averiguar a acurácia das estratégias adotadas no presente caso.

4.1 Quantificação de erro

A acurácia, nesse estudo, foi inferida através das seguintes métricas de erro: RMSE (*Root Mean Squared Error*), MAPE (*Mean Absolute Percentage Error*) e MASE (*Mean Absolute Scaled Error*), onde RMSE é calculado por:

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n}} \quad (4.1)$$

Sendo \hat{y}_i os valores de previsão para o ano de 2019 e y_i os valores disponíveis para o mesmo ano. A importância dessa métrica consiste em servir de base de comparação a diversos outros métodos estatísticos. Sua principal vantagem é não utilizar em seu denominador variáveis que possam tender a números próximos a zero, elevando artificialmente o valor do erro. Sua desvantagem, é amplificar o peso dado a um possível *outlier*, uma vez que eleva ao quadrado as diferenças oriundas entre previsão e valor observado.

Já o MAPE é dado por:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\% \quad (4.2)$$

Sua utilização no presente estudo é justificada por possuir características opostas ao RMSE, ou seja, tem como vantagem não amplificar o seu resultado diante da ocorrência de outliers. Entretanto, retorna valores irrealistas quando os valores observados são próximos de zero. Vale dizer, que o MAPE, assim como qualquer outra métrica de erro, apresentam a vantagem de serem dotadas de independência de uma escala. Isso permite facilmente sua comparação quando há a necessidade de se fazer comparações entre análises realizadas ante conjunto de dados distintos. Por fim, vale lembrar que essa métrica de acurácia apresenta resultados distintos quando aplicadas pra mesma sequência numérica, mas em escalas distintas, por exemplo, séries temporais contendo dados de temperatura, os resultados serão influenciados a depender da escala adotada, Celsius ou Fahrenheit.

Por fim, o MASE é calculado da seguinte forma para séries sem sazonalidade:

$$MASE = \frac{\frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|}{\frac{1}{T-1} \sum_{t=2}^T |y_t - y_{t-1}|} \quad (4.3)$$

Para séries com sazonalidade, o MASE é calculado da seguinte maneira:

$$MASE = \frac{\frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|}{\frac{1}{T-m} \sum_{t=m+1}^T |y_t - y_{t-m}|} \quad (4.4)$$

onde m é o período sazonal e T é o número de observações do conjunto de treinamento. O MASE pode ser entendido como a proporção do erro do método em questão pelo erro do método ingênuo (*naive*). Sua formulação veio no intuito de se suprimir as desvantagens supracitadas, em outras palavras, não ocorrer em desvirtuação nos casos de incorrência de *outliers* ou valores e erros próximos a zero. Ainda assim, em caso de observações iguais, a presente métrica retorna valores sem significado.

Para todos os casos acima, \hat{y}_i são os valores de previsão, y_i são os valores históricos da série temporal, n é o número total de dados no *set* de teste, T é o número de dados totais e m é o período da sazonalidade.

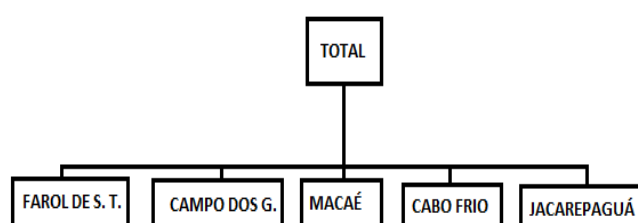
Para mais informações a respeito das métricas de acurácia consultar Hyndman e Koehler (2006).

Passageiros transportados *offshore*

O número total de passageiros *offshore* transportados pela Petrobras no estado do Rio de Janeiro é a soma de todos os passageiros dos aeroportos de Farol de São Tomé, Campo dos Goytacazes, Macaé, Cabo Frio e Jacarepaguá, portanto, configura-se naturalmente como uma estrutura hierárquica.

Nesse contexto, as estratégias de previsão para séries temporais hierárquicas são adequadas, pois possuem a capacidade de produzir resultados coerentes. A figura 2, a seguir, resume a estrutura dos aeroportos.

Figura 2 - Estrutura Hierárquica Aeroportos Rio de Janeiro



Fonte: Elaborado pelo autor, 2020.

O estado do Rio de Janeiro concentra aproximadamente noventa por cento das operações *offshore* da Petrobras, sendo o destino de voos para as bacias de Campos, Santos e Espírito Santo. Nas seções a seguir, serão detalhados os dados para cada aeroporto entre os anos de 2014 e 2019.

5.1

Macaé

O aeroporto de Macaé, de código ICAO SBME, localizado na cidade de mesmo nome, encontra-se a aproximadamente 179 km ao norte do Rio de Janeiro. O aeródromo em questão é base para o transporte de passageiros, majoritariamente, que se destinam a embarcações que se localizam na Bacia de Campos.

Este é um aeródromo público que, até 2019, era administrado pela Empresa Brasileira de Infraestrutura Aeronáutica (Infraero), passando, então, a ser uma concessão pública controlada pelo consórcio Aeroportos do Sudeste do Brasil (Aseb), que pertence ao grupo *Zurich Airports*.

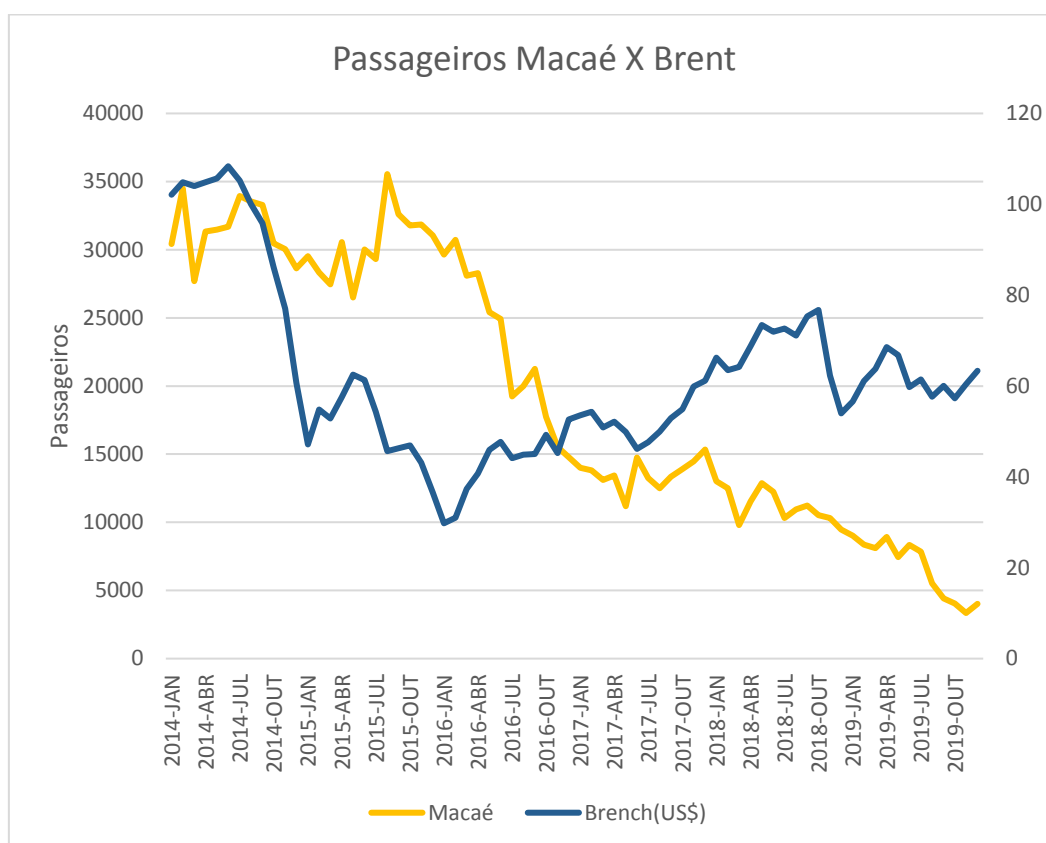
Segundo dados do Rotaer, o aeroporto possui pista asfaltada de 1.200 metros, tendo capacidade de receber aeronaves de médio porte e asa fixa (ATR-72). SBME é dotado de torre de controle, Estação Prestadora de Telecomunicação Aeronáutica (EPTA), estando homologada para operações noturnas e *Instrumented Flight Rules* (IFR). Atualmente, encontram-se operando neste aeroporto as quatro grandes companhias de táxi aéreo *offshore*: Aeróleo, CHC, Líder e Omni. Também estão presentes as empresas de aviação comercial Passaredo e Azul. Vale mencionar, ainda, que o aeródromo não possui nenhuma limitação para ampliação de suas operações no que tange a hangares, pista e processamento de passageiros.

