



**Thaisa de Freitas**

**Modelos e Aplicações para Séries Temporais Hierárquicas:  
abordagens de Reconciliação Ótima e Proporções de  
Previsão**

**Dissertação de Mestrado (Opção acadêmica)**

Dissertação apresentada como requisito parcial para obtenção do título 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  
Fevereiro de 2016



**Thaisa de Freitas**

**Modelos e Aplicações para Séries Temporais Hierárquicas:  
abordagens de Reconciliação Ótima e Proporções de Previsão**

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

**Prof. Fernando Luiz Cyrino Oliveira**

Orientador

Departamento de Engenharia Industrial - PUC-Rio

**Prof. Eugenio Kahn Epprecht**

Departamento de Engenharia Industrial - PUC-Rio

**Prof. Reinaldo Castro Souza**

Departamento de Engenharia Elétrica - PUC-Rio

**Prof. Marcio da Silveira Carvalho**

Coordenador Setorial do Centro Técnico Científico - PUC-Rio

Rio de Janeiro, 29 de Fevereiro de 2016

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

### **Thaiza de Freitas**

Graduada em Engenharia de Produção na Universidade Federal Fluminense (UFF) em 2013.

#### Ficha Catalográfica

Freitas, Thaiza de

Modelos e aplicações para séries temporais hierárquicas: abordagens de reconciliação ótima e proporções de previsão / Thaiza de Freitas ; orientador: Fernando Luiz Cyrino Oliveira. – 2016.

91 f. : il. color. ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial, 2016.

Inclui bibliografia

1. Engenharia Industrial – Teses. 2. Séries Temporais Hierárquicas. 3. Previsão Hierárquica. 4. Reconciliação ótima. 5. Top-Down baseado nas Proporções da Previsão. I. Oliveira, Fernando Luiz Cyrino. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Industrial. III. Título.

CDD: 658.5

## Agradecimentos

Primeiramente agradeço a Deus por guiar os meus caminhos.

À minha mãe Edna que sempre me incentivou a buscar meus objetivos. Obrigada por sempre apoiar minhas decisões e sempre ter a palavra certa na hora exata em que preciso.

Ao meu noivo Vladimir pelo apoio, paciência e incentivo nos momentos mais difíceis.

Ao meu professor e orientador Fernando Cyrino, que mesmo nas pequenas atitudes sempre se mostrou um exemplo de pessoa e profissional, sempre atencioso e paciente. Obrigada por todos os ensinamentos transmitidos.

Aos professores Eugenio Epprecht e Reinaldo Souza por terem aceitado o convite para fazer parte da minha banca.

A todos os professores e funcionários do Departamento de Engenharia Industrial da PUC-Rio.

Ao CNPq pelo auxílio financeiro que possibilitou meu desenvolvimento acadêmico.

A todos os amigos e colegas que conheci ao longo desses dois anos na PUC - Rio e que de alguma maneira me ajudaram neste trajeto.

## Resumo

Freitas, Thaisa; Oliveira, Fernando Luiz Cyrino. **Modelos e Aplicações para Séries Temporais Hierárquicas: abordagens de Reconciliação Ótima e Proporções de Previsão**. Rio de Janeiro, 2016. 91p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

Séries Temporais que podem ser organizadas em níveis de acordo com, por exemplo, o tipo de produto, região geográfica, classe de consumo, dentre outros, são chamadas de Séries Temporais Hierárquicas (ou agrupadas, quando possuem mais de uma variável de agregação). Informações referentes à previsão destas séries são fundamentais para a tomada de decisão seja no nível gerencial ou operacional de todo tipo de negócio. Para atender a essas informações, são utilizadas técnicas de previsão hierárquica, que têm como foco reduzir os custos e melhorar a acurácia da previsão. O objetivo deste trabalho é estudar abordagens para agregar/desagregar previsões feitas para Séries Temporais Hierárquicas ou Agrupadas. Como resultado do trabalho destaca-se a apresentação das abordagens que representam o estado da arte em previsão hierárquica: Reconciliação Ótima (também chamada de Combinação Ótima) e Top-Down baseada na Proporção das Previsões. Ainda referente aos resultados destaca-se a análise das diversas técnicas de previsão hierárquica encontradas na literatura aplicadas a duas séries clássicas do contexto brasileiro: a série agrupada de consumo de energia elétrica agregada por região do país e classe de consumo, e a série hierárquica de demanda de transporte aéreo representada pela variável RPK (*Revenue Passenger Kilometers*). O desempenho preditivo das abordagens foi avaliado com base na métrica MAPE, e o teste de Diebold-Mariano foi aplicado para verificar se a diferença no desempenho das abordagens novas e tradicionais é significativa.

## Palavras-chave

Séries Temporais Hierárquicas; Previsão Hierárquica; Reconciliação ótima; Top-Down baseado nas Proporções da Previsão.

## Abstract

Freitas, Thaisa; Oliveira, Fernando Luiz Cyrino (Advisor). **Models and Applications to Hierarchical Time Series: approaches of Reconciliation Optimal and Forecast Proportions**. Rio de Janeiro, 2016. 91p. MSc. Dissertation – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

Time Series which can be arranged in levels according to, for example, the type of product, geography, consumption class, among others, are called Hierarchical Time Series (or grouped, if they have more than one aggregation variable). Information relating these series prediction is fundamental for decision-making at the management or operational level of all types of business. To meet these information, hierarchical forecasting techniques are used, which are focused on reducing costs and improving the accuracy of prediction. The objective of this work is to study approaches to aggregate / disaggregate predictions for Hierarchical or Grouped Time Series. As a result of the work there is the presentation of the approaches that represent the state of the art hierarchical forecast: Optimal Reconciliation approach (also called the Optimal Combination) and Top-Down Forecast Proportions approach. Still on the results highlight the analysis of the various hierarchical forecasting techniques found in the literature applied to two classic series of the Brazilian context: a grouped series of electricity consumption aggregated by region of the country and consumer class, and the hierarchical series air transport demand represented by the variable RPK (Revenue Passenger Kilometers). The predictive performance of the approaches was evaluated based on the metric MAPE and the Diebold-Mariano test was used to verify that the difference in performance of new and traditional approaches is significant.

## Keywords

Hierarchical Time Series; Hierarchical Forecast; Optimal Reconciliation; Optimal Combination; Top-Down Forecast Proportions.

## Sumário

1	Introdução	13
1.1.	Contextualização da Pesquisa	13
1.2.	Motivação	14
1.3.	Objetivo	15
1.4.	Classificação da pesquisa	15
1.5.	Organização do trabalho	16
2	Revisão da literatura	17
2.1.	Metodologia	17
2.1.1.	Análise quantitativa dos resultados da revisão da literatura	19
2.1.2.	Síntese da literatura	22
2.1.3.	Análise descritiva dos resultados da revisão da literatura	26
2.1.4.	Referências adicionais	38
2.2.	Métodos de Previsão	39
2.2.1.	Amortecimento Exponencial	39
2.2.2.	Modelos ARIMA	42
2.3.	Métricas de desempenho	43
2.3.1.	MAPE	43
2.3.2.	Teste de Diebold-Mariano	43
3	Detalhamento matemático das abordagens de previsão hierárquica	45
3.1.	Notação	45
3.2.	Abordagem <i>Bottom-up</i>	48
3.3.	Abordagem <i>Top-Down</i>	49
3.4.	Abordagem <i>Top Down</i> baseada na proporção das previsões	51
3.5.	Abordagem de Combinação ótima	52
3.5.1.	Inversão da Matriz (S'S) para o caso de séries hierárquicas	55
3.5.2.	Inversão da Matriz (S'S) para o caso de séries agrupadas	57
4	Resultados	59

4.1. Análise das bases de dados	59
4.1.1. Séries de consumo de energia	60
4.1.2. Séries de demanda de transporte aéreo	62
4.2. Aplicação das abordagens de previsão hierárquica e Previsões	63
4.2.1. Séries de energia	64
4.2.2. Séries RPK	71
5 Conclusões	75
6 Referências Bibliográficas	77
Apêndice A	83
Apêndice B	84

## Lista de figuras

Figura 1 – Total de artigos por ano .....	21
Figura 2 – Estrutura Hierárquica .....	45
Figura 3 – Exemplo de séries agrupada.....	46
Figura 4 - Exemplos de Hierarquias baseadas em 2 variáveis de agregação .....	47
Figura 5 – Exemplo do cálculo das proporções.....	49
Figura 6 – Árvore hierárquica .....	56
Figura 7 – Consumo Total de energia elétrica no Brasil .....	60
Figura 8 – Consumo de energia elétrica por Classe de Consumo .....	61
Figura 9 – RPK doméstico e internacional.....	62
Figura 10 - Estruturas das séries de energia .....	64
Figura 11 – Previsão para o Total de consumo de energia elétrica .....	70
Figura 12 - Hierarquia da série RPK .....	71
Figura 13 – Previsões 12 passos-à-frente para as séries de demanda de transporte de passageiros em voos domésticos e internacionais .....	74

## Lista de tabelas

Tabela 1 – Total de publicações por revista .....	21
Tabela 2 – Principais autores .....	22
Tabela 3 – Classificação do método de amortecimento exponencial .....	39
Tabela 4 – Espaços de estados dos modelos de amortecimento exponencial .....	40
Tabela 5 – MAPE <i>out-of-sample</i> da previsão com amortecimento exponencial das séries de energia .....	66
Tabela 6 – MAPE <i>out-of-sample</i> da previsão com modelos ARIMA das séries de energia.....	67
Tabela 7 – Teste de Diebold-Mariano para a hierarquia 1 .....	69
Tabela 8 – Teste de Diebold-Mariano para a hierarquia 2 .....	69
Tabela 9 – MAPE <i>out-of-sample</i> da previsão com amortecimento exponencial das séries RPK.....	72
Tabela 10 - MAPE <i>out-of-sample</i> da previsão com modelos ARIMA das séries RPK .....	73
Tabela 11 – Estatísticas descritivas das 20 séries de energia do nível mais baixo da hierarquia .....	84
Tabela 12– Previsões 12 passos-à-frente para a Hierarquia 1 .....	86
Tabela 13 – Previsões 12 passos-à-frente para a Hierarquia 2 .....	88
Tabela 14 – Estatísticas Descritivas das 2 séries de RPK do nível mais baixo da hierarquia .....	90
Tabela 15– Previsões 12 passos-à-frente para a série RPK.....	91

## Lista de quadros

Quadro 1 – Total de artigos encontrados com as palavras-chave.....	18
Quadro 2 – Autores por ano, revista e número de citações .....	20
Quadro 3 - Descrição dos artigos .....	27
Quadro 4 - Resumo dos artigos que comparam abordagens de previsão hierárquica .....	36
Quadro 5– Matriz de pesos para a hierarquia da Figura 2.....	53
Quadro 6 – Proporções de consumo de energia elétrica no país .....	61

*"Os únicos limites das nossas realizações de amanhã são as nossas dúvidas  
e hesitações de hoje."*

Franklin Roosevelt

# 1 Introdução

## 1.1. Contextualização da Pesquisa

Uma Série Temporal é uma sequência de valores observados ao longo do tempo, em intervalos iguais (Morettin & Toloi, 2006). Um dos principais objetivos do estudo de Séries Temporais é criar modelos que demonstrem o comportamento do fenômeno estudado e, a partir daí, gerar previsões. Modelos de previsão são aplicados em diversas áreas da Engenharia, Economia, Ciências Médicas, dentre outros, servindo como base para o planejamento, permitindo avaliar antecipadamente a demanda, projetar capacidade e necessidade de recursos, além de outras atividades.

Em todo tipo de negócio são necessárias informações diferenciadas de acordo com o setor e nível da tomada de decisão, por exemplo, no nível estratégico são necessárias informações agregadas de previsão de vendas que serão utilizadas pelo setor financeiro para um planejamento orçamentário anual, enquanto que no nível operacional são necessárias informações desagregadas, que serão utilizadas para o planejamento da produção mensal.

Existem diversos tipos de informações referentes ao negócio que podem ser coletadas ao longo do tempo, em intervalos iguais, e que podem ser organizadas em uma estrutura hierárquica, agregando ou desagregando os dados com base, por exemplo, em uma família de produtos ou serviços, distribuição geográfica, dentre outros. Essas observações estruturadas são chamadas de Séries Temporais hierárquicas ou agrupadas.

As séries agrupadas se diferenciam das séries hierárquicas em relação ao total de variáveis de agrupamento. Uma série hierárquica é estruturada com base em uma única variável, ou seja, só há um modo de organizar a hierarquia; já nas séries agrupadas, existem duas ou mais variáveis de agrupamento, assim a estrutura pode ser organizada de formas diferentes (Hyndman et al., 2014a). Um exemplo de Série Temporal agrupada é o consumo de energia elétrica agregado

por regiões do país e por classe de consumo (residencial, industrial, comercial). Neste caso há duas variáveis de agregação, sendo possível escolher qual variável de agregação será aplicada para agrupar os dados do primeiro nível, conseqüentemente, pode-se organizar a hierarquia de dois modos diferentes (agregando os dados primeiramente por região, ou por classe de consumo). Já um exemplo de série hierárquica seria a demanda de um produto subdividida em famílias de produtos e posteriormente em produtos individuais. Neste caso só existe uma maneira de organizar a estrutura da hierarquia, não faz sentido apresentar os produtos individuais no nível intermediário e os produtos agrupados por família no nível mais baixo da hierarquia.

Previsões geradas de maneira independente para cada nível da série hierárquica devem ser evitadas, pois têm como consequência que a previsão em um nível superior não é igual a soma das previsões do nível inferior, e o mesmo pode ser percebido no sentido contrário. Esse problema é solucionado com o uso da previsão hierárquica (*Hierarchical Forecasting* – HF), na qual a previsão é feita de maneira centralizada, garantindo assim que as previsões estejam de acordo com a hierarquia (Fliedner, 1999). Ao longo do capítulo 2 e 3 serão apresentadas diversas abordagens de previsão hierárquica encontradas na literatura e as principais vantagens e desvantagens de cada uma.

As principais vantagens em utilizar a previsão hierárquica são a capacidade de atender às necessidades de informações nos diversos níveis do negócio e também a melhora na acurácia da previsão (Fliedner, 2001; Kourentzes et al., 2014).

## **1.2. Motivação**

O desenvolvimento desta dissertação foi motivado pela gama de possibilidades de aplicação de previsões de Séries Temporais hierárquicas/agrupadas no contexto da Engenharia de Produção e também pela leitura do *working paper* de Hyndman et al. (2014a), no qual foi apresentado um algoritmo de uma nova abordagem para realizar a combinação entre previsões de diversos níveis de uma hierarquia. Essa nova abordagem será detalhada nos

capítulos 2 e 3. O trabalho dos autores Hyndman et al. (2014a) foi publicado recentemente e pode ser encontrado em Hyndman et al. (2016).

### **1.3. Objetivo**

O objetivo geral deste trabalho é realizar um estudo sobre previsão hierárquica, ou seja, sobre abordagens para agregar previsões feitas para Séries Temporais hierárquicas/agrupadas, de modo a encontrar as técnicas mais populares e o estado da arte, tendo como ênfase apontar os pontos fortes e fracos de cada abordagem.

Ainda tem-se como objetivo realizar um estudo empírico aplicando as diversas técnicas encontradas na literatura, utilizando previsão de curto prazo, em duas séries com comportamentos distintos, de modo que seja possível comparar o desempenho das diversas abordagens e verificar qual técnica gerou os melhores resultados, com base em métricas amplamente conhecidas e utilizadas na área de Séries Temporais. Por fim, tem-se como objetivo também apresentar previsões de curto prazo para ambas as séries.

É válido ressaltar que o foco da pesquisa bibliográfica é encontrar técnicas para agregar/desagregar previsões de uma série que possui comportamento hierárquico, ou seja, não faz parte do escopo da pesquisa realizar um levantamento sobre o estado da arte de métodos de previsão de Séries Temporais. Para realizar a previsão serão utilizados métodos tradicionais vastamente discutidos na literatura, como Amortecimento Exponencial e modelos de *Box & Jenkins*.

As aplicações serão implementadas no *software* R (R Core Team, 2013), com linguagem de programação orientada a objetos, e que é livre, o que garante a viabilidade de desenvolvimento da pesquisa.

### **1.4. Classificação da pesquisa**

Quanto aos fins a pesquisa pode ser classificada como descritiva e aplicada.

A pesquisa pode ser classificada como descritiva, pois um dos objetivos da pesquisa é analisar e conseqüentemente descrever diversos métodos utilizados

para modelagem e previsão de Séries Temporais hierárquicas. Além disso, as Séries Temporais serão analisadas com base em diversas características que a descrevem: como tendência, autocorrelação, dentre outros.

A pesquisa também pode ser classificada como aplicada, pois diversos procedimentos de modelagem estocástica serão aplicados em Séries Temporais reais, de modo a comparar os resultados dos mesmos com base em métricas de avaliação do desempenho preditivo dos modelos.

Quanto aos meios de investigação esta dissertação pode ser classificada como bibliográfica e documental. Inicialmente serão realizados estudos para entender melhor as características das diversas abordagens de previsão hierárquica encontradas na literatura, e conseqüentemente a pesquisa pode ser classificada como bibliográfica. Por fim a pesquisa pode ser classificada também como documental, pois serão utilizados diversos dados disponíveis para o público em geral, no caso representado em forma de Séries Temporais, divulgados por diversas entidades pública e/ou privadas para aplicar as técnicas encontradas.

Finalmente, em relação ao enfoque a pesquisa, esta é classificada como quantitativa.

## **1.5. Organização do trabalho**

Essa dissertação está dividida em 5 capítulos, iniciando por esta Introdução geral. No segundo capítulo serão apresentados uma revisão sistemática da literatura sobre previsão hierárquica. Já no terceiro capítulo do trabalho serão apresentados os detalhamentos matemáticos de cada uma das abordagens de previsão hierárquica encontradas e também algumas vantagens e desvantagens de cada uma. No quarto capítulo serão apresentados uma descrição das séries utilizadas para aplicação do estudo e também uma discussão dos resultados encontrados utilizando o *software* R (R Core Team, 2013). O código utilizado para aplicação constará no Apêndice A. Por fim, no quinto capítulo serão apresentadas as conclusões, discussões e propostas de novas pesquisas.

## 2 Revisão da literatura

Neste capítulo será apresentada uma revisão sistemática da literatura sobre previsão hierárquica, constando a metodologia utilizada para a execução da pesquisa, uma análise descritiva dos artigos encontrados e uma síntese com os principais resultados encontrados na literatura.

### 2.1. Metodologia

Primeiramente foi realizada uma revisão sistemática da literatura sobre previsão hierárquica, seguindo uma metodologia de seis passos que foi adaptada de Thomé et al. (2012). Os seis passos utilizados para seleção dos artigos foram: (i) seleção da base de dados, (ii) identificação das palavras-chave para a pesquisa, (iii) leitura dos resumos dos artigos encontrados, (iv) criação de critérios de exclusão, (v) leitura completa dos artigos selecionados, (vi) inserção de outras referências consideradas relevantes. Foram preparadas duas tabelas para análise dos artigos, uma para selecionar os artigos com base nos critérios de seleção e encontrar o total de artigos excluindo as duplicações das bases de dados, e outra apresentando uma análise descritiva dos artigos selecionados.

Duas bases de dados foram selecionadas para a realização da pesquisa: a *Scopus* e *Web of Science* (WoS), estas bases são vastamente conhecidas por possibilitarem a pesquisa de artigos das principais revistas de diversas áreas de conhecimento científico. Estas bases também são utilizadas em outros trabalhos da área de previsão, como se pode observar em Weron (2014). A consulta às bases de dados foi realizada em abril de 2015 e dos resultados retornados foram filtrados somente em artigos e *reviews* publicados em inglês. Já com o intuito de direcionar a pesquisa, as palavras-chave foram pesquisadas somente dentre o título, resumo e palavras-chave do artigo.

O tema referente à previsão de Séries Temporais hierárquicas não é novo e devido a isto muitos trabalhos são encontrados na literatura, o que dificultou a

definição das palavras-chave, que deveriam ser abrangentes o suficiente para que artigos relevantes fossem encontrados, mas também que limitassem o número total de resultados de maneira a não inviabilizar a pesquisa. Assim, três combinações de palavras-chave foram testadas em ambas as bases, como mostra o Quadro 1, com as combinações, o total de artigos encontrados em cada base e o total de artigos encontrados excluindo os artigos repetidos, encontrados em ambas as bases:

**Quadro 1 – Total de artigos encontrados com as palavras-chave**

	Combinação de Palavras-Chave	Total de artigos na WoS	Total de artigos na Scopus	Total de artigos excluindo repetições
(i)	"Hierarchical Forecast*" OR "Hierarchical time series" OR (forecast* AND ("Bottom-up" OR "Top-Down" OR "middle-out"))	229	374	446
(ii)	"Hierarchical Forecast*" OR "Hierarchical time series" OR (forecast* AND (("Bottom-up" OR "Top-Down" OR "middle-out") AND (combin* OR aggregat*)))	88	120	137
(iii)	"Hierarchical Forecast*" OR "Hierarchical time series" OR "Bottom-up forecast*" OR "Top-Down forecast*" OR "middle-out forecast"	35	40	47

**Fonte:** a autora (2016)

As abordagens “*Bottom-up*”, “*Top-Down*” e “*middle-out*” são tradicionalmente conhecidas na área de previsão hierárquica e por isso tais palavras-chave foram consideradas relevantes para a pesquisa e constam nas três combinações. Já observando o número total de resultados, conclui-se que a combinação (i) gerou resultados muito abrangentes e foi descartada, visto que a análise de 446 artigos seria demasiadamente demorada e o presente trabalho não tem como único objetivo realizar uma revisão sistemática da literatura.