Os números de passageiros transportados por mês foram retirados da base de dados do sistema SAP Petrobras e encontram-se na tabela 1. A figura 3 compara os dados de passageiros transportados com o preço do petróleo Brent.

Tabela 1 - Número de Passageiros Transportados Macaé

Macaé	2019	2018	2017	2016	2015	2014
JAN.	9.032	13.011	13.990	29.642	29.518	30.428
FEV.	8.367	12.490	13.808	30.725	28.317	34.494
MAR.	8.095	9.806	13.110	28.100	27.464	27.694
ABR.	8.932	11.527	13.443	28.288	30.554	31.335
MAI.	7.447	12.868	11.187	25.417	26.487	31.469
JUN.	8.337	12.242	14.750	24.916	30.020	31.690
JUL.	7.841	10.311	13.254	19.238	29.306	33.939
AGO.	5.531	10.945	12.507	19.987	35.549	33.540
SET.	4.408	11.239	13.349	21.248	32.611	33.286
OUT.	4.035	10.520	13.912	17.743	31.775	30.477
NOV.	3.345	10.304	14.468	15.537	31.853	30.051
DEZ.	4.024	9.458	15.328	14.764	31.060	28.643

Fonte: Elaborado pelo autor, 2020.

Figura 3 - PAX Macaé X Brent

Fonte: Elaborado pelo autor, 2020.

Como é possível observar, entre 2014 e 2015, o número de passageiros *offshore* transportados no aeroporto de Macaé se manteve na faixa de 30 mil a 35 mil por mês. Já ao observar a figura 3 no mesmo período, o preço do petróleo saiu de sua máxima histórica, aproximadamente 100 dólares o barril, para 40 dólares.

A partir de então, a companhia adotou severamente medidas de contenção de despesas, reduzindo praticamente a zero o investimento em exploração – o que leva à diminuição de voos, uma vez que grande parte deles é de troca de turno de pessoal para sondas exploratórias. Houve também diminuição do número de voos devido a venda de ativos e descomissionamento de plataformas.

No período de 2015 até 2016, a redução de passageiros no aeroporto de Macaé continuou, saindo de 30 mil passageiros para cerca de 15 mil ao final de 2016. A redução ocorreu, em grande parte, devido à política de venda de ativos e de contenção de despesas com exploração. O preço do petróleo, no mesmo período, saiu dos 40 dólares, passando por um valor de mínima abaixo dos 30 dólares e subindo novamente para 40.

Nos anos de 2017 e 2018, o número de passageiros saiu da ordem de 15 mil para 10 mil passageiros por mês. Mesmo com a recuperação do preço do petróleo, que subiu para cerca de 70 dólares o barril ao final de 2016, a política da empresa retira a prioridade dos investimentos na Bacia de Campos, colocando como objetivo estratégico o foco no pré-sal (Bacia de Santos) e a redução do endividamento.

No ano de 2019, o aeroporto de Macaé passa por uma redução significativa dos passageiros transportados, saindo de 10 mil para 4 mil passageiros por mês. O preço do petróleo, nesse período, teve uma leve queda, saindo dos 70 dólares para cerca de 60 dólares o barril. O maior responsável por essa diminuição é a transferência de parte dos voos para o aeroporto de Farol de São Tomé, onde a proximidade com as plataformas da Bacia de Campos é maior. Nesse período, o pátio de Farol de São Tomé foi redesenhado, proporcionando uma maior capacidade de aeronaves.

5.2

Farol de São Tomé

O aeroporto de Farol de São Tomé, de código ICAO SBFS, localizado na cidade de Campos dos Goytacazes, encontra-se a aproximadamente 298 km ao norte do Rio de Janeiro. O aeródromo em questão é base para o transporte de passageiros, majoritariamente, que se destinam a embarcações que se localizam na Bacia de Campos e no Espírito Santo.

Este foi um aeródromo privado da Petrobras até 2019, não possuindo pista para pousos e decolagens, o que pôde resultar em certas limitações de desempenho para os helicópteros. SBFS não é dotado de torre de controle, sendo o controle de pátio feito por um colaborador em solo. Existe o serviço de EPTA, estando homologada para operações IFR. Atualmente, encontra-se operando neste aeroporto apenas uma companhia de táxi aéreo *offshore*: Omni.

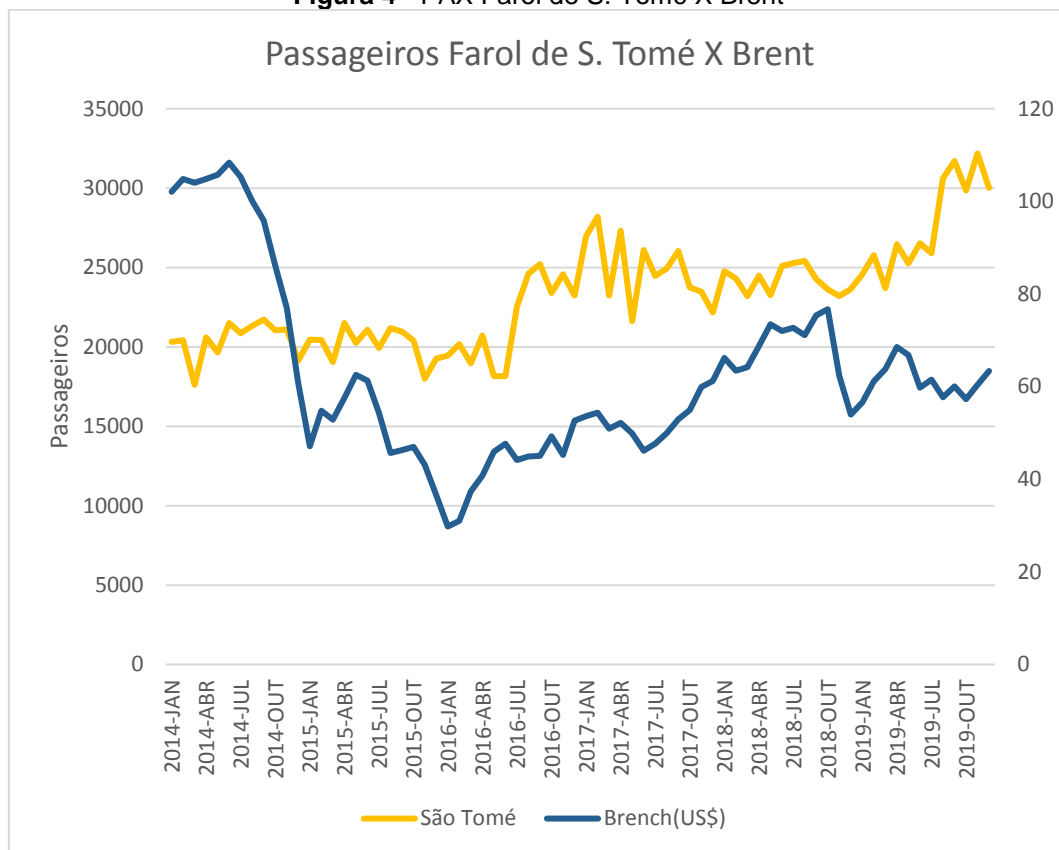
Vale dizer, ainda, que o aeródromo possui limitação de hangares e pátio, e qualquer aumento na capacidade de aeronaves será através de incremento de infraestrutura que dependem de vultosos investimentos. Além disso, é importante salientar que Farol de São Tomé é a melhor localização, em termos de distância, em relação às embarcações da Bacia de Campos.

Os números de passageiros transportados por mês foram retirados da base de dados do sistema SAP Petrobras e encontram-se na tabela 2 e a comparação com os preços do petróleo estão dispostas na Figura 4.

Tabela 2 - Número de Passageiros Transportados Farol de São Tomé

São Tomé	2019	2018	2017	2016	2015	2014
JAN.	24.564	24.775	26.969	19.440	20.456	20.308
FEV.	25.771	24.306	28.207	20.175	20.431	20.417
MAR.	23.709	23.190	23.248	18.967	19.044	17.617
ABR.	26.469	24.494	27.303	20.714	21.502	20.591
MAIO	25.258	23.253	21.619	18.157	20.264	19.651
JUN.	26.531	25.096	26.107	18.165	21.078	21.498
JUL.	25.901	25.275	24.478	22.539	19.929	20.857
AGO.	30.611	25.425	24.938	24.604	21.177	21.323
SET.	31.705	24.271	26.048	25.215	20.957	21.730
OUT.	29.857	23.619	23.734	23.405	20.393	21.045
NOV.	32.182	23.201	23.472	24.570	17.991	21.103
DEZ.	30.005	23.638	22.170	23.243	19.277	19.151

Fonte: Elaborado pelo autor, 2020.

Figura 4 - PAX Farol de S. Tomé X Brent

Fonte: Elaborado pelo autor, 2020.

Como é possível observar na Figura 4, entre 2014 e 2016, o número de passageiros *offshore* transportados no aeroporto de Farol de São Tomé se manteve na faixa de 20 mil por mês. No mesmo período, o preço do petróleo saiu de sua máxima histórica, cerca de 100 dólares o barril para 40, passando por uma mínima de 35 dólares.

No período de 2016 até 2018, houve um incremento no número de passageiros transportados em Tomé para aproximadamente 25 mil passageiros por mês. Isso se explica, majoritariamente, pela mudança no regime de utilização das aeronaves, que passaram a não depender mais do regime dos pilotos, ou seja, a aeronave fica disponível pelo período de doze horas e há mais de uma tripulação para guarnecê-la. O preço do petróleo, nessa janela de tempo, saiu dos 40 dólares, atingindo 70 dólares o barril.

De 2018 até 2019, o número de passageiros saiu da ordem de 25 mil, saltando para 30 mil passageiros por mês. No mesmo período, parte dos voos do aeroporto de Macaé foi transferida para Farol de São Tomé, principalmente por razões de

eficiência logística. O preço do petróleo teve uma leve queda saindo dos 70 dólares para 58.

5.3 Campos do Goytacazes

O aeroporto de Campos dos Goytacazes, de código ICAO SBCP, localizado na cidade de Campos dos Goytacazes, encontra-se a aproximadamente 277 Km ao norte do Rio de Janeiro. O aeródromo serve de base para o transporte de passageiros, majoritariamente, que se destinam a embarcações que localizam-se na Bacia de Campos e Espírito Santo.

Este é um aeródromo público da prefeitura de Campos que é administrado pela Infraero. SBCP não é dotado de torre de controle, sendo o controle de pátio feito por um colaborador em solo. Existe o serviço de EPTA, estando homologada para operações IFR. Atualmente, encontra-se operando neste aeroporto apenas uma companhia de táxi aéreo *offshore*: Omni. Além disso, vale ressaltar que também existem companhias aéreas de asa fixa operando no local, utilizando aeronaves de médio porte (ATR-72).

O aeroporto é dotado de 1.544 metros de pista asfaltada, não existindo nenhum tipo de procedimento de decolagem que limite o desempenho dos helicópteros. Vale dizer, ainda, que o aeródromo não possui limitação de hangares e pátio, suportando um possível aumento de demanda sem necessidade de investimentos elevados.