Já na combinação (ii) buscou-se restringir um pouco mais os resultados adicionando as palavras *combin\** ou *aggregat\** com o objetivo de focar em artigos que falassem sobre agregação ou combinação de previsões em uma hierarquia, entretanto após uma análise em relação à relevância dos artigos foi possível perceber que ainda eram exibidos diversos artigos que não tinham relação direta com o tema, nos quais as palavras *Top-Down* e *Bottom-up* não estavam ligadas a técnicas de previsão de Séries Temporais.

Finalmente chegou-se a combinação (iii), na qual foi incluída a palavra *forecast* junto do nome das abordagens tradicionais já conhecidas para eliminar ao máximo os artigos que não fossem relacionados ao tema.

Já na terceira e quarta etapa da metodologia adotada, após a leitura dos 47 resumos encontrados, foram criados os dois critérios de exclusão: (a) O artigo não trata sobre previsão de Séries Temporais, (b) O artigo não trata sobre abordagens de previsão hierárquica. Com base nestes critérios de exclusão foram selecionados 24 artigos para a leitura completa durante a quinta etapa da metodologia, entretanto não foi possível acessar 2 artigos.

Os resultados da quinta etapa estão apresentados nos subitens 2.1.2. , 2.1.3. e 2.1.4. A sexta etapa da metodologia será apresentada no subitem 2.1.5.

### **2.1.1.**

#### **Análise quantitativa dos resultados da revisão da literatura**

Com base nos 22 artigos selecionados para a leitura completa, inicialmente foi realizada uma análise quantitativa dos mesmos, de modo a antever alguns resultados. O Quadro 2 abaixo mostra todos os autores, ano de publicação, revista e número de citações de todos os artigos selecionados.

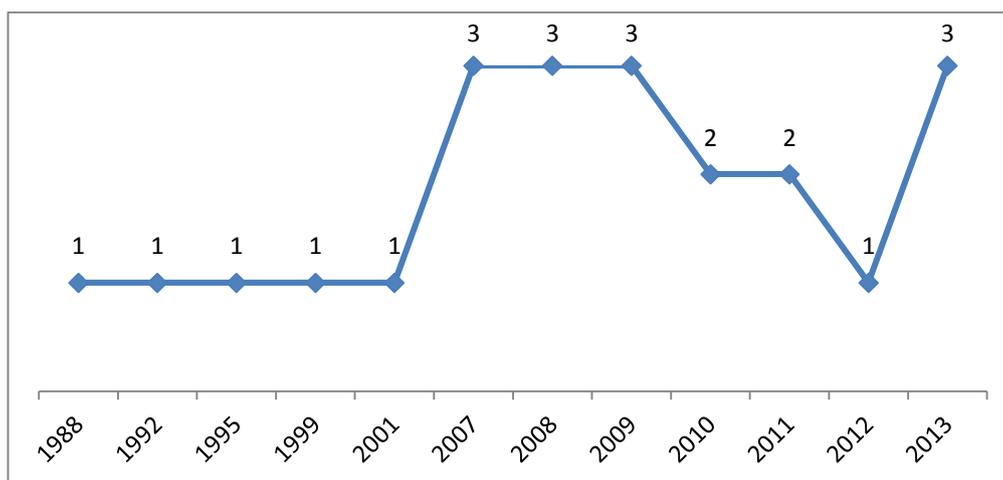
É possível notar, no Quadro 2, que a maioria dos artigos selecionados para leitura completa possui um número significativo de citações, o que corrobora para a validação da combinação de palavras-chaves utilizada na segunda etapa da metodologia de revisão, pois a mesma exhibe resultados relevantes.

**Quadro 2 – Autores por ano, revista e número de citações**

<b>Autor(es)</b>	<b>Ano</b>	<b>Revista</b>	<b>Nº de Citações</b>
Schwarzkopf, A.B.; Tersine, R.J.; Morris, J.S.	1988	International Journal Of Production Research	18
Dangerfield, B.J.; Morris, J.S.	1992	International Journal Of Forecasting	23
Fliedner, E.B.; Lawrence, B.	1995	Journal Of Operations Management	15
Fliedner, G.	1999	Computers And Operations Research	19
Fliedner, G.	2001	Industrial Management And Data Systems	27
Wanke, P.; Saliby E.	2007	Pesquisa Operacional	2
Widiarta, H.; Viswanathan, S.; Piplani, R.	2007	Naval Research Logistics	9
Sohn, S.Y.; Lim, M.	2007	European Journal Of Operational Research	10
Lo, S.-L.; Wang, F.-K.; Lin, J.T.	2008	Journal Of Forecasting	3
Viswanathan, S.; Widiarta, H.; Piplani, R.	2008	IMA Journal Of Management Mathematics	7
Widiarta, H.; Viswanathan, S.; Piplani, R.	2008	Ima Journal Of Management Mathematics	5
Chen, H.; Boylan, J.E.	2009	Advances In Business And Management Forecasting	2
Widiarta, H.; Viswanathan, S.; Piplani, R.	2009	International Journal Of Production Economics	12
Athanasopoulos, G.; Ahmed, R.A.; Hyndman, R.J.	2009	International Journal Of Forecasting	16
Chen, A.; Blue, J.	2010	International Journal Of Production Economics	9
Capistran, C.; Constandse, C.; Ramos-Francia, M.	2010	Economic Modelling	2
Hyndman, R.J.; Ahmed, R.A.; Athanasopoulos, G.; Shang, H.L.	2011	Computational Statistics And Data Analysis	14
Williams, B.D.; Waller, M.A.	2011	Journal Of Business Logistics	4
Moon, S.; Hicks, C.; Simpson, A.	2012	International Journal Of Production Economics	5
Lenort, R.; Besta, P.	2013	Fibres And Textiles In Eastern Europe	0
Sbrana, G.; Silvestrini, A.	2013	International Journal Of Production Economics	1
Moon, S.; Simpson, A.; Hicks, C.	2013	International Journal Of Production Economics	4

**Fonte:** a autora (2016)

A Figura 1 abaixo apresenta a evolução dos artigos utilizados na pesquisa ao longo dos anos, entretanto não deve ser confundida com uma demonstração da evolução de publicações sobre o tema ao passar dos anos.



**Figura 1 – Total de artigos por ano**

**Fonte:** a autora (2016)

A Tabela 1 apresenta uma análise do total de artigos em relação à revista/periódico publicado, na qual se pode verificar que a maioria dos trabalhos considerados na pesquisa foram publicados em revistas/periódicos considerados relevantes. Por exemplo, 5 artigos foram publicados na revista *International Journal of Production Economics*.

**Tabela 1 – Total de publicações por revista**

Revista/Periódico	Total de Publicações
International Journal Of Production Economics	5
IMA Journal Of Management Mathematics	2
International Journal Of Forecasting	2
Advances In Business And Management Forecasting	1
Computational Statistics And Data Analysis	1
Computers And Operations Research	1
Economic Modelling	1
European Journal Of Operational Research	1
Fibres And Textiles In Eastern Europe	1
Industrial Management And Data Systems	1
International Journal Of Production Research	1
Journal Of Business Logistics	1
Journal Of Forecasting	1
Journal Of Operations Management	1
Naval Research Logistics	1
Pesquisa Operacional	1

**Fonte:** a autora (2016)

Já a Tabela 2, apresentada abaixo, pode ser utilizada para direcionar a leitura dos artigos, pois indica os principais autores da área, com maior número de citações e publicações.

**Tabela 2 – Principais autores**

<b>Autor</b>	<b>Total de Publicações</b>	<b>Total de citações</b>
Fliedner, G.	2	46
Morris, J. S.	2	41
Piplani, R.	4	33
Viswanathan, S.	4	33
Widiarta, H.	4	33
Ahmed, R.A.	2	30
Athanasopoulos, G.	2	30
Hyndman, R.J.	2	30

**Fonte:** a autora (2016)

### 2.1.2.

#### Síntese da literatura

Os estudos referentes à direção de previsão hierárquica não são novos, um dos primeiros trabalhos encontrados na literatura foi realizado por Theil (1954) apud<sup>1</sup> (Fliedner, 2001). As duas abordagens mais citadas na literatura são as abordagens *Bottom-up* e *Top-Down*.

Na abordagem *Bottom-up* são feitas previsões com base nas séries do nível mais baixo da hierarquia, ou seja, com os dados mais detalhados (algumas vezes ao longo do trabalho essas séries serão chamadas de desagregadas, ou séries no nível dos itens, fazendo uma analogia à séries hierárquicas de linhas de produtos, por exemplo), e, posteriormente, essas previsões são somadas, agregadas, de acordo com a hierarquia, até chegar no nível mais alto que é a soma total.

Já na abordagem *Top-Down* é feita uma previsão utilizando os dados consolidados, do nível mais alto da hierarquia e, posteriormente, essa previsão é desagregada para os níveis mais baixos com base na proporção dos dados históricos.

Brown (1962) apud Dangerfield & Moris (1992) propõe que ao aplicar a abordagem *Top-Down* a previsão seja desagregada através de uma formulação de

<sup>1</sup>O ‘apud’ será utilizado, pois a síntese da literatura foi baseada somente nos artigos selecionados para a leitura completa na etapa (iv) da metodologia.

Amortecimento Exponencial com base na proporção do item de período anterior e a proporção do item mais recente, ou seja, que a previsão seja desagregada através da aplicação de proporções suavizadas.

Gross & Sohl (1990) apud Lo et al. (2008) examinaram 21 técnicas diferentes que poderiam ser aplicadas na abordagem *Top-Down* para fazer o rateio da previsão do nível mais alto, e concluíram que uma proporção simples da demanda de cada item em relação ao total, com base nos dados históricos, é suficiente para desagregar a previsão feita no nível mais alto.

Hyndman et al. (2011) afirmam que muitas indústrias utilizam na prática uma mistura das abordagens *Top-Down* e *Bottom-up*, chamada de abordagem *middle-out*. Nesta, a previsão é realizada em um nível intermediário da hierarquia e depois é desagregada para os níveis mais baixos, e agregada para os níveis mais altos.

As pesquisas encontradas na literatura basicamente seguem dois paradigmas: (i) consideram que o processo gerador das séries é totalmente conhecido, ou seja, conhece-se todas as propriedades estatísticas da série a priori, ou (ii) consideram estudos com séries reais, onde essas propriedades não são conhecidas. Dentro desses paradigmas, a maioria dos autores compara o desempenho das abordagens de previsão hierárquica, de modo a verificar sob quais circunstâncias uma técnica é preferível em relação à outra. Um dos primeiros trabalhos realizados com este objetivo foi o de Grunfeld & Griliches (1960) apud Hyndman et al. (2011).

Não existe uma unanimidade na literatura em relação ao desempenho das abordagens. Wetherby (1984) apud Fliedner (1999) já apontava que não existia preferência de um método de agregação em relação a outro.

O principal argumento que defende a superioridade da abordagem *Top-Down* parte da premissa que a variabilidade das séries é reduzida ao agregar diversos itens (Viswanathan et al., 2008). Em contrapartida, uma das fraquezas da abordagem *Top-Down* é que ela gera previsões viesadas (Athanasopoulos et al., 2009). Já outra fraqueza da abordagem *Top-Down* é a perda de informação referente às séries de nível mais baixo (Dangerfield & Moris, 1992).

Alguns estudos apontam que não há diferença significativa entre o desempenho das abordagens *Top-Down* e *Bottom-up* (Schwarzkopf et al., 1988; Widiarta et al., 2008; Widiarta et al., 2009)

Diversos autores apontaram que a qualidade da previsão hierárquica depende da correlação entre as séries desagregadas (Widiarta et al., 2007; Sohn & Lim, 2007; Chen & Boylan, 2009; Widiarta et al., 2009). Flidner (1999) utiliza simulação para verificar tal impacto e conclui que a correlação altamente positiva melhora o desempenho de previsão hierárquica (seja aplicando *Top-Down* ou *Bottom-up*), e também que uma correlação altamente negativa melhora o desempenho de previsão hierárquica no nível agregado. O autor explica que apesar de inicialmente tal afirmação parecer controversa, o desempenho da previsão está sendo medido no nível mais agregado, onde teoricamente a série apresenta uma volatilidade menor, o que leva a um melhor desempenho da previsão.

Já Sohn & Lim (2007) também utilizam simulação para entender o impacto da correlação no desempenho da previsão e concluem que quando a correlação é fortemente negativa a abordagem *Bottom-up* é melhor. Por outro lado, quando as observações são independentes ou a correlação é fracamente positiva, a abordagem *Top-Down* gera melhores resultados, já quando a correlação é fracamente negativa ou altamente positiva os autores não encontraram evidências para avaliar o impacto.

Outras conclusões importantes também foram encontradas na revisão da literatura. Flidner & Lawrence (1995) aplicaram várias técnicas de agrupamento estatístico para definir as séries do nível intermediário em sistemas de previsão hierárquica e concluíram que tais técnicas não geram uma melhora no desempenho da previsão quando comparadas à determinação aleatória do nível intermediário da hierarquia.

Já Flidner & Mabert (1992) apud Flidner (1999) concluíram em seu estudo que o total de itens agrupados dentro de uma família não impacta o desempenho da previsão hierárquica.

Apesar da área de previsão hierárquica não ser nova, por muitos anos constavam na literatura somente as técnicas já citadas (*Top-Down* e *Bottom-up*), consideradas tradicionais, até Athanasopoulos et al. (2009) que propuseram uma notação geral que pode ser utilizada em quaisquer abordagens de previsão hierárquica e também duas novas abordagens, chamadas de “Combinação Ótima” (CO) ou “Reconciliação Ótima” (RO) e “*Top-Down* Baseado na Proporção das Previsões” (*TDFP* – do inglês *Top Down Forecast Proportions*). Em ambas as

abordagens, inicialmente são feitas previsões independentes para todas as séries, de todos os níveis da hierarquia, e os autores indicam que esta é uma das principais vantagens das abordagens uma vez que não há perda de informação.

A abordagem *Top-Down* baseada na Proporção das Previsões, cujo nome é autoexplicativo, tem ainda a vantagem de ter proporções (utilizadas para desagregar a previsão do nível mais alto da hierarquia) dinâmicas, pois estas variam de acordo com a realização de novas previsões, diferentemente do que ocorre na abordagem *Top-Down* tradicional que utiliza as proporções históricas (Athanasopoulos et al. 2009).

Na abordagem de Combinação Ótima (ou Reconciliação Ótima), é aplicado um modelo de regressão, com base na variância mínima, para combinar as previsões de todos os níveis. Em outras palavras, na abordagem de Combinação Ótima a previsão para cada nível é dada por uma média ponderada das previsões feitas para todos os níveis. Os autores citam mais duas vantagens desta abordagem: (i) gera previsões não viesadas e (ii) os pesos utilizados para combinar as previsões dependem somente da estrutura hierárquica, ou seja, uma vez calculados podem ser replicados para outras séries com a mesma estrutura (Athanasopoulos et al., 2009).

Por fim, Athanasopoulos et al. (2009) apresentam uma aplicação das novas abordagens propostas em duas hierarquias formadas por séries de turismo da Austrália. Em ambas as hierarquias a abordagem *Top-Down* baseada na Proporção das Previsões obteve o melhor desempenho, seguida pela abordagem de Combinação Ótima em um dos casos, e em outro, seguido pela abordagem *Bottom-up*. Os autores afirmam que o bom desempenho da abordagem *Bottom-up* é justificado pelo comportamento dos dados, que mesmo no nível mais baixo tem componentes de sazonalidade bem definidas.

Já Hyndman et al. (2011) fazem uma extensão do trabalho de Athanasopoulos et al. (2009), no qual discutem os principais problemas computacionais que impactam a abordagem de combinação ótima. Um dos principais problemas de utilização da abordagem seria necessidade de inversão da matriz que representa a estrutura da hierarquia, pois dependendo do tamanho da mesma, a inversão não seria trivial e teria um alto custo computacional. Os autores indicam possíveis métodos que poderiam ser utilizados para sanar tais problemas, dependendo de características da estrutura hierárquica (número de

séries na estrutura, números de nós em cada nível, dentre outros), entretanto esses métodos não foram implementados no trabalho dos autores e não são aplicáveis para quaisquer estruturas. Maiores detalhes sobre essas questões serão apresentados no capítulo 3.

Por fim Hyndman et al. (2011) apresentam um estudo simulado e um estudo empírico. No estudo simulado, a abordagem Combinação ótima apresentou um melhor desempenho para o nível mais alto e intermediário da hierarquia, enquanto que a abordagem *Bottom-up* apresentou um desempenho melhor no nível mais baixo e também na média geral da hierarquia. Já no estudo empírico, os melhores desempenhos foram apresentados pela abordagem *Top-Down* no nível mais alto, Combinação Ótima no nível intermediário e *Bottom-up* no nível mais baixo.

Moon et al. (2012) propuseram em seu estudo uma combinação dos resultados gerados pela previsão *Top-Down* e previsão direta do nível mais baixo da hierarquia, baseados no fato quando um modelo de previsão não é altamente superior a outro, a combinação dos resultados gerará um resultado superior. Entretanto, essa combinação foi baseada numa média e os autores não apresentaram nenhuma metodologia, ou embasamento teórico, e por isso essa combinação não foi considerada como uma nova abordagem de previsão hierárquica.

O capítulo 3 dessa dissertação apresentará maiores detalhes matemáticos de todas as abordagens de previsão hierárquica citadas na literatura.

### **2.1.3. Análise descritiva dos resultados da revisão da literatura**

O Quadro 3 abaixo apresenta uma breve descrição dos artigos analisados na revisão da literatura, incluindo os principais objetivos, conclusões, abordagens de previsão hierárquica e as métricas de desempenho utilizadas pelos mesmos.

Já o Quadro 4 apresenta, de forma mais objetiva, as questões principais somente dos artigos que fazem uma comparação entre o desempenho das diversas abordagens de previsão hierárquica encontradas na literatura.

### Quadro 3 - Descrição dos artigos

(continua)

Referência	Objetivo	Descrição	Conclusão	Abordagem	Métrica de Avaliação
Schwarzkopf et al. (1988)	Apresentar as vantagens de cada uma das abordagens e a situação na qual uma é preferível à outra	Primeiramente simula um estudo com 10.000 itens para comparar o desempenho computacional das duas abordagens. Posteriormente compara analiticamente a influência do erro de previsão, viés e <i>outlier</i> no desempenho das abordagens.	O custo computacional para aplicação das duas abordagens é praticamente igual. Prova analiticamente que não existem evidências para afirmar que uma abordagem é melhor do que outra.	BU e TD	-
Dangerfield & Moris (1992)	Avaliar o desempenho das abordagens TD e BU utilizando o método de amortecimento exponencial em uma hierarquia simples compostas por 2 itens.	Compara o desempenho das abordagens com base em um estudo empírico utilizando 15.753 séries hierárquicas, construídas através da combinação de diferentes pares das 178 séries selecionadas da base de dados <i>M-competition</i> , cada combinação com diferente correlação e proporção entre os itens individuais.	A abordagem BU apresentou um desempenho melhor em quase 3/4 dos casos.	BU e TD	MAPE
Fliedner & Lawrence (1995)	Realizar um estudo empírico para avaliar o impacto que diversos parâmetros, utilizados para formar as famílias dentro de uma estrutura hierárquica, geram no desempenho da previsão hierárquica.	Primeiramente analisa o impacto de 4 fatores no desempenho da previsão hierárquica: método de previsão ((a) média móvel de n-períodos, (b) amortecimento exponencial simples, (c) amortecimento Exponencial de Winters), uso de técnicas estatísticas para agrupar as famílias ((a) método de Ward, (b) método do centroide, (c) UPGMA), característica de agrupamento (unidade de volume ou monetária), e total de itens que compõe uma família (2, 5 ou 10). Posteriormente compara o melhor resultado gerado na fase anterior, com uma estrutura hierárquica definida aleatoriamente.	Somente o método de previsão teve um impacto significativo no desempenho da previsão hierárquica, e que maiores esforços para determinar o agrupamento da estrutura hierárquica não geram uma melhoria no desempenho.	-	MPE e MAPE

### Quadro 3 - Descrição dos artigos

(continuação)

Referência	Objetivo	Descrição	Conclusão	Abordagem	Métrica de Avaliação
Fliedner (1999)	Analisar estratégias para agregar famílias de produtos dentro de um sistema de previsão hierárquica	Utiliza simulação como ferramenta para analisar o impacto do nível da correlação entre as séries do nível desagregado, no desempenho da previsão hierárquica. Compara os resultados das abordagens TD e BU, utilizando os métodos de média-móvel e amortecimento exponencial, simulando duas séries (que compõe o nível mais baixo da hierarquia) com 5 valores diferentes e fixos de correlação entre elas.	O artigo reforça conclusões já existentes na literatura, que quanto mais positiva a correlação das séries do nível desagregado (que compõe uma família), melhor será o desempenho da previsão hierárquica (seja da abordagem BU ou TD). O estudo conclui também que correlação altamente negativa também gera um desempenho melhor da previsão hierárquica, o que parece contra intuitivo, entretanto no nível mais agregado a série será menos volátil, logo mais fácil de prever. O artigo conclui também que abordagem BU gera resultados melhores do que TD.	BU e TD	Desvio Padrão e MAPE
Fliedner (2001)	Resumir as diretrizes existentes na literatura sobre projetos de sistemas de previsão hierárquica.	Inicialmente é fornecida uma explicação sobre o processo de previsão hierárquica, identificando opções estratégicas e parâmetros importantes que são destacados na literatura. E por fim, são apresentadas orientações práticas para a utilização de previsão hierárquica, e sugestões de pesquisas futuras.	Como orientações práticas, são apresentados estudos que comprovam que: (i) uma simples média representando parte da demanda total de cada item é suficiente para fazer o rateio da previsão agregada na abordagem TD (Gross & Sohl, 1990), (ii) o uso de métodos estatísticos de agrupamento para identificar as famílias dentro da hierarquia não apresenta melhoria significativa no resultados (Fliedner & Lawrence, 1995), (iii) as famílias podem ser definidas por itens com características semelhantes ou contrárias (Fliedner, 1999). Como propostas de pesquisas futuras ressaltou-se que não foram encontradas investigações formais que avaliassem o desempenho das abordagens de previsão hierárquica considerando estruturas com mais de dois níveis, e que existe a necessidade de maiores estudos que avaliem o impacto do critério de agrupamento no desempenho da previsão, e também maior estudos para avaliar os custos e benefícios ao usar a previsão hierárquica.	BU e TD	-