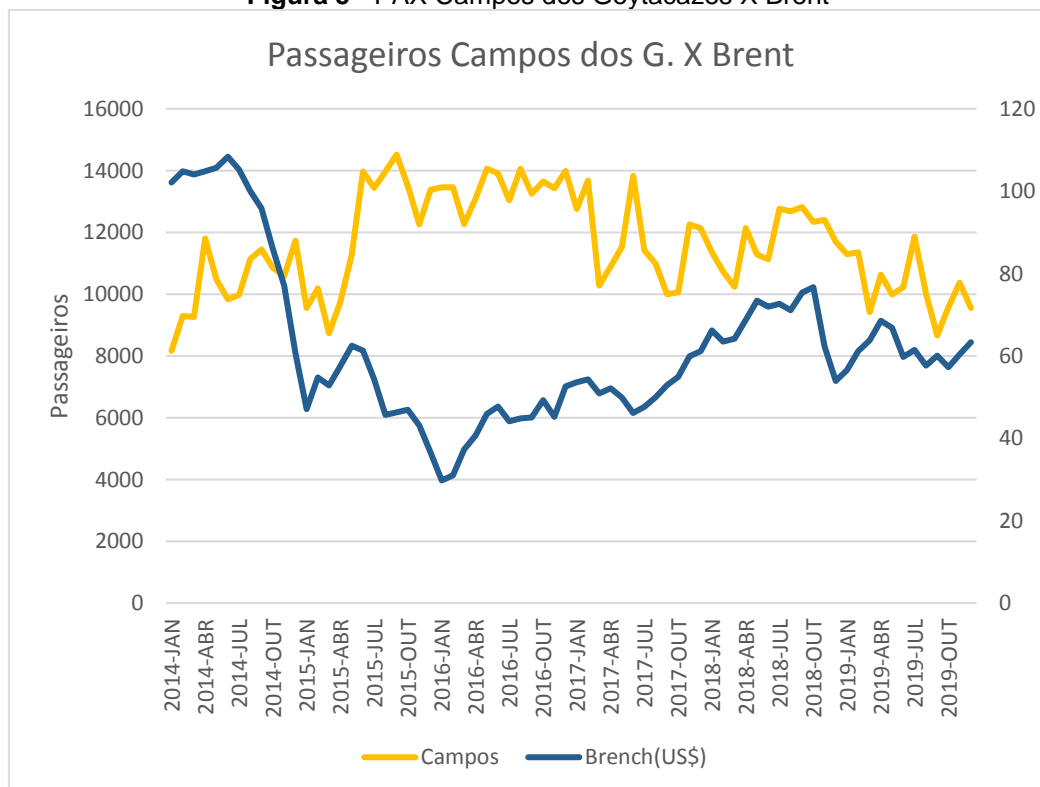
Outrossim, é importante salientar que depois de Farol de São Tomé, Campos dos Goytacazes é a melhor localização, em termos de distância, em relação as embarcações da Bacia de Campos.

Os números de passageiros transportados por mês foram retirados da base de dados do sistema SAP Petrobras e encontram-se na tabela 3 e na Figura 5 as comparações com os preços do petróleo Brent.

Tabela 3 - Passageiros Transportados Campos dos Goytacazes

Campos	2019	2018	2017	2016	2015	2014
JAN.	11.294	11.361	12.757	13.463	9.552	8.163
FEV.	11.355	10.734	13.675	13.461	10.182	9.295
MAR.	9.418	10.233	10.278	12.268	8.740	9.268
ABR.	10.635	12.143	10.896	13.097	9.740	11.805
MAIO	9.993	11.283	11.535	14.059	11.287	10.475
JUN.	10.224	11.126	13.831	13.904	13.965	9.836
JUL.	11.861	12.769	11.420	13.037	13.442	9.976
AGO.	10.023	12.678	10.994	14.064	13.966	11.137
SET.	8.659	12.822	10.002	13.251	14.516	11.451
OUT.	9.570	12.340	10.063	13.648	13.509	10.855
NOV.	10.370	12.401	12.256	13.424	12.255	10.590
DEZ.	9.549	11.703	12.131	13.985	13.376	11.727

Fonte: Elaborado pelo autor, 2020.

Figura 5 - PAX Campos dos Goytacazes X Brent

Fonte: Elaborado pelo autor, 2020.

No ano de 2014, conforme a Figura 5, o número de passageiros transportados saiu da faixa de 8 mil por mês para quase 12 mil. No mesmo período, o preço do petróleo sofreu uma grande desvalorização, saindo de 100 dólares para 85. No ano de 2015, depois de uma baixa de passageiros, em que foram transportados 9 mil, há um grande aumento de demanda, passando a ser movimentados de 13 a 14 mil trabalhadores, valor que se manteve até 2017. O preço do petróleo continuou em queda agressiva até janeiro de 2016, atingindo o valor de 30 dólares o barril. Após isso, houve uma recuperação gradativa no preço do *brench* até 50 dólares. O aumento no número de passageiros nesse período pode ser interpretado como uma busca de eficiência logística, pois alguns voos passaram a ser realizados naquela localidade devido a melhor proximidade com as plataformas.

De 2017 até o começo de 2019, o número de passageiros transportados oscilou na faixa entre 12 mil e 10 mil. Enquanto isso, o preço do petróleo se elevou até a faixa de 80 dólares e depois retornou para 60. Essa significativa queda é resultado da diminuição das atividades exploratórias na Bacia de Campos, bem como do descomissionamento de algumas unidades marítimas.

Por fim, os demais meses de 2019 são caracterizados pelo transporte de 9 mil a 10 mil trabalhadores por mês. O preço do petróleo se manteve na faixa dos 60 dólares o barril. A redução na quantidade de passageiros transportados ocorreu basicamente pelos mesmos motivos explicitados no parágrafo anterior.

5.4 Cabo Frio

O aeroporto internacional de Cabo Frio, de código ICAO SBCB, localizado na cidade de Cabo Frio, encontra-se a aproximadamente 156 km ao norte do Rio de Janeiro, na Região dos Lagos. O aeródromo serve de base para o transporte de passageiros, majoritariamente, que se destinam a embarcações que se localizam na Bacia de Santos.

Este é um aeródromo público federal, atualmente administrado pela Empresa Costa do Sol. SBCB não é dotado de torre de controle, sendo o controle de pátio feito por um colaborador em solo. Existe o serviço de EPTA, estando homologada para operações IFR. Atualmente, encontram-se operando naquele aeroporto duas companhias de táxi aéreo *offshore*: Aeróleo e CHC.

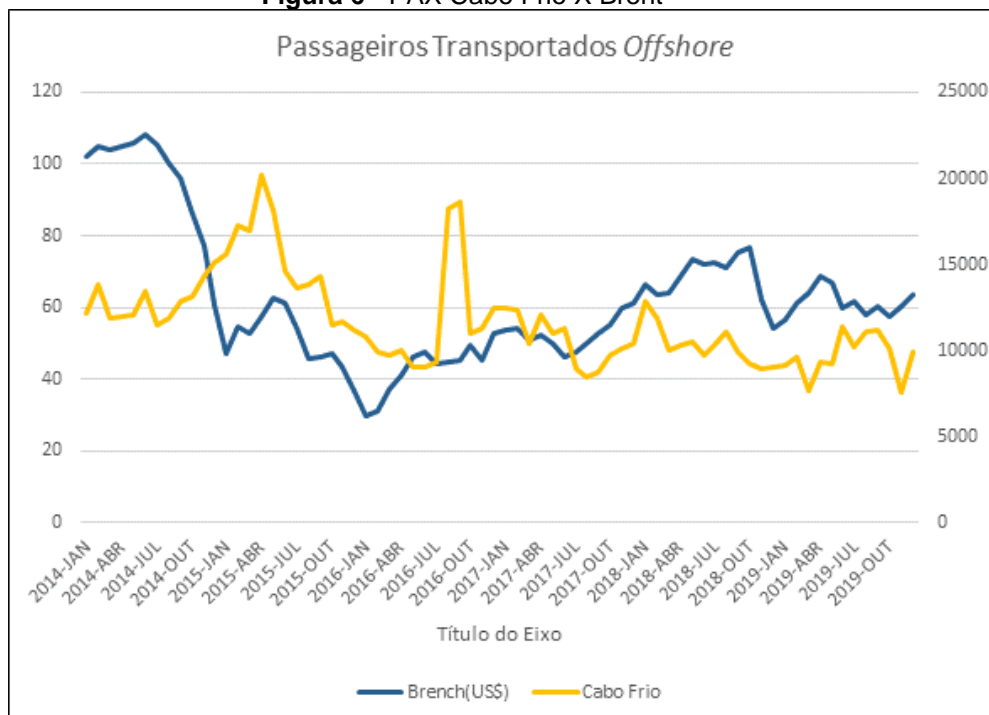
Vale dizer que o aeródromo não possui limitação de hangares e pátio, suportando um possível aumento de demanda sem necessidade de investimentos elevados. Além disso, é importante salientar que os custos de operação neste aeródromo são elevados, em razão dos preços impostos pela operadora e também devido ao acordo de fornecimento único de querosene de aviação que deu exclusividade à fornecedora BP. Tal característica comercial pode fazer com que a Petrobras busque soluções economicamente mais viáveis no médio prazo.

Os números de passageiros transportados por mês foram retirados da base de dados do sistema SAP Petrobras e encontram-se na tabela 4 e na Figura 6 encontram-se as comparações com o petróleo brent.

Tabela 4 - Passageiros Transportados Cabo Frio

Cabo Frio	2019	2018	2017	2016	2015	2014
JAN.	9.093	12.815	12.407	10.765	15.567	12.197
FEV.	9.627	11.843	12.363	9.940	17.210	13.808
MAR.	7.621	9.990	10.356	9.739	16.987	11.859
ABR.	9.352	10.327	12.070	10.006	20.153	11.948
MAIO	9.230	10.454	10.946	9.057	18.139	12.097
JUN.	11.416	9.684	11.326	9.070	14.615	13.394
JUL.	10.244	10.309	8.945	9.334	13.626	11.444
AGO.	11.070	11.055	8.433	18.192	13.804	11.841
SET.	11.141	9.951	8.759	18.628	14.308	12.821
OUT.	10.068	9.257	9.670	10.944	11.494	13.135
NOV.	7.528	8.930	10.150	11.307	11.689	14.267
DEZ.	9.861	9.045	10.404	12.479	11.212	15.051

Fonte: Elaborado pelo autor, 2020.

Figura 6 - PAX Cabo Frio X Brent

Fonte: Elaborado pelo autor, 2020.

De acordo com a Figura 6, no ano de 2014, o número de passageiros transportados saiu da faixa de 12 mil por mês para quase 18 mil no meio de 2015. No mesmo período, o preço do petróleo sofreu uma grande desvalorização, saindo de 100 dólares para 45. O aumento se deu, majoritariamente, devido à maior necessidade de atividade exploratória na região do pré-sal. O ano de 2015 deu início a uma tendência de baixa que foi até outubro de 2016, quando o transporte de passageiros atingiu valores próximos a 11 mil passageiros por mês. O preço do petróleo, no mesmo período, passou por uma mínima de 30 dólares o barril, se recuperando para valores levemente acima de 40 dólares. De 2016 até 2019, mesmo com um pico por dois meses em 2016, houve uma estabilização no número de passageiros transportados entre 11 e 12.5000, apesar da recuperação dos preços do petróleo para valores acima de 60 dólares.

5.5 Jacarepaguá

O aeroporto Roberto Marinho, também denominado Jacarepaguá, de código ICAO SBJR, está localizado na cidade do Rio de Janeiro. O aeródromo serve de

base para o transporte de passageiros, majoritadamente, que se destinam a embarcações que se localizam na Bacia de Santos.

Este é um aeródromo público administrado pela Infraero. SBJR é dotado de torre de controle, conta com o serviço de EPTA, não estando homologado para operações IFR. Atualmente, encontram-se operando neste aeroporto as companhias de táxi aéreo *offshore* Omni, Líder e CHC.

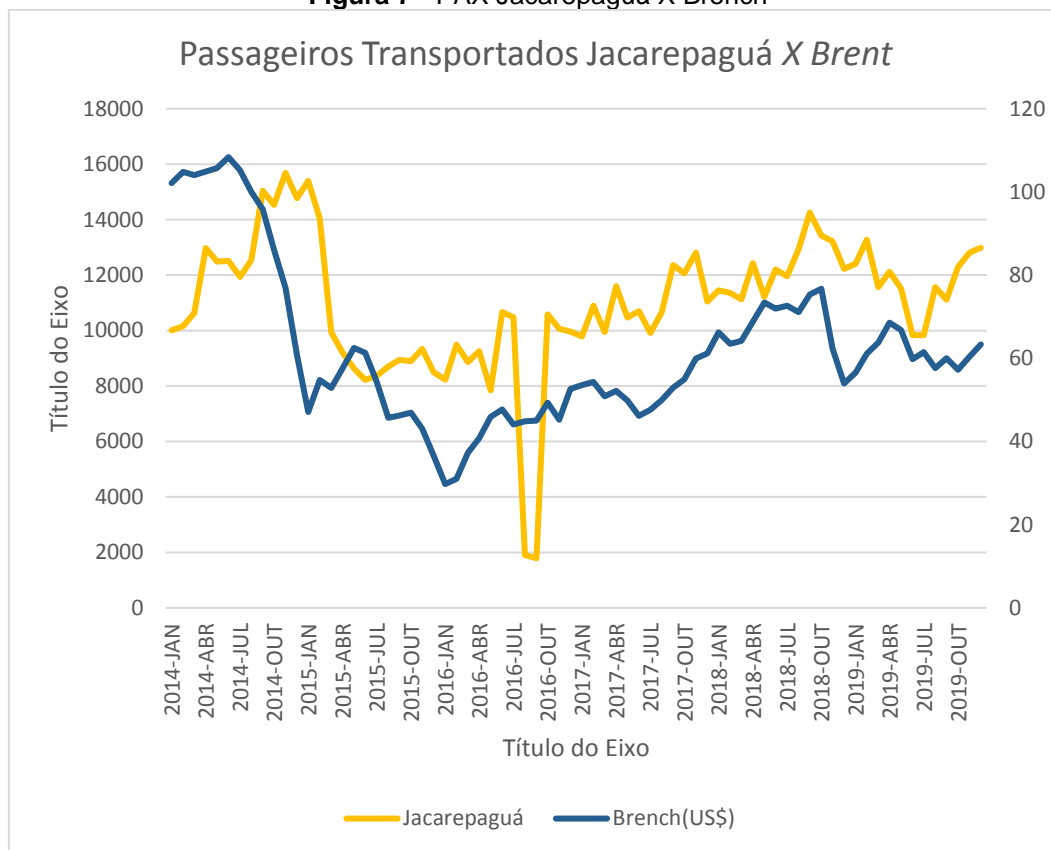
O aeroporto possui pista asfáltica de 900 metros por 30 metros de largura. Vale dizer, ainda, que o aeródromo possui limitação de hangares e pátio e que qualquer tipo de expansão de tal infraestrutura é extremamente difícil, uma vez que se encontra em região de grande densidade urbana.