### Quadro 3 - Descrição dos artigos

(continuação)

Referência	Objetivo	Descrição	Conclusão	Abordagem	Métrica de Avaliação
Sohn & Lim (2007)	Simular taxas de retorno do índice 'Dow Jones Industrial Average' com objetivo de comparar o desempenho das previsões BU e TD considerando diferentes valores de correlação entre as 30 componentes do índice.	Utiliza o método AR-GARCH para realizar as previsões e testa a hipótese de que existe efeito de interação entre a abordagem de previsão e o grau de correlação dos componentes individuais. Também testa a hipótese que o ruído das componentes individuais influencia no desempenho da abordagem de previsão	Quando a correlação é fortemente negativa é melhor aplicar a abordagem BU e que quando as observações são independentes ou a correlação é fracamente positiva a previsão TD é melhor. Não há evidências para mostrar o impacto quando a correlação é fracamente negativa ou altamente positiva. Conclui também que o ruído das componentes independentes não influencia significativamente o desempenho do método de previsão.	BU e TD	MRE
Widiarta et al. (2007)	Avaliar analiticamente o desempenho das abordagens TD e BU considerando que as séries de nível mais baixo são geradas por um processo AR(1).	Apresenta um estudo analítico do desempenho das abordagens, quando a auto correlação de lag 1 das duas séries do nível mais baixo são idênticas, com base na variância do erro de previsão. Posteriormente realiza um estudo simulando diferentes valores para a auto correlação e parâmetros geradores da série e aplica o método de amortecimento exponencial para realizar as previsões em ambas as abordagens.	Quando a auto correlação de lag 1 das Séries Temporais de demanda do item é menor ou igual a 1/3, a diferença máxima no desempenho das duas estratégias de previsão é de apenas 1%. No entanto, se a auto correlação de lag 1 da demanda para a pelo menos um dos itens for maior do que 1/3 a abordagem BU é superior.	BU e TD	Variância do erro de previsão e RMSE
Wanke & Saliby (2007)	Analisar o comportamento do erro de previsão de vendas para as abordagens TD e BU a fim de identificar quando uma abordagem é preferível em relação à outra, sob a ótica do impacto no estoque mínimo de segurança.	Apresenta resultados analíticos para a variância do erro de previsão de vendas durante o lead-time para ambas as abordagens, utilizando método de amortecimento exponencial. Durante o lead-time a previsão de venda é mantida constante.	Os resultados indicam que os produtos com pequena participação no total de vendas e negativamente correlacionados com as vendas agregadas dos produtos restantes (ou itens individuais) tendem a apresentar menor variância do erro de previsão com a abordagem TD. TD também apresenta melhores resultados quando a variância da venda de um produto é suficientemente maior do que a variância da venda agregada dos produtos restantes	BU e TD	Variância do erro de previsão

### Quadro 3 - Descrição dos artigos

(continuação)

Referência	Objetivo	Descrição	Conclusão	Abordagem	Métrica de Avaliação
Lo et al. (2008)	Realizar um estudo empírico sobre previsão hierárquica, utilizando uma série de demanda de monitores de LCD	Primeiramente define uma hierarquia de três níveis para o consumo de monitores de LCD e identifica fatores exógenos que podem afetar a demanda. Utiliza: (i) análise de regressão múltipla (MR), (ii) função de transferência (TF) e (iii) modelos de equação simultânea (SEM) para realizar as previsões.	MO ( <i>middle-out</i> ) foi a abordagem que apresentou o melhor desempenho.	BU, TD e MO	Erro de previsão
Viswanathan et al. (2008)	Avaliar o desempenho das abordagens TD e BU no nível agregado quando as séries do nível mais baixo da hierarquia são intermitentes.	As séries de demanda desagregadas foram geradas de acordo com as distribuições normal, log normal e gama. O intervalo entre as demandas segue a distribuição uniforme ou gama. Utiliza método de Croston para prever as séries intermitentes (desagregadas) e amortecimento exponencial para prever as séries agregadas. Três cenários são simulados e avaliados de acordo com a abordagem de previsão hierárquica: (i) variabilidade do tamanho da demanda, (ii) variação entre os intervalos de demanda e (iii) variação dos itens i e ii juntos. O impacto dessa variabilidade na abordagem de previsão hierárquica e no custo de inventário são estudados através de 3.000 simulações e avaliadas com base na razão entre o MAD das duas abordagens.	Quando (ii) é baixa, a abordagem BU (que utiliza o método de Croston) gera melhores resultados. Quando (ii) e (i) são altos ou quando a série agregada é composta por muitas componentes, a abordagem TD é melhor. Na maioria dos casos as previsões com TD (feitas com amortecimento exponencial) superam as previsões BU.	BU e TD	MAD
Widiarta et al. (2008)	Avalia analiticamente o desempenho das abordagens TD e BU para séries desagregadas geradas por um processo MA(1).	Apresenta um estudo analítico do desempenho das abordagens quando as duas séries do nível mais baixo apresentam termos de correlação serial iguais. Posteriormente realiza um estudo simulando diferentes valores para a correlação e parâmetros geradores da série e aplica o método de amortecimento exponencial para realizar as previsões em ambas as abordagens.	Mostram analiticamente que não há nenhuma diferença no desempenho relativo das abordagens TD e BU quando a correlação serial das séries de nível mais baixo, geradas por um processo MA(1), são iguais. O estudo de simulação mostra que a diferença no desempenho das duas estratégias é relativamente insignificante quando a correlação entre as componentes desagregadas é pequena ou moderada.	BU e TD	Variância do erro de previsão e RMSE

### Quadro 3 - Descrição dos artigos

(continuação)

Referência	Objetivo	Descrição	Conclusão	Abordagem	Métrica de Avaliação
Athanasopoulos et al. (2009)	Apresentar previsões de curto prazo para a série hierárquica de turismo da Austrália, aplicando o método de amortecimento exponencial e utilizando 5 abordagens de previsão hierárquica, sendo duas destas novas propostas: (i) <i>Top-Down</i> baseado na proporção da previsão (TDFP) e (ii) Combinação ótima (CO).	Primeiramente é apresentada uma notação geral utilizada para quaisquer abordagens de previsão hierárquica e em seguida apresenta todas as abordagens. Apresentam análises para duas possíveis estruturas hierárquicas para as séries.	Concluem que para ambas as hierarquias a abordagem TDFP teve o melhor desempenho, seguido pela abordagem de combinação ótima na hierarquia 1. O teste de Diebold-Mariano também foi aplicado para verificar a significância da diferença no desempenho da previsão. Por fim apresenta as previsões para o turismo na Austrália e uma análise para futuras políticas de mercado.	BU, TD, TDFP e CO	MAPE
Widiarta et al. (2009)	Avaliar analiticamente o desempenho das abordagens TD e BU para séries desagregadas geradas por um processo MA(1).	Apresenta um estudo analítico do desempenho das abordagens quando as duas séries do nível mais baixo tem coeficientes de correlação serial iguais. Posteriormente realiza um estudo simulando diferentes valores para a correlação e parâmetros geradores da série e aplica o método de amortecimento exponencial para realizar as previsões em ambas as abordagens. Apresentam ainda gráficos em 3D mostrando o impacto do desempenho de cada abordagem de acordo com a variação da correlação e proporção das séries do nível mais baixo	Mostram analiticamente que não há nenhuma diferença no desempenho relativo das abordagens TD e BU quando a correlação serial das séries de nível mais baixo, geradas por um processo MA(1), são iguais. O estudo de simulação mostra que a diferença no desempenho das duas estratégias é relativamente insignificante quando a correlação entre as componentes desagregadas é pequena ou moderada.	BU e TD	Variância do erro de previsão e RMSE
Chen & Boylan (2009)	Identificar como a correlação afeta o desempenho da previsão hierárquica	Considera que a demanda é gerada por um modelo sazonal aditivo e outro multiplicativo, que não tem componente de tendência.	A abordagem BU gera melhores resultados do que a abordagem TD quando a correlação entre duas séries do nível mais baixo é altamente positiva. Já quando a correlação é altamente negativa a abordagem TD é melhor.	TD e BU	MSE

### Quadro 3 - Descrição dos artigos

(continuação)

Referência	Objetivo	Descrição	Conclusão	Abordagem	Métrica de Avaliação
Chen & Blue (2010)	Estudar propriedades estatísticas (auto correlação, correlação, previsibilidade da tendência) das demandas de modo a desenvolver estratégias e um guia para planejadores de demanda.	Avalia os resultados da previsão TD feita com 5 técnicas de previsão diferentes, duas na qual a demanda é considerada invariante, duas considerando que a demanda segue um processo AR(1) e uma considerando o processo VAR(1) para caracterizar demanda de 2 itens inter-relacionados. Investiga os efeitos de 12 cenários para a auto correlação da série agregada.	A abordagem TD deve ser usada quando a demanda agregada é muito previsível através da sua auto correlação positiva. Além disso, a correlação alta positiva entre as demandas pode aumentar a previsibilidade e, portanto, resultar em previsões mais precisas. Fornece um guia indicando em qual situação cada uma das 5 técnicas de previsão devem ser utilizadas, entretanto o modelo AR(1) gera melhores resultados na maioria dos casos	TD	FSE
Capistran et al. (2010)	Realizar a previsão da inflação do México 12 passos a frente	Gera 4 modelos sazonais (2 assumindo raiz unitária sazonal e 2 assumindo que a sazonalidade é determinística) para prever cada uma das 16 séries de índices de preço do consumidor do México, posteriormente agrega as melhores previsões usando duas abordagens de modo a prever a inflação no México 12 passos a frente. Por fim compara seus resultados com resultados gerados por especialistas.	A abordagem de combinação ótima gerou melhores resultados. O teste de Diebold-Mariano-West indicou que não há diferença significativa entre as previsões geradas no artigo e as previsões geradas por especialistas.	BU e CO	MSFE

### Quadro 3 - Descrição dos artigos

(continuação)

Referência	Objetivo	Descrição	Conclusão	Abordagem	Métrica de Avaliação
Hyndman et al. (2011)	Apresentar uma nova abordagem de previsão hierárquica, a combinação ótima, e o melhor desempenho dos resultados gerados por essa abordagem através de um estudo com séries simuladas, e também com uma série real de turismo na Austrália.	Inicialmente são apresentadas a notação utilizada e a abordagem de combinação ótima e ainda como calcular o intervalo de confiança para todas as abordagens de previsão hierárquica. Também são apresentados problemas computacionais para o uso da abordagem de combinação ótima, e possíveis meios de solução. Posteriormente um processo ARIMA é utilizado para simular 8 séries (com 100 observações) do nível mais baixo de uma estrutura hierárquica de 3 níveis, com 15 séries no total. Para cada uma das 4 abordagens de previsão hierárquica aplicada, o processo de previsão foi simulado 1000 vezes e métricas de comparação de desempenho foram apresentadas para cada nível da hierarquia. Por fim apresenta as previsões para a série real de turismo da Austrália e conclusões.	No estudo simulado a abordagem de combinação ótima apresentou um melhor resultado para o nível mais alto e intermediário da hierarquia, sendo que a abordagem <i>Bottom-up</i> apresentou um melhor desempenho para o nível mais baixo, e melhor na média de todos os níveis também. Já no estudo empírico, <i>Bottom-up</i> teve o melhor desempenho no nível mais baixo da série, combinação ótima no nível intermediário, e a <i>Top-Down</i> no nível mais alto.	BU, TD, CO e previsão independente	MAE e RMSE para séries simuladas e MAPE para série real
Williams & Waller (2011)	Comparar o desempenho das abordagens de previsão hierárquica de acordo com uso de informações compartilhadas de vendas reais de distribuidoras norte-americanas.	Utiliza dados reais semanais de ponto-de-venda (POS) de 18 centros de distribuição de mercados americanos e 10 tipos de cereais. O artigo apresenta uma comparação do desempenho da abordagens e se o uso de dados de POS compartilhados aumentam o desempenho das previsões	A abordagem BU gerou melhores resultados, e o uso de informações de venda compartilhada dentro da cadeia de suprimentos melhora o desempenho da previsão de ambas as abordagens.	BU e TD	MAPE e MSE

### Quadro 3 - Descrição dos artigos

(continuação)

Referência	Objetivo	Descrição	Conclusão	Abordagem	Métrica de Avaliação
Moon et al. (2012)	Estabelecer uma abordagem de previsão adequada para o caso da demanda de peças de reposição da Marinha Sul Coreana.	Primeiramente o artigo apresenta uma análise da demanda de peças de reposição, que é considerada não-normal. Utiliza uma estrutura hierárquica de 3 níveis, com 300, 36 e 8 séries em cada nível respectivamente e faz previsões utilizando o método de amortecimento exponencial. Utilizam 3 tipos de agregações temporais diferentes nas séries (mensal, quadrimestral e anual), e 4 métodos de ajuste dos dados (sem ajuste, tendência linear, sazonalidade aditiva e tendência linear e sazonalidade aditiva). O desempenho das abordagens é medido no nível mais baixo da hierarquia.	A abordagem de combinação simples entre as previsões feitas com a abordagem TD e direta apresentou os melhores resultados. Um estudo de simulação também mostrou que esse modelo minimiza os custos de estoque.	Previsão Direta, TD, Combinação Simples e ponderada das previsões TD e direta.	MAD e RMSE
Moon et al. (2013)	Identificar as características da demanda de itens de reposição da Marinha Sul Coerana que impactam no desempenho da previsão de modo a fornecer um guia para previsão deste tipo de demanda.	Compara os resultados da melhor previsão direta e combinada de Moon et al. (2012) utilizando modelo de regressão logística para analisar as características: (i) correlação, (ii) coeficiente de variação do volume de demanda, (iii) número de períodos com demanda nula e (iv) volume médio da demanda.	O modelo de previsão combinada gera melhores resultados quando a correlação ou a variabilidade do volume da demanda aumenta enquanto as outras características se mantêm constantes.	Previsão direta e combinação simples	MAD e RMSE
Lenort e Besta (2013)	Apresentar um estudo de caso sobre previsão de demanda de uma indústria de vestuário europeia.	Os dados são organizados em uma hierarquia de 3 níveis com 4 séries no nível mais baixo. As previsões são feitas utilizando modelos ARIMA.	Apresenta uma previsão 12 passos a frente utilizando abordagem TD.	TD	-

### Quadro 3 - Descrição dos artigos

(conclusão)

Referência	Objetivo	Descrição	Conclusão	Abordagem	Métrica de Avaliação
Sbrana & Silvestrini (2013)	Apresentar uma comparação analítica das abordagens TD e BU	Apresenta um estudo simulado para verificar as conclusões analíticas. Usa amortecimento exponencial multivariado como gerador do processo estocástico das séries.	A acurácia das abordagens depende dos parâmetros do modelo de previsão. Quanto maior a diferença entre as constantes de amortecimentos do modelo, maior será a variabilidade das previsões TD. Já o desempenho da abordagem BU depende da magnitude dos parâmetros da matriz de correlação	BU e TD	MSE e MAE
<p>Legenda = MR: <i>Multiple Regression analysis</i>; TF: <i>transfer function</i>; SEM: <i>simultaneous equations model</i>; UPGMA: <i>Unweighted Pair Group Method with Arithmetic Mean</i>; RMSE: <i>root-mean-square error</i>; MRE: <i>mean of relative error</i>; MAE: <i>mean absolute error</i>; MAD: <i>mean absolute deviation</i>; MPE: <i>mean percent error</i>; MAPE: <i>mean absolute percent error</i>; MSFE: <i>mean squared forecast errors</i>; MSE: <i>mean square error</i>; FSE: <i>Forecasting mean Square Error</i>; PB: <i>the error measure and report the percentage best result</i>; BU: <i>Bottom-up</i>; TD: <i>Top-Down</i>; TDFP: <i>Top-Down baseado na Proporção das Previsões</i>; CO: <i>Combinação ótima</i>; MO: <i>middle-out</i>.</p>					

Fonte: a autora (2016)

**Quadro 4 – Resumo dos artigos que comparam abordagens de previsão hierárquica**

(continua)

Referência	Contexto	Tipo de Estudo	Fonte dos dados	Nº de itens	Nº de níveis	Abordagem	Melhor abordagem
Schwarzkopf et al. (1988)	Gestão de Estoque	A / S	Modelos com Tendência e Sazonalidade	10.000	2	BU e TD	Não há abordagem superior
Dangerfield & Moris (1992)	Linhas de produção	E	<i>M-competition</i>	2	2	BU e TD	BU
Fliedner & Lawrence (1995)	Distribuição de peças	E	Demanda de peças de reposição automotivas	95 - 477	3	BU e TD	BU
Fliedner (1999)	Linhas de produção	S	As mesmas séries utilizadas em Wei & Abraham (1981) apud Fliedner & Lawrence (1995)	2	2	BU e TD	BU
Sohn & Lim (2007)	Finanças	S	AR(2)-GARCH(1,1)	30	2	BU e TD	Varia de acordo com a correlação dos itens
Widiarta et al. (2007)	Linhas de produção	A / S	AR(1)	2	2	BU e TD	Varia de acordo com a correlação dos itens
Wanke & Saliby (2007)	Previsão de Vendas	A	Amortecimento Exponencial	2	2	BU e TD	Varia de acordo com as características estatística dos itens
Lo et al. (2008)	Previsão de demanda	E	Demanda de uma indústria de monitores de LCD	24	3	BU, TD e MO	MO
Viswanathan et al. (2008)	Previsão de demanda	S	Distribuição Normal, Log Normal e gama	2/4/6/8/10	2	BU e TD	TD
Widiarta et al. (2008)	Linhas de produção	A / S	MA(1)	2	2	BU e TD	Não há abordagem superior
Athanasopoulos et al. (2009)	Previsão de demanda	E	Demanda de passageiros aéreas da Austrália	56 – 82	4	BU, TD, TDFP e CO	TDFP

**Quadro 4 – Resumo dos artigos que comparam abordagens de previsão hierárquica**

(conclusão)

Referência	Contexto	Tipo de Estudo	Fonte dos dados	Nº de itens	Nº de níveis	Abordagem	Melhor abordagem
Widiarta et al. (2009)	Planejamento da Produção	A / S	MA(1)	2	2	BU e TD	Varia de acordo com a correlação dos itens
Chen & Boylan (2009)	Gestão de Estoque	S	Modelos Sazonais	02—64	2	BU e TD	Varia de acordo com a correlação dos itens
Capistran et al. (2010)	Finanças	E	Índices de Preço do Consumidor do México	16	2	BU e CO	CO
Hyndman et al. (2011)	Previsão de demanda	E / S	Demanda de passagens aéreas da Austrália (E) - ARIMA (S)	56(E) - 8(S)	4(E) - 3(S)	BU, TD, CO e previsão independente	CO
Williams & Waller (2011)	Cadeia de Suprimentos	E	Centros de distribuição de mercados Norte-Americanos	180	3	BU e TD	BU
Moon et al. (2012)	Gestão de Estoque	E	Demanda de peças de reposição da Marinha Sul Coreana	300	3	Previsão Direta, TD, Combinação Simples e ponderada das previsões TD e direta.	Previsão Combinada
Moon et al. (2013)	Gestão de Estoque	E	Demanda de peças de reposição da Marinha Sul Coreana	300	3	Previsão direta e combinação simples	Previsão Combinada
Sbrana & Silvestrini (2013)	Planejamento da Produção	A / S	Amortecimento Exponencial Multivariado	2	2	BU e TD	Varia de acordo com os parâmetros do modelo

E: empírico; S: simulação; A: analítico; BU: *Bottom-up*; TD: *Top-Down*; TDFP: *Top-Down* baseado na Proporção das Previsões; CO: Combinação ótima; MO: *Middle-out*.

**Fonte:** a autora (2016)

#### 2.1.4. Referências adicionais

Neste subitem será apresentada uma breve descrição de outros artigos que foram utilizados ao longo dessa dissertação, mas que não constavam na pesquisa inicial realizada na base de dados.

A abordagem de combinação ótima foi inicialmente proposta em um *working paper* de Hyndman et al. (2007), que foi publicado posteriormente em duas partes, dividido entre os artigos de Athanasopoulos et al. (2009) e Hyndman et al. (2011) que já foram citados anteriormente. O artigo de Hyndman et al. (2007) também foi consultado, pois algumas demonstrações foram ocultadas nas publicações posteriores.

Como citado na introdução, uma das motivações do desenvolvimento desta dissertação foi o *working paper* de Hyndman et al. (2014a), neste trabalho os autores apresentam soluções anteriormente inexistentes para os problemas apresentados em Hyndman et al. (2011).

Em Hyndman et al. (2014a) é apresentado um algoritmo para aplicação da abordagem de Combinação Ótima, que pode ser utilizado em hierarquias de tamanho ilimitados, tornando a aplicação da abordagem de combinação ótima viável para as mais diversas estruturas hierárquicas. Inicialmente são apresentados os conceitos de séries temporais hierárquicas e séries temporais agrupadas (conforme descrição encontrada na introdução), tal distinção não foi encontrada na literatura antes deste trabalho. Os autores aproveitam as características diferentes entre a estrutura das séries hierárquicas e agrupadas para tratar de maneira eficiente os problemas computacionais encontrados para a aplicação da abordagem de Combinação Ótima.