Os números de passageiros transportados por mês foram retirados da base de dados do sistema SAP Petrobras e encontram-se na tabela 5 e na Figura 7 estão dispostas as comparações com o preço do petróleo brent.

Tabela 5 - Passageiros Transportados Jacarepaguá

Jacarepaguá	2019	2018	2017	2016	2015	2014
JAN.	12.405	11.451	9.792	8.231	15.405	10.011
FEV.	13.276	11.361	10.899	9.491	14.041	10.164
MAR.	11.568	11.124	9.948	8.860	9.939	10.631
ABR.	12.119	12.427	11.600	9.256	9.208	12.975
MAIO	11.520	11.223	10.469	7.850	8.640	12.487
JUN.	9.832	12.206	10.703	10.668	8.224	12.513
JUL.	9.832	11.958	9.916	10.484	8.368	11.935
AGO.	11.561	12.945	10.654	1.900	8.707	12.531
SET.	11.115	14.262	12.367	1.791	8.951	15.042
OUT.	12.291	13.424	12.068	10.580	8.896	14.529
NOV.	12.816	13.223	12.818	10.067	9.337	15.685
DEZ.	12.990	12.229	11.047	9.959	8.498	14.780

Fonte: Elaborado pelo autor, 2020.

Figura 7 - PAX Jacarepaguá X Brench

Fonte: Elaborado pelo autor, 2020.

No ano de 2014, conforme a Figura 7 o número de passageiros transportados saiu da faixa de 10 mil por mês para quase 16 mil. No mesmo período, o preço do petróleo sofreu uma grande desvalorização, saindo de 110 dólares para 48. No ano de 2015, há uma grande queda no número de passageiros transportados, saindo dos 15.600 para cerca de 8 mil. Isso ocorreu, principalmente, devido à diminuição da atividade exploratória, numa tentativa da empresa de conter custos e focar na diminuição do endividamento.

De 2015 até 2018, é possível observar uma contínua tendência de crescimento de demanda no aeroporto de Jacarepaguá, o total de passageiros transportados saiu dos 8 mil e ultrapassou os 14 mil em setembro de 2018. Tal crescimento pode ser explicado pelo aumento de atividade – tanto de produção quanto exploratória – na Bacia de Santos, acompanhando a recuperação no preço do petróleo, que alcançou o valor de 75 dólares por barril. Vale dizer que a maior baixa em 2016 ocorreu devido à realização dos jogos olímpicos na cidade do Rio de Janeiro, onde o aerodromo SBJR permaneceu inoperante por alguns dias.

De 2018 até o final de 2019, a demanda apresentou uma leve queda para 13 mil passageiros por mês, passando por um mínimo de 9.800 em julho de 2019. O preço do petróleo, no mesmo período, se manteve a cerca de 60 dólares o barril.

Resultados e discussões

Para a síntese dos resultados, foi utilizada a mesma sistemática adotada por Karmy e Maldonado (2019), em que cada uma das métricas descritas no item 4.1 é calculada por nível, na estrutura hierárquica, para cada uma das estratégias modeladas. Posteriormente, os resultados são ranqueados, sendo atribuído o número 1 para o melhor resultado (menor valor de erro) e assim sucessivamente. Ao cabo, é calculada a média aritmética simples dos *rankings* entre os níveis hierárquicos.

Por fim, foi elaborado um *ranking* para cada um dos métodos, cuja posição final é a média aritmética obtida entre os *rankings* para cada uma das métricas de erro. O menor valor encontrado indica que a estratégia em questão é a mais acurada em termos globais, bem como o maior valor indica a menos acurada.

Os resultados obtidos em termos de RMSE, para cada uma das estratégias propostas no presente estudo, foram representados na tabela 6.

Tabela 6 - Resultados RMSE
RMSE (Número de Passageiros)

MÉTODO	Total	Level 1	RANK		
			Total	Level 1	Média
Bottom-UP	2316,08	2515,78	7	7	7
Top-Down Historical Proportions	1351,13	2477,29	3	4	3,5
Top-Down Forecast Proportions	1351,13	5263,64	3	8	5,5
OLS	1339,47	2469,38	2	1	1,5
WLS	1361,40	2470,72	5	2	3,5
Structural Scalling	1573,79	2480,64	6	6	6
MinT - Sample Covariance	1338,02	2477,30	1	5	3

MinT-Shrinkage Estimator	1352,60	2473,32	4	3	3,5
--------------------------	---------	---------	---	---	-----

Fonte: Elaborado pelo autor, 2020.

Vale ressaltar que, quando as métricas de erro apontam o mesmo resultado para estratégias diferentes, foi adotado que o *ranking* desses também teria o mesmo valor.

Os resultados encontrados em termos de MAPE foram representados na tabela 7.

Tabela 7 - Resultados MAPE

MAPE (%)					
MÉTODO	Total	Level 1	RANK		
			TOTAL	Level 1	Média
Bottom-UP	2,67	23,49	7	7	7
Top-Down Historical Proportions	1,69	22,51	3	5	4
Top-DownForecast Proportions	1,69	57,74	3	8	5,5
OLS	1,68	22,38	2	4	3
WLS	1,71	22,32	5	3	4
Structural Scalling	1,97	22,65	6	6	6
MinT - Sample Covariance	1,67	22,11	1	1	1
MinT-Shrinkage Estimator	1,70	22,18	4	2	3

Fonte: Elaborado pelo autor, 2020.

Já os resultados em termos de MASE estão consolidados na tabela 8.

Tabela 8 - Resultados MASE

MASE					
MÉTODO	TOTAL	Level 1	RANK		
			TOTAL	Level 1	MÉDIA
Bottom-UP	0,28	0,75	7	7	7
Top-Down Historical Proportions	0,18	0,75	5	6	5,5
Top-Down Forecast Proportions	0,18	1,59	5	8	6,5
OLS	0,18	0,75	2	4	3
WLS	0,18	0,74	4	3	3,5
Structural Scalling	0,21	0,75	6	5	5,5
MinT - Sample Covariance	0,18	0,73	1	1	1
MinT-Shrinkage Estimator	0,18	0,74	3	2	2,5

Fonte: Elaborado pelo autor, 2020.

Por fim os resultados consolidados, obtidos através da média aritmética dos resultados anteriores, são representados na tabela 9. Segundo o MASE alguns casos obtiveram resultados melhores do que o NAIVE, mas outros foram piores.

Tabela 9 - Resultados Consolidados por Técnica

MÉTODO	RMSE	MAPE	MASE	MÉDIA
Bottom-UP	7	7	7	7,0
Top-Down Historical Proportions	3,5	4	5,5	4,3
Top-Down Forecast Proportions	5,5	5,5	6,5	5,8
OLS	1,5	3	3	2,5
WLS	3,5	4	3,5	3,7
Structural Scalling	6	6	5,5	5,8
MinT - Sample Covariance	3	1	1	1,7
MinT-Shrinkage Estimator	3,5	3	2,5	3,0

Fonte: Elaborado pelo autor, 2020.

Analisando os resultados globais, observamos que as melhor métricas de erro foram aferidas através do método de minimização de traço *sample covariance*. Em relação ao mesmo método, entretanto, se valendo do estimador *shrinkage* para estimar a covariância, observou-se resultados significativamente inferiores. Tal conclusão já era esperada e confirma o que foi exposto na seção métodos, em que é afirmado que o método mais indicado para estruturas hierárquicas pequenas é aquele que se vale da própria amostra para estimativa de covariância.

Por outro lado, o pior resultado, em termos globais, ocorreu pela estratégia *bottom-up*. As outras estratégias baseadas em interpolações lineares de um nível hierárquico para os demais (*top-down*) tiveram resultados ligeiramente superiores, entretanto, aquém daqueles obtidos através de métodos de reconciliação e minimização de traço.

Em um nível intermediário de acuracidade, se encontram os métodos de reconciliação, com destaque para o OLS, que obteve o segundo melhor resultado

global. Assim como pontuado na seção de metodologia, a principal desvantagem desse método é não levar em conta as possíveis correlações que existem entre os diferentes níveis hierárquicos. Contudo, em nosso caso concreto, a estrutura hierárquica é feita de apenas dois níveis, o que pode explicar os resultados obtidos.

Por fim, vale destacar que, em todas as métricas de erros utilizadas no presente estudo, os valores encontrados foram muito superiores para o maior nível de agregação. Uma possível explicação para esse fenômeno é que a quantidade de passageiros transportados em vários aeroportos sofreu alterações bruscas, devido a decisões estratégico-comerciais por parte da empresa em questão. Grande parte das alterações agressivas em número de passageiros transportados nos aeroportos é consequência da transferência de aeronaves, que por sua vez impacta apenas marginalmente no número total de passageiros transportados.

6.1 Análise comparativa

Atualmente, a Petrobras utiliza como modelo de previsão de demanda o arnotecimento exponencial, diferente daquele implementado no presente estudo. Nenhuma estratégia ou concepção de séries temporais hierárquicas foi implementada até então. Além disso, a companhia utiliza seus modelos apenas no nível mais agregado, o que distorce as previsões no níveis mais desagregados, bem como não permite a obtenção de resultados coerentes.

Em termos de comparação, foram calculadas as métricas de erro RMSE, MAPE e MASE utilizando o método de previsão de amortecimento exponencial, tendo como *set* de teste os valores referentes a 2019 para averiguar as diferenças de acurácia entre o método corrente na empresa e o aplicado nesse estudo.

Quadro 3 - Resultados Modelo Petrobras

MÉTODO	RMSE	MAPE	MASE
Petrobras	2998,43	3,02	0,43
MinT - Sample Covariance	1338,02	1,67	0,18

Fonte: Elaborado pelo autor, 2020.

Ou seja, para todas as métricas de erro, os valores são significativamente piores do que qualquer um dos métodos de séries temporais hierárquicas, o que já justificaria por si só a utilização deste tipo de estratégia. Entretanto, o maior ganho na adoção da abordagem das séries temporais hierárquicas é justamente contemplar em sua análise os níveis desagregados de composição.

Conclusão

A previsão de demanda de passageiros *offshore* é um assunto de grande interesse por parte das empresas petrolíferas, bem como daquelas que estão diretamente envolvidas nessa atividade. Uma série de decisões estratégicas e operacionais são tomadas a partir de previsões, o que justifica a necessidade de se utilizar as melhores ferramentas possíveis para tal.

A atividade de transporte aéreo da Petrobras se concentra majoritariamente no estado do Rio de Janeiro, sendo realizada através de cinco aeroportos: Farol de São Tomé, Campos dos Goytacazes, Macaé, Cabo Frio e Jacarepaguá. O estado da arte da literatura científica, no que se refere às séries temporais, preconiza que quando uma variável de interesse é a composição de outras pode-se adotar as metodologias de séries temporais hierárquicas a fim de que as previsões aferidas sejam coerentes, ou seja, respeitem as condições de agregação.

Nesse, sentido o presente estudo se valeu das principais metodologias de previsão de séries temporais hierárquicas e avaliou a acurácia de suas previsões doze passos à frente através de três métricas de erro (RMSE, MAPE e MASE), destacando como melhor resultado aquele obtido através da estratégia de minimização de traço *sample covariance*.

Também foi realizada a comparação dos métodos supramencionados com a abordagem corrente na Petrobras, apontando para a inquestionável superioridade das estratégias de séries temporais hierárquicas, cuja adoção traria ganhos técnicos e estratégicos que poderiam evitar erros ocorridos no passado, como o investimento no aeroporto de Itanhaém.

Por fim, como sugestão de pesquisas futuras, pode-se propor a avaliação das mesmas técnicas de séries temporais, valendo-se de diferentes estratégias de previsão base, incluindo desde métodos básicos como ARIMA até métodos mais sofisticados, à exemplo daqueles que utilizam técnicas de aprendizado de máquina (*machine learning*), bem como modelos que se valem de variáveis explicativas.

Referências bibliográficas

ALEKSEEV, K. P. G.; SEIXAS, J. M. A multivariate neural forecasting modeling for air transport - Preprocessed by decomposition: a brazilian application. **Journal of Air Transport Management**, [S. l.], v. 15, p. 212-216, 2009. Disponível em: <https://doi.org/10.1016/j.jairtraman.2008.08.008>. Acesso em 9 abr. 2020.

ANDREONI, A.; POSTORINO, M. N. Time series models to forecast air transport demand: a study about a regional airport. *In*: IFAC Symposium on Control in Transportation Systems, 11., 2006, Holanda. **Proceedings** [...], v. 39, n. 12, p. 101-106, 2006. Disponível em: <https://doi.org/10.3182/20060829-3-NL-2908.00018>. Acesso em: 9 abr. 2020.

ATHANASOPOULOS, G. *et al.* Forecasting with temporal hierarchies. **European Journal of Operational Research**, [S. l.], v. 262, p. 60-74, 2017. Disponível em: <https://doi.org/10.1016/j.ejor.2017.02.046>. Acesso em 10 abr. 2020.

ATHANASOPOULOS, G. *et al.* **Hierarchical forecasting**. Working Paper. Department of Econometrics and Business Statistics. Monash University: Austrália, 2019. Disponível em: <https://www.monash.edu/business/ebs/research/publications/ebs/wp02-2019.pdf>. Acesso em: 10 abr. 2020.