Os autores apresentam também o algoritmo já implementado no pacote “*hts*” feito por Hyndman et al. (2014b) para o *software* R (R Core Team, 2013). Maiores detalhes serão apresentados no capítulo 3.

## 2.2. Métodos de Previsão

Os métodos de previsão de Séries Temporais mais utilizados na literatura são os métodos de amortecimento exponencial e modelos ARIMA (Hyndman e Athanasopoulos, 2013). Dado o vasto conhecimento destes métodos, serão apresentadas somente suas principais características neste tópico.

### 2.2.1. Amortecimento Exponencial

No método de amortecimento exponencial as previsões são geradas com base em uma média ponderada das observações passadas, na qual o peso das observações decai exponencialmente de acordo com o período. Assim observações mais recente possuem um peso maior (maior impacto para a previsão), enquanto observações mais antigas possuem um peso menor.

O método de amortecimento exponencial se baseia no comportamento das componentes de tendência e sazonalidade da Série Temporal em estudo. A Tabela 3 abaixo apresenta a classificação das componentes, e a Tabela 4 apresenta as equações para cada estado das componentes do modelo.

**Tabela 3 – Classificação do método de amortecimento exponencial**

Comp. de tendência	Comp. Sazonal		
	N (Nenhum)	A (Aditivo)	M (Multiplicativo)
N (Nenhum)	N, N	N, A	N, M
A (Aditivo)	A, N	A, A	A, M
A <sub>am</sub> (Aditivo amortecido)	A <sub>am</sub> , N	A <sub>am</sub> , A	A <sub>am</sub> , M
M (Multiplicativo)	M, N	M, A	M, M
M <sub>am</sub> (Multiplicativo amortecido)	M <sub>am</sub> , N	M <sub>am</sub> , A	M <sub>am</sub> , M

**Fonte:** adaptado de Hyndman & Athanasopoulos, 2014

Tabela 4 – Espaços de estados dos modelos de amortecimento exponencial

(continua)

Componente de Tendência	Componente Sazonal		
	N (Nenhum)	A (Aditivo)	M (Multiplicativo)
N (Nenhum)	$\hat{y}_{t+h t} = l_t$ $l_t = \alpha y_t + (1 - \alpha)l_{t-1}$	$\hat{y}_{t+h t} = l_t + s_{t-m+h_m^*}$ $l_t = \alpha(y_t - l_t) + (1 - \alpha)l_{t-1}$ $s_t = \gamma(y_t - l_t) + (1 - \gamma)s_{t-m}$	$\hat{y}_{t+h t} = l_t s_{t-m+h_m^*}$ $l_t = \alpha \left( \frac{y_t}{s_{t-m}} \right) + (1 - \alpha)l_{t-1}$ $s_t = \gamma \left( \frac{y_t}{l_{t-1}} \right) + (1 - \gamma)s_{t-m}$
A (Aditivo)	$\hat{y}_{t+h t} = l_t + hb_t$ $l_t = \alpha y_t + (1 - \alpha)(l_{t-1} + b_{t-1})$ $b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}$	$\hat{y}_{t+h t} = l_t + m b_t + s_{t-m+h_m^*}$ $l_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + b_{t-1})$ $b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}$ $s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m}$	$\hat{y}_{t+h t} = (l_t + hb_t)s_{t-m+h_m^*}$ $l_t = \alpha \left( \frac{y_t}{s_{t-m}} \right) + (1 - \alpha)(l_{t-1} + b_{t-1})$ $b_t = \beta(l_t - l_{t-1}) + (1 - \beta)b_{t-1}$ $s_t = \gamma \left( \frac{y_t}{l_{t-1} + b_{t-1}} \right) + (1 - \gamma)s_{t-m}$
A <sub>am</sub> (Aditivo amortecido)	$\hat{y}_{t+h t} = l_t + \varphi_h b_t$ $l_t = \alpha y_t + (1 - \alpha)(l_{t-1} + \varphi b_{t-1})$ $b_t = \beta(l_t - l_{t-1}) + (1 - \beta)\varphi b_{t-1}$	$\hat{y}_{t+h t} = l_t + \varphi_h b_t + s_{t-m+h_m^*}$ $l_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + \varphi b_{t-1})$ $b_t = \beta(l_t - l_{t-1}) + (1 - \beta)\varphi b_{t-1}$ $s_t = \gamma(y_t - l_t - \varphi b_t) + (1 - \gamma)s_{t-m}$	$\hat{y}_{t+h t} = (l_t + \varphi_h b_t)s_{t-m+h_m^*}$ $l_t = \alpha \left( \frac{y_t}{s_{t-m}} \right) + (1 - \alpha)(l_{t-1} + \varphi b_{t-1})$ $b_t = \beta(l_t - l_{t-1}) + (1 - \beta)\varphi b_{t-1}$ $s_t = \gamma \left( \frac{y_t}{l_{t-1} + \varphi b_{t-1}} \right) + (1 - \gamma)s_{t-m}$

Tabela 4 – Espaços de estados dos modelos de amortecimento exponencial

(conclusão)

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

Em cada caso  $l_t$  denota o nível,  $b_t$  denota a tendência e  $s_t$  denota a componente sazonal no período  $t$ ,  $m$  denota a sazonalidade em um ano;  $\alpha$ ,  $\beta$ ,  $\gamma$  e  $\varphi$  são constantes onde  $0 < \alpha, \gamma, \varphi < 1$  e  $0 < \beta < \alpha$ ;  $\hat{y}_{t+h|t}$  representa a previsão  $h$  passos-à-frente;  $\varphi_h = \varphi + \varphi^2 + \dots + \varphi^h$ ;  $h_m^+ = [(h - 1) \text{ mod } m] + 1$ .

**Fonte:** a autora (2016)

Maiores detalhes sobre o método podem ser vistos em Hyndman & Athanasopoulos (2013) e Hyndman et al. (2008).

### 2.2.2. Modelos ARIMA

Os modelos ARIMA, também conhecidos como modelos de Box & Jenkins ou modelos auto regressivos integrado de médias móveis, têm como pressuposto que o processo a ser modelado é estacionário de segunda ordem, ou seja, a média e variância são constantes ao longo do tempo, e a correlação depende somente da defasagem entre as observações.

Se baseiam na estrutura da Função de Autocorrelação (FAC) e da Função de Autocorrelação Parcial (FACP) da Série Temporal aqui representada por  $Z_T$ , nas quais são identificados os parâmetros  $q$  (referente à média-móvel) e  $p$  (referente à parte auto regressiva) do modelo respectivamente.

No modelo auto-regressivo (AR), a série de dados é modelada de acordo com os seus valores passados e pelo ruído aleatório  $\varepsilon_t$  conforme segue abaixo:

$$Z_T = \phi_1 Z_{T-1} + \phi_2 Z_{T-2} + \dots + \phi_p Z_{T-p} + \varepsilon_t \quad (1)$$

onde o parâmetro  $\phi_i$  é responsável por descrever como  $Z_t$  relaciona-se com o valor  $Z_{T-i}$ , sendo  $i = 1, 2, \dots, p$ .

No modelo de médias móveis (MA), a série é modelada através da combinação dos ruídos  $\varepsilon$  defasados.

$$Z_T = \varepsilon_T + \theta_1 \varepsilon_{T-1} + \dots + \theta_q \varepsilon_{T-q} \quad (2)$$

onde parâmetro  $\theta_i$  descreve como  $Z_T$  relaciona-se com o valor de  $\varepsilon_{T-i}$ , onde  $i = 1, 2, \dots, q$ .

Quando a Série Temporal em estudo não é estacionária, primeiramente é necessário integrá-la até que a mesma se torne estacionária. Para tal, a série é transformada através de diferenças sucessivas. O número de diferenças necessárias para tornar a série de dados estacionários representa o valor do parâmetro de integração  $d$ . A primeira diferença pode ser representada por:

$$\nabla Z = Z_T - Z_{T-1}$$

onde  $\Delta = (1 - B)$ , e  $B$  é o operador diferença.

Assim tem-se que a equação geral dos modelos ARIMA(p,d,q) é dada por:

$$\phi(B)\nabla^d Z_t = \theta(B)\varepsilon_t \quad (3)$$

Já os modelos sazonais são conhecidos como SARIMA, ou Arima multiplicativo e têm uma formulação semelhante à dos modelos arima, entretanto adiciona-se os parâmetros  $\Phi$  e  $\Theta$  referentes à sazonalidade, e o parâmetro  $D$  que representa o total de diferenças sazonais necessárias para tornar a série estacionária na sazonalidade. Assim, a equação geral do modelo Sarima é dada por:

$$\Phi(B^s)\phi(B)\nabla_s^D \nabla^d Z_t = \theta(B^s)\theta(B)\varepsilon_t \quad (4)$$

Maiores detalhes sobre os modelos ARIMA e estimação dos parâmetros podem ser encontrados em Souza & Camargo (2004).

### 2.3. Métricas de desempenho

Nesta seção serão apresentadas as métricas que serão utilizadas para comparar o desempenho das abordagens de previsão hierárquica.

#### 2.3.1. MAPE

A métrica MAPE do inglês *Mean Absolute Percentage Error* é uma medida que expressa a acurácia da previsão como uma porcentagem do erro, assim quanto menor o MAPE melhor a previsão. A métrica é dada por:

$$MAPE = \frac{\sum_{t=1}^N \frac{|Y(t) - \hat{Y}(t)|}{Y(t)} * 100}{N}, \quad (5)$$

onde  $N$  é o tamanho da série,  $Y(t)$  o valor original da série no período  $t$  e  $\hat{Y}(t)$  o valor previsto.

#### 2.3.2. Teste de Diebold-Mariano

O teste de Diebold-Mariano, proposto pelos autores Diebold e Mariano (1995), é um teste de hipótese para verificar se existe uma diferença significativa entre duas previsões diferentes, nas quais os erros de previsão podem ser não Gaussianos, com média diferente de zero ou correlacionados.

Diebold e Mariano (1995) propõe que a acurácia de dois diferentes métodos de previsão deve ser avaliada pela diferença entre uma função de perda dos erros, representada por  $d(t)$ .

$$d(t) = g(e_{1t}) - g(e_{2t}), \quad (6)$$

onde  $g(e_{1t})$  representa a função de perda do erro do método 1 no período  $t$  e  $g(e_{2t})$  representa a função de perda do erro do método 2 no período  $t$ .

Assim a hipótese nula do teste,  $H_0$ , diz que os métodos têm uma acurácia igual na média se o valor esperado de  $d(t) = 0$ ,  $\forall t$ , ou seja:

$$H_0: E[d(t)] = 0, \forall t$$

A estatística de teste (DM) é dada por:

$$DM = \frac{\bar{d}}{\sqrt{\frac{2\pi\hat{f}_d(0)}{T}}}, \quad (7)$$

onde  $T$  é total de observações,  $2\pi\hat{f}_d(0)$  é um estimador consistente da variância assintótica de  $\sqrt{T}\bar{d}$  e  $\bar{d} = \frac{1}{T}\sum_{t=1}^T d_t$ .

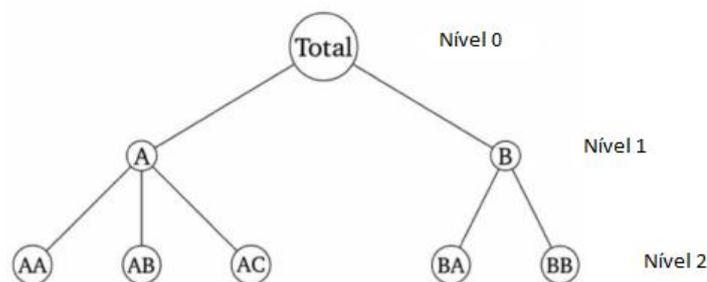
### 3 Detalhamento matemático das abordagens de previsão hierárquica

Neste capítulo será apresentada a notação utilizada e também os principais detalhes matemáticos, incluindo algumas vantagens e desvantagens, de cada abordagem de previsão hierárquica encontrada na literatura.

#### 3.1. Notação

A notação utilizada ao longo do trabalho é baseada na notação geral para séries hierárquica apresentada por Athanasopoulos et al. (2009).

Inicialmente para facilitar o entendimento é mostrada uma estrutura genérica de série hierárquica.



**Figura 2 – Estrutura Hierárquica**

**Fonte:** Adaptado de Hyndman & Athanasopoulos (2014)

O Total, no nível mais agregado, chamado de Nível zero, representa a série temporal com  $n$  observações, que como o nome já diz, é a soma total de todas as séries. A e B são as séries desagregadas em um nível abaixo do total, chamado de Nível 1, também com  $n$  observações cada uma. E assim sucessivamente.

Seja  $Y_{X,t}$  a  $t$ -ésima observação ( $t = 1, \dots, n$ ) da série  $Y_X$ , onde  $X$  representa um nó na hierarquia em determinado nível. Usando como exemplo a Figura 2  $Y_{AB,1}$  representa a primeira observação da série  $Y_{AB}$  que está no nível 2 da

hierarquia. O nível mais baixo é representado pela letra  $K$ , na Figura 2 tem-se  $K = 2$ . O número total de séries em cada nível é representado por  $m_i$ , onde  $i$  denomina o nível,  $i = 0, 1, \dots, K$  e assim o total de séries na hierarquia é representado por  $m = m_0 + m_1 + \dots + m_k$ , na Figura 2 acima tem-se  $m = 1 + 2 + 5 = 8$ .

Também é necessário definir o vetor coluna  $Y_t$ , que contém todas as séries da hierarquia, na Figura 2 temos  $Y_t = (Y_t, Y_{A,t}, Y_{B,t}, Y_{AA,t}, \dots, Y_{BB,t})$ , onde o primeiro termo  $Y_t$  representa a  $t$ -ésima observação da série total no Nível 0.

A partir daí é possível escrever qualquer série hierárquica/agrupada na forma matricial abaixo:

$$Y_t = SY_{K,t}, \quad (8)$$

onde  $S$  é a matriz que soma as séries do nível mais baixo para formar a estrutura da hierarquia.

Considerando que a Figura 2 contenha semente uma variável de agregação, e assim sendo uma série hierárquica, conforme descrito na introdução, a eq. (8) seria representada, por exemplo, pelas matrizes a seguir:

$$\begin{bmatrix} Y_t \\ Y_{A,t} \\ Y_{B,t} \\ Y_{AA,t} \\ Y_{AB,t} \\ Y_{AC,t} \\ Y_{BA,t} \\ Y_{BB,t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Y_{AA,t} \\ Y_{AB,t} \\ Y_{AC,t} \\ Y_{BA,t} \\ Y_{BB,t} \end{bmatrix}$$

Agora considerando uma série agrupada por duas variáveis de agregação, por exemplo, nota-se que é possível organizar os dados conforme apresentado na Figura 3 abaixo:

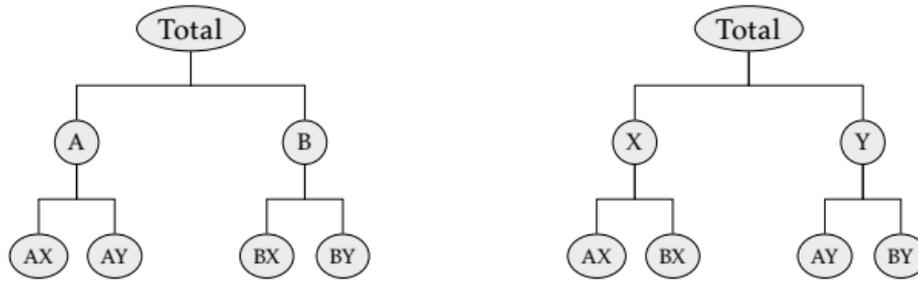
			<b>Soma na linha</b>
	AX	BX	X
	AY	BY	Y
<b>Soma na coluna</b>	A	B	<b>Total</b>

**Figura 3 – Exemplo de séries agrupada**

**Fonte:** Adaptado de Hyndman et al. 2014a

Neste exemplo  $A$  e  $B$  representam uma variável de agregação (tipo de produto, por exemplo) e  $X$  e  $Y$  representam outra (região, por exemplo), sendo

possível notar que existem duas possíveis maneiras de agregar os dados, formando duas hierarquias, conforme mostrado na Figura 4 abaixo.



**Figura 4 - Exemplos de Hierarquias baseadas em 2 variáveis de agregação**

**Fonte:** Hyndman et al. (2014a)

Neste caso não é apropriado analisar somente uma estrutura, logo a matriz  $S$  deve representar todas as maneiras de agregar as séries, se diferenciando um pouco da matriz  $S$  referente às séries hierárquicas (Hyndman et al., 2014a). Utilizando o exemplo da Figura 4 teríamos a eq. (8) representada da seguinte forma:

$$\begin{bmatrix} Y_t \\ Y_{A,t} \\ Y_{B,t} \\ Y_{X,t} \\ Y_{Y,t} \\ Y_{AX,t} \\ Y_{AY,t} \\ Y_{BX,t} \\ Y_{BY,t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Y_{AX,t} \\ Y_{AY,t} \\ Y_{BX,t} \\ Y_{BY,t} \end{bmatrix}$$

Apesar da matriz  $S$  ter uma estrutura diferente, toda a notação pode ser utilizada normalmente, tanto para séries hierárquicas, quanto para séries agrupadas. Como consequência da matriz  $S$ , tem-se que o vetor  $Y_t$  também apresenta uma dimensão diferente, na qual o primeiro termo representa a soma total das séries, posteriormente os termos que representam a soma por coluna, seguidos dos termos que representam a soma por linhas e por fim as séries do nível mais baixo, sempre nesta ordem.

Seja  $\hat{Y}_{X,n}(h)$  a previsão  $h$ -passos-à-frente da série  $X$ , e  $\hat{Y}_n(h)$  um vetor da mesma dimensão de  $Y_t$ , que contém todas as previsões de cada série de cada um

dos níveis da hierarquia, feita de maneira independente, chamado de previsão base. Assim, pode-se escrever:

$$\tilde{Y}_n(h) = SP\hat{Y}_n(h) \quad (9)$$

Na eq. (9)  $\tilde{Y}_n(h)$  é chamada de previsão revisada,  $S$  é a matriz de soma da hierarquia, e  $P$ , que será visto de maneira mais detalhada adiante, é a matriz referente a abordagem de previsão hierárquica a ser utilizada. O efeito da matriz  $P$  será extrair os elementos relevantes da matriz  $\hat{Y}_n(h)$  que serão utilizados para compor a previsão revisada.

Assim, a eq. (9) representa a notação geral proposta por Athanasopoulos et al. (2009) para ser utilizada em previsão hierárquica. Tendo como base a notação e também a diferença entre séries hierárquicas e agrupadas, é possível agora descrever todas as abordagens encontradas na literatura e alguns detalhes matemáticos, conforme segue nos subitens 3.2. até 3.5.

### 3.2. Abordagem *Bottom-up*

Utilizando a notação apresentada anteriormente, as características da matriz  $P$  definem a abordagem de previsão hierárquica que será utilizada. No caso da abordagem *Bottom-up*,  $P$  é dada por uma matriz de zeros nas colunas da esquerda e uma matriz identidade à direita.

$$P = [0_{m_K \times (m-m_K)} | I_{m_K}] \quad (10)$$

Segue abaixo um exemplo numérico simples, no qual a hierarquia é formada somente por duas séries no nível 1 e uma previsão base fictícia representada por  $\hat{Y}_{n(h)} = [5 \quad 2 \quad 2]'$ , assim fica fácil entender o comportamento da matriz  $P$  para o caso *Bottom-up*.

$$\begin{aligned} \tilde{Y}_n(h) &= SP\hat{Y}_{n(h)} \\ \tilde{Y}_n(h) &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ 2 \end{bmatrix} \\ \tilde{Y}_n(h) &= \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ 2 \end{bmatrix} \\ \tilde{Y}_n(h) &= \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix} \end{aligned}$$

Além da facilidade de aplicação, a abordagem *Bottom-up* gera resultados não viesados. É fácil demonstrar esse resultado analiticamente, entretanto essa demonstração é semelhante tanto para o caso da abordagem *Bottom-up* como da *Top-Down*, sendo assim, essa será mostrada somente no próximo subitem, de modo a evitar repetições.