ATHANASOPOULOS, G.; HYNDMAN, R. J. Optimally Reconciling Forecasts in a Hierarchy. **International Journal of Foresight**, [S. l.], p. 42-48, 2014. Disponível em: <https://robjhyndman.com/papers/Foresight-hts-final.pdf>. Acesso em: 10 abr. 2020.

ATHANASOPOULOS, G.; HYNDMAN, R. J.; AHMED, R. A. Hierarchical forecasts for Australian domestic tourism. **International Journal of Forecast**, [S. l.], v. 25, p. 146-166, 2009. Disponível em: <https://doi.org/10.1016/j.ijforecast.2008.07.004>. Acesso em: 10 abr. 2020.

DANTAS, T. M.; OLIVEIRA, F. L. C.; REPOLHO, H. M. V. Air transportation demand forecast through Bagging Holt Winters method. **Journal of Air Transport Management**, [S. l.], v. 59, p. 116-123, 2017. Disponível em: <https://doi.org/10.1016/j.jairtraman.2016.12.006>. Acesso em: 9 abr. 2020.

GELHAUSEN, M. C.; BERSTER, P.; WILKEN, D. A new direct demand model of long-term forecasting air passengers and air transport movements at German airports. **Journal of Air Transport Management**, [S. l.], v. 71, p. 140-152, 2018. Disponível em: <https://doi.org/10.1016/j.jairtraman.2018.04.001>. Acesso em: 9 abr. 2020.

GROSS, C. W.; SOHL, J. E. Disaggregation methods to expedite product line forecasting. **Journal of Forecasting**, [S. l.], v. 9, n. 3, p. 233-254, 1990. Disponível em: <https://doi.org/10.1002/for.3980090304>. Acesso em: 10 abr. 2020.

HALSKAU, O. Offshore helicopter routing in a hub and spoke fashion: minimizing expected number of fatalities. *In: INTERNATIONAL CONFERENCE ON INFORMATION TECHNOLOGY AND QUANTITATIVE MANAGEMENT (ITQM)*, 2., 2014, Molde. **Anais [...]**. Molde: Molde University College, 2014.

HERMETO, N. S. S.; FERREIRA FILHO, V. J. M.; BAHIANSE, L. Logistics net planning for offshore air transport of oil rig crews. **Computers & Industrial Engineering**, [S. l.], v. 75, p. 41-54, 2014. Disponível em: <https://doi.org/10.1016/j.cie.2014.05.021>. Acesso em: 9 abr. 2020.

HUBER, J.; GOSSMANN, A.; STUCKENSCHMIDT, H. Cluster-based hierarchical demand forecasting for perishable goods. **Expert Systems with Applications**, [S. l.], v. 76, p. 140-151, 2017. Disponível em: <https://doi.org/10.1016/j.eswa.2017.01.022>. Acesso em: 10 abr. 2020.

HYNDMAN, R. J.; KOEHLER, A., B. Another look at measures of forecast accuracy. **International Journal of Forecast**, [S. l.], v. 22, p. 679-688, 2006. Disponível em: <https://doi.org/10.1016/j.ijforecast.2006.03.001>. Acesso em 14 abr. 2020.

HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting: principles and practice**. 2. ed. Melbourne: Austrália, 2018.

HYNDMAN, R. J. *et al.* Optimal combination forecasts for hierarchical time series. **Computational and Statistics data Analysis**, [S. l.], v. 55, p. 2579-2589, 2011. Disponível em: <https://doi.org/10.1016/j.csda.2011.03.006>. Acesso em: 10 abr. 2020.

HYNDMAN, R. J. *et al.* **Package ‘hts’: hierarchical and grouped time series: version 5.1.5**. 2018. Disponível em: <https://cran.r-project.org/web/packages/hts/hts.pdf>. Acesso em: 10 abr. 2020.

JIN, F. *et al.* Forecasting air passenger demand with a new hybrid ensemble approach. **Journal of Air Transport Management**, [S. l.], v. 83, 2020. Disponível em: <https://doi.org/10.1016/j.jairtraman.2019.101744>. Acesso em: 9 abr. 2020.

KARMY, J. P.; MALDONADO, S. Hierarchical time series forecasting via Support Vector Regression in the European Travel Retail Industry. **Expert Systems with Applications**, [S. l.], v. 137, p. 59-73, 2019. Disponível em: <https://doi.org/10.1016/j.eswa.2019.06.060>. Acesso em: 10 abr. 2020.

LIU, Z. *et al.* Missing value estimation for hierarchical time series: a study of hierarchical web traffic. *In: International Conference on Data Mining*, 2015, Atlantic City. **Anais [...]**. Atlantic City: IEEE, 2016. p. 895-900. Disponível em: http://people.cs.pitt.edu/~milos/research/ICDM_2015_hierarchical.pdf. Disponível em: 10 abr. 2020.

MARAZZO, M.; SCHERRE, R.; FERNANDES, E. Air transport demand and economic growth in Brazil: a time series analysis. **Transportation Research**

Part E: Logistics and Transportation Review, [S. l.], v. 46, n. 2, p. 261-269, 2010. Disponível em: <https://doi.org/10.1016/j.tre.2009.08.008>. Acesso em: 9 abr. 2020.

RODRIGUES, L. A.; SAUER, I. L. Exploratory assessment of the economic gains of a pre-salt oil field in Brazil. **Energy Policy**, [S. l.], v. 87, p. 486-495, 2015. Disponível em: <https://doi.org/10.1016/j.enpol.2015.09.03>. Acesso em: 9 abr. 2020.

SCHÄFER, J.; STRIMMER, K. A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. **Statistical Applications in Genetics and Molecular Biology**, [S. l.], v. 4, n. 1, p. 1-30, 2005. Disponível em: <http://www.strimmerlab.org/publications/journals/shrinkcov2005.pdf>. Acesso em: 10 abr. 2020.

SUN, S. *et al.* Nonlinear vector auto-regression neural network for forecasting air passenger flow. **Journal of Air Transport Management**, [S. l.], v. 78, p. 54-62, 2019. Disponível em: <https://doi.org/10.1016/j.jairtraman.2019.04.005>. Acesso em: 9 abr. 2020.

WICKRAMASURIYA, S. L.; ATHANASOPOULOS, G.; HYNDMAN, R. J. Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization. **Journal of the American Statistical Association**, [S. l.], v. 1459, p. 1-45, 2018. Disponível em: <https://doi.org/10.1080/01621459.2018.1448825>. Acesso em: 10 abr. 2020.

XIAO, Y. *et al.* A neuro-fuzzy combination model based on singular spectrum analysis for air transport demand forecasting. **Journal of Air Transport Management**, [S. l.], v. 39, p. 1-11, 2014. Disponível em: <https://doi.org/10.1016/j.jairtraman.2014.03.004>. Acesso em: 9 abr. 2020.