### 3.3. Abordagem *Top-Down*

A matriz  $P$  neste caso é formada pelo vetor de proporções  $p = (p_1, p_2, \dots, p_{m_K})'$  na primeira coluna e o restante da matriz é preenchida com zeros.

$$P = [p | 0_{m_K \times (m-1)}] \quad (11)$$

Na abordagem *Top-Down* as proporções podem ser geradas de duas maneiras: (i) pela proporção média dos dados históricos em relação ao total ou (ii) pela média dos dados históricos em relação a média do total. Assim tem-se que:

$$(i) p_j = \left( \sum_{t=1}^n \frac{Y_{j,t}}{Y_t} \right) / n \quad (12)$$

$$(ii) p_j = \sum_{t=1}^n \frac{Y_{j,t}}{n} / \sum_{t=1}^n \frac{Y_t}{n} \quad (13)$$

A Figura 5 abaixo apresenta 3 observações fictícias de uma hierarquia de dois níveis, que será utilizada como exemplo, na qual  $p_j$  foi calculado com base nas maneiras (i) e (ii).

séries hierárquicas			
	Total	A	B
	3	1	2
	3	1	2
	5	2	3
Soma	11	4	7
(i) $p_j$	0,3636	0,6364	
(ii) $p_j$	0,3556	0,6444	

**Figura 5 – Exemplo do cálculo das proporções**

**Fonte:** a autora(2016)

Segue abaixo um exemplo numérico, utilizando a maneira (i) para calcular as proporções, do cálculo da previsão com a abordagem *Top-Down*.

$$\begin{aligned}\tilde{Y}_n(h) &= SP\hat{Y}_n(h) \\ \tilde{Y}_{n(h)} &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0,3636 & 0 & 0 \\ 0,6364 & 0 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ 2 \end{bmatrix} \\ \tilde{Y}_{n(h)} &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1,82 \\ 3,18 \end{bmatrix} \\ \tilde{Y}_{n(h)} &= \begin{bmatrix} 5 \\ 1,82 \\ 3,18 \end{bmatrix}\end{aligned}$$

Agora segue abaixo o mesmo exemplo, agora usando a maneira (ii) para calcular as proporções.

$$\begin{aligned}\tilde{Y}_n(h) &= SP\hat{Y}_n(h) \\ \tilde{Y}_n(h) &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0,3556 & 0 & 0 \\ 0,6444 & 0 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ 2 \end{bmatrix} \\ \tilde{Y}_n(h) &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1,78 \\ 3,22 \end{bmatrix} \\ \tilde{Y}_n(h) &= \begin{bmatrix} 5 \\ 1,78 \\ 3,22 \end{bmatrix}\end{aligned}$$

Uma das desvantagens em aplicar a abordagem *Top-Down* é que ela gera previsões viesadas, mesmo quando a previsão base é não viesada. Hyndman et al. (2007) demonstram essa afirmação assumindo primeiramente que a previsão base é não viesada, ou seja:

$$E[\hat{Y}_n(h)] = E[Y_n(h)] \quad (14)$$

Para que a previsão revisada também seja não viesada é preciso garantir que:

$$E[\tilde{Y}_n(h)] = E[Y_n(h)] = S * E[Y_{k,n}(h)] \quad (15)$$

Agora suponha que  $\beta_n(h) = E[Y_{k,n+h}|Y_1, \dots, Y_n]$ , então tem-se que:

$$E[\tilde{Y}_n(h)] = SPE[\hat{Y}_n(h)] = SPS\beta_n(h) \quad (16)$$

Então para que a previsão revista seja não viesada é necessário que:

$$SPS = S \quad (17)$$

Assim é possível verificar que abordagem *Top-Down* sempre gera previsões viesadas, pois não satisfaz a eq. (17). Já a abordagem *Bottom-up* sempre satisfaz a eq. (17), e assim prova-se que a mesma gera previsões não viesadas.

### 3.4.

#### Abordagem *Top Down* baseada na proporção das previsões

Athanasopoulos et al. (2009) introduziram uma nova abordagem para desagregar as previsões feitas no nível agregado. Nessa abordagem são feitas previsões independentes para todas as séries de todos os níveis da hierarquia, e posteriormente calculam-se as proporções, que serão utilizadas para desagregar a previsão do nível mais alto, com base na proporção das previsões realizadas para os níveis mais baixos.

Assim como na abordagem *Top-Down* tradicional, a matriz  $P$  é composta por um vetor com as proporções na primeira coluna e zeros nas demais posições.

$$P = [p | 0_{m_K \times (m-1)}]$$

A fim de apresentar uma notação geral para o cálculo do vetor  $p$  é necessário introduzir novas variáveis na notação. Seja  $\hat{Y}_{j,n}^{(l)}(h)$  a previsão  $h$  passos-à-frente da série que corresponde ao nó  $(l)$  níveis acima de  $j$ . Seja  $\hat{S}_{j,n}(h)$  a soma das previsões  $h$  passos-à-frente abaixo do nó  $j$  e que estão diretamente ligadas ao nó  $j$ .

Utilizando a Figura 2 como exemplo, temos que:

$$\hat{S}_{AA,n}^{(2)}(h) = \hat{S}_{Total,n}(h) = \hat{Y}_{A,n}(h) + \hat{Y}_{B,n}(h)$$

Assim a previsão revisada será:

$$\tilde{Y}_{A,n}(h) = \left( \frac{\hat{Y}_{A,n}(h)}{\hat{S}_{A,n}^{(1)}(h)} \right) \hat{Y}_{Total,n}(h) = \left( \frac{\hat{Y}_{AA,n}^{(1)}(h)}{\hat{S}_{AA,n}^{(2)}(h)} \right) \hat{Y}_{Total,n}(h)$$

e

$$\tilde{Y}_{AA,n}(h) = \left( \frac{\hat{Y}_{AA,n}(h)}{\hat{S}_{AA,n}^{(1)}(h)} \right) \tilde{Y}_{A,n}(h) = \left( \frac{\hat{Y}_{AA,n}(h)}{\hat{S}_{AA,n}^{(1)}(h)} \right) \left( \frac{\hat{Y}_{AA,n}^{(1)}(h)}{\hat{S}_{AA,n}^{(2)}(h)} \right) \hat{Y}_{Total,n}(h)$$

Consequentemente:

$$p_1 = \left( \frac{\hat{Y}_{AA,n}(h)}{\hat{S}_{AA,n}^{(1)}(h)} \right) \left( \frac{\hat{Y}_{AA,n}^{(1)}(h)}{\hat{S}_{AA,n}^{(2)}(h)} \right)$$

Generalizando o resultado, tem-se que:

$$p_j = \prod_{l=0}^{K-1} \frac{\hat{Y}_{j,n}^{(l)}(h)}{\hat{S}_{j,n}^{(l+1)}(h)} \quad (18)$$

para  $j = 1, 2, \dots, m_K$ .

Uma das desvantagens dessa abordagem é que ela também gera previsões viesadas, assim como a abordagem *Top-Down* tradicional (Hyndman et al., 2007).

### 3.5. Abordagem de Combinação ótima

Na Combinação Ótima a previsão revisada em cada nível será uma média ponderada das previsões feitas para todos os níveis da previsão base. O objetivo é encontrar  $P$ , que neste caso é uma matriz de pesos. Athanasopoulos et al. (2009) consideram que pode-se reescrever a previsão base como uma equação de regressão:

$$\hat{Y}_n(h) = S\beta_n(h) + \varepsilon_h \quad (19)$$

onde  $\beta_n(h) = E[Y_{K,n+h}|Y_1, \dots, Y_n]$  e  $\varepsilon_h$  representa o erro da regressão e não deve ser confundido com erro de previsão. Ressalta-se ainda que o erro da regressão  $\varepsilon_h$  tem média zero e  $V[\varepsilon_h] = \Sigma_h$ .

Utilizando o método de Mínimos Quadrados Ordinários (MQO) para estimar  $\beta_n$  Athanasopoulos et al. (2009) chegam na eq. (20) que representa a abordagem de Combinação Ótima.

$$\tilde{Y}_n(h) = S(S'S)^{-1}S'\hat{Y}_n(h) \quad (20)$$

Na abordagem de Combinação Ótima tem-se que  $P = (S'S)^{-1}S'$ , ou seja, fica claro que a matriz de pesos  $P$  depende somente da estrutura da hierarquia e não dos dados históricos, sendo esta uma das vantagens desta abordagem, pois uma vez calculada essa matriz pode ser reutilizada para outras hierarquias que tenham a mesma estrutura.

Segue abaixo no Quadro 5 um exemplo da matriz de pesos utilizada para ponderar as previsões independentes e formar a previsão revisada, utilizando a

abordagem de combinação ótima. A estrutura da hierarquia é a mesma vista na Figura 2.

**Quadro 5– Matriz de pesos para a hierarquia da Figura 2**

	$\widehat{Total}$	$\widehat{A}$	$\widehat{B}$	$\widehat{AA}$	$\widehat{AB}$	$\widehat{AC}$	$\widehat{BA}$	$\widehat{BB}$
$\widehat{Total}$	0,586	0,31	0,276	0,103	0,103	0,103	0,138	0,138
$\widehat{A}$	0,31	0,517	-0,207	0,172	0,172	0,172	-0,103	-0,103
$\widehat{B}$	0,276	-0,207	0,483	-0,069	-0,069	-0,069	0,241	0,241
$\widehat{AA}$	0,103	0,172	-0,069	0,724	-0,276	-0,276	-0,034	-0,034
$\widehat{AB}$	0,103	0,172	-0,069	-0,276	0,724	-0,276	-0,034	-0,034
$\widehat{AC}$	0,103	0,172	-0,069	-0,276	-0,276	0,724	-0,034	-0,034
$\widehat{BA}$	0,138	-0,103	0,241	-0,034	-0,034	-0,034	0,621	-0,379
$\widehat{BB}$	0,138	-0,103	0,241	-0,034	-0,034	-0,034	-0,379	0,621

**Fonte:** adaptado de Hyndman & Athanasopoulos (2014)

Assim tem-se que a previsão revisada para a série  $A$  seria:

$$\begin{aligned} \widehat{A} = & 0,31\widehat{Total} + 0,517\widehat{A} - 0,207\widehat{B} + 0,172\widehat{AA} + 0,172\widehat{AB} + 0,172\widehat{AC} \\ & - 0,103\widehat{BA} - 0,103\widehat{BB} \end{aligned}$$

No exemplo acima nota-se que os pesos podem ser negativos, para que não haja duplicação de valores. No caso acima, a previsão revisada é composta por uma componente do  $\widehat{Total}$ , que contém componentes do nó  $A$  e  $B$ , conseqüentemente os pesos dos nós associados a  $B$  são negativos.

Athanasopoulos et al. (2009) consideram também que possa existir uma correlação entre os resíduos da regressão da eq. (19). Caso  $\Sigma_h$  fosse conhecida, seria possível aplicar o método de mínimos quadrados generalizados para obter um estimador com variância mínima de  $\hat{\beta}_n(h)$ . Assim teríamos que:

$$\hat{\beta}_n(h) = (S'\Sigma_h^+S)^{-1}S'\Sigma_h^+\hat{Y}_n(h) \quad (21)$$

onde  $\Sigma_h^+$  é a inversa generalizada de Moore-Penrose de  $\Sigma_h$ . Utiliza-se a inversa generalizada porque frequentemente  $\Sigma_h$  é uma matriz singular. Assim teríamos que:

$$P = (S'\Sigma_h^+S)^{-1}S'\Sigma_h^+ \quad (22)$$

A dificuldade em aplicar o método de Mínimos Quadrados Generalizados está na necessidade em conhecer  $\Sigma_h$ , ou pelo menos uma boa estimativa do mesmo, o que pode não ser possível em hierarquias grandes com centenas de séries. Contudo, Hyndman et al. (2011) mostram que os cálculos podem ser simplificados expressando  $\varepsilon_h \approx S\varepsilon_{K,h}$ , onde  $\varepsilon_{K,h}$  é o erro da previsão das séries do

nível mais baixo da hierarquia, ou seja, assumem que o erro da regressão segue a mesma estrutura de agregação dos dados originais. Os autores afirmam que essa é uma aproximação razoável do que acontece na prática, pois ao utilizar um modelo ARIMA com parâmetros fixos, por exemplo, para prever todas as séries, os erros seriam aditivos seguindo a estrutura da hierarquia. Essa premissa conduz ao seguinte resultado:

$$\Sigma_h \approx S\Omega_h S', \quad (23)$$

onde  $\Omega_h = \text{Var}(\varepsilon_{k,h})$ .

Com base nesse resultado, Hyndman et al. (2011) provam também que os resultados encontrados utilizando o método de Mínimos Quadrados Generalizados são iguais aos resultados de mínimos quadrados ordinários, como segue abaixo.

Seja  $\Sigma_h = BC$  onde  $B = S\Omega_h$  e  $C = S'$ . A generalizada inversa de  $\Sigma_h$  é dada por:

$$\Sigma_h^+ = C'(CC')^{-1}(B'B)^{-1}B' = S(S'S)^{-1}(\Omega_h' S' S \Omega_h)^{-1} \Omega_h' S' \quad (24)$$

Logo:

$$(S'\Sigma_h^+ S)^{-1} S'\Sigma_h^+ = (S'S)^{-1} S' \quad (25)$$

Hyndman et al. (2011) também indicam que uma das principais desvantagens da aplicação da abordagem de combinação ótima é a dificuldade em inverter a matriz  $(S'S)$ , principalmente para o caso de grandes estruturas hierárquicas. Os autores indicam três possíveis soluções, entretanto nenhuma delas é prática para aplicações reais.

A primeira solução indicada pelos autores é que como  $S$  é uma matriz esparsa, então pode-se usar um armazenamento de matriz esparsa de Duff et al. (2002) apud Hyndman et al. (2011) para economizar memória e tempo computacional, e então aplicar o algoritmo implementado em Koenker and Ng (2010) apud Hyndman et al. (2011) para resolver a eq. (20), entretanto esses métodos não são adequados para grandes hierarquias.

Outra solução indicada pelos autores, agora para grandes hierarquias, é a aplicação do algoritmo iterativo de Paige e Saunders (1982) apud Hyndman et al. (2011) para resolver problemas lineares esparsos de Mínimos Quadrados, entretanto, apesar de bons, os resultados são uma aproximação à solução direta.

Por fim, Hyndman et al. (2011) propõem que seja feita uma reparametrização da eq. (19), e fornecem as novas equações onde os estimadores

são reespressos em equações que envolvem somatórios, mas não envolvem inversão de matrizes, entretanto essa proposta só pode ser aplicada em hierarquias equilibradas, ou seja, em que dentro de um determinado nível cada nó tem o mesmo número de sub-séreis. No caso da Figura 2, por exemplo, essa proposta não poderia ser aplicada.

Hyndman et al. (2014a) propõe ainda que o método de Mínimos Quadrados Ponderados (MQP) seja aplicado para resolver a eq. (19) devido a dificuldade em se estimar  $\sum_h$ , assim tem-se:

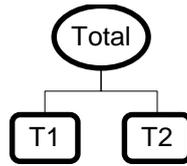
$$\tilde{Y}_h = S(S'\Lambda_h S)^{-1}S'\Lambda_h \hat{Y}_h, \quad (26)$$

onde  $\Lambda_h$  é a matriz diagonal com elementos iguais ao inverso da variância de  $\varepsilon_h$ . Entretanto, as estimativas de  $\Lambda_h$  também não são facilmente disponíveis. Nesses casos os elementos de  $\Lambda_h$  são definidos como o inverso da variância da previsão base,  $Var(Y_{T+1} - \hat{Y}_1)$ , que estão prontamente disponíveis como as variâncias residuais de cada um dos modelos da previsão base (Hyndman et al., 2014a).

Contudo, para todos os métodos de resolução da regressão da eq.(19), o problema da inversão da matriz  $(S'S)$  persiste. Isto posto, Hyndman et al. (2014a), analisando a estrutura da matriz  $S$ , tanto das séries hierárquicas quanto agrupadas, demonstram como a inversão dessa matriz pode ser realizada e propõem algoritmos recursivos eficientes que fazem a inversão da mesma e o cálculo da previsão revisada. Nos subitens 3.5.1. e 3.5.2 abaixo serão apresentadas as demonstrações para a inversão da matriz  $(S'S)$  propostas por Hyndman et al. (2014a).

### 3.5.1. Inversão da Matriz (S'S) para o caso de séries hierárquicas

Para o caso de séries hierárquicas, Hyndman et al. (2014a) considera que a estrutura pode ser representada por uma hierarquia de somente dois níveis, como uma árvore de árvores, composta pelo total no nível zero, e por  $K$  sub-árvores que representam todos os níveis mais baixos da hierarquia. As séries hierárquicas da Figura 2, por exemplo, seriam representadas da seguinte maneira:



**Figura 6 – Árvore hierárquica**

**Fonte:** a autora (2016)

Na Figura 6, T1 e T2 representam a sub árvore do nó A e B respectivamente.

Assim a matriz  $S$  agora passa a representar todas as sub-árvores da hierarquia, e é dada por:

$$S = \begin{bmatrix} 1'_{n_1} & 1'_{n_2} & \dots & 1'_{n_K} \\ S_1 & 0 & \dots & 0 \\ 0 & S_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & S_K \end{bmatrix} \quad (27)$$

Onde  $1_n$  é um vetor de números 1 e tamanho  $n$ , e  $S_1, S_2, \dots, S_K$  representam a matriz de soma  $S$  de cada uma das sub-árvores.

Assim, pode-se escrever:

$$S'AS = \lambda_0 J_n + \begin{bmatrix} S'_1 A_1 S_1 & 0 & \dots & 0 \\ 0 & S'_2 A_2 S_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & S'_K A_K S_K \end{bmatrix} \quad (28)$$

onde  $J_n$  é uma matriz  $n \times n$  de números 1, e  $n = \sum_k n_k$ . A partir daí os autores somam a  $S'AS$  uma matriz de posto 1 (operação conhecida em inglês como *rank-one update*) e invertem a matriz, conforme segue no resultado abaixo:

$$(S'AS)^{-1} = \begin{bmatrix} (S'_1 A_1 S_1)^{-1} & 0 & \dots & 0 \\ 0 & (S'_2 A_2 S_2)^{-1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (S'_K A_K S_K)^{-1} \end{bmatrix} - c S_o \quad (29)$$

onde  $S_o$  é uma matriz  $n \times n$  que pode ser dividida em  $K^2$  blocos, e o bloco  $j, k$  é dado por  $(S'_j A_j S_j)^{-1} J_{n_j, n_k} (S'_k A_k S_k)^{-1}$ , e  $c$  é dado por:

$$c = \lambda_0^{-1} + \sum_j 1'_{n_j} (S'_j A_j S_j)^{-1} 1_{n_j} \quad (30)$$

Dadas as bases para inverter a matriz  $(S'AS)$ , Hyndman et al. (2014a) apresentam todas as demonstrações para o cálculo recursivo da matriz inversa, e

posteriormente, o cálculo da previsão ajustada. Por fim, os autores apresentam um algoritmo recursivo de 3 passos para calcular a previsão revisada dada pela aplicação da abordagem de combinação ótima. Maiores detalhes sobre o algoritmo podem ser verificados Hyndman et al. (2014a). Esse algoritmo foi implementado por Hyndman et al. (2014b) no pacote “*hts*” para o software R (R Core Team, 2013). Esse pacote será utilizado para realizar as aplicações exibidas no capítulo 4.

### 3.5.2. Inversão da Matriz ( $S'S$ ) para o caso de séries agrupadas

Para o casos das séries agrupadas, Hyndman et al. (2014a) reescrevem a matriz  $S$  utilizando o produto direto, conforme abaixo:

$$S = \begin{bmatrix} 1'_m \otimes 1'_n \\ 1'_m \otimes I_n \\ I_m \otimes 1'_n \\ I_m \otimes I_n \end{bmatrix}, \quad (31)$$

onde, assim como para o caso das séries hierárquicas,  $1_n$  é um vetor de números 1 e tamanho  $n$ , e  $I_m$  é uma matriz identidade de tamanho  $m \times m$ . Utilizando como exemplo da Figura 3 mostrada anteriormente, sabe-se que:

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

Segue abaixo, somente para ilustrar, os cálculos da primeira linha da matriz  $S$  utilizando o produto direto:

$$1'_m \otimes 1'_n = [1 \ 1][1 \ 1] = [1 \ 1 \ 1 \ 1]$$

Também utilizando o produto direto pode-se escrever a matriz  $S'S$ , conforme segue abaixo:

$$S'S = [1_m \otimes 1_n | 1_m \otimes I_n | I_m \otimes 1_n | I_m \otimes I_n] \times \begin{bmatrix} 1'_m \otimes 1'_n \\ 1'_m \otimes I_n \\ I_m \otimes 1'_n \\ I_m \otimes I_n \end{bmatrix} \quad (32)$$

$$S'S = (1_m \otimes 1_n)(1'_m \otimes 1'_n) + (1_m \otimes I_n)(1'_m \otimes I_n) \quad (33)$$

$$+ (I_m \otimes 1_n)(I_m \otimes 1'_n) + (I_m \otimes I_n)(I_m \otimes I_n)$$

Aplicando a seguinte propriedade de produto direto,  $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$  tem-se que:

$$S'S = J_m \otimes J_n + J_m \otimes I_n + I_m \otimes J_n + I_m \otimes I_n \quad (34)$$

em que  $J_m$  é uma matriz de números 1 de dimensão  $m \times m$ .

Aplicando a propriedade  $(A \otimes B) + (A \otimes C) = A \otimes (B + C)$  tem-se que:

$$S'S = J_m \otimes (J_n + I_n) + I_m \otimes (J_n + I_n) \quad (35)$$

$$S'S = (I_m + J_m) \otimes (I_n + J_n) \quad (36)$$

Aplicando a propriedade  $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ , tem-se:

$$(S'S)^{-1} = [(I_m + J_m) \otimes (I_n + J_n)]^{-1} \quad (37)$$

$$(S'S)^{-1} = (I_m + J_m)^{-1} \otimes (I_n + J_n)^{-1} \quad (38)$$

$$(S'S)^{-1} = (I_m - J_m/(m + 1)) \otimes (I_n - J_n/(n + 1)) \quad (39)$$

Dadas as bases necessárias para calcular a inversa para o caso das séries agrupadas, Hyndam et al. (2014a) apresentam as demonstrações para a equação da previsão revisada, uma extensão dos cálculos feitos acima, agora utilizando Mínimos Quadrados Ponderados, e por fim um algoritmo iterativo de seis passos para o cálculo da previsão revisada utilizando a abordagem de Combinação Ótima para o caso de séries agrupadas. Assim, como no caso das séries hierárquicas, maiores detalhes sobre o algoritmo podem ser verificados Hyndman et al. (2014a). O algoritmo para o caso das séries agrupadas também foi implementado por Hyndman et al. (2014b) no pacote “*hts*” para o software R (R Core Team, 2013).

## 4 Resultados

Este capítulo tem como objetivo apresentar os resultados das aplicações das abordagens de previsão hierárquica apresentadas no capítulo 3. Serão realizadas duas aplicações diferentes, uma com a utilização de uma série temporal hierárquica, e outra com uma série temporal agrupada. O capítulo está organizado em dois subitens, nos quais serão apresentados, respectivamente, uma breve descrição das bases de dados, os resultados encontrados com a aplicação das abordagens de previsão e as previsões 12 passos-à-frente para ambas as aplicações.

### 4.1. Análise das bases de dados

A primeira aplicação apresentada neste trabalho é referente às séries de consumo de energia elétrica no Brasil que estão disponíveis no Sistema gerenciador de séries temporais (SGS) do Banco Central do Brasil (BCB, 2015). O SGS é uma ferramenta que tem como objetivo consolidar, manter a uniformidade e tornar públicas as informações econômico-financeiras de documentos produzidos com base em séries temporais nele armazenadas (BCB, 2015).

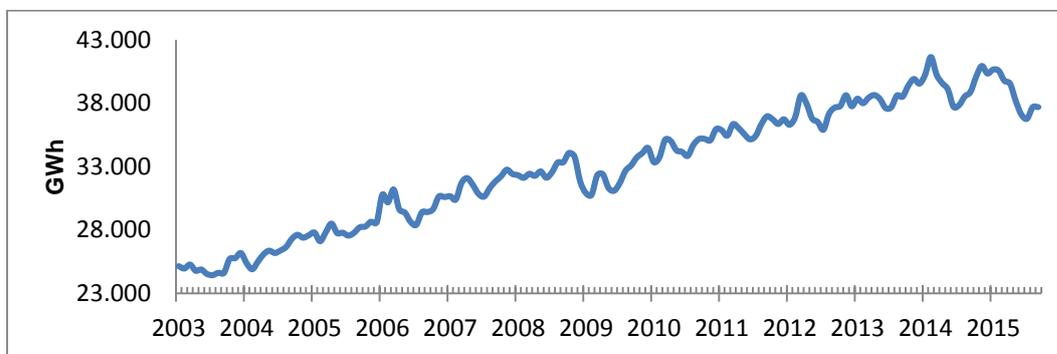
Já a segunda aplicação é referente às séries de demanda de passagem aérea no Brasil. A demanda por transporte aéreo é medida em RPK (*Revenue Passenger Kilometers*), e os dados são disponibilizados mensalmente pela Agência Nacional de Aviação Civil (ANAC, 2015).

Ambas as séries serão detalhadas nos subitens 4.1.1. e 4.1.2. respectivamente.

#### 4.1.1. Séries de consumo de energia

São disponibilizados pelo SGS dados referentes ao consumo de energia elétrica do Brasil desde janeiro de 1979. Essas séries são disponibilizadas pela Eletrobras e é válido ressaltar que o SGS não se responsabiliza por quaisquer deficiências, defasagens ou erros nas informações prestadas como séries temporais cujas fontes são externas ao Banco Central do Brasil (BCB, 2015).

Neste trabalho foram utilizados dados históricos desde janeiro de 2003 a setembro de 2015. A série histórica foi cortada no ano de 2003 devido à crise energética que afetou o país no ano de 2001 com reflexos ainda no ano de 2002. A Figura 7 apresenta a série do total de energia elétrica consumida no país.



**Figura 7 – Consumo Total de energia elétrica no Brasil**

**Fonte:** a autora (2016)

Essas séries tem frequência mensal, sua unidade de medida é *GWh*. O SGS também as fornece subdividas de acordo com a classe de consumo (comercial, industrial, residencial e outros) e também de acordo com as regiões do país (Centro-Oeste, Nordeste, Norte, Sudeste e Sul). Sendo assim essas séries podem ser consideradas como séries agrupadas, pois podem ser agregadas de acordo com duas variáveis de agregação (classe de consumo e região).

O Quadro 6 abaixo apresenta as proporções de consumo de energia elétrica no país de acordo com a classe de consumo e região de acordo com os dados históricos. Nele é possível perceber a proporção de consumo da região nordeste se assemelha a proporção da região sul, enquanto que a proporção de consumo da região norte se assemelha a região centro-oeste, e que a região sudeste é responsável pelo consumo mais elevado.

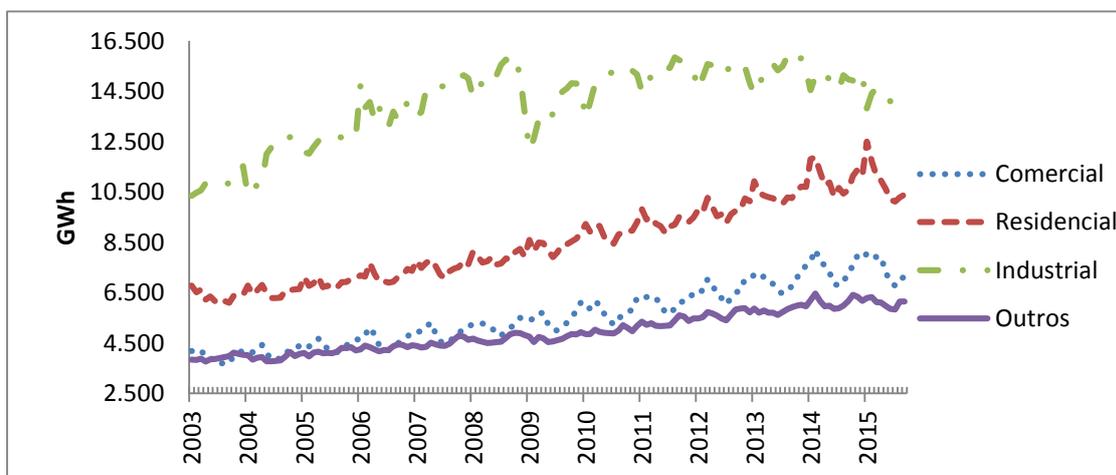
**Quadro 6 – Proporções de consumo de energia elétrica no país**

		Classe de Consumo				Total
		Comercial	Residencial	Industrial	Outros	
Região	Centro-Oeste	1,32%	1,99%	1,72%	1,44%	6,46%
	Nordeste	2,47%	4,54%	6,90%	2,93%	16,84%
	Norte	0,84%	1,44%	3,23%	0,84%	6,35%
	Sudeste	9,39%	13,60%	23,44%	6,63%	53,06%
	Sul	2,87%	4,16%	7,20%	3,05%	17,29%
	Total	16,90%	25,74%	42,48%	14,88%	100%

**Fonte:** a autora (2016)

As estatísticas descritivas das 20 séries referentes ao nível mais baixo da hierarquia são apresentadas no Apêndice B na Tabela 11.

Na Figura 8 é apresentado o gráfico das séries agregadas por classe de consumo. É possível notar que as séries são bem comportadas e apresentam sazonalidade bem definida, com exceção da série de consumo industrial, que apresenta um corte brusco em meados de 2008 quando ocorreu a crise econômica mundial.

**Figura 8 – Consumo de energia elétrica por Classe de Consumo**

**Fonte:** a autora (2016)

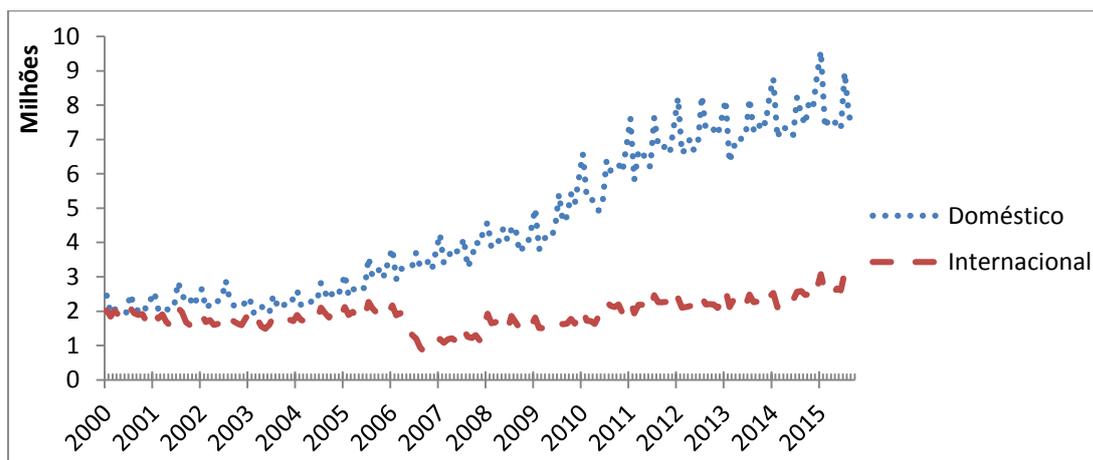
O gráfico das séries agregadas por região do país não foi exibido, pois todas as séries são bem comportadas e apresentam sazonalidade bem definida e noções do comportamento de consumo de cada região podem ser verificadas no Quadro 6.

#### 4.1.2. Séries de demanda de transporte aéreo

A medida RPK é um indicador internacionalmente conhecido do setor de aviação e mede a soma dos produtos entre o total de passageiros pagantes e a distância em quilômetros percorridos pelo voo. É considerada como uma variável que mede a demanda de transporte aéreo de passageiros (ANAC, 2015).

A ANAC (2015) disponibiliza esta série com dados históricos desde janeiro de 2000, tanto para a demanda de voos domésticos, como também para a demanda de voos internacionais. Sendo assim estes dados podem ser considerados como uma série hierárquica, representando a demanda total de transporte aéreo no país, desagregada de acordo com o mercado (doméstico e internacional).

Neste trabalho foram utilizados os dados de janeiro de 2000 até setembro de 2015. A Figura 9 abaixo mostra o comportamento das séries RPK para o mercado doméstico e internacional



**Figura 9 – RPK doméstico e internacional**

**Fonte:** a autora (2016)

Na Figura 9 é possível verificar uma forte queda na demanda de voos internacionais em meados de 2006 devido à crise no setor aéreo brasileiro, conhecida como “apagão aéreo”.

Conforme exibido na Figura 9 é notável que a demanda por voos domésticos é bem maior do que a demanda internacional. Analisando os dados foi verificado que os voos domésticos representam 70% da demanda de transporte aéreo de passageiros, enquanto os voos internacionais representam os outros 30%.

As estatísticas descritivas das séries RPK referente aos voos domésticos e internacionais está apresentada no Apêndice B na Tabela 14.

#### 4.2.

#### Aplicação das abordagens de previsão hierárquica e Previsões

Neste item serão apresentadas as aplicações de cinco abordagens de previsão hierárquica encontradas na literatura, tanto para a série agrupada de consumo de energia elétrica, quanto para a série hierárquica de demanda de transporte aéreo. As abordagens que serão apresentadas são: (i) *Bottom-up*, (ii) *Top-down* (cálculo feito com base na metodologia (i) apresentada em 3.3.), (iii) *Top-Down* baseado a Proporção da Previsão, (iv) Combinação ótima utilizando Mínimos Quadrados Ordinários e (v) Combinação Ótima utilizando Mínimos Quadrados Ponderados.

Para realizar as previsões das séries foram utilizados os métodos de Amortecimento Exponencial e modelos ARIMA, pois além de serem métodos vastamente conhecidos na área de Séries Temporais, também foram os métodos mais utilizados nos trabalhos encontrados na revisão da literatura.

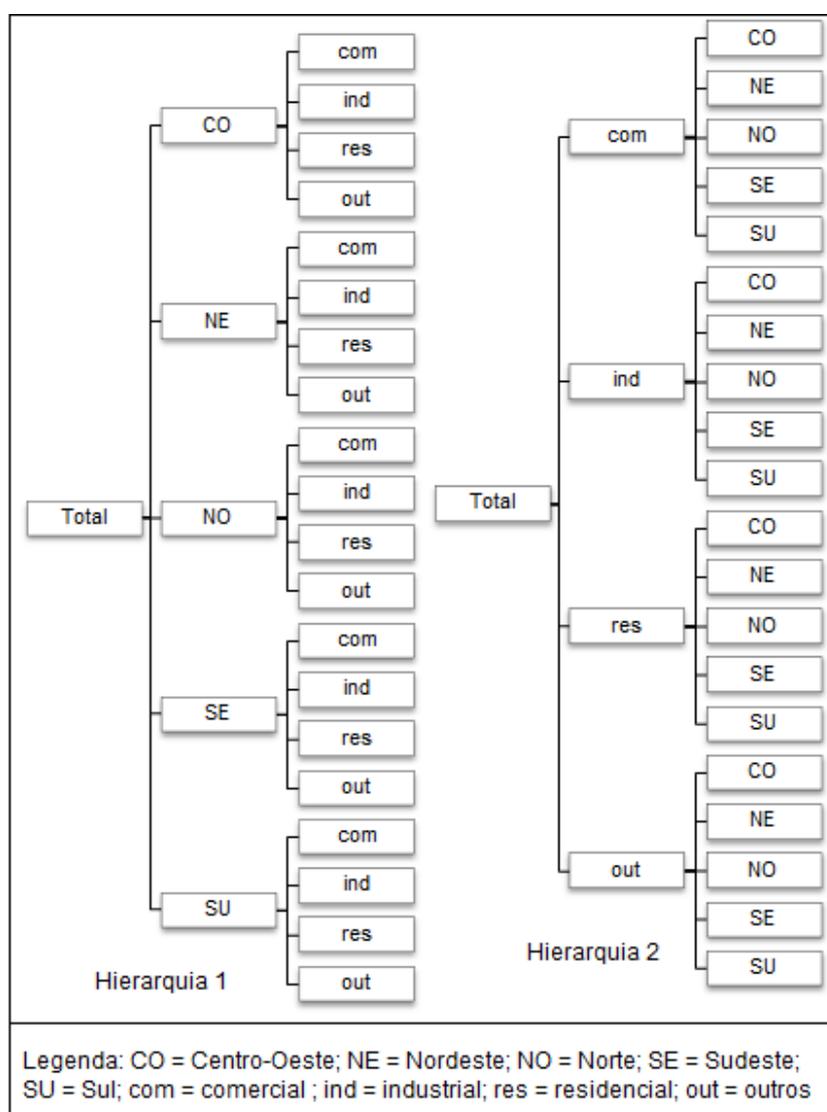
As aplicações foram implementadas no software R (R Core Team, 2013), utilizando os pacotes “*hts*” (Hyndman et al., 2014b) e o pacote “*forecast*” (Hyndman, 2015; Hyndman & Khandakar, 2008). O código utilizado será apresentado no Apêndice A.

O pacote “*hts*” organiza as séries temporais em uma estrutura hierárquica ou agrupada, e nele, estão implementadas todas as abordagens de previsão hierárquica supracitadas. Para realizar as previsões foi utilizada o pacote “*forecast*”, para o caso dos modelos ARIMA os parâmetros  $p$  e  $q$  são gerados automaticamente, enquanto para o método de Amortecimento Exponencial é utilizado uma abordagem de Espaços de Estados conforme exposto na Tabela 4 anteriormente.

Para comparar o desempenho das 5 abordagens de previsão hierárquicas testadas será utilizada a métrica de desempenho MAPE (*Mean Absolute Percentage Error*), e também foi aplicado o teste de Diebold-Mariano, que está implementado no pacote *forecast*, para avaliar se a diferença entre o desempenho das previsões é significativa.

### 4.2.1. Séries de energia

A Figura 10 abaixo mostra as duas hierarquias que podem ser montadas com as séries de consumo de energia elétrica, que é considerada como uma série agrupada. A Hierarquia 1 é composta por 26 séries no total, sendo 1 no nível 0 que representa o consumo total de energia elétrica no país, 5 no nível 1 que representam as regiões do país e 20 no nível desagregado. Já a Hierarquia 2 é composta por 25 séries no total com a série total no nível 0, no nível 1 estão 4 séries agregadas por classe de consumo, e no nível 2, as 20 séries desagregadas por classe de consumo e região.



**Figura 10 - Estruturas das séries de energia**

**Fonte:** a autora (2016)

Para comparar o desempenho de cada uma das abordagens de previsão hierárquica citadas acima, foram utilizados os dados históricos de janeiro de 2003 até dezembro de 2014, e posteriormente foram calculadas as previsões de 1 até 9 passos-à-frente para cada uma das séries. Os nove últimos dados históricos (de janeiro a setembro de 2015) foram utilizados para calcular o MAPE *out-of-sample* das previsões.

As Tabela 5 e 6 apresentam respectivamente, um resumo dos resultados encontrados utilizando o método de amortecimento exponencial e modelos ARIMA para realizar as previsões. As proporções apresentadas no Quadro 6 foram utilizadas para calcular uma média ponderada do MAPE de cada uma das séries dos níveis 1 e 2, para as Hierarquias 1 e 2 mostradas anteriormente na Figura 10. A última coluna das Tabelas apresenta uma média do desempenho dos 9 horizontes de previsão para cada uma das abordagens.

**Tabela 5 – MAPE *out-of-sample* da previsão com amortecimento exponencial das séries de energia**

Método de Previsão: Amortecimento Exponencial										
MAPE	Horizonte de Previsão (h)									Média
	1	2	3	4	5	6	7	8	9	
Nível 0 – Total										
BU	0,86	0,53	1,74	2,25	2,88	<b>3,50</b>	<b>4,10</b>	<b>4,40</b>	<b>4,70</b>	2,77
TD	0,12	0,25	1,53	2,21	3,14	4,00	4,80	5,30	5,80	3,02
TDFP	0,12	0,25	1,53	2,21	3,14	4,00	4,80	5,30	5,80	3,02
CO – MQO	<b>0,002</b>	<b>0,18</b>	<b>1,41</b>	2,01	2,81	3,58	4,30	4,70	5,10	2,68
CO – MQP	0,08	0,19	1,42	<b>2,01</b>	<b>2,79</b>	3,54	4,20	4,60	5,00	<b>2,65</b>
Nível 1 - Agregação por região										
BU	1,11	1,11	2,16	2,57	3,15	<b>3,72</b>	<b>4,30</b>	<b>4,55</b>	<b>4,91</b>	3,06
TD	4,08	4,15	4,85	5,27	6,01	6,75	7,54	8,06	8,73	6,16
TDFP	1,08	1,14	2,12	2,65	3,50	4,36	5,04	5,58	6,08	3,51
CO – MQO	0,76	0,86	1,87	2,37	3,11	3,87	4,50	4,92	5,38	3,07
CO – MQP	<b>0,59</b>	<b>0,70</b>	<b>1,76</b>	<b>2,27</b>	<b>3,01</b>	3,73	4,37	4,75	5,22	<b>2,93</b>
Nível 1 - Agregação por classe de consumo										
BU	4,71	3,58	3,79	3,80	4,14	<b>4,58</b>	<b>4,99</b>	<b>5,20</b>	<b>5,48</b>	4,47
TD	19,23	17,52	16,02	15,21	14,52	13,91	13,46	13,22	13,16	15,14
TDFP	<b>3,79</b>	<b>2,94</b>	<b>3,31</b>	3,55	4,24	4,95	5,60	6,01	6,46	4,54
CO – MQO	4,05	3,12	3,35	<b>3,49</b>	<b>4,00</b>	4,59	5,14	5,49	5,86	<b>4,34</b>
CO – MQP	4,33	3,32	3,50	3,60	4,07	4,61	5,13	5,47	5,81	4,43
Nível 2 - Séries desagregadas por região e classe										
BU	5,77	4,70	4,66	4,48	4,76	<b>5,13</b>	<b>5,50</b>	<b>5,66</b>	<b>5,94</b>	5,18
TD	20,02	18,17	17,07	16,56	16,31	16,04	15,90	15,81	16,07	16,88
TDFP	5,82	4,73	4,56	4,52	5,04	5,63	6,21	6,59	7,02	5,57
CO – MQO	<b>5,30</b>	<b>4,34</b>	<b>4,29</b>	<b>4,21</b>	<b>4,63</b>	5,14	5,64	5,95	6,31	<b>5,09</b>
CO – MQP	5,91	4,83	4,64	4,47	4,80	5,24	5,72	5,96	6,29	5,32
Nível 2 - Séries desagregadas por classe e região										
BU	5,77	4,70	4,66	4,48	4,76	<b>5,13</b>	<b>5,50</b>	<b>5,66</b>	<b>5,94</b>	5,18
TD	20,02	18,17	17,07	16,56	16,31	16,04	15,90	15,81	16,07	16,88
TDFP	<b>5,24</b>	4,39	4,42	4,42	4,98	5,59	15,90	6,56	6,99	6,50
CO – MQO	5,30	<b>4,34</b>	<b>4,29</b>	<b>4,21</b>	<b>4,63</b>	5,14	5,64	5,95	6,31	<b>5,09</b>
CO – MQP	5,91	4,83	4,64	4,47	4,80	5,24	5,72	5,96	6,29	5,32

Legenda: BU = Bottom-up; TD = Top-Down; TDFP = Top-Down baseado na proporção da previsão; CO-MQO = Combinação ótima por mínimos quadrados ordinários; CO-MQP = Combinação ótima por mínimos quadrados ponderados.

**Fonte:** a autora (2016)

Utilizando o método de amortecimento exponencial para realizar as previsões nota-se que a abordagem de Combinação Ótima, seja utilizando MQO

ou MQP, apresentou na média, resultados melhores do que todas as outras abordagens de previsão hierárquica.

**Tabela 6 – MAPE *out-of-sample* da previsão com modelos ARIMA das séries de energia**

Método de Previsão: ARIMA										
MAPE	Horizonte de Previsão (h)									Média
	1	2	3	4	5	6	7	8	9	
Nível 0 – Total										
BU	1,10	1,81	2,50	2,71	3,42	4,10	4,80	5,10	5,50	3,45
TD	<b>0,01</b>	<b>0,82</b>	<b>1,28</b>	<b>1,47</b>	<b>2,20</b>	<b>2,94</b>	<b>3,63</b>	<b>3,90</b>	<b>4,20</b>	<b>2,27</b>
TDFP	<b>0,01</b>	<b>0,82</b>	<b>1,28</b>	<b>1,47</b>	<b>2,20</b>	<b>2,94</b>	<b>3,63</b>	<b>3,90</b>	<b>4,20</b>	<b>2,27</b>
CO – MQO	0,41	1,17	1,78	2,01	2,75	3,48	4,20	4,50	4,90	2,80
CO – MQP	0,51	1,26	1,87	2,10	2,84	3,55	4,30	4,60	4,90	2,88
Nível 1 – Hierarquia 1 (agregação por região)										
BU	1,13	1,86	2,54	2,73	3,46	4,12	4,80	5,14	5,52	3,48
TD	4,08	4,69	5,01	5,19	5,66	6,22	6,89	7,23	7,60	5,84
TDFP	<b>0,40</b>	<b>1,18</b>	<b>1,53</b>	<b>1,82</b>	<b>2,50</b>	<b>3,20</b>	<b>3,91</b>	<b>4,18</b>	<b>4,49</b>	<b>2,58</b>
CO – MQO	0,50	1,34	1,90	2,15	2,89	3,59	4,33	4,70	5,03	2,94
CO – MQP	0,58	1,40	2,00	2,23	2,96	3,67	4,38	4,77	5,17	3,02
Nível 1 – Hierarquia 2 (agregação por classe de consumo)										
BU	4,33	3,57	3,67	3,64	4,17	4,68	5,26	5,59	5,88	4,53
TD	19,17	17,58	16,01	15,25	14,48	13,77	13,18	12,90	12,81	15,02
TDFP	4,37	3,24	<b>2,90</b>	<b>2,77</b>	<b>3,23</b>	<b>3,77</b>	<b>4,32</b>	<b>4,53</b>	<b>4,73</b>	<b>3,76</b>
CO – MQO	<b>4,12</b>	<b>3,20</b>	3,13	3,08	3,61	4,18	4,78	5,03	5,35	4,06
CO – MQP	4,41	3,46	3,34	3,26	3,76	4,31	4,88	5,17	5,45	4,23
Nível 2 – Hierarquia 1										
BU	4,85	3,93	3,96	3,97	4,45	4,94	5,53	5,81	6,11	4,84
TD	19,97	18,32	17,01	16,32	16,01	15,56	15,37	15,17	15,23	16,55
TDFP	4,84	3,72	<b>3,35</b>	<b>3,37</b>	<b>3,81</b>	<b>4,33</b>	<b>4,88</b>	<b>5,13</b>	<b>5,39</b>	<b>4,31</b>
CO – MQO	<b>4,68</b>	<b>3,66</b>	3,48	3,51	3,99	4,52	5,13	5,38	5,68	4,45
CO – MQP	4,98	3,93	3,74	3,75	4,20	4,74	5,33	5,62	5,92	4,69
Nível 2 – Hierarquia 2										
BU	4,85	3,93	3,96	3,97	4,45	4,94	5,53	5,81	6,11	4,84
TD	19,97	18,32	17,01	16,32	16,01	15,56	15,37	15,17	15,23	16,55
TDFP	4,94	<b>3,74</b>	<b>3,38</b>	<b>3,30</b>	<b>3,67</b>	<b>4,18</b>	<b>4,69</b>	<b>4,90</b>	<b>5,12</b>	<b>4,22</b>
CO – MQO	<b>4,68</b>	3,66	3,48	3,51	3,99	4,52	5,13	5,38	5,68	4,45
CO – MQP	4,98	3,93	3,74	3,75	4,20	4,74	5,33	5,62	5,92	4,69

Legenda: BU = Bottom-up; TD = Top-Down; TDFP = Top-Down baseado na proporção da previsão; CO-MQO = Combinação ótima por mínimos quadrados ordinários; CO-MQP = Combinação ótima por mínimos quadrados ponderados.

**Fonte:** a autora (2016)

Já observando os resultados gerados pelos modelos ARIMA pode-se verificar que a abordagem *Top-Down* baseada na Proporção das Previsões apresentou o melhor desempenho médio, seguido pela abordagem de Combinação Ótima.

Ao observar os resultados no nível 0 das Tabelas 5 e 6 nota-se que as abordagens *Top-Down* e *Top-Down* baseada na Proporção das Previsões apresentam os mesmos resultados, o que faz sentido visto que as previsões são as mesmas para o nível superior, o que diferencia uma abordagem da outra são as proporções dos níveis mais baixos.

Já observando somente o nível 2 na Tabelas 5 ou na Tabela 6 é possível notar que o MAPE das abordagens *Bottom-up* e de Combinação Ótima são iguais, o que já era esperado, pois as apesar de organizadas de maneiras diferentes, as previsões das séries do nível 2 são iguais para ambas as Hierarquias. O mesmo ocorre com a abordagem *Top-Down*, pois como as proporções são baseadas nos dados históricos, elas não mudarão da hierarquia 1 para a hierarquia 2. Pode-se notar que o mesmo não ocorre para a abordagem *Top-Down* baseada na Proporção da Previsão, pois como as proporções para desagregar as previsões são diferentes, logo o MAPE também será diferente.

A abordagem *Top-Down* tradicional apresentou na média o pior desempenho dentre as demais abordagens de previsão hierárquica, tendo gerado o melhor resultado somente no nível mais agregado, e somente utilizando os modelos ARIMA para previsão. Esse resultado já era esperado, pois a maioria dos trabalhos encontrados na literatura apresentaram resultados melhores para outras abordagens de previsão hierárquica.

Apesar do bom desempenho no nível zero da abordagem *Top-Down* tradicional, é válido ressaltar que a abordagem *Top-Down* baseada na Proporção da Previsão sempre apresentará o mesmo desempenho neste nível, e um desempenho possivelmente melhor nos outros, sendo assim pode-se concluir que se for necessário escolher aplicar somente uma dessas duas abordagens, é melhor aplicar a abordagem *Top-Down* baseada na Proporção da Previsão.

Agora comparando os resultados em relação às hierarquias 1 e 2 nota-se que a agregação por região no nível 1 gera previsões mais acuradas do que as previsões feitas quando a agregação é feita por classe de consumo, pois apresenta

uma MAPE menor. Entretanto a agregação por classe de consumo é outro ponto de vista e não pode ser desprezado, principalmente considerando que ela pode ser o objetivo da previsão. Sendo assim, uma das vantagens de considerarmos uma série como agrupada, e não como hierárquica, é podermos ter todas as perspectivas de previsão no nível intermediário, não sendo necessário descartar quaisquer variáveis de agregação.

Dado o desempenho das abordagens de previsão hierárquica mostrado acima, foi realizado o teste de Diebold-Mariano para verificar se a diferença no desempenho entre as abordagens tradicionais e as novas é significativa. O teste foi realizado somente com os erros gerados pelos modelos ARIMA, visto que esses apresentaram um melhor desempenho que as previsões realizadas com o método de amortecimento exponencial.

O teste foi realizado comparando o desempenho da abordagem *Bottom-up* com o desempenho que cada uma das demais abordagens. As Tabelas 7 e 8 abaixo apresentam um resumo dos resultados com a proporção de vezes que as respectivas abordagens tiveram desempenho significativamente melhor ou pior do que a abordagem *Bottom-up*, considerando os erros das previsões de 1 até 9 passos-à-frente.

**Tabela 7 – Teste de Diebold-Mariano para a hierarquia 1**

	Hierarquia 1							
	CO - MQO melhor	CO - MQO pior	CO - MQP melhor	CO - MQP pior	TDFP melhor	TDFP pior	TD melhor	TD pior
nível 0	0,89	0,00	1,00	0,00	1,00	0,00	1,00	0,00
nível 1	0,53	0,00	0,56	0,02	0,56	0,00	0,07	0,38
nível 2	0,48	0,11	0,42	0,18	0,53	0,04	0,07	0,47

**Fonte:** a autora (2016)

**Tabela 8 – Teste de Diebold-Mariano para a hierarquia 2**

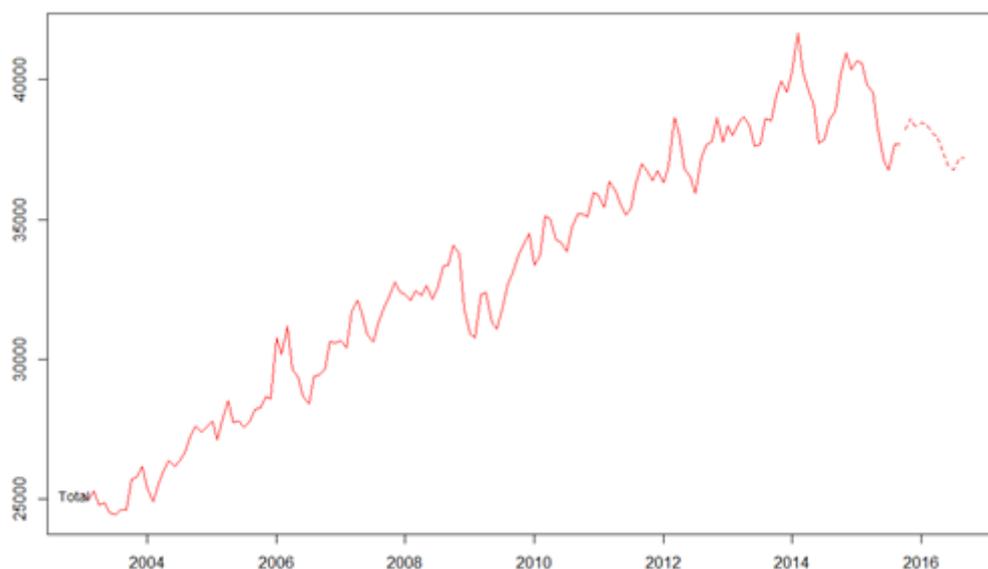
	Hierarquia 2							
	CO - MQO melhor	CO - MQO pior	CO - MQP melhor	CO - MQP pior	TDFP melhor	TDFP pior	TD melhor	TD pior
nível 0	0,89	0,00	1,00	0,00	1,00	0,00	1,00	0,00
nível 1	0,75	0,00	0,50	0,19	0,64	0,00	0,14	0,25
nível 2	0,48	0,11	0,42	0,18	0,56	0,04	0,06	0,47

**Fonte:** a autora (2016)

Inicialmente comparando somente o desempenho das abordagens tradicionais (*Bottom-up* e *Top-Down*) é possível verificar que no nível zero a abordagem *Top-Down* sempre apresenta um desempenho significativamente superior ao da abordagem *Bottom-up* conforme já era esperado. Já observando os níveis inferiores, a abordagem *Top-Down* apresenta na maioria das vezes um desempenho igual ou pior do que a abordagem *Bottom-up*, o que está de acordo com os resultados encontrados neste trabalho e na literatura.

Já em relação às abordagens mais recentes de previsão hierárquica (Combinação Ótima e *Top-Down* baseado na Proporção da Previsão) nota-se que essas apresentaram na maioria das vezes um desempenho significativamente melhor ou igual à abordagem tradicional *Bottom-up*.

Segue abaixo, na Figura 11, um gráfico da previsão 12 passos-à-frente pontilhadas utilizando um modelo ARIMA e a abordagem *Top-Down* baseada na Proporção das Previsões, que apresentou o melhor desempenho neste caso.



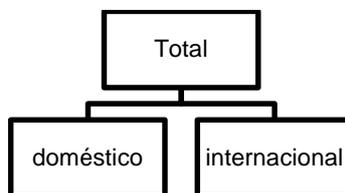
**Figura 11 – Previsão para o Total de consumo de energia elétrica**

**Fonte:** a autora (2016)

As tabelas com as previsões detalhadas de todas as séries de ambas as hierarquias será exibida no Apêndice B nas Tabelas 12 e 13.

#### 4.2.2. Séries RPK

Nesta aplicação serão utilizadas 3 séries distribuídas conforme a estrutura hierárquica apresentada abaixo na Figura 12.



**Figura 12 - Hierarquia da série RPK**

**Fonte:** a autora (2016)

Foram utilizados dados históricos da série RPK de janeiro de 2000 até dezembro de 2014 para gerar as previsões de 1 até 9 passos-à-frente. Os nove últimos dados históricos (de janeiro a setembro de 2015) ocultados foram utilizados para calcular o MAPE *out-of-sample* das previsões. As proporções supracitadas de 70% e 30% foram utilizadas para calcular uma média ponderada do MAPE de cada uma das séries do níveis 1. As Tabela 9 e 10 apresentam, respectivamente, um resumo dos resultados encontrados utilizando o método de amortecimento exponencial e modelos ARIMA para realizar as previsões. A última coluna das Tabelas 9 e 10 apresenta uma média do desempenho dos 9 horizontes de previsão para cada uma das abordagens

**Tabela 9 – MAPE *out-of-sample* da previsão com amortecimento exponencial das séries RPK**

Método de Previsão: Amortecimento Exponencial										
<u>MAPE</u>	Horizonte de Previsão (h)									Média
	1	2	3	4	5	6	7	8	9	
Nível 0 – Total										
BU	0,88	2,80	4,50	5,18	5,52	6,20	6,30	6,50	6,80	4,96
TD	0,07	2,52	4,35	4,83	4,84	5,35	5,32	5,46	5,60	4,26
TDFP	<b>0,07</b>	<b>2,52</b>	<b>4,35</b>	<b>4,83</b>	<b>4,84</b>	<b>5,35</b>	<b>5,32</b>	<b>5,46</b>	<b>5,60</b>	<b>4,26</b>
CO – MQO	0,20	2,54	4,35	4,91	5,01	5,58	5,60	5,80	6,00	4,44
CO – MQP	0,25	2,56	4,36	4,93	5,04	5,62	5,60	5,80	6,00	4,46
Nível 1 - Agregação por mercado										
BU	2,79	<b>3,68</b>	<b>5,09</b>	5,57	6,17	6,86	7,56	8,33	8,90	6,11
TD	17,63	15,43	15,29	15,22	14,46	13,85	13,65	12,79	12,23	14,50
TDFP	3,21	4,02	5,35	<b>5,50</b>	<b>5,98</b>	<b>6,53</b>	<b>7,14</b>	<b>7,87</b>	<b>8,41</b>	<b>6,00</b>
CO – MQO	2,94	3,86	5,23	5,50	6,05	6,67	7,35	8,15	8,69	6,05
CO – MQP	<b>2,70</b>	3,73	5,15	5,45	6,13	6,67	7,48	8,29	8,92	6,06

**Fonte:** a autora (2016)

Assim como o observado no caso das séries de consumo de energia, a abordagem *Top-Down* apresentou um bom desempenho somente para o nível mais alto da hierarquia.

Observando os resultados na Tabela 9 vemos que abordagem de previsão *Top-Down* baseada na Proporção da Previsão apresentou um resultado médio melhor do que as demais, seguida pela abordagem de Combinação Ótima.

**Tabela 10 - MAPE *out-of-sample* da previsão com modelos ARIMA das séries RPK**

Método de Previsão: ARIMA										
MAPE	Horizonte de Previsão (h)									Média
	1	2	3	4	5	6	7	8	9	
Nível 0 – Total										
BU	<b>3,47</b>	<b>3,09</b>	<b>3,88</b>	<b>4,18</b>	<b>4,40</b>	<b>4,70</b>	<b>4,19</b>	<b>4,14</b>	<b>4,40</b>	<b>4,05</b>
TD	3,50	3,12	4,08	4,45	4,71	5,07	4,40	4,46	4,81	4,29
TDFP	3,50	3,12	4,08	4,45	4,71	5,07	4,40	4,46	4,81	4,29
CO – MQO	3,49	3,11	4,02	4,40	4,61	4,96	4,37	4,37	4,70	4,23
CO – MQP	3,49	3,11	4,01	4,40	4,60	4,94	4,36	4,36	4,69	4,22
Nível 1 - Agregação por mercado										
BU	<b>3,68</b>	<b>3,12</b>	<b>3,95</b>	<b>4,21</b>	<b>4,37</b>	<b>4,69</b>	<b>4,47</b>	<b>5,01</b>	<b>5,45</b>	<b>4,33</b>
TD	18,47	15,86	15,97	15,73	14,86	14,20	14,11	13,33	12,79	15,04
TDFP	3,70	3,15	4,15	4,50	4,68	5,01	4,78	5,33	5,81	4,57
CO – MQO	3,69	3,14	4,09	4,42	4,60	4,90	4,68	5,22	5,67	4,49
CO – MQP	3,69	3,14	4,09	4,42	4,60	4,90	4,65	5,18	5,60	4,47

**Fonte:** a autora (2016)

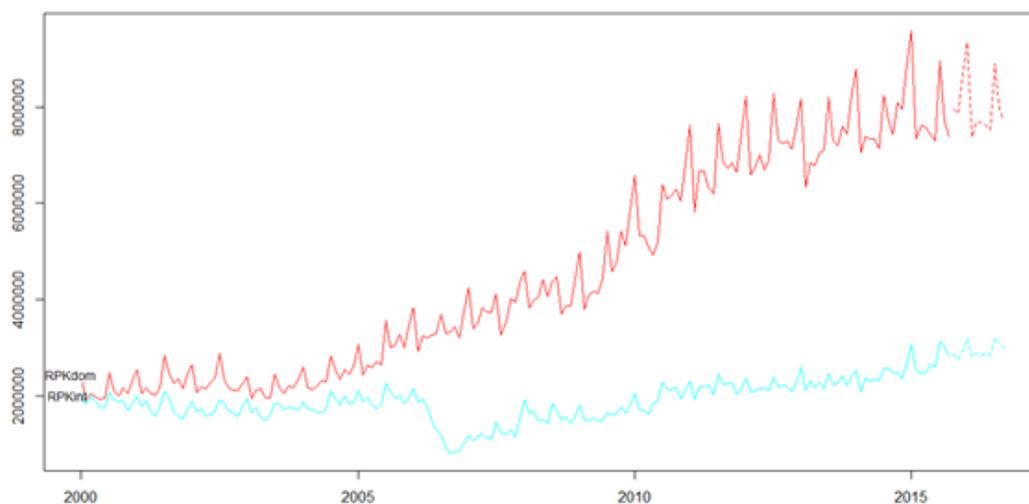
Já observando os resultados da Tabela 10 verifica-se que a abordagem *Bottom-up* gerou os melhores resultados, o que foi uma surpresa, entretanto acredita-se que esse desempenho superior ocorreu devido às características das séries de demanda de transporte aéreo. As séries do nível 1 claramente tem um comportamento de tendência muito diferenciado, como pode ser observado na Figura 9, o que leva a crer que a previsão considerando as séries separadas gere melhores resultados.

Esperava-se que as novas abordagens apresentassem um desempenho superior às tradicionais, entretanto o desempenho superior da abordagem *Bottom-up* está de acordo com alguns resultados encontrados na literatura. Nos estudos apresentados por Hyndman et al. (2011) e Athanasopoulos et al. (2009) as novas abordagens não geram os melhores resultados sempre. De acordo com as conclusões encontradas por Chen & Boylan (2009) quando as séries do nível mais baixo da hierarquia apresentam correlação alta, a abordagem *Bottom-up* gera bons resultados, e as séries do nível mais baixo da hierarquia RPK apresentam uma correlação de 0,7.

Apesar do desempenho da abordagem *Bottom-up*, destaca-se que a abordagem que apresentou o segundo melhor desempenho foi a de Combinação

Ótima, ou seja, mesmo não sendo a melhor deve-se ressaltar que a abordagem tem apresentado bons resultados na média, e não deve ser desconsiderada.

Segue abaixo na Figura 13 um gráfico da previsão 12 passos-à-frente das séries do nível 1, RPK doméstico e internacional, feita com modelos ARIMA e abordagem Bottom-up visto que essa gerou os melhores resultados



**Figura 13 – Previsões 12 passos-à-frente para as séries de demanda de transporte de passageiros em voos domésticos e internacionais**

**Fonte:** a autora (2016)

As previsões detalhadas para a série total de demanda de transporte aéreo de passageiros, e para as séries de demanda por voos domésticos e internacionais está exibida na Tabela 15 no Anexo B.

## 5 Conclusões

Neste trabalho foram apresentadas as diversas abordagens de previsão hierárquica existentes na literatura, dando destaque para o estado da arte que são as abordagens: Combinação Ótima e *Top-Down* baseado na Proporção das Previsões.

Esse trabalho apresentou os resultados gerados pelas diversas abordagens de previsão hierárquica, sejam as tradicionais ou novas, para séries hierárquicas e agrupadas clássicas do contexto brasileiro. Acredita-se que este trabalho contribuiu com a disseminação das novas abordagens de previsão hierárquica, visto que foram encontrados na literatura somente quatro trabalhos referentes às mesmas, sendo três dos autores que as propuseram, e somente um outro estudo empírico realizado por Capistran et al. (2010) considerando a abordagem de Combinação Ótima. Já a abordagem *Top-Down* baseada na Proporção das Previsões teve uma aplicação apresentada somente em Athanasopoulos et al. (2009) e Hyndman et al. (2016).

Outro diferencial deste trabalho em relação aos encontrados na literatura foi o desenvolvimento de uma aplicação com uso de uma Série Temporal Agrupada, com duas variáveis de agregação, diferentemente dos outros trabalhos encontrados na literatura que consideraram somente Séries Temporais Hierárquicas.

Em ambos os estudos empíricos apresentados o desempenho preditivo das novas abordagens de previsão hierárquica apresentou bons resultados. Mesmo quando uma abordagem tradicional apresentou o melhor resultado, as novas abordagens apresentaram resultados bem próximos à tradicional.

Tem-se como proposta de trabalho futuro expandir os estudos sobre os fatores que impactam o desempenho das abordagens de previsão hierárquicas tradicionais (como por exemplo, a correlação das séries no nível mais baixo da hierarquia) para as novas abordagens encontradas na literatura. Ainda como proposta de trabalho futuro, tem-se a combinação de previsões geradas por duas

ou mais abordagens de modo a verificar se ocorrerá melhora no desempenho da previsão.

Uma das lacunas encontradas na literatura é que a maioria dos estudos que comparam o desempenho das abordagens consideram situações muito específicas, por exemplo, Chen e Boylan (2009) consideram situações em que as séries tem somente componentes de tendência, Widiarta (2009) considera somente séries geradas por um processo MA(1). Existe a necessidade de desenvolvimento de um estudo mais abrangente que compare o desempenho das abordagens de previsão considerando séries com características variadas.

Todas as abordagens de previsão hierárquica citadas neste trabalho foram implementadas no pacote “*hts*”, conforme citado anteriormente, sendo a praticidade e facilidade de aplicação uma das maiores vantagens do pacote. Entretanto algumas deficiências foram encontradas no mesmo, o que gerou algumas limitações no trabalho.

Uma das maiores limitações encontradas no pacote, que também impactaram os resultados da dissertação, é que apesar das previsões das séries da hierarquia serem feitas de maneira independente, o pacote só permite que seja aplicado um único método de previsão para todas as séries. O fato de que métodos diferentes podem gerar melhores resultados para séries diferentes dentro da mesma hierarquia não foi considerado.

Outra deficiência encontrada no pacote “*hts*” é que os modelos utilizados para ajustar as séries temporais são gerados automaticamente, entretanto seus parâmetros não são exibidos. Além disso, séries agrupadas com mais de duas variáveis de agregação também ainda não podem ser analisadas no pacote “*hts*”.

Outra proposta de trabalho futuro é analisar o código do pacote “*hts*” e implementar as possíveis soluções para as limitações supracitadas no *software* R.

É válido ressaltar que até o momento não foi encontrado na literatura nenhum trabalho que comprove que o desempenho de uma determinada abordagem de previsão hierárquica é sempre superior às outras. Sendo assim todas as abordagens devem ser sempre testadas, tanto as tradicionais como as mais recentes. O nível de detalhamento da informação de interesse também deve ser considerado, pois determinada abordagem pode ser melhor que outra em determinado nível.

## 6 Referências Bibliográficas

AGÊNCIA NACIONAL DE AVIAÇÃO CIVIL (ANAC). Dados Estatísticos do Transporte Aéreo – Empresas Brasileiras, Setembro de 2015. Brasília, DF, 2015. Disponível em: <http://www2.anac.gov.br/ofertaedemanda/ofertaedemanda.asp>. Acesso em: 13 de novembro de 2015.

ATHANASOPOULOS, G.; AHMED, R. A.; HYNDMAN, R. J. Hierarchical forecasts for Australian domestic tourism. **International Journal of Forecasting**, v. 25, n. 1, p. 146-166, Jan-Mar 2009.

BANCO CENTRAL DO BRASIL (BCB). Economia e finanças: indicadores de conjuntura: sistema gerenciador de séries temporais: economia regional: setor real. Brasília, DF, 2015. Disponível em: <https://www3.bcb.gov.br/sqspub/consultarvalores/telaCvsSelecionarSeries.paint>. Acesso em: 09 de novembro de 2015.

CAPISTRAN, C.; CONSTANDSE, C.; RAMOS-FRANCIA, M. Multi-horizon inflation forecasts using disaggregated data. **Economic Modelling**, v. 27, n. 3, p. 666-677, May 2010.

CHEN, A.; BLUE, J. Performance analysis of demand planning approaches for aggregating, forecasting and disaggregating interrelated demands. **International Journal of Production Economics**, v. 128, n. 2, p. 586-602, Dec 2010.

CHEN, H.; BOYLAN, J. E. The effect of Correlation between demands on hierarchical forecasting. **Advances in Business and Management Forecasting**, v. 6, p. 173-188, 2009.

DANGERFIELD, B. J.; MORRIS, J. S. Top-down or bottom-up: Aggregate versus disaggregate extrapolations. **International Journal of Forecasting**, v. 8, n. 2, p. 233-241, Oct 1992.

DIEBOLD, F. X.; MARIANO, R. S. Comparing predictive accuracy. **Journal of Business and Economic Statistics**, v. 13, p. 253-263, 1995.

FLIEDNER, E. B.; LAWRENCE, B. Forecasting system parent group formation: an empirical application of cluster analysis. **Journal of Operations Management**, v. 12, n. 2, p. 119–130, 1995.

FLIEDNER, G. An investigation of aggregate variable time series forecast strategies with specific subaggregate time series statistical correlation. **Computers & Operations Research**, v. 26, n. 10-11, p. 1133-1149, Sep 1999.

\_\_\_\_\_. Hierarchical forecasting: issues and use guidelines. **Industrial Management & Data Systems**, v. 101, n. 1-2, p. 5-12, 2001.

HYNDMAN, R. J. forecast: Forecasting functions for time series and linear models. R package version 5.9. <http://github.com/robjhyndman/forecast>. 2015.

HYNDMAN, R. J.; AHMED, R. A.; ATHANASOPOULOS, G. Optimal combination forecasts for hierarchical time series. Working paper. Department of Econometrics and Business Statistics, Monash University. Sep 2007.

HYNDMAN, R.J.; ATHANASOPOULOS, G. Forecasting: principles and practice. Oct 2013. Disponível em: <http://otexts.org/fpp/>. Acesso em: 29 de abril de 2015.

HYNDMAN, R. J.; ATHANASOPOULOS, G. Optimally Reconciling Forecasts in a Hierarchy. *Foresight: The International Journal of Applied Forecasting*. 35. ed., p. 42-50. Fall, 2014.

HYNDMAN, R. J.; KHANDAKAR, Y. Automatic time series forecasting: the forecast package for R. **Journal of Statistical Software**, v. 26, n. 3, p. 1-22, 2008.

HYNDMAN, R. J.; LEE, A. J.; WANG, E. Fast computation of reconciled forecasts for hierarchical and grouped time series. Working paper. Department of Econometrics and Business Statistics, Monash University. Jun 2014a.

HYNDMAN, R. J.; LEE, A. J.; WANG, E. com contribuições de AHMED, R. A. e SHANGHTS, H. L. Hierarchical and grouped time series. R package version 4.4. <http://CRAN.R-project.org/package=hts>. 2014b.

HYNDMAN, R. J.; LEE, A. J.; WANG, E. Fast computation of reconciled forecasts for hierarchical and grouped time series. **Computational Statistics and Data Analysis**, v. 97, p. 16–32, 2016.

HYNDMAN, R. J. et al. **Forecasting with exponential smoothing: the state space approach**. 1. ed. [S.1]: Springer, 2008.

HYNDMAN, R. J. et al. Optimal combination forecasts for hierarchical time series. **Computational Statistics & Data Analysis**, v. 55, n. 9, p. 2579-2589, Sep 2011.

KOURENTZES, N.; PETROPOULOS, F.; TRAPERO, J. R. Improving forecasting by estimating time series structural components across multiple frequencies. **International Journal of Forecasting**, v. 30, p. 291–302, 2014.

LENORT, R.; BESTA, P. Hierarchical Sales Forecasting System for Apparel Companies and Supply Chains. **Fibres & Textiles in Eastern Europe**, v. 21, n. 6, p. 7-11, Nov-Dec 2013.

LO, S.-L.; WANG, F.-K.; LIN, J. T. Forecasting for the LCD monitor market. **Journal of Forecasting**, v. 27, n. 4, p. 341-356, Jul 2008.

MOON, S.; HICKS, C.; SIMPSON, A. The development of a hierarchical forecasting method for predicting spare parts demand in the South Korean Navy-A case study. **International Journal of Production Economics**, v. 140, n. 2, p. 794-802, Dec 2012.

MOON, S.; SIMPSON, A.; HICKS, C. The development of a classification model for predicting the performance of forecasting methods for naval spare parts demand. **International Journal of Production Economics**, v. 143, n. 2, p. 449-454, Jun 2013.

MORETTIN, P. A.; TOLÓI, C. M. C. **Análise de séries temporais**. 2. ed. São Paulo: Edgard Blücher, 2006. 564 p. (Coleção Projeto Fisher).

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>. 2013.

SBRANA, G.; SILVESTRINI, A. Forecasting aggregate demand: Analytical comparison of top-down and bottom-up approaches in a multivariate exponential smoothing framework. **International Journal of Production Economics**, v. 146, n. 1, p. 185-198, Nov 2013.

SCHWARZKOPF, A. B.; TERSINE, R. J.; MORRIS, J. S. Top-down versus Bottom-up forecasting strategies. **International Journal of Production Research**, v. 26, n. 11, p. 1833-1843, Nov 1988.

SOHN, S. Y.; LIM, M. Hierarchical forecasting based on AR-GARCH model in a coherent structure. **European Journal of Operational Research**, v. 176, n. 2, p. 1033-1040, Jan 2007.

SOUZA, R. C.; CAMARGO, M. E. **Análise e Previsão de Séries Temporais: os Modelos ARIMA**. 2 ed. Rio de Janeiro, 2004.

STRIJBOSCH, L. W. G.; HEUTS, R. M. J.; MOORS, J. J. A. Hierarchical estimation as a basis for hierarchical forecasting. **Ima Journal of Management Mathematics**, v. 19, n. 2, p. 193-205, Apr 2008.

THOMÉ, A. M. T. et al. Sales and operations planning: A research synthesis. **International Journal of Production Economics**, v. 138 p. 1–13, 2012.

VISWANATHAN, S.; WIDIARTA, H.; PIPLANI, R. Forecasting aggregate time series with intermittent subaggregate components: top-down versus bottom-up forecasting. **Ima Journal of Management Mathematics**, v. 19, n. 3, p. 275-287, Jul 2008.

WANKE, P.; SALIBY, E. Top-down or bottom-up forecasting? **Pesquisa Operacional**, v. 27, n. 3, p. 591-605, 2007.

WIDIARTA, H.; VISWANATHAN, S.; PIPLANI, R. On the effectiveness of top-down strategy for forecasting autoregressive demands. **Naval Research Logistics**, v. 54, n. 2, p. 176-188, Mar 2007.

\_\_\_\_\_. Forecasting item-level demands: an analytical evaluation of top-down versus bottom-up forecasting in a production-planning framework. **Ima Journal of Management Mathematics**, v. 19, n. 2, p. 207-218, Apr 2008.

\_\_\_\_\_. Forecasting aggregate demand: An analytical evaluation of top-

down versus bottom-up forecasting in a production planning framework. **International Journal of Production Economics**, v. 118, n. 1, p. 87-94, 2009.

WERON, R. Electricity price forecasting: A review of the state-of-the-art with a look into the future. **International Journal of Forecasting**, v. 30, p. 1030–1081, 2014.

WILLIAMS, B. D.; WALLER, M. A. Top-Down Versus Bottom-Up Demand Forecasts: The Value of Shared Point-of-Sale Data in the Retail Supply Chain. **Journal of Business Logistics**, v. 32, n. 1, p. 17-26, 2011.

## Apêndice A

Devido ao tamanho dos arquivos, os códigos utilizados para o desenvolvimento desta dissertação poderão ser acessados pelos links abaixo.

- Séries de energia:

<https://www.dropbox.com/s/eswl8xwp2crw5n8/C%C3%B3digo%20utilizado%20para%20previs%C3%A3o%20das%20s%C3%A9ries%20de%20energia.docx?dl=0&preview=C%C3%B3digo+utilizado+para+previs%C3%A3o+das+s%C3%A9ries+de+energia.docx>

- Séries RPK:

<https://www.dropbox.com/s/vn5yzo1xal09qty/C%C3%B3digo%20utilizado%20para%20previs%C3%A3o%20das%20s%C3%A9ries%20RPK.docx?dl=0&preview=C%C3%B3digo+utilizado+para+previs%C3%A3o+das+s%C3%A9ries+RPK.docx>

## Apêndice B

**Tabela 11 – Estatísticas descritivas das 20 séries de energia do nível mais baixo da hierarquia**

(continua)

	<b>COcom</b>	<b>COres</b>	<b>COind</b>	<b>COout</b>	<b>NEcom</b>	<b>NEres</b>	<b>NEind</b>	<b>NEout</b>	<b>NOcom</b>	<b>NOres</b>	<b>NOind</b>	<b>NOout</b>
Média	435,8	656,8	566,9	475,4	817,7	1499,8	2279,4	966,6	279,1	476,5	1067,9	276,1
Mediana	426	626	529	473	780	1414	2314	939	264	443	1062	277
Desvio-padrão	108,9	142,5	133,5	92,7	197	392,6	189,1	180,6	73,7	131,5	132,8	59,9
Coefficiente de variação	0,3	0,2	0,2	0,2	0,2	0,3	0,1	0,2	0,3	0,3	0,1	0,2
Assimetria	0,3	0,5	0,4	0,2	0,4	0,4	-0,8	0,2	0,4	0,7	-0,3	0,2
Curtose	1,8	2,1	1,8	2,1	1,9	1,8	2,9	1,8	1,9	2,5	2,2	1,9
Mínimo	267	450	319	302	516	943	1769	645	170	305	737	177
Máximo	663	1006	838	683	1200	2299	2575	1327	443	820	1289	403

**Tabela 11 – Estatísticas descritivas das 20 séries de energia do nível mais baixo da hierarquia**

(conclusão)

	<b>SEcom</b>	<b>SEres</b>	<b>SEind</b>	<b>SEout</b>	<b>SUcom</b>	<b>SUres</b>	<b>SUind</b>	<b>SUout</b>
Média	3103,8	4495,1	7744,1	2190,2	946,9	1376,1	2380	1009,4
Mediana	2975	4466	8004	2149	915	1350	2468	979
Desvio-padrão	618	717,4	894,3	286,8	221,1	239,8	290,9	199
Coeficiente de variação	0,2	0,2	0,1	0,1	0,2	0,2	0,1	0,2
Assimetria	0,4	0,2	-0,9	0,2	0,4	0,6	-0,6	0,7
Curtose	2,2	2,1	2,9	1,9	2,3	3,0	2,4	2,7
Mínimo	2056	3295	5569	1686	607	1034	1648	714
Máximo	4572	6451	9000	2840	1553	2166	2832	1541

**Fonte:** a autora (2016)

**Tabela 12– Previsões 12 passos-à-frente para a Hierarquia 1**

(continua)

<b>Data</b>	<b>Total</b>	<b>CO</b>	<b>NE</b>	<b>NO</b>	<b>SE</b>	<b>SU</b>	<b>COcom</b>	<b>COres</b>	<b>COind</b>	<b>COout</b>	<b>NEcom</b>	<b>NEres</b>	<b>NEind</b>	<b>NEout</b>	<b>NOcom</b>	<b>NOres</b>	<b>NOind</b>	<b>NOout</b>	<b>SEcom</b>	<b>SERes</b>
out/15	38223	3008	6563	2948	19042	6662	636	970	751	651	1195	2112	1957	1299	445	819	1284	400	3955	5241
nov/15	38593	2944	6527	2859	19307	6957	634	950	744	616	1207	2121	1914	1285	427	803	1247	382	4110	5370
dez/15	38333	2778	6467	2840	19204	7045	602	881	732	563	1206	2117	1881	1262	426	794	1244	377	4091	5302
jan/16	38468	2748	6497	2729	19300	7194	583	906	726	533	1191	2199	1856	1252	388	772	1219	350	4137	5834
fev/16	38429	2746	6263	2589	19221	7609	592	863	737	554	1182	2123	1745	1213	373	740	1136	340	4199	5595
mar/16	38083	2694	6424	2682	19030	7252	591	846	750	508	1191	2125	1897	1211	382	738	1215	348	4076	5445
abr/16	37989	2780	6339	2661	19118	7091	597	857	756	570	1192	2118	1823	1205	383	741	1183	354	4112	5376
mai/16	37400	2751	6299	2699	18882	6769	575	849	767	560	1176	2118	1801	1204	388	739	1216	355	3890	5311
jun/16	36925	2691	6104	2677	18799	6655	549	805	767	570	1140	2033	1736	1195	392	745	1180	359	3778	5187
jul/16	36781	2656	6074	2763	18703	6584	533	770	779	574	1105	1997	1797	1175	402	774	1221	366	3708	5191
ago/16	37175	2748	6166	2749	18757	6754	555	806	789	599	1126	1971	1842	1227	406	770	1209	364	3766	5235
set/16	37180	2841	6166	2805	18738	6628	576	857	794	614	1149	2006	1774	1237	423	800	1205	378	3778	5202

**Tabela 12 – Previsões 12 passos-à-frente para a Hierarquia 1**

(conclusão)

<b>Data</b>	<b>SEind</b>	<b>SEout</b>	<b>SUcom</b>	<b>SUres</b>	<b>SUind</b>	<b>SUout</b>
out/15	7327	2519	1224	1600	2586	1251
nov/15	7355	2472	1346	1694	2614	1303
dez/15	7399	2411	1407	1734	2497	1406
jan/16	6988	2340	1461	1982	2205	1546
fev/16	7061	2366	1550	1967	2533	1560
mar/16	7131	2379	1445	1812	2489	1506
abr/16	7237	2394	1398	1709	2578	1406
mai/16	7278	2404	1283	1651	2500	1335
jun/16	7391	2442	1227	1664	2495	1268
jul/16	7379	2426	1194	1686	2444	1259
ago/16	7304	2452	1250	1710	2523	1270
set/16	7264	2495	1238	1652	2474	1264

**Fonte:** a autora (2016)

**Tabela 13 – Previsões 12 passos-à-frente para a Hierarquia 2**

(continua)

<b>Data</b>	<b>Total</b>	<b>Com</b>	<b>Res</b>	<b>ind</b>	<b>out</b>	<b>comCO</b>	<b>comNE</b>	<b>comNO</b>	<b>comSE</b>	<b>comSU</b>	<b>resCO</b>	<b>resNE</b>	<b>resNO</b>	<b>resSE</b>	<b>resSU</b>	<b>indCO</b>	<b>indNE</b>	<b>indNO</b>
out/15	38223	7468	10680	13887	6189	636	1192	442	3991	1207	964	2091	808	5251	1567	749	1946	1272
nov/15	38593	7745	10791	13925	6132	641	1196	426	4161	1321	944	2070	788	5352	1637	752	1900	1245
dez/15	38333	7736	10661	13871	6065	620	1196	432	4121	1367	892	2065	792	5254	1658	760	1881	1271
jan/16	38468	7785	11497	13121	6066	600	1172	391	4225	1397	914	2121	763	5842	1858	751	1836	1235
fev/16	38429	7923	11043	13342	6121	610	1166	380	4281	1486	867	2042	735	5560	1839	764	1731	1164
mar/16	38083	7712	10763	13560	6049	616	1178	393	4139	1386	863	2055	741	5405	1699	783	1880	1251
abr/16	37989	7713	10590	13638	6047	631	1189	395	4147	1352	884	2061	745	5289	1612	800	1820	1220
mai/16	37400	7332	10442	13592	6034	611	1182	403	3892	1244	880	2074	748	5179	1560	816	1810	1262
jun/16	36925	7083	10181	13626	6036	593	1150	411	3733	1195	848	1999	761	4994	1580	833	1761	1244
jul/16	36781	6955	10192	13599	6034	578	1120	419	3672	1166	814	1974	787	5012	1605	843	1817	1270
ago/16	37175	7135	10260	13644	6136	603	1138	429	3741	1224	851	1937	791	5055	1626	852	1851	1272
set/16	37180	7213	10314	13498	6155	627	1165	444	3769	1209	907	1976	817	5044	1569	859	1786	1257

**Tabela 13 – Previsões 12 passos-à-frente para a Hierarquia 2**

(conclusão)

<b>Data</b>	<b>indSE</b>	<b>indSU</b>	<b>outCO</b>	<b>outNE</b>	<b>outNO</b>	<b>outSE</b>	<b>outSU</b>
out/15	7375	2544	659	1310	402	2570	1248
nov/15	7459	2570	630	1289	385	2534	1294
dez/15	7512	2446	585	1264	386	2451	1380
jan/16	7178	2121	554	1245	356	2416	1494
fev/16	7239	2443	580	1216	353	2452	1520
mar/16	7253	2391	539	1219	364	2457	1470
abr/16	7304	2494	613	1224	371	2457	1383
mai/16	7281	2424	612	1243	379	2471	1330
jun/16	7343	2444	636	1245	389	2491	1275
jul/16	7290	2380	645	1234	395	2487	1272
ago/16	7215	2455	671	1278	397	2510	1281
set/16	7196	2400	680	1277	404	2535	1258

**Fonte:** a autora (2016)

**Tabela 14 – Estatísticas Descritivas das 2 séries de RPK do nível mais baixo da hierarquia**

	<b>RPKdom</b>	<b>RPKint</b>
Média	4527304	1901038
Mediana	3870494	1876194
Desvio-padrão	2179091	414707
Coefficiente de variação	0,48	0,22
Assimetria	0,43	0,05
Curtose	1,73	3,40
Mínimo	1932886	809423
Máximo	9579996	3133260

**Fonte:** a autora (2016)

**Tabela 15– Previsões 12 passos-à-frente para a série RPK**

<b>Data</b>	<b>Total</b>	<b>RPKdom</b>	<b>RPKint</b>
out/15	10.847.168	7.962.777	2.884.391
nov/15	10.586.075	7.852.211	2.733.863
dez/15	11.615.748	8.677.471	2.938.277
jan/16	12.528.116	9.324.847	3.203.269
fev/16	10.160.748	7.369.828	2.790.920
mar/16	10.586.965	7.696.522	2.890.443
abr/16	10.516.825	7.672.283	2.844.542
mai/16	10.522.417	7.605.905	2.916.512
jun/16	10.376.125	7.515.200	2.860.925
jul/16	12.123.049	8.903.891	3.219.158
ago/16	11.041.371	7.948.633	3.092.738
set/16	10.717.692	7.707.949	3.009.742

**Fonte:** a autora (2016)